

TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**UZUNLAMASINA ÇALIŞMALARIN
ANALİZİNDE KARMA ETKİ MODELLERİ**

Beyza DOĞANAY

BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ

DANIŞMAN
Yrd.Doç.Dr. S. Kenan KÖSE

2007-ANKARA

İÇİNDEKİLER

Kabul ve Onay.....	ii
İçindekiler	iii
Önsöz	v
Simgeler ve Kısaltmalar	vi
Şekiller	vii
Çizelgeler	viii
1.GİRİŞ	1
1.1. Karma Etki Modellerinin Tarihçesi	3
1.2. Tekrarlı Ölçümlerin Analizinde Yaygın Olarak Kullanılan Diğer Yöntemler	4
1.2.1. Tekrarlı Ölçümlerde Tek Değişkenli Varyans Analizi	7
1.2.2. Tekrarlı Ölçümlerde Çok Değişkenli Varyans Analizi.....	9
1.3. Doğrusal Karma Etki Modelleri.....	11
1.3.1. Rasgele ve Sabit Etkiler	11
1.3.2. Karma Etki Modellerinin Avantajları	12
1.3.2.1. Tamamlanmamış Veri Yapısı	13
1.3.3. Genel Doğrusal Karma Etki Modeli	14
1.3.4. Basit Doğrusal Regresyon Modeli	15
1.3.5. Rasgele Kesim Noktası Modeli.....	16
1.3.5.1. Tam Simetri ve Sınıf İçi Korelasyon	20
1.3.6. Rasgele Kesim Noktası ve Rasgele Eğim Modeli.....	21
1.3.7. Matris Gösterimi	24
1.3.7.1. İki Aşamalı Modelin Matris Gösterimi	29
1.3.7.1.1. İlk Aşama Modeli.....	29
1.3.7.1.2. İkinci Aşama Modeli.....	31
1.3.8. Rasgele Etkiler Kovaryans Yapısının Tekrarlı Ölçümler Arası Kovaryansa Katkısı	33
1.4. Kestirim Yöntemleri.....	35
1.4.1. En Çok Olabilirlik Kestirimi (EÇOK)	36
1.4.1.1. Bağımsız Ölçümler İçin EÇOK	37

1.4.1.2. İlişkili Ölçümler İçin EÇOK	39
1.4.2. Kayıp Gözlem Durumları	42
1.4.3. Sınırlı En Çok Olabilirlik Kestirimi (SEÇOK)	44
1.5. Rasgele Etkiler İçin Önkestirim	48
1.5.1. Koşullu Beklenen Değer	49
1.5.1.1. Koşullu Beklenen Değer ve Normal Dağılım	50
1.5.2. En İyi Doğrusal Yansız Önkestirici (EİDYÖK)	51
1.5.2.1. Bireysel Ortalamanın Kestirilmesi	53
1.5.2.2. Ağırlıklı Ortalama Olarak EİDYÖK	55
2. GEREÇ VE YÖNTEM	58
2.1. Uygulama Verileri	58
2.2. Uygulama Yöntemleri	59
2.2.1. PROC MIXED İşlemi	60
2.2.2. PROC GLM İşlemi	63
2.2.3. Doğrusal Karma Etki Modelleri ve Tekrarlı Ölçümlerde Varyans Analizi Algoritmaları	64
3. BULGULAR	67
3.1. Karma Etki Modellerine Ait Sonuçlar ve Model Seçimi	69
3.1.1. Karma Modellerine İlişkin Uyum İstatistikleri ve Sabit Etkilerin Testleri	69
3.1.2. Model Seçimi	72
3.1.3. Model V'e İlişkin Sonuçlar	72
3.2. Tekrarlı Ölçümlerde Varyans Analizi Sonuçları	80
4. TARTIŞMA	83
5. SONUÇ VE ÖNERİLER	86
ÖZET	88
SUMMARY	89
KAYNAKLAR	90
ÖZGEÇMİŞ	93

ÖNSÖZ

Tezim süresince benden ilgisini ve desteğini esirgemeyen değerli danışman hocam Yrd.Doç.Dr. S. Kenan Köse'ye, çalışmanın gerçekleşmesi için gerekli ortamı hazırlayan A.Ü.Biyostatistik Bölümü Başkanı değerli hocam sayın Prof.Dr. Ersöz Tüccar'a,

Bana vakit ayıran ve hiçbir zaman yardımlarını esirgemeyen değerli hocam Doç.Dr. Atilla Halil Elhan'a, ilgisinden dolayı değerli hocam Yrd.Doç.Dr. Yasemin Genç'e, çalışmanın uygulama aşamasında verdiği desteği ve sabrından dolayı sevgili Eren Demirhan'a, ilgi ve anlayışlarından dolayı tüm Biyoistatistik Anabilim Dalı çalışma arkadaşlarıma,

Tez için uygulama verilerini sağlayan Ankara Numune Eğitim ve Araştırma Hastanesi, Fizik Tedavi ve Rehabilitasyon Kliniği'nden Uzm.Dr. Dilek Keskin'e,

Her zaman beni destekleyen, gösterdikleri sabır ve sevgiden dolayı AİLEM'e, tez çalışmamın her anında bana destek olan Emre'ye içtenlikle teşekkür ederim.

SİMGELER ve KISALTMALAR

AIC	: Akaike bilgi kriteri
ANOVA	: Tek deęişkenli varyans analizi
BIC	: Schwartz Bayesçi bilgi kriteri
EÇOK	: En çok olabilirlik kestirimi
EİDYÖK	: En iyi doğrusal yansız ön kestirici
EKNKYK	: En küçük normlu karesel yansız kestirici
GEKK	: Genelleştirilmiş en küçük kareler
MANOVA	: Çok deęişkenli varyans analizi
SEÇOK	: Sınırlı en çok olabilirlik kestirimi

ŞEKİLLER

Şekil 1.1. Rasgele kesim noktası modelinde herhangi iki bireye ve popülasyona ait zaman içindeki yanıt eğilimleri.....	19
Şekil 1.2. Rasgele ölçüm hataları eklendikten sonra zaman içindeki marjinal ve koşullu ortalama yanıt profilleri	20
Şekil 1.3. Ölçüm hataları eklendikten sonra zaman içindeki marjinal ve koşullu yanıtlara ait profiller	23
Şekil 3.1. Hastanede tedavi alanlarda fleksiyon değerlerinin zaman içindeki değişimi	68
Şekil 3.2. Ev programı alanlarda fleksiyon değerlerinin zaman içindeki değişimi	68

ÇİZELGELER

Çizelge 2.1. SAS programında karma etki modelleri analizi için veri girişi	59
Çizelge 2.2. SAS programında tekrarlı ölçümlerde varyans analizi için veri girişi.....	60
Çizelge 2.3. Proc mixed işlemi	61
Çizelge 2.4. Proc glm işlemi	63
Çizelge 2.5. SAS programında doğrusal karma etki algoritmaları	65
Çizelge 2.6. SAS programında tekrarlı ölçümlerde varyans analizi algoritması	66
Çizelge 3.1. Gruplara göre tanımlayıcı istatistikler	67
Çizelge 3.2. Model I: zaman sürekli, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.....	69
Çizelge 3.3. Model II: zaman sürekli, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.....	70
Çizelge 3.4. Model III: zaman sürekli, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.....	70
Çizelge 3.5. Model IV: zaman kategorik, rasgele kesim noktası modeli için sonuçlar.....	71
Çizelge 3.6. Model V: zaman hem sürekli hem de kategorik, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.....	71
Çizelge 3.7. Model V: zaman hem sürekli hem de kategorik, rasgele kesim noktası ve rasgele eğim modeli için algoritma.....	72
Çizelge 3.8. Kovaryans parametre kestirimleri.....	73
Çizelge 3.9. Sabit etkiler için SEÇOK kestirimeleri.....	74
Çizelge 3.10. Sabit etkilerin testi	75
Çizelge 3.11. Marjinal beklenen değerlere ilişkin sonuçlar	75
Çizelge 3.12. Marjinal beklenen değerlerin farklarına ilişkin sonuçlar	76
Çizelge 3.13. Her birey için rasgele kesim noktası ve eğim önkestirimleri.....	77
Çizelge 3.14. Her bireyin her zaman noktasındaki beklenen fleksiyon yanıt önkestirimleri	79
Çizelge 3.15. Küresellik testi	80
Çizelge 3.16. Zaman etkisi için çok değişkenli varyans analizi sonuçları.....	80

Çizelge 3.17. Zaman*grup etkileşimi için çok değişkenli varyans analizi sonuçları	80
Çizelge 3.18. Grup etkisi için varyans analizi sonuçları	81
Çizelge 3.19. Her zaman noktasında gruplar arası farkın testi	81
Çizelge 3.20. Gruplar için zamanlara ait çoklu karşılaştırmalar	82

1. GİRİŞ

Uzunlamasına veriler, deney biriminden gözlemlerin farklı zaman noktalarında, tekrarlı ölçümler biçiminde toplanması ile ortaya çıkar. Özellikle sağlık alanında yapılan çalışmalarda toplanan veriler, tekrarlı ölçümler biçimindedir. Bu tür çalışmaların amacı genellikle, ortalama yanıt profillerinin denemeler arasında nasıl farklılık gösterdiğinin ve ayrıca yanıtların zaman içerisindeki seyirlerinin incelenmesidir. Bu soruların yanıtları çeşitli istatistiksel modeller yardımı ile bulunabilir. Uzunlamasına verilerin en iyi şekilde betimlenmesi ve yorumlanmasında, son yıllarda yapılan çalışmalar oldukça artmış ve hız kazanmıştır. Karma etki modelleri uzunlamasına veri analizinde güçlü araçlardır (Wu ve Zhang, 2006). Karma etki modelleri ayrıca, rasgele etki modelleri, rasgele katsayılar modelleri, hiyerarşik modeller ve çok düzeyli modeller gibi başka adlarla da anılır (Armitage ve Colton, 1998).

Uzunlamasına çalışmaların en iyi biçimde çözümlenmesi ve yorumlanması ile ilgili çalışmalar geçtiğimiz yıllarda büyük bir ilerleme göstermiş ve bu anlamda güçlü birçok yöntem geliştirilmiştir. Ülkemizde yaygın olarak kullanılmamasına karşın, karma etki modelleri, son yıllarda üzerinde durulan ve ilgili yazılımların da gelişmesiyle kullanımı günden güne artan modelleme yöntemlerindedir.

“Uzunlamasına veriler” terimi daha çok, zaman içerisinde toplanan verileri vurgulasa da, bu tezin kapsamında anlatılacak yöntem ve modeller, genelde her çeşit tekrarlı ölçüm verilerinin çözümünü kapsamaktadır. Tekrarlı ölçümler; bir deney biriminden zaman içinde birden fazla ölçüm alınması biçiminde ortaya çıkıyor olsa da, yine bir deney biriminden farklı koşullar altında toplanan veriler şeklinde de olabilir. Örneğin; her bireyin farklı ilaç dozlarını aldığı bir çalışmada, yine her birey için aldıkları ilaç dozlarında, diyastolik kan basıncındaki azalma değerleri kaydedilebilir. Bu durumda tekrarlı ölçümler ilaç dozları üzerinden alınmış olur. Ya da, doğurmak üzere olan farelerin deney birimleri olduğu bir çalışmada, her farenin bir batında doğurduğu yavrularının doğum ağırlıkları kaydedilir. Bu durumda fareler için alınan tekrarlı ölçümler doğurduğu yavrular üzerinden olacaktır. İkinci örnek

için, tekrarlı gözlemlerin herhangi bir sıra takip etmediği burada vurgulanmalıdır. Ancak yapılan çoğu çalışmada, tekrarlı ölçümlerin alındığı koşul, genellikle zaman noktalarıdır.

Bu çalışmada, doğrusal karma etki modellerinin uzunlamasına verilerin analizinde kullanımını incelenecek ve sağlık alanındaki bir veri seti üzerinde uygulamasını yapılacaktır. Ayrıca aynı veri seti, tekrarlı ölçümlerin analizinde yaygın olarak kullanılan tekrarlı ölçümlerde varyans analizi ile çözümlenecek ve her iki yöntemle bulunan sonuçlar birbirleriyle karşılaştırılacaktır. Tezde, toplanan ölçümlerden “yanıt” olarak bahsedilecektir. Verilen örneklerde, toplanan yanıt değişkeni süreklidir. Ancak, yanıt değişkeninin kesikli ya da ikili (binary) olduğu durumlar da olabilir. Örneğin, epilepsi hastaları üzerinde yapılan bir çalışmada uygulanan iki farklı tedavinin hastalardaki epilepsi nöbeti sayılarına etkisi incelenebilir. Nöbet sayıları farklı zaman noktalarında kaydedilir. Bu durumda hastalardan alınan yanıtlar nöbet sayıları, yani kesikli değerler olacaktır. Ya da, sigara kullanan annelerin çocuklarında nefes alma güçlüğüne incelendiği bir çalışmada, çocuklardan alınan yanıtlar nefes alma güçlüğü olup olmadığı biçiminde olabilir. Bu durumda yanıt değişkeni ikili sonuç veren bir değişkendir (var/yok gibi). Bu gibi durumlar için geliştirilmiş yöntemler de mevcuttur. Ancak bu tezde ilgilenilen yanıt değişkeninin sürekli ve uzunlamasına toplanmış olduğu varsayılacaktır. Yanıt değişkeninin kesikli ve ikili sonuç olduğu durumlar tez kapsamı dışında tutulacaktır.

Çalışmanın ilk bölümünde, tekrarlı ölçümlerin analizinde yaygın olarak kullanılan yöntemlere ve kısıtlılıklarına değinilmiş, doğrusal karma etki modellerinin avantajlarından bahsedilmiş, rasgele kesim noktası modeli ile rasgele kesim noktası ve rasgele eğim modeli gibi karma etki modelleri tanıtılmıştır. Son olarak kestirim yöntemleri hakkında bilgi verilmiştir. İkinci bölümde ise, meme kanseri nedeniyle radikal mastektomi ameliyatı olan ve lenf bezleri alınan kadınlardan tekrarlı ölçümler biçiminde toplanan veri seti üzerinde karma etki modellerinin ve klasik tekrarlı ölçümlerde varyans analizinin SAS paket programında uygulaması, PROC MIXED ve PROC GLM işlemleri hakkında bilgiler yer almaktadır. Çalışmanın üçüncü

bölümünde, analiz sonuçları üzerinde durulmuştur. Son olarak tartışma ve sonuç bölümlerinde, bulgular üzerinden elde edilen sonuçlar yorumlanmıştır.

1.1. Karma Etki Modellerinin Tarihçesi

Karma etki modelleri tıp, tarım, biyoloji, ekonomi ve jeofizik gibi birçok alanda karşılaşılan tekrarlı ölçümlerin analizinde çok kullanışlı bir araç olmuştur. Bu modellerin artan önemliliği, tekrarlı gözlemlerde ortaya çıkan denek-içi korelasyon yapısını modellemedeki esnekliği, hem dengeli hem de dengesiz veri yapılarında kullanılabilirliği ve artık günümüzde uygun yazılımların da bulunması ile kullanım rahatlığından kaynaklanmaktadır. Karma etki modelleri ayrıca, rasgele etki modelleri, hiyerarşik modeller, çok düzeyli modeller gibi isimlerle de bilinmektedir. Rasgele etki modelleri ile ilgili literatürde görülen ilk uygulamalar 19. yüzyılın ortalarında astronomi alanında yapılmıştır (Searle ve ark., 1992). Doğrusal modellerin ve özellikle karma etki modellerinin gelişimi genellikle genetik alanında yapılan uygulamalarla bağlantılıdır. Yirminci yüzyılın başlarında genetik ve karma etki modelleri paralel olarak gelişmeye başlamıştır. K. Pearson ve F. Galton, seleksiyon indeksi teorisi ve kalıtım teorisi üzerine yaptıkları çalışmalarda, regresyon ve korelasyon analizinin gelişmesinde katkıda bulunmuşlardır. 1918 yılında R. A. Fisher varyans analizinin temellerini çalışmış ve aynı çalışmada, G. Mendel ve F. Galton tarafından yapılan genetik çalışmalarını varyans analizine dayandırmıştır. Birkaç yıl sonra da varyans bileşenleri kestirimi için ilk yöntemi elde etmiştir. Hayvancılık ve genetik üzerinde çalışan C.R. Henderson ve S.R. Searle 20. yüzyılın ikinci yarısında karma etki modellerinin teorisinin gelişiminde önemli katkıda bulunmuşlardır (Kaart, 2005).

Fisher tarafından çalışılan varyans analizi yöntemi, ortalamadan ayrılış kareler toplamının bileşenlere ayrılması olarak bilinmektedir. Ancak varyans analizi dengeli olmayan verilerde iyi sonuçlar vermediği için son yıllarda varyans analizi için doğrusal modellerin teorisine dayanan yöntemler ortaya çıkmıştır. Khattree 1999 yılında yayınladığı makalesinde, Henderson'ın 1953 yılında dengeli olmayan

verilerde ANOVA yöntemi ile varyans bileşenlerinin kestirilmesinde kullanılmak üzere bulunduğu üç farklı yöntem üzerinde birtakım değişiklikler yapmıştır. Henderson tarafından önerilen yöntemler kestirimler tanım kümesi dışında bulunabilmekte iken (varyansın negatif kestirilebilmesi), Khattree yaptığı değişikliklerle negatif varyans kestirimi yapmayan yöntemler bulmuştur (Khattree, 1999). Dengeli olmayan veriler için geliştirilen kestirim yöntemleri arasında Hartley ve Rao (1967) tarafından geliştirilen en çok olabilirlik (maximum likelihood, ML) ile, 1971'de Patterson ve Thompson (1971) tarafından geliştirilen sınırlı en çok olabilirlik (restricted maximum likelihood, REML) yöntemleri yer almaktadır. Karma etki modellerinde rasgele etkilerin kestiriminde kullanılan en iyi doğrusal yansız önkestirici (Best Linear Unbiased Predictor, BLUP), 1950 yılında Henderson tarafından elde edilmiştir (Robinson, 1991). Bu tezin kapsamında en çok olabilirlik ile sınırlı en çok olabilirlik yöntemlerine değinilecektir.

Sürekli yanıt değişkenleri için uzunlamasına verilerin çözümlenmesinde karma etki modellerinin kullanımı ve bu modellerin iki aşamalı model biçiminde gösterimi ilk defa Laird ve Ware (1982) tarafından yapılmıştır.

1.2. Tekrarlı Ölçümlerin Analizinde Yaygın Olarak Kullanılan Diğer Yöntemler

Uzunlamasına çalışmaların analizinde yaygın olarak kullanılan iki klasik yaklaşım vardır. Bunlardan ilki, çeşitli adlarla bilinen; tek değişkenli karışık model, bölünmüş parseller (split-plot) ya da tekrarlı ölçümlerde ANOVA, ikincisi çok değişkenli ANOVA (MANOVA)'yı temel alır. Her iki modelde de, gruplar arasında homojen olan hataların normal dağıldığı varsayılır. Bazı durumlarda normallik ve varyans homojenliği verilere dönüşüm uygulanarak (örn, doğal logaritmik dönüşüm) sağlanabilir. Her iki model için de, öncelikli amaç, grup ortalamalarının karşılaştırılmasıdır ve bu modellerden hiçbiri bireysel değişim eğrileri (örn, bireye-özel eğilimler gibi) hakkında bilgilendirici değildirler.

Bunun yanı sıra, tekrarlı ölçümlerin alındığı zaman noktaları her birey için aynıdır ve her bireyden aynı sayıda tekrarlı ölçüm alınması gereklidir. Böyle bir veri yapısı dengelidir. Bu durum, farklı bireylerden farklı koşullar altında ölçüm alındığı dengesiz tasarımların analizini engellemiş olur. Özellikle deney birimlerinin insan olduğu durumlarda, çalışma çok özenli tasarlanmış olsa bile, birimlerden alınan gözlem sayılarının aynı olması mümkün olmayabilir. Deney süresince, deneyden ayrılan denekler olabilir. Ya da örneğin, kan örnekleri toplanıyorsa, laboratuarda bazı örnekler düşüp kırılabilir. Buna benzer nedenlerden dolayı veri setinde kayıp gözlemler ortaya çıkabilir. Uzunlamasına çalışmalarda, kayıp gözlemlerle karşılaşma olasılığı çok yüksektir. Klasik yöntemlerde, bazı kayıp veri yapıları için örneğin, tek değişkenli tekrarlı ölçümler çözümlemesinde F testinde düzeltmeye gidilmesi gibi çözüm yolları tartışılırken, bu gibi çözüm yolları sorunun ancak yüzeysel olarak çözümlenmesini sağlayabilir. Bunun dışında, MANOVA modeli herhangi kayıp gözlem verilerini modelleyemez (Davidian, 2007). MANOVA modeli için bütün deneklere ait gözlemlerin tam olması gerekliliği büyük bir kısıtlayıcıdır. MANOVA uygulamasında tamamlanmamış veriye sahip olan denekler çalışmadan çıkartılmalıdır. Böyle bir durumda, çalışmada randomizasyon sırasında alınan tüm deneklerle çalışılmayacağından, önemli ölçüde yanlılık oluşacaktır (Hedeker ve Gibbons, 2006).

ANOVA modeli, yanıt değişkenine ait varyans kovaryans matrisinin tam simetrik (zaman içinde eşit varyans ve kovaryanslara sahip) olduğunu varsayar. Ancak bu yapı, bireylerden alınan gözlemler arası ilişki yapısını yeterli derecede açıklayamaz. Böyle bir yapıya göre aynı denekten alınan gözlemler birbirine ne kadar yakın ya da uzak olursa olsun, ilişkinin hep aynı olacağı düşünülür. Bu nedenle, tek değişkenli tekrarlı ölçümlerde varyans analizi çözümlemesindeki tam simetrik kovaryans yapısı varsayımı uzunlamasına veri yapısı için çok sınırlayıcı olabilmektedir.

MANOVA modelinde, simetrik olması dışında varyans kovaryans matrisine ilişkin herhangi bir varsayım yoktur. Uzunlamasına veriler için, denekler arası rasgele değişim (biyolojik) ve denek içi gözlemlerin değişimi (ölçüm hatası, zaman

içindeki ilişki gibi) olmak üzere iki değişim kaynağı olduğu düşünülür. Oysa burada varsayılan kovaryans matrisi ile bu değişimler tam olarak göz önünde bulundurulamaz. Kovaryans matrisi, bu iki değişim kaynağını içermeyen herhangi bir yapıda olabilir. Bu nedenle çok değişkenli yöntemler de uzunlamasına verilerin çözümlenmesinde yetersiz ve yüzeysel kalabilmektedir (Davidian, 2007).

Bunların yanı sıra, hem tek değişkenli hem de çok değişkenli tekrarlı ölçümlerde varyans analizi yöntemlerinde, tekrarlı ölçümlere ait varyans kovaryans matrisinin bütün gruplar için aynı olduğu varsayılır. Ancak bu varsayım pratikte çok da doğru değildir. Örneğin; yüksek sistolik kan basıncına sahip bireylerden zaman içinde alınan birey içi gözlemler, düşük sistolik kan basıncına sahip bireylerin birey içi gözlemlerine göre daha fazla değişim gösterme eğilimindedirler. Yani, denek içi hata varyansı, yüksek sistolik kan basıncına sahip grupta daha fazla olacaktır. Bu da, tekrarlı gözlemlerin varyansına yansıtacaktır. Büyük yanıt değerleri büyük varyansa, küçük yanıt değerleri küçük varyansa sahip olacaklardır. Böyle bir durumda farklı gruplardaki gözlemlerin aynı varyans kovaryans matrisine sahip olduğunu varsaymak yanlış olacaktır. Yüksek kan basınçlı gruba ait bir gözlem için $\text{var}(y_i) = \Sigma_0$ ve düşük kan basınçlı gruba ait bir gözlem için $\text{var}(y_i) = \Sigma_1$ gibi iki farklı kovaryans matrisi olduğunu varsaymak daha doğru olacaktır. Klasik yöntemler üzerinde bu gibi bir sorunla başa çıkabilmek için değişiklik yapmak mümkün olabilir. Genel bir yaklaşım, varyansların homojenliğini sağlamak amacıyla veriler üzerinde dönüşüm yapmaktır. Örneğin logaritmik dönüşüm yapılmış veriler modellenabilir. Ancak, analiz sonucunda çıkarsamaların gerçek ölçek üzerinden yorumlanması beklendiğinden, dönüşüm yapılmış veri ile yapılan analiz sonuçlarını yorumlamak güç olabilir. Başka bir yol da, her bir grup için farklı bir kovaryans matrisi bulunmasını göz önünde tutarak test istatistiği üzerinde düzeltme yapmak olabilir. Ancak bu durumda da istatistiksel güç azalacaktır (Davidian, 2007).

Klasik yöntemlere dayalı analiz, hipotezlerin test edilmesine odaklıdır. Analiz sonucunda hipotez reddine ya da reddedilmemesine ilişkin yorumlar yapılır. Ancak bazı durumlarda araştırmacının farklı amaçları da olabilir. Araştırmacı, sonuç değişkenine ilişkin zaman içerisindeki ortalama değişimin, gruplar arasında nasıl

değiştirdiğinden daha fazlasıyla ilgilenebilir. Araştırmadan elde edilecek bilgiler ışığında, daha sonraki hastalara nasıl müdahale edilmesi gerektiği tavsiye edilmek istenebilir. Yani daha özelleşmiş sonuçlara ulaşılmak istenebilir. Örneğin farklı ilaçlar denenerek kolesterol seviyesindeki düşüşün araştırıldığı bir çalışmada, her ilacın kolesterol seviyesini hangi oranda düşürdüğü de incelenmek istenebilir. Örneğin, ilk ilaç kolesterol seviyesini her ay 5 birim düşürüyor ve ikinci ilaç da 15 birim düşürüyor ise bu bilgiye dayanarak belirli bir hasta için hangi ilacın kullanılması gerektiği saptanmak istenebilir. Böyle bir durumda, araştırmacı her grup için ortalama yanıtta zaman içerisindeki değişim oranını (ortalama yanıt profilini) kestirmekle ilgilenebilir. Ya da örneğin, belli bir ilacı alan 45 yaşında bir erkek hastanın kolesterol değerleri profili nasıl olacaktır sorusuna yanıt arayabilir. Bu durumda ise araştırmacı uygulanacak müdahaleden ziyade, belli karakteristiklere sahip bir hastadan alınan yanıtların zaman içerisindeki profilini ve izlem süresi sonundaki yanıt seviyesinin başlangıçtaki yanıt seviyesine göre ne durumda olduğunu öğrenmek isteyebilir.

Tekrarlı ölçümlerde varyans analizi yöntemleri hala çok yaygın olarak kullanılan yöntemlerdir. Bu tür yaklaşımlar artık rutin kullanımda tavsiye edilmiyor olsalar da daha gelişmiş modellerin ve daha ileri düzeydeki yöntemlerin oluşturulması için fikirlerin belirlenmesinde önemlidirler (Hedeker ve Gibbons, 2006).

1.2.1. Tekrarlı Ölçümlerde Tek Değişkenli Varyans Analizi

Varyans analizi mantığı 20. yüzyılın başlarında Fisher tarafından geliştirilmiştir. Tekrarlı ölçümlerde varyans analizi, ilişkili verilerin analizinde önerilen eski yöntemlerden biri olup tek değişkenli ya da karışık etkili varyans analizi olarak da bilinir. Tekrarlı ölçümlerde varyans analizi modelinde tekrarlı ölçümler arası korelasyonun, bireye özel rasgele etkinin her ölçüm değerine olan katkısından ortaya çıktığı varsayılır. Yani, her birey zaman içinde devam eden kendi yanıt düzeyine sahiptir ve bu düzey bireyden alınan yanıtlar arası ilişkinin ortaya çıkmasına neden

olur. Bu bireye özel etki, rasgele etki olarak kabul edilir. Tekrarlı ölçümler için ANOVA modeli aşağıdaki biçimde gösterilir (Davidian, 2007).

$$y_{ij} = X'_{ij}\beta + b_i + e_{ij} \quad \text{Denklem (1.1)}$$

Modelde b_i bireye özel rasgele etki ve e_{ij} birey içi ölçüm hatasını göstermektedir. X'_{ij} tasarım vektörü ise yalnızca kesikli (ya da kategorik) ortak değişkenleri içerir. Hem b_i hem de e_{ij} rasgele olsalar da, birbirinden bağımsız oldukları varsayılır. b_i 'nin sıfır ortalama ve $Var(b_i) = \sigma_b^2$ varyansı ile normal dağıldığı ve aynı şekilde, e_{ij} 'lerin de sıfır ortalama ve $Var(e_{ij}) = \sigma_e^2$ ile normal dağıldığı varsayılır. b_i ve e_{ij} 'nin sıfır ortalamaya sahip olduklarından, ortalama yanıt bütün rasgele kaynaklar üzerinden düşünüldüğünde aşağıdaki biçimde yazılabilir.

$$E(y_{ij}) = \mu_{ij} = X'_{ij}\beta \quad \text{Denklem (1.2)}$$

Böylece, tekrarlı ölçümlerde varyans analizi modelinde i . birey için yanıt, popülasyon ortalaması μ_{ij} 'den bireye özel rasgele etki b_i ve birey içi rasgele etki e_{ij} ile farklılık gösterir. Yani tekrarlı ölçümlerde ANOVA modeli, verideki iki ana değişim kaynağını; bireyler arası değişim kaynağı (σ_b^2) ve birey-içi değişim kaynağı (σ_e^2) olmak üzere ikiye ayırır. Bireyler arası varyasyon, bireylerin yanıt düzeylerinin birbirinden farklı olduğunu; birey-içi varyasyon ise tekrarlı ölçümler içinde ölçüm hatasından ya da örnekleme değişkenliğinden kaynaklanan rasgele dalgalanmalar olduğunu varsayar. Tekrarlı gözlemlere ilişkin varsayılan kovaryans matrisi ise aşağıdaki biçimde tanımlanmıştır.

$$Cov(y_i) = \begin{pmatrix} \sigma_b^2 + \sigma_e^2 & \sigma_b^2 & \sigma_b^2 & \cdots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma_e^2 & \sigma_b^2 & \cdots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 + \sigma_e^2 & \cdots & \sigma_b^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \cdots & \sigma_b^2 + \sigma_e^2 \end{pmatrix} \quad \text{Denklem (1.3)}$$

Görüldüğü gibi varyans, $\sigma_b^2 + \sigma_e^2$, ve kovaryanslar, σ_b^2 , her zaman noktasında (ya da her koşul altında) aynı değere sahiptir. Sonuç olarak tekrarlı ölçümler arası korelasyon sabittir ve aşağıdaki gibi gösterilebilir.

$$Corr(y_{ij}, y_{ik}) = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2} \quad \text{Denklem (1.4)}$$

Bu özel kovaryans yapısı tam simetrik olarak da bilinir. Ancak zaman içinde ölçümlerin alındığı bir yapıda, gözlemler arası ilişkinin sabit ve her gözlem çifti için aynı olduğunu söylemek çok doğru olmayacaktır. Birbirine uzak zaman noktalarında alınan gözlemler arası ilişkinin bozulması beklenebilir. Ayrıca zaman içinde varyansın değişmediği varsayımı da çok gerçekçi bir yaklaşım olmayabilir. Bunların yanı sıra, tekrarlı ölçümlerde varyans analizi yaklaşımı, tekrarlı gözlemlerin bütün bireyler için aynı olan zaman noktalarında ya da koşullar altında alındığı ve veride kayıp gözlemin bulunmadığı durumlar için geliştirilmiştir. Ayrıca ortak değişkenler (deneme grupları ve zaman noktaları gibi) kesikli olarak modele alınabilir. Bu nedenle bu yöntem her bireyden farklı zaman noktalarında ölçümlerin alındığı, verilerde eksik gözlemlerin olduğu ve analize sürekli ortak değişkenlerin dahil edilmek istendiği durumlarda yetersiz kalacaktır.

1.2.2. Tekrarlı Ölçümlerde Çok Değişkenli Varyans Analizi

Tekrarlı ölçümlerde varyans analizi için tek değişkenli ya da karışık etkili varyans analizi olarak da bahsedildiği söylenmişti. MANOVA, orijinal olarak yanıt

değişkenlerinin çok değişkenli vektörünün analizi için geliştirildiğinden, uzunlamasına yanıtlar ile daha genel bir durum olan çok değişkenli yanıtlar arası farklılığı vurgulamak gerekebilir. Öncelikle uzunlamasına yapıdaki veriler birbirine yakın özelliklere sahipken, çok değişkenli veriler için bu durum olmayabilir. Yani uzunlamasına veride, tek bir yanıt değişkeni zaman içinde birden çok kez ölçülürken, örneğin kan basıncına ait üç ölçüm alınırken, genel çok değişkenli veri yapısında birden çok yanıt değişkeninin, kan basıncı, kan şekeri ve LDL düzeyi gibi, birer ölçümleri mevcuttur. Ayrıca uzunlamasına veriler doğal olarak çok değişkenli yapıdadırlar ve gözlemler arası kovaryansın belli bir yapıda olması beklenir. Genel çok değişkenli veri yapısında kovaryans matrisi için çok nadir bir gösterge bulunur.

Yani MANOVA, birden çok yanıt değişkeninin eş zamanlı analizinin yapılmasını sağlamak için geliştirilmiş olsa da, bu tür verilerin, uzunlamasına veriler gibi ilişkili oldukları fark edilmiştir. Böylelikle MANOVA, uzunlamasına ölçümlerin çözümlenmesinde de kullanılmaya başlanmıştır (Fitzmaurice ve ark., 2004).

Tekrarlı ölçümlerde çok değişkenli varyans analizinin uzunlamasına verilerin çözümünde uygun olmayan özellikleri mevcuttur. MANOVA formüllerine göre birey-içi ortak değişken düzeylerinin her birey için aynı olması gereklidir. Ancak veri yapısı dengesiz olduğunda, yani her denekten aynı sayıda tekrarlı ölçüm alınmadığında ya da ölçümler farklı zaman noktalarında alındığında, MANOVA yöntemi uygun olmayacaktır. MANOVA ölçümlerin alındığı zaman değişkenini gerektiği gibi hesaplamaya katmadığı için zaman içindeki ortalama yanıt da doğru olarak modellenemeyecektir (Davidian, 2007). Ayrıca MANOVA kayıp gözlem olduğu durumda kullanılmadığından, kayıp gözleme sahip bireylerin analiz dışında tutulması gerekecektir. Böylelikle örnek büyüklüğü düşecek ve bunun yanı sıra eldeki bütün bilgi kullanılmamış olacaktır. Kayıp veriye sahip bireylerin değerlendirme dışında tutulması, zaman içinde ortalama yanıt profillerinin belirlenmesinde yanlış tahminler elde edilmesine neden olacaktır. Ayrıca, analizde kalan ve bütün ölçümleri tam olan bireyler hedef popülasyonu temsil etmeyeceğinden örneklem ortalamaları, varyansları ve kovaryansları da yanlış tahminler verecektir.

Tekrarlı ölçümlerde tek değişkenli varyans analizinde, tekrarlı ölçümler arası kovaryans yapısının tam simetri olduğu varsayılmıyordu. MANOVA’da ise kovaryans matrisinin yapılandırılmamış olduğu varsayılır.

$$Cov(y_i) = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \sigma_{13}^2 & \cdots & \sigma_{1n}^2 \\ \sigma_{21}^2 & \sigma_2^2 & \sigma_{23}^2 & \cdots & \sigma_{2n}^2 \\ \sigma_{31}^2 & \sigma_{32}^2 & \sigma_3^2 & \cdots & \sigma_{3n}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \sigma_{n3}^2 & \cdots & \sigma_n^2 \end{pmatrix} \quad \text{Denklem (1.5)}$$

Kovaryans matrisi belli bir yapıya sahip değildir ve tek değişkenli yöntemde olduğu gibi iki parametre (σ_b^2 ve σ_e^2), yerine daha çok sayıda ($n(n+1)/2$ sayıda) parametrenin kestirilmesi gerekecektir. Böyle bir kovaryans yapısı da eldeki uzunlamasına verinin sahip olduğu kovaryans yapısının doğru olarak tanımlanmasını sağlamayacaktır (Davidian, 2007).

1.3. Doğrusal Karma Etki Modelleri

1.3.1. Rasgele ve Sabit Etkiler

Herhangi bir A değişkeninin düzeyleri α_i olarak düşünüldüğünde, yapılan analizde bu değişkenin düzeyleri önceden belirlenmişse α_i ’ler sabit etkiler olarak bilinir. Analiz sonucunda yapılacak yorum yalnızca seçilen düzeyler için geçerli olacaktır. Eğer etkenin düzeyleri daha geniş bir popülasyondan rasgele olarak seçilmiş iseler bu etkiler rasgele etki olarak düşünülebilirler. Seçilen rasgele etkiler, değişkenin bütün olası düzeyleri arasından seçilmiş rasgele bir örneklem olarak düşünülebilir. Rasgele etkiler belli bir olasılık dağılımına sahiptir ve analiz sonucunda yapılacak yorum bütün popülasyon için genellenebilir. Karma etki modelleri, hem sabit hem de rasgele etkileri bir arada bulundurma özelliğine sahiptir. Karma etki modellerinde sabit etkiler popülasyon etkileri, rasgele etkiler ise bireylerin popülasyon

değerlerinden sapmalarını gösteren birey etkileri olarak alınır. Rasgele etkinin aynı düzeyini paylaşan ölçümler ilişkili olarak modellenirler (Vittinghoff ve ark., 2005).

1.3.2. Karma Etki Modellerinin Avantajları

Karma etki modellerini özellikle uzunlamasına çalışmaların analizinde kullanışlı hale getiren çeşitli özellikleri mevcuttur. Öncelikle, deneklerden aynı sayıda ölçüm alınmış olması zorunlu değildir, bu nedenle tamamlanmamış ölçümlere sahip olan denekler de analize dahil edilebilirler. Bu durum karma etki modellerinin, dengeli veri yapısı isteyen klasik yöntemlere göre önemli bir avantajdır, çünkü eldeki mevcut bütün veriyi analize katarak istatistiksel olarak gücün artmasını sağlayacak ve yanlılığı ortadan kaldıracaktır. Ayrıca, eksik gözleme sahip bireyler analizden çıkartıldığında, geriye kalan bireyler popülasyonu iyi temsil edemeyebilir. Bir diğer önemli özellik ise, karma etki modellerinin, ölçümlerin alındığı zaman noktalarını sürekli bir değişken gibi analize katıyor olmasıdır. Bu nedenle deneklere ait ölçümlerin aynı zaman noktalarında alınmış olması zorunluluğu ortadan kalkar. Hem zamandan bağımsız, hem de zamana bağımlı ortak değişkenler modele dahil edilebilir. Bu durumda yanıt değişkenindeki değişim, hem bireyin değişmez karakteristikleriyle (cinsiyet, ırk gibi), hem de zaman içinde değişen özellikleriyle (kolesterol düzeyi, yaş gibi) tanımlanabilir (Hedeker ve Gibbons, 2006).

Bunların yanısıra, doğrusal karma etki modellerinde, regresyon parametrelerinin bir kısmı bireyden bireye farklılık gösterdiğinden, bu modeller popülasyon içindeki doğal heterojenlik kaynaklarını hesaba katar. Yani, popülasyondaki bireyler kendi ortalama yanıt profillerine ve rasgele olduğu kabul edilen regresyon parametrelerine sahiptir. Doğrusal karma etki modellerinin ayırt edici bir özelliği, ortalama yanıt düzeyini, bütün bireyler tarafından paylaşılan popülasyon karakteristiklerinin (popülasyon ortak değişkenleri, yani β 'lar) ve bireye özel etkilerin (bireye özel ortak değişkenler, yani ν 'ler) bir birleşimi olarak modelleyebilmesidir. Popülasyon karakteristikleri sabit etkiler iken, bireye özel etkiler rasgele etkilerdir. (Fitzmaurice ve ark., 2004). Karma etki modelleri, birey içi

değişimin ve bireyler arası değişimin ayrı ayrı tanımlanmasına olanak sağlar. Ayrıca model, bireye özel karakteristikleri de içerdiğinden, araştırmacıya deneklerin bireysel yanıt profilleri hakkında da En İyi Doğrusal Yansız Önkestirici (Best Linear Unbiased Predictor, BLUP) ile çıkarsamalarda bulunmasını sağlar. Örneğin; zaman içerisindeki bireysel büyüme eğrilerinin kestirimi doğrusal karma etki modeller yardımıyla yapılabilir. Yürütülen bir klinik çalışmada, çalışmadaki deneklere ait genel ortalama yanıtta çok, bireylere özel çıkarsamalarda bulunmak istenebilir. Bu tür tahminler klinik çalışmada deneklerden, atandıkları grupta aldıkları tedaviye iyi yanıt vermeyenleri belirlemede kullanılabilir. Karma etki modellerinin dikkat çekici bir yanı da, bireylerden alınan tekrarlı gözlemlerin aynı sayıda ve/veya aynı koşullar altında alınmış olmasının zorunlu olmayışıdır. Bu nedenle, bu modeller dengesiz yapıdaki tekrarlı ölçüm verilerinin analizine uygundur (Fitzmaurice ve ark., 2004).

1.3.2.1. Tamamlanmamış Veri Yapısı

Tekrarlı ölçümlerin alındığı zaman noktaları $j = 1, \dots, n_i$ olmak üzere, her bir bireyden n_i tane ölçüm alındığı düşünülün. Gösterimde tekrarlı ölçüm sayısı n yerine n_i ile gösterildiğinden, her bireyden alınan tekrarlı ölçüm sayıları farklı olabilir. Bu nedenle karma etki modellerinde bireylerden alınan tekrarlı ölçüm sayılarının eşit olması gerekmez ve kayıp gözlemi olan bireyler de analize dahil edilirler. Bunun yanı sıra, tekrarlı ölçümlerin alındığı zaman noktaları da t_i ile gösterildiğinden, bireylerden alınan tekrarlı ölçümler, her birey için farklı zaman noktalarında olabilir. Karma etki modelleri kayıp gözlemlerin göz ardı edilebilir olduğu durumda geçerli sonuçlar vermektedir. Kayıp gözlemlerin göz ardı edilebilir olması, kayıp gözlem olma olasılığının yalnızca gözlenen ortak değişkenlere (zaman gibi) ve kayıp gözleme sahip bireylerden elde edilen gözlenen yanıt değişkeni değerlerine bağlı olabileceği anlamına gelmektedir. Yani, eğer bireyin kayıp gözlemleri önceki yanıt değerlerine ve diğer gözlenebilen bireysel karakteristiklere bağlı ise karma etki modelleri, model parametreleri için geçerli kestirimler sağlayacaktır (Hedeker ve Gibbons, 2006). Kayıp gözlemlerin genellikle önceki yanıt

değerlerine ve diğer bireysel özelliklerle ilişkili olduğu varsayılabilirliğinden, karma etki modelleri tekrarlı ölçümlerde kayıp gözlem olduğu durumlarda avantajlı bir çözüm yoludur.

1.3.3. Genel Doğrusal Karma Etki Modeli

Genel olarak, doğrusal bir karma etki modeli aşağıdaki koşulları sağlayan herhangi bir model olabilir.

$$\begin{aligned}
 y_i &= X_i\beta + Z_iv_i + e_i \\
 v_i &\sim N(0, G) \\
 e_i &\sim N(0, R_i) \\
 v_1, \dots, v_N; e_1, \dots, e_N &\text{ bağımsız}
 \end{aligned}
 \tag{Denklem (1.6)}$$

Modelde y_i , i . birey için n_i boyutlu yanıt vektörü, $1 \leq i \leq N$, N : birey sayısı, X_i ve Z_i sırasıyla $(n_i \times p)$ ve $(n_i \times q)$ boyutlu bilinen ortak değişkenler matrisleri, β sabit etkileri içeren p -boyutlu vektör, v_i rasgele etkileri içeren q -boyutlu vektör, e_i hata bileşenlerini içeren n_i -boyutlu vektördür. Son olarak, G , $g_{ij} = g_{ji}$ olmak üzere, $(q \times q)$ boyutlu genel kovaryans matrisi, ve R_i , $(n_i \times n_i)$ boyutlu kovaryans matrisidir (Verbeke ve Molenbergs, 2000). Bölüm 1.3.7’de karma etki modelinin matris biçiminde gösterimine daha ayrıntılı olarak değinilecektir.

Doğrusal karma etki modelleri, yanıt profillerini popülasyon parametreleri (sabit etkiler) ve bireye özel rasgele etkilerin bir birleşimi olarak yorumlansa da, rasgele etkilerin dağılımı üzerinden ortalama alınarak bulunan marjinal ortalama yanıt için bir modele ulaşılabilir (Fitzmaurice ve ark., 2004). Denklem (1.6)’ya göre y_i ’ler, v_i rasgele etkilerine koşullu olarak, $X_i\beta + Z_iv_i$ ortalama vektörü ve Σ_i kovaryans matrisi ile normal dağılıma sahiptir. Ayrıca v_i ’lerin de 0 ortalama ve G kovaryans matrisi ile normal dağıldığı varsayılır. y_i ve v_i için sırasıyla yoğunluk fonksiyonlarının $f(y_i | v_i)$ ve $f(v_i)$ olduğu düşünüldüğünde y_i için marjinal

yoğunluk fonksiyonu; $f(y_i) = \int f(y_i | v_i) f(v_i) dv_i$ biçiminde olacaktır. Bu marjinal yoğunluk fonksiyonu, n_i boyutlu, $X_i \beta_i$ ortalama vektörü ve $\Sigma_i = Z_i G Z_i' + R_i$ kovaryans matrisine sahip normal dağılım yoğunluk fonksiyonudur (Verbeke ve Molenberghs, 2000).

Sonuç olarak marjinal ortalama yanıt aşağıdaki gibi gösterilebilir.

$$E(Y_i) = X_i \beta \quad \text{Denklem (1.7)}$$

Karma etki modelleri, basit doğrusal regresyon modelinin bir uzantısı olarak düşünülebilir. Bu nedenle öncelikle basit doğrusal modelinden başlanarak karma etki modelleri daha detaylı anlatılmaya çalışılacaktır.

1.3.4. Basit Doğrusal Regresyon Modeli

$i = 1, 2, \dots, N$ olmak üzere i . bireyden j . durumda ($j = 1, 2, \dots, n_i$) alınan ölçüm y_{ij} aşağıdaki gibi yazılabilir;

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + e_{ij} \quad \text{Denklem (1.8)}$$

İndisler göz ardı edildiğinde Denklem (1.8), y bağımlı değişkeninin t ile gösterilen bağımsız zaman değişkeni ile olan ilişkisini göstermektedir. İndisler verinin hangi bireye ait olduğu (i) ve hangi zaman noktasında ölçüldüğü (j) gibi özellikleri belirtmektedir. Bağımsız değişken t , zamanın düzeylerini göstermektedir ve gün, hafta ya da ay olarak belirtilebilir. y ve t , i ve j indislerinin her ikisini de taşıdıkları için, hem yanıt değişkeni hem de zaman değişkeni bireyler ve zaman noktaları için değişebilir.

Yukarıdaki model gibi doğrusal regresyon modellerinde e_{ij} hatalarının sıfır ortalama ve σ^2 varyansı ile normal dağıldıkları ve birbirinden bağımsız oldukları varsayılır. Bu varsayım ile yukarıdaki modelin uzunlamasına verilerin analizinde kullanımı olanaksızdır. Çünkü y bağımsız değişkeni, aynı bireyden birden fazla ölçüm alınarak elde edilmiştir. Bu nedenle de aynı bireyden alınan ölçümlere ait hataların birbirleri ile ilişkili olduklarının düşünülmesi daha mantıklı olacaktır.

Bunların yanı sıra yukarıdaki modele göre, regresyon parametreleri her birey için aynı olduğundan, zaman içindeki değişimin bütün bireyler için aynı olduğu varsayılır (Hedeker ve Gibbons, 2006).

Bu nedenlerden dolayı, verinin bağımlılık yapısını göz önünde tutacak ve her birey için farklı zaman eğilimlerini tanımlayabilecek bireye özel etkilerin modele eklenmesi uygun olacaktır. Karma etki modellerinin yaptığı da tam olarak budur. Karma etki modelleri bu nedenle doğrusal regresyon modellerinin geliştirilmiş biçimi olarak düşünülebilir.

1.3.5. Rasgele Kesim Noktası Modeli

Doğrusal karma etki modelinin en basit hali, rasgele değişen birey etkisinin bulunduğu bir doğrusal modeldir. Bu modelde, her bireyin kendi yanıt düzeyine sahip olduğu varsayılır. Birey etkisi, karma etki modeline rasgele etki olarak alınır. Yukarıda verilen regresyon modelinin basit bir uzantısı olarak, her bireyin kendi tekrarlı ölçümleri üzerindeki etkisini modele ekleyerek yapılabilir.

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \nu_{0i} + e_{ij} \quad \text{Denklem (1.9)}$$

Burada ν_{0i} parametresi i . bireyin kendi tekrarlı gözlemleri üzerindeki etkisini göstermektedir. Yani, ν_{0i} Denklem (1.9)'da rasgele birey etkisi olarak yer

almaktadır. Bireylerin kendi ölçümleri üzerinde etkisi olmadığında bütün v_{0i} değerleri sifira eşit olacaktır. Ancak, bireylerin kendi tekrarlı ölçümleri üzerinde pozitif ya da negatif etkiye sahip olmaları daha olası bir durumdur ve bu nedenle v_{0i} değerleri sifirdan farklı olacaktır. Bu modelde i . bireyden j . durum ya da zamanda alınan yanıt popülasyon ortalaması $\beta_0 + \beta_1 t_{ij}$ 'den, birey etkisi v_{0i} ve birey içi ölçüm hatası e_{ij} kadar farklılık gösterir. Burada hem birey etkisi hem de birey içi ölçüm etkisi sıfır ortalama ve sırasıyla $Var(v_{0i}) = \sigma_v^2$ $Var(e_{ij}) = \sigma^2$ varyanslarına sahip rasgele etkilerdir.

Herhangi bir birey etkisi verildiğinde koşullu ortalama aşağıdaki gibi gösterilebilir.

$$E(y_{ij} | v_{0i}) = \beta_0 + \beta_1 t_{ij} + v_{0i} \quad \text{Denklem (1.10)}$$

Popülasyona ait ortalama yanıt ise aşağıdaki marjinal model yardımı ile gösterilebilir.

$$E(y_{ij}) = \mu_{ij} = \beta_0 + \beta_1 t_{ij} \quad \text{Denklem (1.11)}$$

Denklem (1.10)'dan, bireye özel etki verildiğinde y_{ij} 'nin koşullu ortalaması, Denklem (1.11)'den ise y_{ij} 'nin marjinal ortalaması (bireye özel etkilere ait dağılım üzerinden ortalama alınmış) olarak bahsedilebilir (Fitzmaurice ve ark., 2004).

Karma etki modellerini hiyerarşik ya da çok düzeyli yapıda göstermek, bireysel etkilerin daha iyi anlaşılması açısından faydalı olabilir. Doğrusal karma etki modelleri iki aşamalı hiyerarşik yapıda yazılabilir. İki aşamalı model ilk olarak Laird ve Ware (1982) tarafından önerilmiştir. Denek içi ya da ilk aşama modeli;

$$y_{ij} = b_{0i} + b_{1i} t_{ij} + e_{ij} \quad \text{Denklem (1.12)}$$

ve denekler arası ya da ikinci aşama modeli,

$$\begin{aligned} b_{0i} &= \beta_0 + \nu_{0i} \\ b_{1i} &= \beta_1 \end{aligned} \quad \text{Denklem (1.13)}$$

olarak gösterilebilir. Burada ilk aşama modeli, i . bireye ait j . zamandaki yanıt değerinin, aynı bireye ait başlangıç değeri b_{0i} ve zaman içindeki yanıt eğilimi b_{1i} tarafından etkilendiğini gösterir.

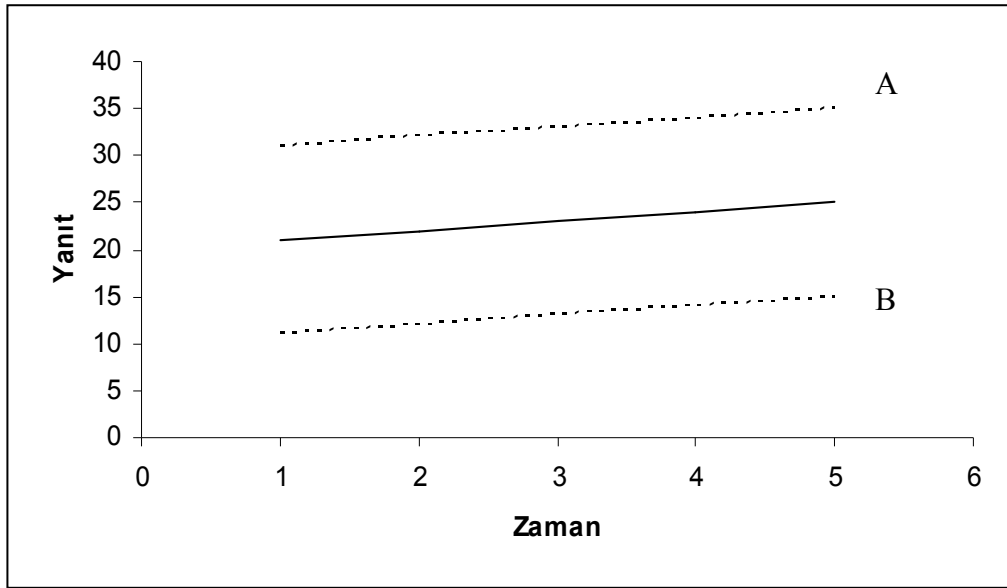
İkinci aşama modelinde ise, i . bireye ait başlangıç değeri b_{0i} , popülasyon başlangıç değeri β_0 ve bireye özel rasgele etki ν_{0i} ile tanımlanmıştır. Yani ilk aşamada kullanılan parametreler ikinci aşamada bağımlı değişken olarak düşünülebilir. Burada gösterilen ikinci aşama modelinde, her birey kendi ayrı başlangıç düzeyine sahiptir. Ancak bunun tersine, yine ikinci aşama modeline göre bütün bireylerin zaman içindeki eğilimleri aynıdır ve hepsi popülasyon eğimi β_1 'e eşittir. Diğer bir ifadeyle, her bireyin zaman içindeki eğilim doğrusu, β_0 ve β_1 ile tanımlanan popülasyon eğilim doğrusuna paralel, ancak popülasyon eğilim doğrusuna göre ν_{0i} kadar yer değiştirmiş olacaktır.

Bu hiyerarşik gösterime göre, ilk aşamada yer alan ortak değişkenler ilgilenilen y_{ij} sonuç değişkenindeki değişimi açıklarken, ikinci aşama modelinde yer alan ortak değişkenler ise ilk aşama modelindeki ortak değişkenlerdeki değişimi açıklamaktadır. Böylelikle ilk ve ikinci aşama modelleri birleştirilerek Denklem (1.9)'a ulaşılabilir.

Seçilen örnekteki bireylerin daha geniş olan popülasyondaki bireyleri temsil ettiklerinden, bireye özel etkiler, ν_{0i} , rasgele etkiler olarak düşünülür. Yani, ν_{0i} 'lerin popülasyondaki birey etkilerinin bir dağılımını temsil ettiği düşünülebilir. Bu dağılım için en genel biçim, sıfır ortalama ve σ_v^2 varyans ile normal dağılımdır.

Bu durumda, Denklem (1.9)'da yer alan hatalar, e_{ij} , sıfır ortalama ve σ^2 varyansı ile normal ve koşullu bağımsız olarak dağıldıkları söylenebilir. Koşullu bağımsızlık burada, ν_{0i} bireye-özel rasgele etkiler üzerine koşullu olduğu anlamındadır. Birey etkileri hatalardan arıtıldığı için koşullu bağımsızlık varsayımı, klasik regresyon modelindeki tam bağımsızlık varsayımından daha mantıklıdır.

Daha önce de belirtildiği gibi, Denklem (1.9)'da yalnızca tek bir birey etkisi vardır ve bu birey etkisi de genel bireylerin regresyon doğrusundan paralel olarak sapmasını göstermektedir. Bu nedenle bu model rasgele kesim noktaları modeli olarak tanımlanır ve her bir ν_{0i} , i . bireyin popülasyon eğiliminden nasıl saptığını gösterir. Şekil (1.1) yardımıyla bu durum daha açık görülebilir.

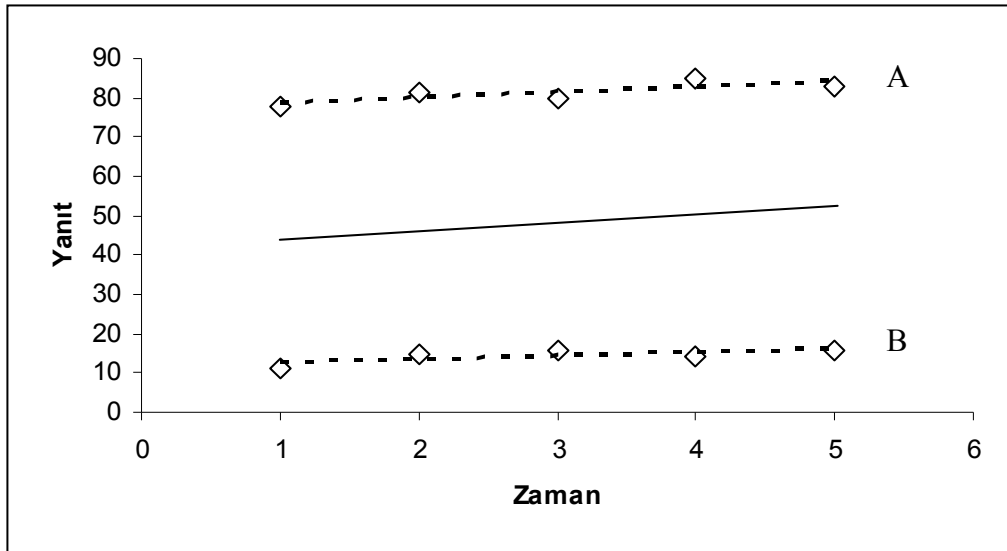


Şekil 1.1. Rasgele kesim noktası modelinde herhangi iki bireye ve popülasyona ait zaman içindeki yanıt eğilimleri.

Şekil 1.1'de, düz çizgi ile gösterilen eğri β_0 ve β_1 'e bağlı popülasyon eğilimini göstermektedir. A ve B bireylerine ait koşullu yanıtlar ise, ortalama (popülasyon) eğilimin üzerinde ve altında, ortalama eğilime paralel olarak yer almaktadır. Seçilen örnekleme N birey için N tane eğri olacaktır. Rasgele birey etkilerine ait varyans terimi, σ_v^2 , bu eğrilerin yayılımını göstermektedir. Eğer σ_v^2 sıfıra yakın ise, bireysel eğilimler ortalama eğilimden farklılık göstermeyecektir.

Bireylerin popülasyon eğiliminden farklılık gösterdiği durumda birey eğrileri popülasyon eğrisinden uzaklaşacak ve σ_v^2 değeri artacaktır.

Bu basit örnekte A'ya ait yanıtların hep popülasyon ortalamasının üzerinde olduğu, yani A bireyinin pozitif etkiye (v_{0A}) sahip olduğu, B'ye ait yanıtların ise hep popülasyon ortalamasının altında olduğu yani B bireyinin negatif bir etkiye (v_{0B}) sahip olduğu görülmektedir. Rasgele ölçüm hataları e_{ij} 'ler de Şekil 1.1'e eklendikten sonra alınan yanıtların bireysel eğriler üzerinde rasgele dağıldığı gözlenecektir (Şekil 1.2).



Şekil 1.2. Rasgele ölçüm hataları eklendikten sonra zaman içindeki marjinal ve koşullu ortalama yanıt profilleri.

1.3.5.1. Tam Simetri ve Sınıf İçi Korelasyon

Rasgele kesim noktası modeli, uzunlamasına verinin varyans ve kovaryansları için tam simetri varsayımına sahiptir. Hem varyans hem de kovaryansların zaman içinde sabit olduğu varsayılır;

$$Var(y_{ij}) = \sigma_v^2 + \sigma^2$$

$$Cov(y_{ij}, y_{ik}) = \sigma_v^2 \quad \text{Denklem (1.14)}$$

Kovaryansın korelasyon biçiminde gösterilmesi, yani bireysel varyans σ_v^2 'nin toplam varyansa, $\sigma_v^2 + \sigma^2$, oranı sınıf içi korelasyonu verecektir. Bu katsayı, birey içi uzunlamasına verinin ilişki derecesini ve aynı zamanda verideki toplam değişimin bireylerden kaynaklanan kısmını göstermektedir.

1.3.6. Rasgele Kesim Noktası ve Rasgele Eğim Modeli

Rasgele kesim noktası modelinde, yalnızca kesim noktasının bireyden bireye rasgele değiştiği varsayılıyordu. Sonuç olarak ortaya çıkan gözlemler arası kovaryans yapısı da tam simetri biçimindeydi.

Uzunlamasına veri için rasgele kesim noktası modeli basit düzeyde kalmaktadır. Öncelikle, bütün bireyler için zaman içindeki değişimin aynı olduğunu varsaymak yanlış olacaktır. Her bireyin aynı oranda değişim göstermemesi, yani bireylerin zaman içindeki eğilimleri bakımından farklılık göstermeleri daha olası bir durumdur. Dahası, varyans kovaryans matrisi için tam simetri varsayımı, çoğu uzunlamasına veri yapısına uygun değildir. Genelde, birbirine yakın zaman noktalarında alınan ölçümler arasındaki ilişkinin daha yüksek olması beklenirken, birbirinden uzak zaman noktalarında alınan ölçümler arası ilişkinin daha az olması beklenir. Ayrıca, yapılan çalışmaların çoğunda bireyler başlangıç değerleri bakımından benzerlik gösterirken, zaman içindeki değişimleri farklı olabilmektedir. Yani zaman içinde değişkenliğin artması beklenebilir (Hedeker ve Gibbons, 2006).

Bu nedenlerle, hem kesim noktasının hem de zaman içindeki eğilimin bireyler arasında değişim gösterdiğini varsayan bir karma etki modeli daha uygun olacaktır. Ayrıca, modele rasgele kesim noktası dışında bazı yeni rasgele regresyon parametreleri eklendiğinde rasgele etkiler kovaryans yapısı da değişecektir. Bunun için ilk aşama modeli yine Denklem (1.12) gibi yazılabilir. Ancak ikinci aşama modeli Denklem (1.15)'te gösterildiği biçimde olacaktır (Hedeker ve Gibbons, 2006).

$$b_{0i} = \beta_0 + v_{0i}$$

$$b_{1i} = \beta_1 + v_{1i} \quad \text{Denklem (1.15)}$$

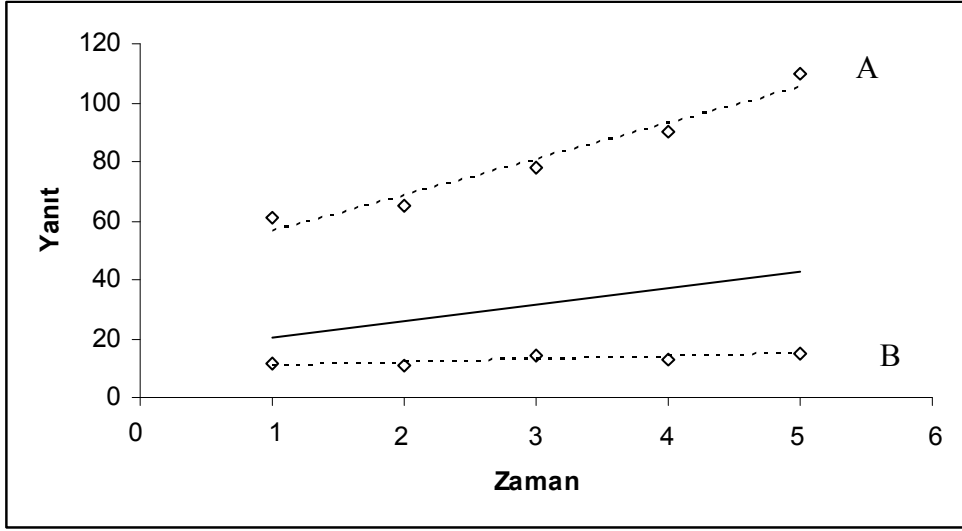
Sonuç olarak, j . zaman noktasında i . bireyden alınan yanıt için karma etki modeli aşağıdaki gibi yazılabilir (Fitzmaurice ve ark., 2004).

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + v_{0i} + v_{1i} t_{ij} + e_{ij} \quad \text{Denklem (1.16)}$$

Bu modelde β_0 popülasyon kesim noktasını, β_1 ise popülasyon eğimini, v_{0i} i . bireyin popülasyon kesim noktasından sapmasını, v_{1i} ise i . bireyin popülasyon eğiminden sapmasını göstermektedir. İlk aşama modelindeki e_{ij} , sıfır ortalama ve σ^2 varyansı ile normal dağılan ve koşullu bağımsız hata terimidir. Yani ilk aşama hata terimleri v_{0i} ve v_{1i} 'ye koşullu olarak bağımsızdır. Bireye özel rasgele etkiler v_{0i} ve v_{1i} 'nin ise sıfır ortalama ve G kovaryans matrisi ile çok değişkenli normal dağıldığı varsayılır.

$$G = \begin{bmatrix} \sigma_{v_0}^2 & \sigma_{v_0 v_1} \\ \sigma_{v_0 v_1} & \sigma_{v_1}^2 \end{bmatrix} \quad \text{Denklem (1.17)}$$

Kesim noktası parametreleri başlangıç yanıt düzeyini, eğim parametreleri ise zaman içindeki değişimi göstermektedir. Popülasyon kesim noktası ve eğim parametreleri genel eğilimi (popülasyon eğilimini) gösterirken, bireye özel parametreler ise bireylerin popülasyondan nasıl saptıklarını gösterir. Yani, böyle bir modelde bireyler arası değişkenlik yalnızca ilk yanıt düzeylerinden değil, zaman içindeki yanıt profillerinin bireyler arası değişiminden de kaynaklanmaktadır. Şekil 1.3 bu modeli grafiksel olarak tanımlamaktadır.



Şekil 1.3. Ölçüm hataları eklendikten sonra zaman içindeki marjinal ve koşullu yanıtlara ait profiller.

Şekil 1.3'de düz çizgi ile gösterilen eğilim popülasyon eğilimi (marjinal ortalama yanıt), kesik çizgiler ise A ve B bireylerinin zaman içindeki eğilim eğrilerini (koşullu ortalama yanıt trendleri) göstermektedir. Bu modelde zaman içindeki eğilim de bireyler arasında farklılık gösterdiği için, bazı bireyler zaman içinde çok değişim göstermezken, bazıları çok belirgin değişim gösterebilirler. Popülasyon eğilimi bütün bireyler için ortalama eğilimi gösterirken, varyans terimleri popülasyonda ne kadar heterojenlik olduğunu göstermektedir. Yani, $\sigma_{v_0}^2$ bireylerin popülasyon kesim noktası etrafında ne kadar değişim gösterdiklerini gösterirken, $\sigma_{v_1}^2$ eğimlerdeki değişimi göstermektedir. Ayrıca, $\sigma_{v_0v_1}$ kovaryans parametresi bireye özel kesim noktası ile eğimin birlikte değişimini göstermektedir. Örneğin pozitif bir korelasyon büyük başlangıç değerine sahip bireylerin daha büyük pozitif eğime sahip olacaklarını, negatif korelasyon ise tam tersi olacağını gösterir (Hedeker ve Gibbons, 2006).

Bu örneğe göre; A ve B bireyleri ayrı ayrı popülasyonla karşılaştırılacak olursa, A bireyine ait başlangıç düzeyinin ($\beta_0 + v_{0A}$), popülasyon ortalamasından (β_0) daha yüksek olduğu, yani pozitif bir bireysel etkiye sahip olduğu görülmektedir. Diğer bir yandan, B bireyine ait başlangıç yanıt düzeyinin ($\beta_0 + v_{0B}$), popülasyon ortalamasından (β_0) daha düşük olduğu, yani negatif bir bireysel etkiye sahip olduğu

görülmektedir. Bunun yanı sıra, A bireyinin popülasyona ait ortalama değişime (β_1) göre daha dik bir değişim gösterdiği ($\beta_1 + v_{1A}$), yani pozitif bir bireysel etkiye sahip olduğu görülmekteyken, B bireyinin daha yatık bir değişim ($\beta_1 + v_{1B}$) göstermekte olduğu, yani negatif bir bireysel etkiye sahip olduğu görülmektedir. Son olarak, ölçüm hataları e_{ij} 'ler de göz önüne alındığında, gözlenen yanıtların bireylere özel eğriler etrafında rasgele olarak değişim gösterdiği görülmektedir. Bu modelde rasgele kesim noktaları ve rasgele eğimler yer almaktadır. Doğrusal karma etkiler modelleri rasgele değişen ek regresyon katsayıları için de genellenebilir. Ayrıca rasgele etkilere ait ortalamaların ortak değişkenlere (tedavi grupları, cinsiyet, ... gibi) bağlı olarak değişmesini sağlayabilir (Fitzmaurice ve ark., 2004).

1.3.7. Matris Gösterimi

Matris gösterimi kullanılarak doğrusal karma etki modeli aşağıdaki gibi yazılabilir.

$$y_i = X_i\beta + Z_iv_i + e_i \quad \text{Denklem (1.18)}$$

Modelde β ($p \times 1$) boyutlu popülasyon parametreleri (sabit etkiler) vektörü; v_i ($q \times 1$) boyutlu bireye özel parametreler (rasgele etkiler) vektörü; X_i ($n_i \times p$) boyutlu ortak değişken matrisi ve Z_i ($n_i \times q$) boyutlu ortak değişken matrisi, $q \leq p$ olmak üzere, olarak tanımlanabilir. Burada Z_i , v_i rasgele etkiler vektörünü y_i 'ye bağlayan tasarım matrisi olarak da bilinir. Aslında Z_i 'nin kolonları, X_i 'ye ait kolonların bir alt kümesidir. Modelde hangi parametrelerin bireyden bireye değişeceği X_i 'nin Z_i 'ye karşılık gelen kolonları ile belirlenebilir. Yani, β vektörünün içinde yer alan herhangi bir parametre, X_i 'de hangi kolonla eşleşiyorsa, rasgele etkiler için tasarım matrisi olan Z_i 'de de karşılık gelen kolon oluşturularak, parametrenin denekler arası değişimi (rasgele olarak) sağlanabilir. Modeldeki rasgele etkilerin, v_i , sıfır ortalama ve G kovaryans matrisi ile çok değişkenli normal

dağılıma sahip oldukları varsayılır. Prensipde v_i 'ler için herhangi bir, çok değişkenli dağılım varsayılabilir; pratikte ise v_i 'ler için çok değişkenli normal dağılım varsayılır. Bu modelde v_i rasgele etkiler vektörünün sıfır ortalamaya sahip olduğu varsayıldığında, rasgele etkiler, i . birey için regresyon parametrelerinin bir alt kümesinin (bireye özel regresyon parametrelerinin) popülasyondan ne kadar saptığı bakımından yorumlanabilir. Daha önce de bahsedildiği gibi β regresyon parametreleri içinden bireyler arası rasgele değişim gösterenler, X_i 'nin Z_i 'yi kapsayan kolonları ile belirlenebilir. Örneğin; yalnızca rasgele değişen kesim noktalarının olduğu bir modelde; Z_i yalnızca 1'lerden oluşan $(n_i \times 1)$ boyutlu vektördür ($X_{ij}=1$, bütün i ve j 'ler için). Doğrusal karma etki modellerinde yapılan önemli bir ayırım y_i 'nin koşullu ve marjinal ortalamaları arasındadır. y_i 'nin koşullu, diğer bir ifadeyle bireye özel ortalaması aşağıdaki gibi gösterilebilir.

$$E(y_i | v_i) = X_i\beta + Z_i v_i \quad \text{Denklem (1.19)}$$

y_i 'nin marjinal ya da, başka bir ifadeyle popülasyon ortalaması, v_i rasgele etkilerine ait dağılım üzerinden ortalama alınarak Denklem (1.20)'deki gibi gösterilebilir.

$$\begin{aligned} E(y_i) &= \mu_i \\ &= E(E(y_i | v_i)) \\ &= E(X_i\beta + Z_i v_i) \\ &= X_i\beta + Z_i E(v_i) \end{aligned} \quad \text{Denklem (1.20)}$$

Rasgele etkilerin beklenen değeri sıfır olarak varsayıldığından eşitlik Denklem (1.21)'deki yazılabilir.

$$\begin{aligned} E(Y_i) &= \mu_i \\ &= X_i\beta \end{aligned} \quad \text{Denklem (1.21)}$$

Yani, doğrusal karma etki modelinde, β regresyon parametreleri (sabit etkiler) vektörünün, bütün bireyler için aynı olduğu ve popülasyon ortalamalı yorumlarının olduğu varsayılır. Örneğin; bütün bireylerin yanıtlarındaki değişim profillerinin ortalaması, popülasyon için ortalama değişim profilinin verecek ve bu değişim profilinin bütün bireyler için aynı olduğu varsayılacaktır. Rasgele etkiler vektörü v_i ise, β 'nin tersine bireye özel regresyon parametrelerini içerir. Bunlar rasgele etkilerdir ve sabit etkilerle birleşince, herhangi bir bireye ait ortalama yanıt profilini verir. Yani, i . birey için ortalama yanıt profili, y_i 'nin koşullu ortalaması yardımıyla aşağıdaki biçimde gösterilebilir.

$$E(y_i | v_i) = X_i\beta + Z_iv_i \quad \text{Denklem (1.22)}$$

Son olarak, Denklem (1.18)'de yer alan e_i , $(n_i \times 1)$ boyutlu hata vektörüdür. Sıfır ortalama ve R_i kovaryans matrisi ile çok değişkenli normal dağılıma sahip oldukları ve v_i rasgele etkilerinden bağımsız oldukları varsayılır. Bölüm 1.3.5.'te e_i 'nin koşullu bağımsız olduğu söylenmişti. Yani herhangi bir v_i rasgele etkisi verildiğinde, e_i 'nin bileşenlerinin birbirinden bağımsız olduğu düşünülebilir. Bu durumda, alışıldığı gibi, R_i matrisi, I_{n_i} $n_i \times n_i$ boyutlu birim matris olmak üzere, $\sigma^2 I_{n_i}$ olarak tanımlanan köşegen matristir. Böyle tanımlandığında denek içi hatayı gösteren e_{ij} ve e_{ik} aynı varyansa sahip ve birbirinden bağımsız olacaktır. e_{ij} 'ler örnekleme ya da ölçüm hataları olarak düşünülebilir. R_i 'nin herhangi bir kovaryans yapısına sahip olduğu düşünüldüğünde, e_{ij} 'ler arası bağımlılık söz konusu olacaktır. Böyle bir durumda artık e_{ij} 'ler basitçe ölçüm ya da örnekleme hatası olarak yorumlanamayabilir. Ayrıca R_i 'nin köşegen olmayan bir matris olduğu varsayıldığında, eldeki veriden faydalanarak hem G 'yi hem de R_i 'yi kestirmek zor olacağından, model tanımlanırken güç algılanan önemli noktalar olabilir. Örneğin, hem G hem de yapılandırılmamış R_i 'yi kestirmek zor olabilir. Bu tez boyunca e_{ij} 'lerin tam anlamıyla ölçüm ya da örnekleme hatası olduğu, yani $R_i = \sigma^2 I_{n_i}$

olduğu varsayılacaktır. Hem v_i 'lerin hem de e_{ij} 'lerin çok değişkenli normal dağılımlı olduğu varsayılsa da, bu dağılımsal varsayımlar model geliştirilirken gerekli değildir. Koşullu ve marjinal ortalamaların biçimleri, yalnızca ölçüm hatalarının rasgele etkilerden bağımsız olduğu ve her ikisinin de sıfır ortalamaya sahip olduğu, $E(v_i) = 0$ ve $E(e_{ij}) = 0$, varsayımını gerektirir (Fitzmaurice ve ark., 2004).

Şimdiye kadar bahsedilen vektör ve matris gösterimlerini daha net anlayabilmek için rasgele kesim noktası ve rasgele eğime sahip bir karma etki modeli göz önüne alınsın. i . bireyden j . durum ya da zamanda alınan yanıt aşağıdaki gibi modellenenabilir.

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + v_{0i} + v_{1i} t_{ij} + e_{ij}, \quad j = 1, \dots, n_i \quad \text{Denklem (1.23)}$$

Vektör gösterimini kullanarak, Denklem (1.18)'de gösterilen aynı model tekrar yazılabilir.

$$y_i = X_i \beta + Z_i v_i + e_i$$

Burada;

$$X_i = Z_i = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{in_i} \end{pmatrix} \quad \text{Denklem (1.24)}$$

$q = p = 2$ ve Z_i matrisi X_i matrisinin kolonlarından oluşmaktadır. Bu modele göre bireyler yalnızca başlangıç noktasındaki yanıtları bakımından değil, yanıtların zaman içindeki değişim biçimleri bakımından da farklılık gösterirler. Modele ortak değişken etkileri de katılabilir. Örneğin kontrol ve tedavi olmak üzere iki gruplu bir çalışmada; eğer, kesim noktaları ve eğimler bakımından gruplara bağlı olarak, ortalama yanıt zaman içinde yaklaşık doğrusal bir yapıda değişim gösteriyorsa model Denklem 1.25'deki gibi yazılabilir.

$$y_{ij} = \beta_{0,K} + \beta_{0,T} + (\beta_{1,K} + \beta_{1,T})t_{ij} + v_{0i} + v_{1i}t_{ij} + e_{ij} \quad \text{Denklem (1.25)}$$

Modelde grup etkisi ile grup ve zaman etkileşimi yer almaktadır (Verbeke ve Molenberghs, 2000; Hedeker ve Gibbons, 2006). Bu modelde, X_i matrisi kontrol grubu için;

$$X_i = \begin{pmatrix} 1 & t_{i1} & 0 & 0 \\ 1 & t_{i2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{in_i} & 0 & 0 \end{pmatrix} \quad \text{Denklem (1.26)}$$

ve tedavi grubu için;

$$X_i = \begin{pmatrix} 0 & 0 & 1 & t_{i1} \\ 0 & 0 & 1 & t_{i2} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & t_{in_i} \end{pmatrix} \quad \text{Denklem (1.27)}$$

biçiminde olacaktır.

Hem tedavi hem de kontrol grubu için Z_i matrisi aynı formdadır ve aşağıdaki gibi yazılabilir.

$$Z_i = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{in_i} \end{pmatrix} \quad \text{Denklem (1.28)}$$

Rasgele kesim noktaları ve rasgele eğimlerin yer aldığı karma etki modeli için aynı bireyden alınan gözlemler arası kovaryans yapısı incelenirken, rasgele etkiler arası kovaryans matrisi G 'nin bileşenleri $Var(v_{0i}) = g_{11}$, $Var(v_{1i}) = g_{22}$ ve $Cov(v_{0i}, v_{1i}) = g_{12}$ olarak düşünülün. Ayrıca hata kovaryans matrisinin $R_i = \sigma^2 I_{n_i}$ olduğu varsayılın. Bu durumda herhangi bir bireyden herhangi bir durum ya da zamanda alınan yanıtı ait varyans Denklem (1.29)'daki biçimde bulunabilir.

$$\begin{aligned}
Var(y_{ij}) &= Var(X'_{ij}\beta + Z'_{ij}v_i + e_{ij}) \\
&= Var(Z'_{ij}v_i + e_{ij}) \\
&= Var(v_{1i} + v_{2i}t_{ij} + e_{ij}) && \text{Denklem (1.29)} \\
&= Var(v_{1i}) + 2t_{ij}Cov(v_{1i}v_{2i}) + t_{ij}^2Var(v_{2i}) + Var(e_{ij}) \\
&= g_{11} + 2t_{ij}g_{12} + t_{ij}^2g_{22} + \sigma^2
\end{aligned}$$

Benzer biçimde aynı bireyden alınan herhangi iki yanıt arasındaki korelasyon da Denklem (1.30)'daki gibi hesaplanabilir.

$$Cov(Y_{ij}Y_{ik}) = g_{11} + (t_{ij} + t_{ik})g_{12} + t_{ij}t_{ik}g_{22} \quad \text{Denklem (1.30)}$$

Görüldüğü gibi bu modelde (rasgele kesim noktası ve rasgele eğim modelinde) herhangi iki bireyden alınan gözlemler arası kovaryans, zamanın (t_{ij}) bir fonksiyonu biçiminde tanımlanabilmektedir (Verbeke ve Molenberghs, 2000). Yanıtlar arası kovaryansa katkısı olan rasgele etkiler arası kovaryans yapısına Bölüm 1.3.8. altında değinilecektir.

1.3.7.1. İki Aşamalı Modelin Matris Gösterimi

1.3.7.1.1. İlk Aşama Modeli

İlk aşamada, her bir bireyin kendi ortalama yanıt profiline sahip olduğu varsayılır. Bu durumda, bireylere ait tekrarlı ölçümlerin, her birey için aynı ortak değişkenlere ancak farklı katsayılara sahip regresyon modelleri ile modellendikleri varsayılır.

i . bireye ait yanıt vektörü y_i olmak üzere, regresyon modeli aşağıdaki gibi yazılabilir.

$$y_i = Z_i b_i + e_i \quad \text{Denklem (1.31)}$$

Bu modelde e_i hatalarının, sıfır ortalama ve $\sigma^2 I_{n_i}$ kovaryans matrisi ile çok deęişkenli normal dağılıma sahip, ölçüm ya da örnekleme hataları oldukları varsayılır. Bireye özel regresyon parametrelerinin sayısı, tekrarlı gözlem sayıları ne olursa olsun, her birey için aynıdır. Yani her birey için b_i parametre vektörü q boyutludur. Bu regresyon katsayıları i . bireyin gerçek regresyon katsayıları olarak ya da $Z_i b_i$ i . bireyin gerçek yanıt profili açıklanabilir (Fitzmaurice ve ark., 2004).

Modelde Z_i matrisi herhangi bir bireye ait ortalama yanıt profilinin, zaman içerisinde nasıl deęiştüğünü ya da bu ortalama profilinin, zamanla deęişen ortak deęişkenler (örneğin sigara içme durumu, yaş gibi) ile nasıl deęişim gösterdiğini gösterir. Örneğin; ortalama yanıt, zamanın doğrusal, karesel ya da daha yüksek dereceli fonksiyonu olabilir. Bireye özel ortalama yanıt eğrisinin zaman içinde doğrusal olduğu bir model düşünülüğünde, ilk aşama modeli, $y_i = Z_i b_i + e_i$, Denklem (1.32)'deki gibi yazılabilir.

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in_i} \end{pmatrix} = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{in_i} \end{pmatrix} \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} + \begin{pmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{in_i} \end{pmatrix} \quad \text{Denklem (1.32)}$$

İlk aşama modelinin temel mantığı, çalışmadaki her birey için ayrı regresyon modelleri oluşturmaktır. Modelde Z_i ortak deęişkenler matrisi, kesim noktası için 1'lerden oluşan kolon dışında, yalnızca birey içi ya da zamana baęlı ortak deęişkenleri içerir. Cinsiyet, deneme grubu gibi bireyler arası ya da zamandan bağımsız ortak deęişkenler bu matris içerisinde yer alamaz, ancak etkileri kesim noktası teriminin içinde yer alacaktır. Bireyler arası ortak deęişkenler ikinci aşama modelinde kullanılacaktır.

1.3.7.1.2. İkinci Aşama Modeli

İkinci aşama modelinde, bireye özel regresyon ilk aşamada yer alan parametrelerinin rasgele oldukları varsayılır. Buna göre b_i 'lerin rasgele sayıldıklarında belli bir ortalama ve kovaryans matrisi ile bir olasılık dağılımına sahip oldukları düşünülür. İkinci aşamada, b_i 'ye ait ortalama ve kovaryans modellenir. Özellikle bireyden bireye b_i 'ler arası değişim, zamandan bağımsız ortak değişkenler yardımı ile modellenir. b_i 'ye ait ortalama, bireyler arası ortak değişkenlerin doğrusal bir fonksiyonu olarak modellenebilir. K_i bireyler arası ortak değişkenleri içeren $(q \times p)$ boyutlu tasarım matrisi olmak üzere,

$$E(b_i) = K_i \beta \quad \text{Denklem (1.33)}$$

biçiminde yazılabilir. b_i 'deki bireyler arası değişim ise kovaryans matrisi G ile tanımlanabilir, $Cov(b_i) = G$.

Kontrol ve tedavi olmak üzere iki gruptan oluşan bir klinik çalışma örneği düşünülün. İlk aşama modeli Denklem (1.32)'deki gibi yazılabilir. İkinci aşamada β_i 'lere (hem kesim noktası hem de eğim) ait ortalamanın gruptan gruba değiştiği düşünüldüğünde;

$$\text{Kontrol grubu için: } b_{0i} = \beta_{0,K} + v_{0i}, \quad b_{1i} = \beta_{1,K} + v_{1i} \quad \text{Denklem (1.34)}$$

$$\text{Tedavi grubu için: } b_{0i} = \beta_{0,T} + v_{0i}, \quad b_{1i} = \beta_{1,T} + v_{1i} \quad \text{Denklem (1.35)}$$

olarak verilebilir. Bu modelde $\beta_{0,K}$ ve $\beta_{1,K}$ kontrol grubu için ortalama kesim noktası ve eğimi gösterirken, $\beta_{0,T}$ ve $\beta_{1,T}$, deneme grubu için ortalama kesim noktası ve eğimi göstermektedir. Bu modelde bireyler arası ortak değişkenleri içeren tasarım matrisi K_i Denklem (1.36)'daki gibi gösterilebilir.

$$\text{Kontrol grubu için } K_i = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\text{Tedavi grubu için } K_i = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{Denklem (1.36)}$$

Yani, kontrol grubu için;

$$E \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{0,K} \\ \beta_{1,K} \\ \beta_{0,T} \\ \beta_{1,T} \end{pmatrix} = \begin{pmatrix} \beta_{0,K} \\ \beta_{1,K} \end{pmatrix} \quad \text{Denklem (1.37)}$$

Aynı şekilde tedavi grubu için;

$$E \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{0,K} \\ \beta_{1,K} \\ \beta_{0,T} \\ \beta_{1,T} \end{pmatrix} = \begin{pmatrix} \beta_{0,T} \\ \beta_{1,T} \end{pmatrix} \quad \text{Denklem (1.38)}$$

biçiminde yazılabilir. Böylelikle ilk ve ikinci aşama modelleri bir araya getirilerek y_i yanıt değişkeni için doğrusal karma etki modeli oluşturulabilir. Bireye özel regresyon parametreleri b_i 'ler için ikinci aşama modeli Denklem 1.39'daki gibi yazılabilir.

$$b_i = K_i \beta + v_i \quad \text{Denklem (1.39)}$$

İlk aşama modelinde b_i yerine ikinci aşamadaki fonksiyonu yazılacak olursa Denklem (1.40)'daki eşitlik elde edilebilir;

$$\begin{aligned}
y_i &= Z_i b_i + e_i \\
&= Z_i (K_i \beta + v_i) + e_i \\
&= (Z_i K_i) \beta + Z_i v_i + e_i \\
&= X_i \beta + Z_i v_i + e_i
\end{aligned}
\tag{Denklem (1.40)}$$

Burada $X_i = Z_i K_i$ olarak gösterilmiştir. Bu model koşullu model olarak da bilinir ve $E(y_i | v_i)$ biçiminde de gösterilebilir. Koşullu model yardımıyla, v_i verildiğinde bireye özel ortama yanıt profili elde edilebilir. Bütün v_i rasgele etkileri üzerinden ortalama alındığında marjinal model elde edilebilir. Marjinal model ile bütün çalışmadaki bireyler ya da çalışmadaki gruplar için ortalama yanıt profilleri belirlenebilir (Verbeke ve Molenberghs, 2000).

$$E(y_i) = Z_i K_i \beta = X_i \beta \tag{Denklem (1.41)}$$

$$Cov(y_i) = Z_i G Z_i' + \sigma^2 I_{n_i} \tag{Denklem (1.42)}$$

1.3.8. Rasgele Etkiler Kovaryans Yapısının Tekrarlı Ölçümler Arası Kovaryansa Katkısı

Denklem (1.18)'de yer alan e_i hata terimlerine ait kovaryans matrisi, $Cov(e_i) = R_i$, belirli bir bireye özel ortalama yanıt profili göz önüne alındığında, o bireye ait tekrarlı gözlemler arası kovaryansı tanımlar. Yani, i . bireye ait tekrarlı gözlemlerin, kendi ortalama yanıt profilinden, $E(y_i | v_i)$, sapmaları arasındaki kovaryansı verecektir. Şekil 1.2 ve Şekil 1.3'te A ve B bireyleri için bu sapmalar bireysel ortalama yanıt eğrilerinin üzerinde ve altında, sıfıra yaklaşık değerler ile rasgele dağılımlı olarak seyretmektedir. Önceden de bahsedildiği gibi, genellikle R_i matrisinin $\sigma^2 I_{n_i}$ biçiminde tanımlı köşegen matris olduğu varsayılır. Bu koşullu bağımsızlık varsayımı olarak da bilinir ve v_i rasgele birey etkisi verildiğinde ölçüm hatalarının genel σ^2 varyansı ile birbirinden bağımsız olarak dağıldığı varsayılır

(Fitzmaurice ve ark., 2004). Denklem (1.19) ve Denklem (1.21)' de gösterilen koşullu ve marjinal ortalamalar gibi aynı biçimde koşullu ve marjinal kovaryanslar da düşünülebilir. v_i verildiğinde y_i 'ye ait koşullu kovaryans aşağıdaki gibi yazılabilir.

$$Cov(Y_i | b_i) = Cov(e_i) = R_i \quad \text{Denklem (1.43)}$$

v_i 'nin dağılımı üzerinden ortalama alındığında y_i 'ye ait marjinal kovaryans ise aşağıdaki biçimde yazılabilir.

$$\begin{aligned} Cov(Y_i) &= Cov(Z_i b_i) + Cov(e_i) \\ &= Z_i Cov(b_i) Z_i' + Cov(e_i) \\ &= Z_i G Z_i' + R_i \end{aligned} \quad \text{Denklem (1.44)}$$

Marjinal kovaryans yapısına bakıldığında, R_i matrisi $\sigma^2 I_{n_i}$ biçiminde köşegen bir matris olsa da, aynı v_i rasgele etkilerinin modele alınmasıyla, aynı bireye ait tekrarlı gözlemler arası bir ilişki ortaya çıkacağı görülebilir. Bunun yanı sıra, tekrarlı gözlemlere ait marjinal kovaryans matrisi, bazıları G bazıları da R_i matrislerine ait kovaryans parametrelerinden oluşmaktadır. Böylelikle karma etki modelleri, yanıtlara ait birey içi (R_i) ve bireyler arası (G) değişkenliğin çözümlenmesini sağlar (Fitzmaurice ve ark., 2004). Ayrıca marjinal kovaryans, Z_i matrisinin kovaryans yapısında yer almasından dolayı, zamanın bir fonksiyonu biçiminde yazılmış olur. Örneğin; hem rasgele kesim noktaları hem de rasgele eğimlerin olduğu modelin gösterildiği Denklem (1.16) için varyans ve kovaryans tekrarlı ölçümlerin alındığı zamanların birer fonksiyonu olarak yazılabilir.

$$Var(Y_{ij}) = g_{11} + 2t_{ij} g_{12} + t_{ij}^2 g_{22} + \sigma^2 \quad \text{Denklem (1.45)}$$

$$Cov(Y_{ij}, Y_{ik}) = g_{11} + (t_{ij} + t_{ik}) g_{12} + t_{ij} t_{ik} g_{22}$$

Böylelikle aynı model için ortaya çıkacak varyans kovaryans matrisi aşağıdaki biçimde yazılabilir;

$$Cov(Y_i) = Z_i G Z_i' + \sigma^2 I_{n_i} \quad \text{Denklem (1.46)}$$

Böyle bir kovaryans yapısında birey içi ve bireyler arası varyasyon kaynakları açık olarak yazılmıştır. Bunun yanı sıra bu yapıda, bireylere ait tekrarlı gözlem sayılarının eşit olması ve/veya gözlemlerin alındığı zaman noktalarının aynı olması şart değildir. Ayrıca rasgele etkiler kovaryans yapısında zaman içerisinde varyans homojenliğinin varsayımı gerekli değildir. Varyans ya da kovaryanslar zaman içinde artabilir ya da azalabilirler.

1.4. Kestirim Yöntemleri

Önceden de belirtildiği gibi, ortalama yanıt vektörü y_i 'nin $E(y_i) = X_i \beta$ ortalama ve $Cov(y_i) = \Sigma_i = \Sigma_i(\theta)$ kovaryans matrisi ile çok değişkenli normal dağılım gösterdiği varsayılmaktadır. Burada θ , $q \times 1$ boyutlu Σ_i 'yi oluşturan kovaryans parametreleri vektörüdür. Örneğin, $n_i = n$ olmak üzere dengeli bir veri yapısı için, tekrarlı ölçümler için yapılandırılmamış kovaryans matrisi varsayıldığında, θ 'nın bileşenleri n sayıda varyans ve $n(n-1)/2$ kovaryans parametresinden oluşacaktır (toplam parametre sayısı $q = n(n+1)/2$ olacaktır). Diğer bir yandan, eğer kovaryans matrisi tam simetrik olarak varsayıldıysa, $q = 2$ olacak ve θ 'nun iki parametresi varyanslar, genel bir değeri ve kovaryanslar, genel bir değeri gösterecektir.

Bu bölümde, bilinmeyen parametreler β ve θ (veya Σ_i) için en çok olabilirlik (EÇOK) ve sınırlı en çok olabilirlik (SEÇOK) kestirim yöntemleri üzerinde durulacaktır. Analiz için kullanılan SAS programında PROC MIXED işlemi altında en çok olabilirlik ve sınırlı en çok olabilirlik yöntemlerinin yanı sıra, Rao ve LaMotte tarafından 1972'de geliştirilen en küçük normlu karesel yansız kestirim, EKNKYK, (Minimum Norm Quadratic Unbiased Estimation, MINQUE) yöntemi de

yer almaktadır. Ancak Swallow ve Monahan (1984) tarafından yapılan simülasyon çalışmasının sonuçları, EÇOK ve SEÇOK yöntemlerinin daha iyi sonuçlar verdiğini göstermiştir. Ancak çok geniş veri setinde, EÇOK ve SEÇOK yöntemlerinin yakınsayamadığı durumlarda EKNKYK'in kullanılması önerilmiştir.

1.4.1. En Çok Olabilirlik Kestirimi (EÇOK)

Eldeki veri setine dayanan modeli karakterize edecek parametreler kestirilmek istendiğinde, yanıt vektörü y_i 'nin çok değişkenli normal dağılım gösterdiği varsayımı altında en çok olabilirlik (EÇOK) yöntemi genel bir yaklaşım olarak kullanılmaktadır. EÇOK kestiriminin temelinde yatan mantık, eldeki veri için β ve θ 'nın en olabilir (en muhtemel) kestirimlerinin bulunmasıdır. Yani eldeki veri tarafından en iyi desteklenen parametre değerinin bulunmasıdır (Davidian, 2007). β ve θ 'nın en çok olabilirlik kestirimleri, yanıt değişkeninin gözlenen değerlerinin bileşik olasılık fonksiyonunun değerini maksimum yapan β ve θ değerlerdir. Bu yöntem ile tahminde hareket noktası olabilirlik fonksiyonudur. Olabilirlik fonksiyonu, $L(\beta, \Sigma_i | y_i)$, yanıt vektöründe yer alan değişkenlerin bileşik olasılık yoğunluk fonksiyonudur. β ve θ en çok olabilirlik kestirimleri olabilirlik fonksiyonu maksimum yapan değerlerdir ve genellikle $\hat{\beta}$ ve $\hat{\Sigma}_i$ (ya da $\Sigma_i(\hat{\theta})$) ile gösterilir (Fitzmaurice ve ark, 2004).

Bu yöntemin anlaşılır olması açısından, öncelikle bütün gözlemlerin bağımsız olduğu, yani tek değişkenli normal dağılıma sahip, bağımsız hata terimlerinin olduğu standart doğrusal regresyon modeli göz önüne alınacaktır. Sonrasında ise, gözlemler bağımlı olduğunda EÇOK yönteminin işleyişi anlatılacaktır.

1.4.1.1. Bağımsız Ölçümler İçin EÇOK

Toplanan verilerin, n farklı durumda tekrarlanan bir kesitsel çalışmalar dizisi sonucu ortaya çıktığı düşünölsün. Her durumda veriler N farklı bireyden elde edilmiş olsun. Böyle bir durumda, her bireyden yalnızca bir ölçüm alındığı için verilerin birbirinden bağımsız olduğu söylenebilir. Ayrıca gösterimde kolaylık açısından varyansın sabit olduğu, σ^2 düşünölsün. Böylelikle ortalama yanıt ortak değişkenlerle ilişkili olarak Denklem (1.47)'deki doğrusal regresyon modeli ile gösterilebilir.

$$E(y_{ij}) = X'_{ij} \beta \quad \text{Denklem (1.47)}$$

β 'ya ait en çok olabilirlik kestirimlerini elde etmek için, olabilirlik fonksiyonunu maksimum yapan regresyon parametrelerinin değerlerinin bulunması gereklidir. y_{ij} için tek değişkenli normal olasılık yoğunluk fonksiyonu aşağıdaki gibi yazılabilir;

$$f(y_{ij}) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2}(y_{ij} - \mu_{ij})^2 / \sigma^2\right\}, \quad -\infty < y_{ij} < \infty \quad \text{Denklem (1.47)}$$

Bütün gözlemler birbirinden bağımsız olduğu durumda, olabilirlik fonksiyonu bütün bireylere ait y_{ij} için tek değişkenli normal olasılık yoğunluk fonksiyonlarının çarpımı biçiminde olacaktır (Denklem 1.48).

$$\prod_{i=1}^N \prod_{j=1}^n f(y_{ij}) \quad \text{Denklem (1.48)}$$

Toplamsal bir modelle çalışmak daha kolay olduğundan bu çarpımsal modelin logaritmasının alınarak toplamsal bir model haline getirilmesi çok genel bir yaklaşımdır. Bu durumda, olabilirliğin logaritmasını maksimum yapmak, olabilirliği maksimum yapmaya eşdeğer olarak düşünölebilir. Olabilirliğin logaritması l ve

toplam gözlem sayısı $K = n \times N$ ile gösterildiğinde, Denklem (1.49)'un β regresyon parametreleri açısından maksimum yapılması amaçlanır.

$$l = \log \left\{ \prod_{i=1}^N \prod_{j=1}^n f(y_{ij}) \right\} = -\frac{K}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^n (y_{ij} - X'_{ij}\beta)^2 / \sigma^2 \quad \text{Denklem(1.49)}$$

Dikkat edilecek olursa log-olabilirliğin ilk teriminde β yer almamaktadır. Dahası, ikinci terim negatif işaretli olduğu için log-olabilirliği β 'ya göre maksimum yapmak aşağıdaki ifadeyi minimum yapmak demektir.

$$\sum_{i=1}^N \sum_{j=1}^n (y_{ij} - X'_{ij}\beta)^2 \quad \text{Denklem (1.50)}$$

Herhangi bir fonksiyonu maksimum ya da minimum yapmak genel matematiksel bir problemdir. β 'nın en çok olabilirlik kestirimi log-olabilirlik fonksiyonunun türevini sıfıra eşitleyerek edilebilir. Ancak burada ele alınan durum için, türeve başvurmaya gerek yoktur. β 'nın en çok olabilirlik kestirimini elde etmek, β 'nın hata kareler toplamını (Denklem 1.50) minimum yapan sıradan en küçük kareler kestirimini elde etmek ile aynıdır. Vektör gösterimi kullanılarak en küçük kareler çözümü Denklem (1.51)'deki gibi yazılabilir (Fitzmaurice ve ark., 2004);

$$\hat{\beta} = \left\{ \sum_{i=1}^N \sum_{j=1}^n (X_{ij} X'_{ij}) \right\}^{-1} \sum_{i=1}^N \sum_{j=1}^n (X_{ij} y_{ij}) \quad \text{Denklem (1.51)}$$

Buraya kadar gözlemlerin bağımsız olduğu durum göz önüne alınarak, yalnızca β kestirimi üzerinde durulmuş ve σ^2 için bir kestirimden bahsedilmemiştir. Bundan sonraki bölümde gözlemler bağımlı olduğunda kestirim problemi üzerinde durulacak, β 'nın yanı sıra kovaryans matrisi de göz önünde tutulacaktır.

1.4.1.2. İlişkili Ölçümler İçin EÇOK

Aynı bireyden alınan n_i gözlem olduğunda, bu gözlemlerin birbirleri ile bağımsız oldukları varsayılmaz. Bu durumda, olabilirlik fonksiyonunun oluşturulabilmesi için, bireylere ait tekrarlı gözlem vektörlerinin bileşik olasılık yoğunluk dağılımları düşünülür. Burada yanıt vektörleri birbirlerinden bağımsız, ancak vektör içi bileşenler bağımlı olacaktır. Her bir yanıt vektörü y_i 'nin çok değişkenli normal dağılıma sahip olduğu varsayılır. Sonuç olarak, log-olabilirlik fonksiyonu l , $y_i = (y_{i1}, y_{i2}, \dots, y_{in_i})'$ yanıt vektörlerine ait çok değişkenli normal dağılımların toplamları biçiminde olacaktır.

Tekrarlı gözlemler yapısında, β 'nin en çok olabilirlik kestirimini bulmak için öncelikle Σ_i 'nin (ya da θ 'nin) bilindiği varsayılacaktır. Sonrasında, daha gerçekçi bir durum olan varyansın bilinmediği durum ele alınacaktır. β 'nin en çok olabilirlik kestirimini bulmak için log-olabilirlik fonksiyonunu maksimum yapan değer bulunması gereklidir. $y_i = (y_{i1}, y_{i2}, \dots, y_{in_i})'$ çok değişkenli normal dağılıma sahip olduğu verildiğinde ve $K = \sum_{i=1}^N n_i$ olarak bütün gözlem sayısını gösterdiğinde, log-olabilirlik fonksiyonu Denklem (1.52)'de görüldüğü gibi yazılabilir (Fitzmaurice ve ark., 2004).

$$l = -\frac{K}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^N \log|\Sigma_i| - \frac{1}{2} \left\{ \sum_{i=1}^N (y_i - X_i \beta)' \Sigma_i^{-1} (y_i - X_i \beta) \right\} \quad \text{Denklem (1.52)}$$

Log-olabilirlik fonksiyonunun ilk iki bileşeni β 'yi içermediğinden olabilirliği β açısından maksimum yaparken dikkate alınmaz. Ayrıca, son bileşen negatif işaretli olduğu için log-olabilirliği β açısından maksimum yapmak, aşağıdaki ifadeyi minimum yapmakla aynı sonucu verecektir.

$$\sum_{i=1}^N (y_i - X_i \beta)' \Sigma_i^{-1} (y_i - X_i \beta) \quad \text{Denklem (1.53)}$$

Yukarıdaki ifadeyi minimum yapan β kestiricisi, β 'nin genelleştirilmiş en küçük kareler kestiricisi olarak da bilinir ve aşağıdaki gibi gösterilebilir (Littell ve ark., 2006).

$$\hat{\beta} = \left\{ \sum_{i=1}^N (X_i' \Sigma_i^{-1} X_i) \right\}^{-1} \sum_{i=1}^N (X_i' \Sigma_i^{-1} y_i) \quad \text{Denklem (1.54)}$$

Varyansın bilindiği varsayıldığında, Σ_i 'nin herhangi bir formu için β 'nin ağırlıklı en küçük kareler kestiricisi yansızdır. Yani; $E(\hat{\beta}) = \beta$ olarak gösterilebilir. Ayrıca, büyük örneklerde (ya da asimptotik olarak) $\hat{\beta}$ 'nin örneklem dağılımı β ortalama ve $Cov(\hat{\beta}) = \left\{ \sum_{i=1}^N X_i' \Sigma_i^{-1} X_i \right\}^{-1}$ kovaryans matrisi ile çok değişkenli normal dağılım gösterir. Bu durum y_i yanıt vektörü çok değişkenli normal dağılıma sahip olduğunda kesinlikle doğrudur. Ayrıca büyük örneklerde y_i çok değişkenli normal dağılım göstermese bile doğrudur. Büyük örneklem ile kastedilen, tekrarlı gözlem sayısı ve parametre sayısı değişmezken denek sayısı N 'in arttığı durumdur. Yani, β 'nin genelleştirilmiş en çok olabilirlik kestiricisinin önemli bir özelliği, y_i çok değişkenli normal dağılım göstermese de geçerli olacaktır (Fitzmaurice ve ark., 2004).

Bunların yanı sıra, eğer Σ_i bütün diyagonal öğeleri σ^2 sabit varyansına eşit olan ve korelasyonların sıfır olduğu diyagonal matris ise, bütün hata kareleri aynı değere bölüneceğinden, genelleştirilmiş en küçük kareler kestiricisi standart en küçük kareler kestiricisine eşit olacaktır. Son olarak, Σ_i 'nin herhangi bir seçimi için β 'nin genelleştirilmiş en çok olabilirlik kestiricisi yansız olacaktır ancak, en küçük varyansa sahip kestirici Σ_i 'nin doğru değeri kullanıldığında bulunabilecektir (Fitzmaurice ve ark., 2004).

Şimdiye kadar varyansın, Σ_i ya da θ , bilindiği varsayımı altında β 'nin kestiriminden bahsedildi. Ancak gerçekte varyans bilinmediğinden eldeki veri seti kullanılarak kestirilmesi gereklidir. θ 'nın en çok olabilirlik kestirimi için izlenen yol β 'nin kestirimi için izlenen yolun aynısıdır. Yani, θ 'nın en çok olabilirlik kestirimi log-olabilirlik fonksiyonunun θ açısından maksimum yapılması ile bulunabilir. Daha önce de bahsedildiği gibi bu değer log-olabilirlik fonksiyonunun θ 'ya göre türevinin sıfıra eşitlenmesi ile bulunabilir. Ancak, genelde θ 'ya göre türev alındığında bu fonksiyon doğrusal olmadığından ve θ 'nın en çok olabilirlik kestiricisinin basit kapalı bir formu yazılmadığından, iterasyon yöntemleri kullanılarak çözülmeye çalışılır. Bunun için bilgisayar yazılımları kullanılmaktadır. θ için en çok olabilirlik kestirimi elde edildiğinde β 'nin genelleştirilmiş kareler kestiricisinde yerine yazılarak β için en çok olabilirlik kestirimi elde edilmiş olur.

$$\hat{\beta} = \left\{ \sum_{i=1}^N (X_i' \hat{\Sigma}_i^{-1} X_i) \right\}^{-1} \sum_{i=1}^N (X_i' \hat{\Sigma}_i^{-1} y_i) \quad \text{Denklem (1.55)}$$

Büyük örneklerde (asimptotik olarak) bu kestiricinin, varyansın bilindiği durumda elde edilen kestirici ile aynı özelliklere sahip olduğu söylenebilir (Laird ve Ware, 1982). Yani, büyük örneklerde $\hat{\beta}$, β 'nin tutarlı bir kestiricisidir. Eğer hataların, e_i , dağılımının normal ya da en azından simetrik olduğu varsayılıyor ise $\hat{\beta}$, β 'nin yansız bir kestiricisidir, $E(\hat{\beta}) = \beta$. Ayrıca Σ_i eldeki veri setinden kestirildiğinde $\hat{\beta}$ 'nin örneklem dağılımı β ortalama ve $Cov(\hat{\beta}) = \left\{ \sum_{i=1}^N X_i' \hat{\Sigma}_i^{-1} X_i \right\}^{-1}$ kovaryans matrisi ile yaklaşık olarak çok değişkenli normal dağılım olacaktır.

Ayrıca, kayıp gözlem olmamak koşulu ile y_i çok değişkenli normal dağılıma sahip olmasa bile, büyük örneklerde bu özellikler $\hat{\beta}$ kestiricisi için geçerli olacaktır. Böylelikle, β 'nin en çok olabilirlik kestiriminin önemli bir özelliği, çok değişkenli normal dağılım varsayımı sağlanmasa bile geçerli bir kestirim olmasıdır. β 'nin en çok olabilirlik kestiriminin ve ayrıca genelleştirilmiş kareler kestiriminin

bu özelliği bazı kayıp gözlem durumlarında da geçerlidir (β 'nin en çok olabilirlik kestirimi, $\Sigma_i(\theta)$ 'nin EÇOK kestiriminin kullanıldığı genelleştirilmiş en küçük kareler kestirimine eşdeğer olduğundan bahsedilmiştir). Ancak örnek sayısı az olduğu durumlarda β 'nin kestirimi çok sayıda kovaryans parametresinin kestirilmesinden (dolayısıyla seçilen kovaryans yapısından) olumsuz olarak etkilenebilir (Fitzmaurice ve ark., 2004).

1.4.2. Kayıp Gözlem Durumları

Sağlık alanında yapılan uzunlamasına çalışmalarda, yüksek ihtimalle kayıp gözlemler ortaya çıkacaktır. Kayıp gözlemlerin olması durumunda, her birey aynı sayıda gözleme sahip olmayacağından kullanılan istatistiksel analizler, kayıp gözleme sahip bireyleri analiz dışında bırakmadan yapılabilmelidir. Ayrıca kayıp gözlemlerin olması veride bilgi eksikliğine yol açacağından, zaman içindeki ortalama yanıtta değişim hakkında yapılan çıkarımların doğruluğunda azalma olacaktır. Bu azalma ise kayıp gözlem miktarından ve kayıp gözlem durumunun nasıl ele alındığından etkilenecektir. Örneğin, kayıp gözleme sahip bireylerin analizden çıkartılması genellikle en az etkili yöntemdir. Ayrıca kullanılan yöntemlerin geçerliliği açısından kayıp gözlemlerin ortaya çıkış nedenleri, kayıp gözlem mekanizması hakkında belli varsayımların da sağlanması gerekebilir. Sonuç olarak, kayıp gözlemlerin ortaya çıkış nedenleri dikkatlice ele alınmalıdır.

Burada iki genel kayıp gözlem mekanizması üzerinde durulacaktır. İki mekanizma da, kayıp gözlemlerin gözlenen yanıtlarla ilişkili olup olmadığına göre farklılık gösterir.

Kayıp gözlem mekanizması, herhangi bir durum için verinin kayıp ya da gözlenmiş olması olasılığını tanımlayan bir model olarak düşünülebilir. Tamamen rasgele ve rasgele olarak bilinen iki kayıp gözlem mekanizması arasında önemli farklılık vardır. Bu iki mekanizma arasındaki fark, yanıtlar için çok değişkenli normal dağılım varsayımı altında en çok olabilirlik kestirimiyle, dağılımın şekli

hakkında varsayıma sahip olmayan genelleştirilmiş en küçük kareler kestiriminin uygunluğunu belirler.

Kısaca tanımlanacak olursa; eğer kayıp gözlemler hem gözlenen hem de gözlenmeyen değerlerden bağımsız ise kayıp gözlem mekanizması tamamen rasgeledir. Eğer gözlenmeyen değerlerden bağımsız, yalnızca gözlenen değerlere bağımlı ise rasgeledir. Hem gözlenen hem de gözlenmeyen değerlere bağımlı olduğunda ise kayıp gözlem mekanizmasından rasgele olmayan biçimde bahsedilebilir. Tamamen rasgele yapısının özelliği elde var olan verinin tamamlanmış verinin bir rasgele örnekleme olarak düşünülebilir olmasıdır. Sonuç olarak ortalama, varyans ve kovaryans gibi momentler ve eldeki verinin dağılımı tamamlanmış verinin karşılık gelen moment değerlerinden ve yine tamamlanmış verinin dağılımından farklı olmayacaktır (Molenberghs ve Verbeke, 2001).

Tamamen rasgele mekanizmasının uzunlamasına veri analizi açısından önemli sonuçları vardır. Kayıp gözlem olmadığı durumda geçerli sonuçlar veren bir yöntem, kayıp gözlem mekanizması tamamen rasgele olduğunda ve eldeki bütün veri analizde kullanıldığında yine geçerli sonuçlar verecektir. Ortalamalar, varyanslar ve kovaryanslar için geçerli kestirimler elde edilebildiği bilindiğinde, genelleştirilmiş en küçük kareler yöntemi y_i 'nin dağılımı hakkında varsayımlar gerektirmeden β için geçerli tahminler verebilir. β 'nin GEKK kestiricisi ortalama yanıt doğru modellendiği takdirde geçerli olacaktır ve uzunlamasına verilerin bileşik dağılımı hakkında varsayımlar gerektirmeyecektir. Önce de belirtildiği gibi, y_i 'nin çok değişkenli normal dağılım gösterdiği varsayımı altında β 'nin en çok olabilirlik kestirimi, genelleştirilmiş en küçük kareler kestiricisine eşittir ($\Sigma_i(\theta)$ yerine, EÇOK kestirimi $\hat{\Sigma}_i = \Sigma_i(\hat{\theta})$ yazıldığında, Denklem 1.55). Yani, bu çerçevede EÇOK ve genelleştirilmiş en küçük kareler kestiricileri tamamen aynı özelliklere sahip olacaklardır (Fitzmaurice ve ark., 2004).

Rasgele kayıp gözlem yapısında, kayıp gözlem mekanizması gözlenen değerlere bağımlı olarak değişecektir. Eğer bireyler verdikleri benzer yanıtlar temel

alınarak tabakalara ayrılacak olurlarsa, her bir tabakada kayıp gözlem olması gözlenmeyen değerlere bağlı olmayan bir şans mekanizması olarak düşünülebilir. Ancak bu durumda, kayıp gözlem yapıları tarafından tanımlanan her bir ayrı tabakada y_i 'nin dağılımı, hedef popülasyondaki y_i 'nin dağılımı ile aynı olmayacaktır (Molenberghs ve Verbeke, 2001). Böylece, eldeki veri setinden elde edilen örneklem ortalamaları, varyansları ve kovaryansları karşılık gelen popülasyon parametrelerinin yanlı kestirimleri olacaktır. Sonuç olarak, genelleştirilmiş en küçük kareler kestirimi, uzunlamasına yanıtların bileşik olasılık fonksiyonu için doğru varsayımlar yapılmadıkça β 'nin kestirimi için geçerli sonuçlar vermeyecektir. Diğer yandan, kayıp gözlem yapısı rasgele iken, β 'nin EÇOK kestirimi, çok değişkenli normal dağılım doğru olarak tanımlanması durumunda geçerli olacaktır. Bu durum yalnızca ortalama yanıtın değil, yanıtlar arası kovaryans yapısının da düzgün biçimde belirlenmiş olmasını gerektirir (Fitzmaurice ve ark., 2004).

Kısacası, tamamen rasgele ve rasgele arasındaki fark, çok değişkenli normal dağılım altında en çok olabilirlik kestiriminin ve herhangi bir dağılım biçimi varsaymayan genelleştirilmiş en küçük kareler kestiriminin uygunluğunu belirler. GEKK ile bulunan kestirimlerin geçerli olabilmesi için verinin kayıp gözlem içermemesi ya da kayıp gözlem mekanizmasının tamamen rasgele olması gereklidir. Kayıp gözlem mekanizması rasgele ise, β için bulunan GEKK kestirimleri yanlı olacaktır. Tersine, EÇOK yöntemi β için kayıp gözlem mekanizması tamamen rasgele ya da rasgele olduğu durumlarda geçerli kestirimler verecektir. Ancak, kayıp gözlem mekanizması rasgele olduğunda yanıt vektörlerinin bileşik dağılımı doğru olarak tanımlanmalıdır (Fitzmaurice ve ark., 2004).

1.4.3. Sınırlı En Çok Olabilirlik Kestirimi (SEÇOK)

Bölüm 1.4.1.2.'de anlatıldığı gibi, β ve θ (ya da Σ_i)'nin en çok olabilirlik kestirimi Denklem (1.52)'de verilen log-olabilirlik fonksiyonu kullanılarak bulunuyordu. β ve Σ_i 'nin EÇOK kestirimleri büyük örneklerde istenen sonuçları veriyor olsalar

da, sınırlı örneklemlerde Σ_i 'nin EÇOK kestirimi yanlı sonuçlar verebilmektedir (Fitzmaurice ve ark., 2004). Örneğin, Σ_i 'nin köşegen elemanları gerçekte olduğundan daha düşük kestirilebilir. Bu sorunu göstermek amacıyla n tekrarlı kesitsel çalışmanın yapıldığı ve her tekrarda farklı bireylerden ölçüm alındığı düşünölsün. Burada ölçümlerin birbirinden bağımsız olduğu düşünölebilir ve gösterim kolaylığı açısından sabit varyans varsayılabilir, σ^2 . Bölüm 1.4.1.1.'de açıklandığı gibi β ve σ^2 'nin EÇOK kestirimleri aşağıdaki ifadenin maksimum yapılması ile elde edilebilir.

$$-\frac{K}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^n (y_{ij} - X'_{ij}\beta)^2 / \sigma^2$$

Buradan β 'nin EÇOK kestiricisi, Denklem (1.51)'den;

$$\hat{\beta} = \left\{ \sum_{i=1}^N \sum_{j=1}^n (X_{ij}X'_{ij}) \right\}^{-1} \sum_{i=1}^N \sum_{j=1}^n (X_{ij}y_{ij}),$$

aynı şekilde σ^2 'nin EÇOK kestiricisi;

$$\hat{\sigma}^2 = \sum_{i=1}^N \sum_{j=1}^n (y_{ij} - X'_{ij}\hat{\beta})^2 / K \quad \text{Denklem (1.56)}$$

olacaktır. Burada $K = n \times N$ olmak üzere toplam gözlem sayısını göstermektedir. p , β 'nin boyutunu göstermek üzere, $\hat{\sigma}^2$ 'nin beklenen değeri Denklem (1.57)'deki gibi gösterilebilir.

$$E(\hat{\sigma}^2) = \left(\frac{K-p}{K} \right) \sigma^2 \quad \text{Denklem (1.57)}$$

Bu ifadeden de görüldüğü gibi σ^2 'nin EÇOK kestirimi küçük örneklerde yanı olacaktır ve σ^2 'den daha küçük bir değer bulunacaktır. Yansız kestirici, Denklem (1.56)'da paydadaki K yerine $K - p$ yazılarak elde edilebilir. Böylelikle yeni kestirici;

$$\hat{\sigma}^2 = \sum_{i=1}^N \sum_{j=1}^n (y_{ij} - X'_{ij} \hat{\beta})^2 / (K - p)$$

biçiminde yazılabilir. Bu kestirici σ^2 'nin sınırlı en çok olabilirlik (SEÇOK) kestiricisi olarak bilinir (Fitzmaurice ve ark., 2004). Varyans kestiriminde yanlılığın ortaya çıkmasının nedeni, EÇOK kestiriminin β 'nin de eldeki veriden yararlanarak kestirilmesini hesaba katmamasından kaynaklanmaktadır (Littell ve ark., 2006). Yani, σ^2 'nin kestiriminde β yerine $\hat{\beta}$ yazıldığı, fakat bu kestirimin de eldeki veri kullanılarak yapıldığı bilgisi hesaba katılmadığından yanlılık ortaya çıkacaktır. Aynı mantıkla, bağımsız gözlemler için σ^2 'nin kestiriminde yanlılık problemi ortaya çıkıyorsa, bağımlı gözlemler için de Σ_i 'nin (ya da θ 'nın) kestiriminde aynı sorunun ortaya çıkabileceği düşünülebilir.

SEÇOK, bu sorunu çözmek için geliştirilmiş bir yöntemdir. SEÇOK'nin ana düşüncesi, verinin Σ_i 'nin kestiriminde kullanılan kısmından, β 'nin kestiriminde kullanılan kısmını ayırmaktır (Fitzmaurice ve ark., 2004). Böylelikle, Σ_i 'nin kestirimi eldeki verinin yalnızca ilgili kısmına dayanacaktır. Yani, Σ_i 'nin sınırlı en çok kestirimindeki temel düşünce, β 'yi olabilirlikten çıkarmaktır. Böylelikle kestirici yalnızca Σ_i bakımından tanımlanır. Bunu yapmanın çeşitli yolları olabilir. Sınırlı olabilirliği elde etmek için mümkün olan bir yollardan biri; verinin, gözlemlerin dağılımı β 'ya bağlı olmayan doğrusal birleşimlerinin bir kümesine dönüştürülmesidir. Örneğin, β 'nin en küçük kareler kullanılarak kestiriminden sonra hatalar Σ_i 'nin (ya da θ 'nın) kestiriminde veri olarak kullanılabilirler. Bu hatalar için olabilirlik β 'ya değil, yalnızca θ 'ya bağlı olacaktır. Yani,

$$-\frac{1}{2} \sum_{i=1}^N \log|\Sigma_i| - \frac{1}{2} \left\{ \sum_{i=1}^N (y_i - X_i \hat{\beta})' \Sigma_i^{-1} (y_i - X_i \hat{\beta}) \right\}$$

ile gösterilen log-olabilirliği maksimum yapmak yerine, SEÇOK aşağıdaki değiştirilmiş log-olabilirliği maksimum yapacaktır (Fitzmaurice ve ark., 2004).

$$-\frac{1}{2} \sum_{i=1}^N \log|\Sigma_i| - \frac{1}{2} \sum_{i=1}^N (y_i - X_i \hat{\beta})' \Sigma_i^{-1} (y_i - X_i \hat{\beta}) - \frac{1}{2} \log \left| \sum_{i=1}^N X_i' \Sigma_i^{-1} X_i \right| \quad \text{Denklem (1.58)}$$

Denklem (1.58) ile gösterilen hata olabilirliğinin maksimum yapılması ile, β 'nin da kestirilmiş olduğu gerçeği dikkate alınarak, θ 'nin kestirimi elde edilmiş olur. Dikkat edilecek olursa, SEÇOK kestiricisindeki ek terim, $\hat{\beta}$ 'nin kovaryansı biçiminde gösterilebilir (Denklem 1.59).

$$\begin{aligned} -\frac{1}{2} \log \left| \sum_{i=1}^N X_i' \Sigma_i^{-1} X_i \right| &= \frac{1}{2} \log \left| \left(\sum_{i=1}^N X_i' \Sigma_i^{-1} X_i \right)^{-1} \right| \\ &= \log |Cov(\hat{\beta})|^{1/2} \end{aligned} \quad \text{Denklem (1.59)}$$

Sonuç olarak, gözlemlerin bağımsız olduğu durumda σ^2 kestirimindeki düzeltme gibi burada da, en çok olabilirlik kestiriminde düzeltme yapılarak sınırlı en çok olabilirlik kestirimi elde edilmiş olur. Fitzmaurice ve ark. (2004), Σ_i 'nin kestiriminde SEÇOK kullanılmasını önermiş ve Σ_i için SEÇOK kestiricisinin EÇOK kestiricisine göre daha az yanlı olacağını belirtmişlerdir. EÇOK ile SEÇOK arasındaki farkın, örnek büyüklüğü, N , β vektörünün boyutu p 'den yeteri kadar büyük olduğunda, daha az önemli olacağı da aynı çalışmada belirtilmiştir.

Σ_i 'nin kestiriminde SEÇOK kullanıldığında, β 'nin kestirimi genelleştirilmiş en küçük kareler kestiricisi (Denklem 1.55) kullanılarak yapılır. Denklem (1.55)'te, $\hat{\Sigma}_i$, Σ_i 'nin SEÇOK kestirimidir. Yine EÇOK kestiriminde olduğu gibi, Σ_i 'nin

kestirimi iteratif yöntemler kullanılarak yapılır ve Σ_i için sınırlı en çok olabilirlik kestirimi elde edildiğinde β 'nın genelleştirilmiş kareler kestiricisinde yerine yazılarak β için en çok olabilirlik kestirimi elde edilmiş olur.

1.5. Rasgele Etkiler İçin Önkestirim

Uzunlamasına veri analizinin yer aldığı uygulamalarda çıkarsamalar çoğunlukla β sabit etkileri için yapılır. Bu regresyon parametreleri, ortalama yanıtın zaman içindeki değişimi ve bu değişimin ortak değişkenlerle ilişkisini tanımlar. Ancak bazı çalışmalarda amaç, bireye özel yanıt profillerinin kestirilmesi olabilir. Örneğin, büyümeyi takip eden çalışmalarda bireysel büyüme eğrileri ile ilgilenilebilir, ya da uygulanan tedaviye zaman içerisinde en yüksek ya da en düşük yanıtı verecek bireyler tanımlanmak istenebilir. Karma etki modelleri sabit ve rasgele etkileri ayırdığı için bireye özel yanıt eğrileri de kestirilebilir. Yani, bireysel yanıt davranışları tanımlanmak isteniyorsa, β 'nın yanı sıra v_i 'nin de kestirilmesi gereklidir. Böylelikle belli ortak değişken düzeylerine sahip herhangi bir birey için b_i elde edilebilir. Teknik olarak v_i , β gibi sabit bir değer olmayıp bireyler arasında rasgele değiştiğinden, kestirim deyimi tam olarak anlamını bulmayacaktır. Ancak burada, b_i için “kestirim” ile rasgele olan bir değeri tanımlamak kastedilecektir. Bu nedenle rasgele etkiler için “kestirim” yerine “önkestirim” deyimi uygun olacaktır.

Sıradan bir regresyon analizinde kestirim problemi, veri setinde yer almayan bağımsız değişken değerleri için ortaya çıkabilecek bağımlı değişken değerini kestirmek gerektiğinde ortaya çıkabilir. Bu durumda x_0 gibi ortak değişkenlerin belirli bir setini içeren bir ortak değişken vektörüne karşılık gelecek bağımlı değişken değeri için en iyi tahmin, yine aynı x_0 vektörü için gözlemlenebilecek ortalama y değeri, $x_0'\beta$ olabilir. β değeri bilinmediğinden, ortalama da bilenemeyeceği için, β 'nın tahmini $\hat{\beta}$ kullanılır. Aynı mantık izlenecek olursa, v_i 'lerin kestirimini bulabilmek için v_i 'lerin popülasyonuna ait ortalama değerini kestirmek

düşünülebilir. Ancak, karma etki modelinde yer alan varsayım, v_i 'lerin sıfır ortalama ve G kovaryans matrisi ile çok değişkenli normal dağılıma sahip olduklarıdır. Yani, bütün i 'ler için $E(v_i) = 0$ olacaktır. Bu durumda herhangi bir birey için v_i 'nin kestirimi olarak 0 kullanılacaktır. Bu da b_i gibi bireye özel değerlerin, bütün bireyler için, aynı kestirime sahip olmasına neden olacaktır. Ancak asıl düşünce bireylerin birbirinden farklılık gösterdiğidir ve bu çözüm yolu her birey için aynı sonucu veriyor olmasından dolayı kullanışlı değildir. Bu nedenle v_i 'lere ait popülasyonun ortalamasını kullanmak uygun bir sonuç vermeyecektir. Bu çözüm yolu hakkında olumsuz diğer bir nokta da, eldeki veri setine ait ekstra bilgiden yararlanmıyor oluşudur. Genel model, $y_i = X_i\beta + Z_iv_i + e_i$, altında yanıt değişkeni, y_i , ve rasgele etkilerin, v_i , ilişkili olduğu görülmektedir. Bu da, y_i içinde v_i hakkında kullanılabilir bilginin var olması gerektiğini gösterir. Böylelikle v_i 'nin kestiricisi olarak kullanılabilir, y_i verisini kullanan bir fonksiyon olacağı düşünülebilir. y_i rasgele olduğundan, bu fonksiyon da rasgele olacaktır (Davidian, 2007).

1.5.1. Koşullu Beklenen Değer

Daha anlaşılır olması açısından, bir an için v_i vektörünün tek bir değer içerdiği durum düşünülün, $q = 1$ gibi. Aynı mantık $q > 1$ için de geçerli olacaktır. Bu değere rasgele etki v_i denilsin. Eğer $c(y_i)$ kestirim için kullanacağımız veriye ait fonksiyon olarak tanımlanırsa, $c(y_i)$ ile v_i arasındaki uzaklık, $\{v_i - c(y_i)\}^2$ olarak ölçülebilir. Hem v_i , hem de y_i rasgele olduklarından, bu uzaklığın v_i ve y_i 'nin alabileceği bütün olası değerler için küçük olması istenir. Bu nedenle, uzaklığın beklenen değeri, $E\{v_i - c(y_i)\}^2$ düşünülebilir. Seçilebilecek bütün olası fonksiyonlar arasında bu beklenen değeri minimum yapan $c(y_i)$ fonksiyonu seçilmelidir. Beklenen uzaklığı minimum yapan $c(y_i)$ fonksiyonu, y_i verildiğinde v_i 'nin koşullu beklenen değeri olarak adlandırılır ve $E(v_i | y_i)$ biçiminde gösterilebilir (Davidian, 2007). Rasgele vektör y_i 'nin bir fonksiyonu olan koşullu beklenen değer rasgele bir niceliktir. Bu

nedenle, gösterimden bu deęerin sabit bir deęer olduęu düşünülmemelidir. Bunlar, $q > 1$ için, v_i bir vektör olduęu durum için de geçerli olacaktır.

1.5.1.1. Koşullu Beklenen Deęer ve Normal Dağılım

Genelde, rasgele bir deęiřkeni önkestirmek, eldeki veri seti verildięinde, deęiřkenin koşullu ortalamasını önkestirmek olarak gösterilebilir (Fitzmaurice ve ark., 2004). y_i ve v_i normal dağılıma sahip olduklarında, koşullu beklenen deęer için belirgin bir ifade bulmak mümkün olabilir. İlk olarak karma etki modellerinin en basit hali olan, v_i vektörünün tek bir deęerden oluřtuęu, rasgele kesim noktası modeli ele alınsın. Anlařılır olması aısından, zamanın da etkisinin olmadıęı (bir batında doęan fare yavrularının aęırlıkları gibi), $X_i = Z_i = 1_{n_i}$ olduęu durum düşünülecek olursa, herhangi bir i . bireye ait en basit karma etki modeli Denklem (1.60)'daki gibi yazılabilir.

$$y_i = 1_{n_i} \beta + 1_{n_i} v_i + e_i \quad \text{Denklem (1.60)}$$

Modelde yer alan $y_i = (y_{i1}, \dots, y_{in_i})'$, $e_i = (e_{i1}, \dots, e_{in_i})'$, $v_i \sim N(0, G)$ ve $e_i \sim N_{n_i}(0, \sigma^2 I_{n_i})$ biçiminde gösterilebilir. Böylelikle herhangi bir j . ölçüm için,

$$y_{ij} = \beta + v_i + e_i \quad \text{Denklem (1.61)}$$

olacaktır. Ayrıca i . bireyden j . ölçümde alınan yanıt deęeri $y_{ij} \sim N(\beta, G + \sigma^2)$ olacaktır. Bu model altında v_i için koşullu beklenen deęer ařağıdaki biçimde yazılabilir (Davidian, 2007).

$$E(v_i | y_i) = \frac{n_i G}{n_i G + \sigma^2} (\bar{y}_i - \beta) \quad \text{Denklem (1.62)}$$

Eşitlikte, \bar{y}_i , y_i vektörü içindeki n_i tane y_{ij} değerinin ortalamasıdır. Yukarıdaki eşitliğe denk bir ifade olarak $E(v_i | \bar{y}_i)$ yazıldığında v_i ile ilgili bütün bilgi bireysel ortalama \bar{y}_i 'nin içinde özetlendiği anlaşılmaktadır. Yani, $E\{v_i - c(y_i)\}^2$ değerini minimum yapmak amacıyla, y_i 'nin v_i 'ye en yakın fonksiyonunu bulmak için, i . bireye ait verinin örneklem ortalamasının bilinmesi yeterli olacaktır. Eğer v_i büyük (ve pozitif) bir değer ise, \bar{y}_i 'nin de büyük (ve β 'dan daha büyük), eğer v_i küçük (ve negatif) bir değer ise \bar{y}_i 'nin de küçük (ve β 'dan daha küçük) olması beklenir. Ayrıca, yukarıdaki eşitliğin \bar{y}_i nedeniyle y_i vektörüne ait verilerin doğrusal bir fonksiyonu olduğu söylenebilir. Yukarıdaki kestiriciye bakıldığında dikkat edilmesi gereken bir nokta, bu kestiricinin, bilinmeyen fakat kestirilebilen G , σ^2 ve β parametrelerine bağlı olduğunun görülmesidir. Eğer, yukarıdaki kestiriciye ait beklenen değer, yani y_i 'nin olası bütün değerleri üzerinden bir ortalaması alınacak olursa aşağıdaki eşitlik elde edilebilir.

$$E\{E(v_i | y_i)\} = \frac{n_i G}{n_i G + \sigma^2} E(\bar{y}_i - \beta) = 0 \quad \text{Denklem (1.63)}$$

Denklem (1.63) için, $E(\bar{y}_i) = \beta$ olduğundan, bu değer sıfıra eşit olduğu görülmektedir. Bu kestiricinin, verinin olası bütün değerleri üzerinden ortalaması, yani beklenen değeri, sıfırdır. Ayrıca kestirilmeye çalışılan v_i 'nin beklenen değerinin de sıfır olduğu unutulmamalıdır. Bu da kestiricinin yansız olduğunu gösterecektir (Davidian, 2007).

1.5.2. En İyi Doğrusal Yansız Önkestirici (EİDYÖK)

v_i için koşullu beklenen değer hesaplanırken, G , σ^2 ve β değerlerinin biliniyor olması gereklidir. G ve σ^2 'nin (yani varyansın) bilindiği ideal bir durumda, $q = 1$ için koşullu beklenen değer Denklem (1.64)'deki gibi yazılabilir.

$$E(v_i | y_i) = \frac{n_i G}{n_i G + \sigma^2} (\bar{y}_i - \hat{\beta}) \quad \text{Denklem (1.64)}$$

Denklem (1.64) \bar{y}_i nedeniyle hala y_i vektörüne ait verilerin doğrusal bir fonksiyonudur. Ayrıca bu fonksiyonun, v_i 'yi kestirmek için kullanılabilen y_i 'nin herhangi bir doğrusal fonksiyonundan daha küçük varyansa sahip olduğu Davidian (2007) tarafından belirtilmiştir. Böylelikle, kestiricinin “en iyi” olduğu, yani en güvenilir kestirici olduğu söylenebilir. Ayrıca bu kestiricinin yansız olduğu, yani $E(v_i | y_i) = E(v_i) = 0$, söylenebilir. Sonuç olarak bu kestirici, v_i için En İyi Doğrusal Yansız Kestirici (Best Linear Unbiased Predictor, BLUP) olarak adlandırılır.

Ancak, pratikte G , σ^2 değerleri bilinmediği için genellikle kestirimleri kullanılır. $\hat{\beta}$ için, GEKK kestiricisi kullanıldığında, pratikte kullanılan kestirici aşağıdaki biçimde yazılabilir.

$$\hat{v}_i = \frac{n_i \hat{G}}{n_i \hat{G} + \hat{\sigma}^2} (\bar{y}_i - \hat{\beta}) \quad \text{Denklem (1.65)}$$

Denklem (1.65)'te gerçek değerler yerine kestirilen değerler yazıldığı için yansızlık ve en küçük varyanslılık özelliği daha fazla tam olarak sağlanmayacaktır. Ancak en azından yaklaşık olarak sağladığından, pratikte kullanılan kestiriciden genellikle deneysel EİDYÖK ya da deneysel Bayes kestiricisi olarak söz edilir (Littell ve ark., 2006).

Anlatılanlar, genel bir doğrusal karma etki modeli için de genelleştirilebilir. y_i verildiğinde v_i 'nin koşullu beklenen değeri,

$$E(v_i | y_i) = GZ_i' \Sigma_i^{-1} (y_i - X_i \beta) \quad \text{Denklem (1.66)}$$

biçiminde gösterilebilir. Varyansın bilindiği ideal durumda v_i için EİDYÖK Denklem (1.67)'deki gibidir.

$$E(v_i | y_i) = GZ_i' \Sigma_i^{-1} (y_i - X_i \hat{\beta}) \quad \text{Denklem (1.67)}$$

Daha gerçekçi olan varyansın bilinmediği durumda ise v_i için yaklaşık EİDYÖK;

$$\hat{v}_i = \hat{G} \hat{Z}_i' \hat{\Sigma}_i^{-1} (y_i - X_i \hat{\beta}) \quad \text{Denklem (1.68)}$$

olacaktır. Bu kestirici v_i için deneysel EİDYÖK ya da deneysel Bayes kestiricisi olarak adlandırılır (Fitzmaurice ve ark., 2004; Davidian, 2007).

1.5.2.1. Bireysel Ortalamannın Kestirilmesi

Karma etki modellerinin dayandığı mantık, her bir bireyin kendine ait bir regresyon modeline sahip olduğudur, ve bu regresyon modeline göre bireye özel ortalama yanıt, $X_i \beta + Z_i v_i$ biçimde tanımlanabilir. Buradan da görüldüğü üzere, her birey için bireye özel ortalamaların tanımlanabilmesi için v_i 'nin bulunması gereklidir.

Teknik olarak bireye özel ortalama yanıtlar, v_i verildiğinde y_i 'nin koşullu beklenen değeri olarak da düşünülebilir. Yine anlaşılır olması açısından, $X_i = Z_i = 1_{n_i}$ olduğu durum düşünülecek olursa, herhangi bir i . bireye ait en basit karma etki modeli Denklem (1.60)'da gösterildiği biçimde tanımlanabiliyordu. Modele göre bireysel ortalama, $1_{n_i} \beta + 1_{n_i} v_i$ ifadesi ile gösterilebilir. Herhangi bir i . birey için koşullu ortalama,

$$b_i = \beta + v_i \quad \text{Denklem (1.69)}$$

biçiminde olacaktır. Teknik olarak b_i , v_i verildiğinde y_i 'nin koşullu beklenen değeri olarak yazılabilir (Denklem 1.70).

$$b_i = E(y_{ij} | v_i) \quad \text{Denklem (1.70)}$$

b_i 'nin kestirimi için izlenecek yol v_i ile benzerdir. Bu durumda, b_i 'yi kestirmek için y_i yanıt vektörü verildiğinde koşullu beklenen değerinin bulunması düşünülebilir. Denklem (1.69)'da v_i yerine kestirimi yazılırsa, bireye özel ortalamının koşullu beklenen değeri Denklem (1.71)'deki gibi olacaktır.

$$E(b_i | y_i) = \beta + E(v_i | y_i) \quad \text{Denklem (1.71)}$$

Bu eşitlik, b_i için en iyi yansız önkestirici olarak kullanılabilir (Davidian, 2007).

Varyansın bilindiği durumda kestirici aşağıdaki gibi yazılabilir.

$$E(b_i | y_i) = \hat{\beta} + \frac{n_i G}{n_i G + \sigma^2} (\bar{y}_i - \hat{\beta}) \quad \text{Denklem (1.72)}$$

Ancak pratikte, varyans bilinmediği için G ve σ^2 ise en çok olabilirlik ya da sınırlı en çok olabilirlik kestirimleri ile değiştirilir. $\hat{\beta}$ ise genelleştirilmiş en küçük kareler kestirimi olacaktır. Bireysel ortalamının bu kestiricisi yine deneysel EİDYÖK ya da deneysel Bayes kestiricisi olarak adlandırılır.

Bireye özel ortalama kestirimi genel bir doğrusal karma etki modeli için genelleştirilecek olursa, i . birey için ortalamının koşullu beklenen değeri,

$$E(y_i | v_i) = X_i \beta + Z_i b_i \quad \text{Denklem (1.73)}$$

ve bu değer için en iyi doğrusal yansız önkestiricisi varyans bilinmediğinde,

$$X_i \hat{\beta} + Z_i \hat{G} Z_i' \hat{\Sigma}_i^{-1} (y_i - X_i \hat{\beta}) \quad \text{Denklem (1.74)}$$

olacaktır. Bu kestirici bireye özel ortalama $X_i\beta + Z_ib_i$ için deneysel EİDYÖK ya da deneysel Bayes kestiricisi olarak adlandırılır.

1.5.2.2. Ağırlıklı Ortalama Olarak EİDYÖK

Öncelikle varyansın bilindiği durum için Denklem (1.72)'de gösterilen EİDYÖK alternatif olarak Denklem (1.75)'deki gibi yazılabilir.

$$\left(\frac{\sigma^2 / n_i}{G + \sigma^2 / n_i} \right) \hat{\beta} + \left(\frac{G}{G + \sigma^2 / n_i} \right) \bar{y}_i, \quad \text{Denklem (1.75)}$$

Bu biçimde yazıldığında, EİDYÖK kestiricisinin, i . bireye ait ortalama \bar{y}_i ve örneklem ortalaması $\hat{\beta}$ değerlerinin birbirlerine göre yorumlanmasını sağlayabilir. \bar{y}_i değeri yalnızca i . bireye ait verilere dayanan b_i için en iyi tahmindir. Bunun tersine, $\hat{\beta}$ değeri örneklemdeki bütün bireylere ait verilere dayanan genel ortalamanın en iyi tahminidir. Tekrar hatırlanacak olursa, G bireyler arası değişimin, σ^2 ise birey içi değişimin ölçüsüdür. n_i 'nin ise belirli bir birey için erişilebilir bilgi miktarını tanımladığı düşünülecek olursa, σ^2 / n_i hem birey içi değişim kaynağını hem de kaç gözleme sahip olduğunu dikkate alarak, i . birey hakkındaki bilgi kalitesini ölçtüğü söylenebilir. Bu düşünce ile, eğer G büyük, yani bireyler arası değişim büyük ise popülasyona ait fazla bilgi olsa da, belirli bir birey hakkında bilgi sahibi olmak için yeterli olmayacaktır. Tersine, eğer G küçük, yani bireyler birbirine oldukça benzer, ise popülasyon hakkında fazla bilgiye sahip olmak, belirli bir birey hakkında bilgi sahibi olmak için yeterli olacaktır.

Sonuç olarak, eğer G değeri σ^2 / n_i değerine göre büyük ise, herhangi bir i . birey hakkında bir kestirimde bulunmak için, aynı bireyden elde edilen bilgi, popülasyondan elde edilenden daha güvenilir olacaktır. Böyle bir durumda, EİDYÖK kestiriminin Denklem (1.75)'te gösterilen biçimi düşünülürse, eşitlikte

yer alan $G/(G + \sigma^2/n_i)$ değeri 1'e yaklaşırken, $(\sigma^2/n_i)/(G + \sigma^2/n_i)$ değeri 0'a yaklaşacaktır. Yani, $EIDYÖK(b_i) \approx \bar{y}_i$ gibi düşünülebilir. b_i 'yi kestirmek için \bar{y}_i içinde sahip olunan bilgi, kestirilen popülasyon ortalaması $\hat{\beta}$ içinde sahip olunan bilgiden daha iyidir.

Diğer yandan, eğer G değeri σ^2/n_i değerine göre küçük ise, herhangi bir i . birey hakkında popülasyondan elde edilen bilgi, aynı bireye ait veriden elde edilen bilgiden daha iyi olacaktır. Bunların yanı sıra, eğer i . bireye ait gözlem sayısı, n_i , çok az ise birey hakkında çıkarsamada bulunmak için çok az veri var demektir. Burada, yukarıda anlatılanların tam tersi söylenebilir. Yani $EIDYÖK(b_i) \approx \hat{\beta}$ olarak düşünülebilir.

Bu sonuçlar b_i 'nin en iyi doğrusal yansız önkestiricisinin sadece i . bireye ait bilgi ile bütün örneklemden elde edilen bilgi arasında bir kıyaslama yaptığını da göstermektedir. EİDYÖK, Denklem (1.75)'te hem G hem de σ^2/n_i terimlerinin bir kombinasyonu biçiminde yazılmıştır. Böylelikle EİDYÖK kullanılarak, yalnızca o bireye ait bilgi dışında, bireyin ait olduğu örneklemden de destekleyici bilgi alınmaktadır.

Yine bu durum genelleştirilecek olursa, Denklem (1.75), varyansın bilinmediği durumda, bilinmeyen kovaryans parametrelerinin yerine EÇOK ya da SEÇOK kestirimleri yazıldığında eşitlik Denklem (1.76)'daki gibi olacaktır.

$$\hat{y}_i = X_i \hat{\beta} + Z_i \hat{v}_i = (\hat{R}_i \hat{\Sigma}_i^{-1}) X_i \hat{\beta} + (I_{n_i} - \hat{R}_i \hat{\Sigma}_i^{-1}) y_i \quad \text{Denklem (1.76)}$$

Aynı şekilde, R_i denek içi değişimi tanımlarken, Σ_i denek içi ve denekler arası değişimi birlikte içerir. Sonuç olarak, R_i daha büyük olduğunda, yani denek içi değişim denekler arası değişime göre büyük olduğunda $X_i \hat{\beta}$ 'nin sahip olduğu ağırlık daha fazla olacaktır. Yani, i . birey hakkında popülasyondan elde edilen bilgi,

aynı bireye ait veriden elde edilen bilgiden daha iyi olacaktır. Diğer yandan, denekler arası değişim denek içi değişime göre fazla olduğunda, i . bireyin gözlenen değerlerinin sahip olduğu ağırlık daha fazla olacaktır. Aynı şekilde bu denge i . bireyden elde edilen gözlem sayısı n_i 'ye göre de değişecektir. n_i küçük olduğunda popülasyondan elde edilen bilginin sahip olduğu ağırlık daha fazla olacaktır (Fitzmaurice ve ark., 2004).

2. GEREÇ VE YÖNTEM

2.1. Uygulama Verileri

Meme kanserinin görülme sıklığı dünyada giderek artmakta ve konu ile ilgili tedavi yöntemleri de sürekli bir gelişme göstermektedir. Memeye total ya da parsiyel cerrahiyle birlikte uygulanmış ve aksiller diseksiyon yapılmış olgularda omuz hareket açıklığında kısıtlılığın önlenmesi amacıyla tutulan kolun mobilizasyonu önerilmektedir. Bu nedenle olası en erken dönemde aktif yardımcı ve nazik pasif hareket açıklığı egzersizleri başlatılmalıdır. Ayrıca omuzun tam hareket açıklığının sağlanması postoperatif radyoterapi uygulaması sırasında hastaya pozisyon vermek için de gereklidir. Uygulanacak egzersiz izlencesi hem postoperatif dönem, hem de radyoterapi uygulamasını izleyen dönemde gelişebilecek lenf ödeminin önlenmesinde de etkilidir. Ev bakım prensipleri ve olası komplikasyonlar konusunda hasta eğitimi ise rehabilitasyonun önemli öğelerindendir (Özkan ve ark., 1998). Mastektomi göğsün ameliyat ile alınması işlemidir. Günümüzde, cerrahi müdahale meme kanserinin en yaygın tedavi şeklidir. Pek çok durumda, mastektominin hemen sonrasında ya da daha sonra göğsün yeniden yapılanması için operasyon yapılması mümkündür. Radikal mastektomi ise göğsün tamamının, göğüs ucunun ve göğüs ucu çevresindeki renkli bölgenin (areola), göğüs kaslarının (küçük ve büyük pektoral kaslar) ve koltuk altı lenf bezlerinin alınması işlemidir. Çalışmada kullanılacak veriler, Ankara Numune Hastanesi, Fizik Tedavi ve Rehabilitasyon Kliniği tarafından sağlanmıştır. 2005 yılında Ankara Numune Hastanesi'nde meme kanseri nedeniyle radikal mastektomi ameliyatı olan ve lenf bezleri alınan 29 yaş ve üzeri kadınlar rasgele iki gruba ayrılmışlardır. Birinci gruba hastanede tedavi uygulanırken, ikinci grup ev programı almıştır. Çalışmada, hastanede tedavi gören 27, ev programı alan 30 hasta olmak üzere toplam 57 birey vardır. Hastaların ameliyat öncesi (0. gün), ameliyat sonrası 5. gün, 30. gün, 90. gün ve 180. gün omuz hareket açıklığı ölçümlerinden fleksiyon (bükülme, bu hareket ile eklemi oluşturan iki kemik arasındaki açı daralır) açı değerleri alınmıştır. Ayrıca, bütün bireylerin yaş ve vücut kitle indeksi (vki) ölçümleri de kaydedilmiştir. Tekrarlı ölçümler, her

bireyden aynı zaman noktalarında alınmıştır ve her birey eşit ölçüm sayısına sahiptir.

Fleksiyon ölçümleri çalışmada bağımlı değişken olarak alınacaktır. Grup değişkeni, tedavi:0 ve ev programı:1 olarak kodlanmıştır. Tekrarlı ölçümlerin alındığı zaman noktaları gün bazında ve kodlanmadan gerçek değerleri ile analize alınmıştır. Ameliyat öncesi ölçüm zamanı 0. gün olarak alınmıştır. Başlangıç düzeyi yaş ve başlangıç düzeyi vki ise sürekli değişkenler olarak alınmıştır.

2.2. Uygulama Yöntemleri

Bu çalışmada, tekrarlı ölçümlerde varyans analizi çözümlemesi ve karma etki modellerinin uygulaması SAS paket programı ile yapılmıştır. Klasik tekrarlı ölçümlerde varyans analizi için PROC GLM işlemi kullanılmış, karma etki modelleri için ise PROC MIXED işlemi kullanılmıştır. Bağımlı değişken ve ortak değişkenler için tanımlayıcı istatistikler PROC MEANS işlemi ile elde edilmiştir. Karma etki modelleri analizi için SAS programında veri girişi Çizelge 2.1’de görüldüğü gibidir.

Çizelge 2.1. SAS programında karma etki modelleri analizi için veri girişi

Gözlem	hasta	grup	yas	vki	zaman	fleks
1	1	0	54	26.44	0	180
2	1	0	54	26.44	5	110
3	1	0	54	26.44	30	170
4	1	0	54	26.44	90	180
5	1	0	54	26.44	180	180
6	2	0	49	31.25	0	180
7	2	0	49	31.25	5	168
8	2	0	49	31.25	30	180
.....						
.....						
.....						
280	56	1	42	27.34	180	168
281	57	1	68	31.06	0	180
282	57	1	68	31.06	5	132
283	57	1	68	31.06	30	146
284	57	1	68	31.06	90	164
285	57	1	68	31.06	180	165

Tekrarlı ölçümlerde varyans analizi yapmak için ise veriler Çizelge 2.2'deki gibi girilmiştir.

Çizelge 2.2. SAS programında tekrarlı ölçümlerde varyans analizi için veri girişi.

Gözlem	grup	fleks1	fleks2	fleks3	fleks4	fleks5
1	0	180	110	170	180	180
2	0	180	168	180	180	180
3	0	180	95	163	165	170
4	0	173	135	170	170	170
5	0	180	165	180	180	180
6	0	180	170	180	180	180
7	0	180	150	180	180	180
8	0	180	150	180	180	180
.....						
.....						
47	1	180	105	130	142	155
48	1	180	132	148	170	170
49	1	180	140	140	150	160
50	1	180	140	145	155	170
51	1	180	160	180	180	180
52	1	170	90	115	132	157
53	1	180	90	125	134	168
54	1	180	95	115	128	145
55	1	165	92	119	134	148
56	1	180	125	135	152	168
57	1	180	132	146	164	165

PROC MIXED işlemi altında, farklı sabit etki ve rasgele etki yapılarına sahip modeller oluşturulmuş ve bu modeller arasından, en küçük Akaike Bilgi Kriteri (AIC) ya da Schwartz Bayesyen Bilgi Kriteri (BIC)'ne dayanarak model seçimi yapılmıştır. Oluşturulan doğrusal karma etki modeli ile klasik tekrarlı ölçümlerde varyans analizinden elde edilen sonuçlar karşılaştırılmıştır.

2.2.1. PROC MIXED İşlemi

Analizde kullanılacak PROC MIXED işleminin karma etki modellerinin analizinde kullanılan genel biçimi ve içindeki önemli bazı komutlar aşağıda özetlenmiştir;

Çizelge 2.3. Proc mixed işlemi.

```

PROC MIXED <options>;
CLASS variable-list;
MODEL dependent=fixed effects</ options>;
RANDOM random effects< / options>;
REPEATED repeated effects </ options>;
LSMEANS fixed effects < / options>;
RUN;

```

Aşağıda PROC MIXED işlemi altında sıkça kullanılan komutlar özetlenmiştir.

PROC MIXED <options>;

DATA: Analiz edilecek veri kümesi burada tanımlanır.

METHOD: En çok olabilirlik, sınırlı en çok olabilirlik ve en küçük normlu karesel yansız kestirim olmak üzere üç farklı kestirim yöntemi tanımlanabilir (METHOD=REML/ METHOD=ML/ METHOD=MINQUE0). Eğer METHOD yazılmazsa, sınırlı en çok olabilirlik yöntemi kullanılmaktadır.

CLASS variables;

Analizde kullanılacak kategorik veriler tanımlanır.

MODEL dependent = fixed effects </options>;

Sabit etkiler tasarım matrisini, X_i , tanımlamak için kullanılır.

NOINT: X_i matrisinde kesim noktasının yer almaması isteniyor ise, seçeneği kullanılabilir.

SOLUTION: Sabit etkiler için parametre kestirimlerini verir.

OUTPRED: Her zaman noktasında, her bireye ait koşullu ortalamaların, $E(y_i | v_i)$, EİDYÖ değerlerini verir.

RANDOM random effects </options>;

Rasgele etkiler bu komut altında tanımlanır. Diğer bir deyişle, v_i rasgele etkileri için Z_i tasarım matrisindeki ortak değişkenleri tanımlamak amacıyla kullanılır. Bu komut altında, kesim noktasının tasarım matrisine dahil edilmesi için bunun belirtilmesi gerekmektedir. "INTERCEPT" ya da "INT" komutu ile kesim noktasının

tasarım matrisine ekleneceği belirtilmelidir. MODEL komutu altında buna gerek yoktur, çünkü kesim noktası tasarım matrisine zaten dahil edilir.

TYPE: Rasgele etkiler için kovaryansın, G , yapısı tanımlanır. En çok kullanılan yapılar varyans bileşenleri (VC) ve yapılandırılmamıştır (UN). Ancak bir ya da birden fazla varyans bileşeni negatif kestirildi ise MIXED işlemi bunları sıfır olarak alacak ve G matrisinin pozitif tanımlı olmaması hatasını verecektir (“Estimated G matrix is not positive definite”). Bu durumda SAS yarı tanımlı pozitif kovaryans matrisi oluşturmadığından, yarı tanımlı pozitif kovaryans matrisi oluşturmak için faktör analitik yapı kullanılabilir (TYPE=FA0(q)) (McCulloch, 2005).

SOLUTION: Rasgele etkiler, v_i , için EİDYÖ değerlerini verir.

REPEATED repeated effects </ options>;

Hata terimlerinin, e_i , kovaryans yapısını tanımlamak için kullanılır. Tekrarlanan etki kategorik ve CLASS komutu altında tanımlanmış olmalıdır. Örneğin; her bireyden beş tekrarlı ölçüm eşit zaman aralıkları ile alınmış olsun. Bu durumda, zaman tekrarlanan etki olacaktır. Eğer zaman değişkeni aynı zamanda modelde sürekli bir değişken olarak kullanılmış ise, bu durumda örneğin t gibi, yeni bir değişken oluşturulmalıdır. Bu yeni değişken CLASS komutu altında tanımlanmalı ve REPEATED komutuna bu biçimde yazılmalıdır.

TYPE: Hata korelasyon yapısı tanımlanır. Bu seçenek ile çok sayıda korelasyon yapısı tanımlanabilir. En çok kullanılanları otokorelasyon (AR(1)), tam simetri (CS), Toeplitz (TOEP), varyans bileşenleri (VC) ve yapılandırılmamıştır (UN).

LSMEANS fixed effects < / options>;

Sabit etkiler için genelleştirilmiş en küçük kareler kestirimlerini verir.

DIFF: Olası bütün ikili karşılaştırmaları verir.

ADJUST: p değerindeki çoklu karşılaştırma düzeltmesini belirlemek için kullanılır. BON(Bonferroni), DUNNET, SCHEFFE, SIDAK, SIMULATE, SMM|GT2 ve TUKEY düzeltmeleri belirlenebilir.

2.2.2. PROC GLM İşlemi

Tekrarlı ölçümlerde tek değişkenli ve çok değişkenli varyans analizi çözümlenmeleri aşağıdaki işlem yardımı ile gerçekleştirilebilir.

Çizelge 2.4. PROC GLM işlemi.

```
PROC GLM < options > ;
CLASS variable-list ;
MODEL dependents=independents < / options > ;
LSMEANS effects < / options > ;
REPEATED factor-specification < / options > ;
RUN;
```

Aşağıda PROC GLM işlemi altında sıkça kullanılan komutlar özetlenmiştir.

PROC GLM <options>;

DATA: Analiz edilecek veri kümesi burada tanımlanır.

CLASS variables;

Analizde kullanılacak kategorik veriler tanımlanır.

MODEL dependent = independents </options>;

Bireyler arası etkiler (grup, cinsiyet gibi) burada tanımlanır.

NOUNI: Her tekrarlı ölçüm noktasında bireyler arası etkiler için standart tek yönlü ANOVA yapılmaması için kullanılır.

LSMEANS effects </ options>;

Ana etkiler ve etkileşim terimine ait standart en küçük kareler kestirimlerini verir.

PDIFF: İkili karşılaştırma testlerine ait p değerlerinin görünmesini sağlar.

ADJUST: p-değerlerinde yapılacak düzeltme belirlenebilir.

STDERR: standart hata kestirimlerini verir.

REPEATED repeated effects </ options>;

Birey içi etki (tekrarlı ölçümlerin alındığı zaman etkisi gibi) burada tanımlanır. Tekrarlı etki MODEL altında tanımlanmaz ve bireyler arası etki ile olan etkileşimi PROC GLM tarafından oluşturulur.

PRINTE: Küresellik varsayımının testi için kullanılır.

POLINOMIAL: Birey içi etkinin polinomiyal etkilerinin kestirimi ve önemlilik testlerini verir. Polinom derecesi tekrarlı ölçüm sayısının bir eksigidir. Örneğin, dört zaman noktasında ölçüm alındığında, zamanın birinci dereceden, karesel ve kübik etkilerine ait kestirimler ve önemlilikler yer alacaktır. Sonuçların görüntülenmesi için SUMMARY seçeneği kullanılır.

2.2.3. Doğrusal Karma Etki Modelleri ve Tekrarlı Ölçümlerde Varyans Analizi Algoritmaları

PROC MIXED içinde yazılan karma etki modelleri Çizelge 2.5.'te verilmiştir. Bu işlem altında beş model oluşturulmuştur. Model I, II ve III'te zaman faktörü modele sürekli değişken olarak alınmış ve rasgele kesim noktası ve eğim modeli oluşturulmuştur. Model IV'te, zaman faktörü modele kategorik olarak alınmış ve rasgele kesim noktası modeli oluşturulmuştur. Model V'te ise zamanın hem sürekli hem de kategorik etkisi modele alınarak rasgele kesim noktası ve eğim modeli oluşturulmuştur. Zamanı kategorik olarak da modele almak için “zaman2” gibi bir değişken oluşturularak, bu değişken PROC MIXED işlemi altında CLASS kısmında kategorik olarak tanımlanmış ve MODEL kısmına yazılmış, sürekli olarak alınan “zaman” değişkeni ise RANDOM kısmına yazılmıştır. Önceden de belirtildiği gibi, tez kapsamında e_{ij} terimleri ölçüm hatası olarak varsayıldığı için, denek içi kovaryans matrisi diagonal ($R_i = \sigma^2 I_5$) olarak tanımlanmış ve her bir grup için ayrı bir diagonal matris oluşturulması (repeated / group=grup) seçeneği ile belirtilmiştir. REPEATED komutunda TYPE tanımlanmadığında, SAS varyans kovaryans matrisi için $R_i = \sigma^2 I_{n_i}$ sonucunu verecektir. Rasgele kesim noktası ve eğim modellerinde (Model I, II, III ve V) yakınsama sorunu yaşandığı için, G matrisi faktör analitik, FA0(q), olarak tanımlanmıştır (McCulloch, 2005). Model IV'te ise tek bir rasgele etki tanımlandığından (rasgele kesim noktası), G için tek bir varyans bileşeni kestirilmiştir.

Tekrarlı ölçümlerde varyans analizi algoritması ise, Çizelge 2.6.'da yer almaktadır.

Çizelge 2.5. SAS programında doğrusal karma etki algoritmaları.

```

*****
title 'Model I: Zaman Sürekli, Rasgele Kesim Noktası & Eğim Modeli
';
run;
proc mixed method=reml data=fleksiyon ;
  class grup id;
  model fleks = grup zaman zaman*zaman grup*zaman grup*zaman*zaman
              yas vki/ noint;
  random intercept zaman / subject=id type=fa0(2) ;
  repeated / group=grup subject=id;
run;

*****

title 'Model II: Zaman Sürekli, Rasgele Kesim Noktası & Eğim
Modeli';
run;
proc mixed method=reml data=fleksiyon ;
  class grup id;
  model fleks = grup grup*zaman grup*zaman*zaman/ noint;
  random intercept zaman / subject=id type=fa0(2);
  repeated / group=grup subject=id;
run;

*****

title 'Model III: Zaman Sürekli, Rasgele Kesim Noktası & Eğim
Modeli';
run;
proc mixed method=reml data=fleksiyon ;
  class grup id;
  model fleks = grup grup*zaman/ noint;
  random intercept zaman / subject=id type=fa0(2);
  repeated / group=grup subject=id;
run;

*****

title 'Model IV: Zaman kategorik, Rasgele Kesim Noktası Modeli';
run;
proc mixed method=reml data=fleksiyon ;
  class grup id zaman;
  model fleks = grup zaman grup*zaman/ noint;
  random intercept / subject=id ;
  repeated / group=grup subject=id;
run;

*****

```

Çizelge 2.5. Devam. SAS programında doğrusal karma etki algoritmaları.

```

*****
title 'Model V: Zaman Hem Sürekli Hem de Kategorik, Rasgele Kesim
Noktası & Eğim Modeli ';
run;
data fleksiyon2;
  set fleksiyon;
  zaman2=zaman;
run;
proc mixed method=reml data=fleksiyon2;
  class grup id zaman2 ;
  model fleks = grup zaman2 grup*zaman2 / noint;
  random intercept zaman / subject=id type=fa0(2) ;
  repeated / group=grup subject=id;
run;
*****

```

Çizelge 2.6. SAS programında tekrarlı ölçümlerde varyans analizi algoritması

```

title 'Tekrarlı Ölçümlerde Varyans Analizi'
run;
proc glm data=fleksglm;
  class grup;
  model fleks1-fleks5=grup / nouni;
  repeated zaman 5 (0 5 30 90 180) polynomial / printe summary;
  lsmeans grup /stderr adjust=bon pdiff;
run;

```

3. BULGULAR

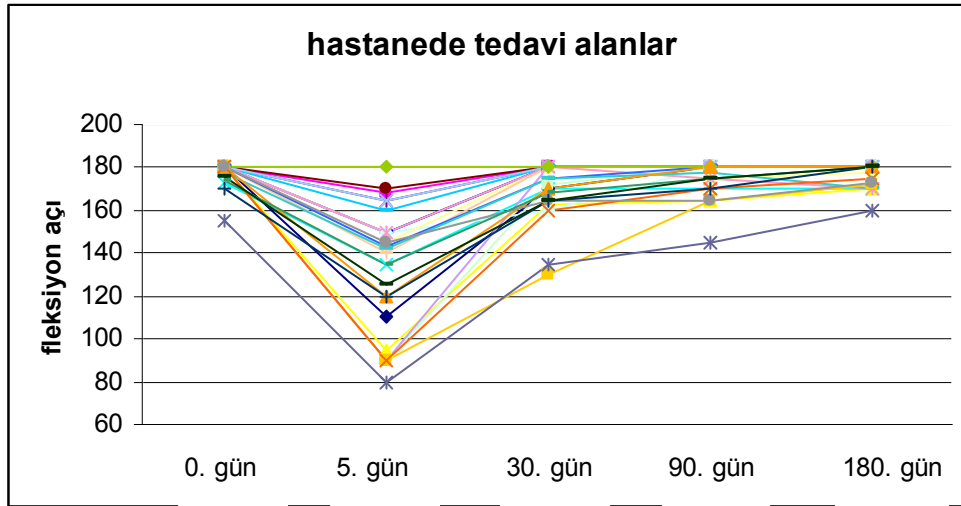
Bu bölümde, bireylere ait fleksiyon ölçümlerinin zaman içinde gösterdikleri değişimin doğrusal karma etki modelleri yardımı ile modellenmesi, bulunan model sonuçları ve klasik tekrarlı ölçümlerde varyans analizi sonuçları üzerinde durulacaktır. Çizelge 3.1’de hastanede tedavi gören (grup=0) ve ev programı alan (grup=1) bireyler için tanımlayıcı istatistikler yer almaktadır.

Çizelge 3.1. Gruplara göre tanımlayıcı istatistikler.

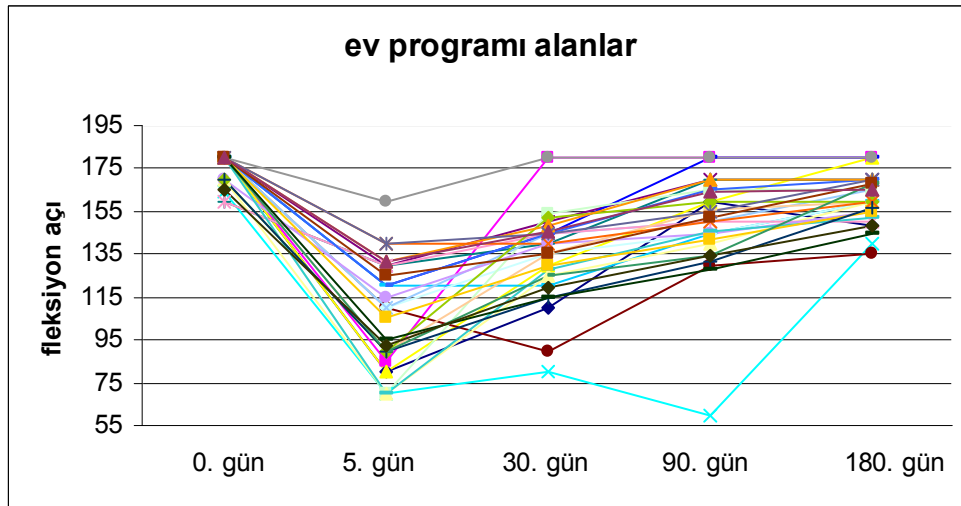
	Hastanede tedavi görenler n=27		Ev programı alanlar n=30	
	$\bar{X} \pm SS$	Min-Maks	$\bar{X} \pm SS$	Min-Maks
Yaş	52,6±12,2	29,0-76,0	51,1±13,0	31,0-78,0
Vki	28,7±2,5	24,0-33,5	27,8±3,2	23,1-35,2
Fleksiyon 0. gün	178,0±5,2	155,0-180,0	176,3±6,6	160,0-180,0
Fleksiyon 5. gün	133,4±28,9	80,0-180,0	106,7±24,7	70,0-160,0
Fleksiyon 30. gün	170,4±13,0	130,0-180,0	134,9±21,2	80,0-180,0
Fleksiyon 90. gün	175,1±7,9	145,0-180,0	150,6±22,7	60,0-180,0
Fleksiyon 180. gün	176,9±5,1	160,0-180,0	161,5±11,7	135,0-180,0

SS: standart sapma, \bar{X} : aritmetik ortalama, Min: minimum, Maks: maksimum

Görüldüğü gibi yaş ve vki değerleri bakımından gruplar arasında önemli bir fark görülmemektedir (sırasıyla p=0,665; p=0,247). Ayrıca herhangi bir analiz yapmadan yalnızca ortalamalar üzerinden konuşulacak olursa, ameliyat öncesi fleksiyon değerlerinin (fleksiyon 0. gün) her iki grupta benzer olduğu görülmektedir. Ancak ameliyat sonrası 5. gün fleksiyon değerlerindeki düşüş ve sonrasında 30, 90 ve 180. günlerdeki artışın gruplarda farklı olduğu görülmektedir. Hastanede tedavi görenlerin başlangıç fleksiyon değerlerine dönüşü daha hızlı iken, evde tedavi alan grupta başlangıç değerlerine yaklaşma daha yavaştır. Buna göre zaman ve grup arasında bir etkileşimin varlığı düşünülebilir. Bunun yanı sıra, her iki grupta 5. gün ölçüm değerlerinin düştüğü ve sonrasında tekrar artmaya başladığı görülmektedir. Bu durum modelleme sırasında, zamanın karesel etkisinin de göz önüne alınması gerektiğini göstermektedir. Çalışmadaki her bireye ait fleksiyon değerlerinin gruplara göre zaman içindeki değişimini gösteren grafikler, Şekil 3.1. ve Şekil 3.2.’de yer almaktadır.



Şekil 3.1. Hastanede tedavi alanlarda fleksiyon değerlerinin zaman içindeki değişimi.



Şekil 3.2. Ev programı alanlarda fleksiyon değerlerinin zaman içindeki değişimi.

Şekillere bakılarak da, hastanede tedavi görenlerin ameliyat öncesi değerlerine daha hızlı yaklaştığı görülmektedir. Veriyi en iyi tanımlayabilecek modeli oluşturabilmek için beş farklı doğrusal karma etki modeli kurulmuştur. Modellerden elde edilen popülasyon parametrelerinin tahminine ilişkin önemlilikler ve bu modeller arasından hangisinin seçilmesi gerektiği ile ilgili açıklamalara Bölüm 3.1'de yer verilmiştir.

3.1. Karma Etki Modellerine Ait Sonuçlar ve Model Seçimi

Bölüm 2.2.3’de de belirtildiği gibi, eldeki veriye en uygun modeli seçebilmek için beş farklı doğrusal karma etki modeli kurulmuştur. Bu modellere ilişkin sonuçlar sırasıyla Çizelge 3.2, 3.3, 3.4, 3.5 ve 3.6’da yer almaktadır.

3.1.1. Karma Modellerine İlişkin Uyum İstatistikleri ve Sabit Etkilerin Testleri

Model I, hem kesim noktası hem de eğimin rasgele olduğu bir doğrusal karma etki modelidir. Modelde zaman sürekli değişken olarak modellenmiş, yaş ve vki de sürekli ortak değişkenler olarak modele alınmıştır.

Çizelge 3.2. Model I: zaman sürekli, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.

Fit Statistics				
-2 Res Log Likelihood				2665.6
AIC (smaller is better)				2673.6
AICC (smaller is better)				2673.7
BIC (smaller is better)				2681.7
Null Model Likelihood Ratio Test				
DF	Chi-Square		Pr >	ChiSq
3	13.77			0.0032
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
grup	2	169	34.49	<.0001
zaman	1	55	1.64	0.2062
zaman*zaman	1	55	0.00	0.9479
zaman*grup	1	169	4.38	0.0378
zaman*zaman*grup	1	169	4.88	0.0286
yas	1	169	0.81	0.3685
vki	1	169	2.13	0.1463

Sonuçlara göre yalnızca grup etkisinin, zaman*grup ve zaman*zaman*grup etkileşimlerinin anlamlı olduğu görülmektedir (sırasıyla, $p < 0,001$; $p = 0,038$; $p = 0,029$).

Model II, hem kesim noktası hem de eğimin rasgele olduğu bir doğrusal karma etki modelidir. Modelde zaman sürekli değişken olarak modellenmiştir.

Çizelge 3.3. Model II: zaman sürekli, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.

Fit Statistics				
-2 Res Log Likelihood				2666.4
AIC (smaller is better)				2674.4
AICC (smaller is better)				2674.6
BIC (smaller is better)				2682.6
Null Model Likelihood Ratio Test				
DF	Chi-Square		Pr > ChiSq	
3	15.68		0.0013	
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
grup	2	169	2003.31	<.0001
zaman*grup	2	169	4.10	0.0183
zaman*zaman*grup	2	169	2.60	0.0772

Kurulan bu modelde grup ve zamanın karesel etkileşimi (zaman*zaman*grup) anlamsız bulunmuştur ($p=0,077$).

Model III, hem kesim noktası hem de eğimin rasgele olduğu bir doğrusal karma etki modelidir. Modelde zaman sürekli değişken olarak modellenmiştir.

Çizelge 3.4. Model III: zaman sürekli, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.

Fit Statistics				
-2 Res Log Likelihood				2645.9
AIC (smaller is better)				2653.9
AICC (smaller is better)				2654.0
BIC (smaller is better)				2662.0
Null Model Likelihood Ratio Test				
DF	Chi-Square		Pr > ChiSq	
3	15.24		0.0016	
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
grup	2	171	2602.05	<.0001
zaman*grup	2	171	14.27	<.0001

Model III'te yer alan grup etkisi ve zaman*grup etkileşimi anlamlı bulunmuştur ($p < 0,001$).

Model IV, yalnızca kesim noktasının rasgele olduğu bir doğrusal karma etki modelidir. Modelde zaman, kategorik değişken olarak modellenmiştir.

Çizelge 3.5. Model IV: zaman kategorik, rasgele kesim noktası modeli için sonuçlar.

Fit Statistics				
-2 Res Log Likelihood				2315.7
AIC (smaller is better)				2321.7
AICC (smaller is better)				2321.8
BIC (smaller is better)				2327.8
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
grup	1	220	42.27	<.0001
zaman	4	220	158.55	<.0001
grup*zaman	4	220	13.30	<.0001

Kurulan modelde grup ve zaman ana etkileri ile grup*zaman etkileşimi anlamlı bulunmuştur ($p < 0,001$).

Model V, hem kesim noktası hem de eğimin rasgele olduğu bir doğrusal karma etki modelidir. Modelde zaman hem sürekli hem de kategorik değişken olarak modellenmiştir.

Çizelge 3.6. Model V: zaman hem sürekli hem de kategorik, rasgele kesim noktası ve rasgele eğim modeli için sonuçlar.

Fit Statistics				
-2 Res Log Likelihood				2309.1
AIC (smaller is better)				2317.1
AICC (smaller is better)				2317.2
BIC (smaller is better)				2325.3
Null Model Likelihood Ratio Test				
DF	Chi-Square		Pr > ChiSq	
3	65.86		<.0001	
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
grup	1	165	43.08	<.0001
zaman2	4	165	159.58	<.0001
grup*zaman2	4	165	13.66	<.0001

Modelde grup ve zaman ana etkisi ile etkileşim terimi anlamlı bulunmuştur ($p < 0,001$).

3.1.2. Model Seçimi

Model seçimi için AIC ve BIC değerlerine bakılacak olursa, en küçük değerlere sahip modelin Model V olduğu görülür. Bu nedenle, veri setinin çözümlenmesi ve klasik tekrarlı ölçümlerde varyans analizi sonuçları ile karşılaştırmak için Model V'e ait sonuçlar göz önünde tutulacaktır.

3.1.3. Model V'e İlişkin Sonuçlar

Sabit etkilerin, β , çalışmadaki her bir bireye ait rasgele etkilerin, v_i , ve yine her zaman noktasında her birey için beklenen ortalama yanıt değerlerinin, $E(y_i | v_i)$, EİDYÖ ile önkestirilmesi için Bölüm 2.2.3'te Model V için verilen algoritma Çizelge 3.7'deki gibi yeniden yazılabilir.

Çizelge 3.7. Model V: zaman hem sürekli hem de kategorik, rasgele kesim noktası ve rasgele eğim modeli için algoritma.

```

title 'Model V: Zaman Hem Sürekli Hem de Kategorik, Rasgele Kesim
Noktası & Eğim Modeli ';
run;
data fleksiyon2;
  set fleksiyon;
  zaman2=zaman;
run;
proc mixed method=reml data=fleksiyon2;
  class grup id zaman2 ;
  model fleks = grup zaman2 grup*zaman2 / noint solution outp=pdata;
  random intercept zaman / subject=id type=fa0(2) solution;
  repeated / group=grup subject=id;
  lsmeans grup*zaman2 / adjust=bon diff;
run;
proc print data=pdata;
run;

```

Algoritmada sabit etki kestirimleri MODEL kısmında SOLUTION ile, her bireye ait koşullu ortalamaların önkestirimleri yine MODEL kısmında OUTF ile, rasgele etkilerin önkestirimleri ise RANDOM kısmındaki SOLUTION seçenekleri ile alınmıştır. Model V'e ait sonuçlar aşağıda tablolar halinde özetlenmiştir.

Çizelge 3.8. Kovaryans parametre kestirimleri.

Covariance Parameter Estimates			
Cov Parm	Subject	Group	Estimate
FA(1,1)	id		12.4913
FA(2,1)	id		-0.03423
FA(2,2)	id		6.94E-17
Residual	id	grup 0	132.37
Residual	id	grup 1	215.93

Hastanede tedavi görenler ve ev programı alanlarda bireyler arası değişimin aynı olduğu varsayıldığından her iki grup için de aynı G matrisi kestirilmiştir.

Sonuçlara göre, $G_0 = G_1 = \begin{bmatrix} 12,49 & -0,034 \\ -0,034 & 6,94E-17 \end{bmatrix}$ olarak bulunmuştur. Buna göre

bireylerin başlangıç yanıt düzeyleri arasındaki değişimin (12,49) zaman içinde sahip oldukları eğilimlerin değişiminden (6,94E-17) çok daha fazla olduğu görülmektedir. Bu modelde her iki grup için de G matrisinin aynı olduğu varsayılmıştır. Bunun tersine, denek içi değişimlerin her iki grupta farklılık gösterdiği, yani her iki grup

için iki ayrı R matrisi varsayılmıştır. Sonuçlara göre $R_0 = \begin{bmatrix} 132,37 & 0 \\ 0 & 132,37 \end{bmatrix}$ ve

$R_1 = \begin{bmatrix} 215,93 & 0 \\ 0 & 215,93 \end{bmatrix}$ olarak gösterilebilir. Burada, hastanede tedavi alanların

(grup0) denek içi değişimlerinin ev programı alanlara (grup1) göre daha düşük olduğu görülmektedir (sırasıyla $\sigma_0^2 = 132,37$; $\sigma_1^2 = 215,93$). Bu modelde, R matrislerinin diyagonal matris olarak varsayıldığı ve G matrisinin faktör analitik yapıda varsayıldığından Bölüm 2.2.3'te bahsedilmiştir.

Model V:

$$y_{ij} = \beta_1 grup_0 + \beta_2 grup_1 + (\beta_3 grup_0 + \beta_4 grup_1) zaman2_0 + (\beta_5 grup_0 + \beta_6 grup_1) zaman2_5 + (\beta_7 grup_0 + \beta_8 grup_1) zaman2_{30} + (\beta_9 grup_0 + \beta_{10} grup_1) zaman2_{90} + v_{0i} + v_{0i} zaman$$

biçiminde kurulmuştur. Bu modele ait parametre kestirimleri Çizelge 3.9'dan elde edilebilir.

Çizelge 3.9. Sabit etkiler için SEÇOK kestirimleri.

Solution for Fixed Effects							
Effect	grup	zaman2	Estimate	Standard Error	DF	t Value	Pr > t
grup	0		176.96	2.5271	165	70.03	<.0001
grup	1		161.53	2.9212	165	55.30	<.0001
zaman2		0	14.7667	3.9574	165	3.73	0.0003
zaman2		5	-54.8333	3.9486	165	-13.89	<.0001
zaman2		30	-26.6333	3.9082	165	-6.81	<.0001
zaman2		90	-10.9333	3.8356	165	-2.85	0.0049
zaman2		180	0
grup*zaman2	0	0	-13.7296	5.1838	165	-2.65	0.0089
grup*zaman2	0	5	11.3148	5.1697	165	2.19	0.0300
grup*zaman2	0	30	20.1519	5.1045	165	3.95	0.0001
grup*zaman2	0	90	9.0815	4.9868	165	1.82	0.0704
grup*zaman2	0	180	0
grup*zaman2	1	0	0
grup*zaman2	1	5	0
grup*zaman2	1	30	0
grup*zaman2	1	90	0
grup*zaman2	1	180	0

Burada 180. gün referans olarak alınmıştır. Bu nedenle, grup=0 için bulunan 176,96 değeri aslında grup0'da 180. gün değeri kestirimidir. Aynı şekilde, grup=1 için bulunan 161,53 değeri grup1 için 180. gün kestirimidir. Diğer parametre kestirimleri ise 180. gün ile diğer günler arasındaki farkları göstermektedir. Buna göre parametre kestirimleri sırasıyla;

$$\beta_1 = 176,96$$

$$\beta_2 = 161,53$$

$$\beta_3 = 14,7667 - 13,7296 = 1,037$$

$$\beta_4 = 14,7667$$

$$\beta_5 = -54,8333 + 11,3148 = -43,5185 ;$$

$$\beta_6 = -54,8333$$

$$\beta_7 = -26,6333 + 20,1519 = -6,4814$$

$$\beta_8 = -26,6333$$

$$\beta_9 = -10,9333 + 9,0815 = -1,8518$$

$$\beta_{10} = -10,9333 \text{ olacaktır.}$$

Çizelge 3.10. Sabit etkilerin testi.

The Mixed Procedure				
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
grup	1	165	43.08	<.0001
zaman2	4	165	159.58	<.0001
grup*zaman2	4	165	13.66	<.0001

Çizelge 3.10'da sırasıyla;

grup için, $H_0 : \beta_1 + \beta_3 + \beta_5 + \beta_7 + \beta_9 = \beta_2 + \beta_4 + \beta_6 + \beta_8 + \beta_{10} = 0$;

zaman2 için: $H_0 : \beta_1 + \beta_2 = \beta_3 + \beta_4 = \dots = \beta_9 + \beta_{10} = 0$ ve son olarak,

grup*zaman2 için: $H_0 : \beta_1 = \beta_2 = \dots = \beta_{10} = 0$ hipotezlerinin testleri yer almaktadır.

Çizelge 3.11. Marjinal beklenen değerlere ilişkin sonuçlar.

Least Squares Means							
Effect	grup	zaman2	Estimate	Standard Error	DF	t Value	Pr > t
grup*zaman2	0	0	178.00	3.2683	165	54.46	<.0001
grup*zaman2	0	5	133.44	3.2441	165	41.13	<.0001
grup*zaman2	0	30	170.48	3.1258	165	54.54	<.0001
grup*zaman2	0	90	175.11	2.8605	165	61.22	<.0001
grup*zaman2	0	180	176.96	2.5271	165	70.03	<.0001
grup*zaman2	1	0	176.30	3.5212	165	50.07	<.0001
grup*zaman2	1	5	106.70	3.5010	165	30.48	<.0001
grup*zaman2	1	30	134.90	3.4028	165	39.64	<.0001
grup*zaman2	1	90	150.60	3.1858	165	47.27	<.0001
grup*zaman2	1	180	161.53	2.9212	165	55.30	<.0001

Çizelge 3.11'de marjinal beklenen değerlere, $E(y_i) = \mu_i$, ilişkin regresyon kestirimleri ve önemlilik testleri yer almaktadır. Burada test edilen hipotezler sırasıyla;

$$H_0 : \beta_1 + \beta_3 = 0$$

$$H_0 : \beta_1 + \beta_5 = 0$$

$$H_0 : \beta_1 + \beta_7 = 0$$

$$H_0 : \beta_1 + \beta_9 = 0$$

$$H_0 : \beta_1 = 0$$

$$H_0 : \beta_2 + \beta_4 = 0$$

$$H_0 : \beta_2 + \beta_6 = 0$$

$$H_0 : \beta_2 + \beta_8 = 0$$

$$H_0 : \beta_2 + \beta_{10} = 0$$

$$H_0 : \beta_2 = 0 \text{ şeklindedir.}$$

Sonuçlara göre her iki grupta da ikinci ölçüm zamanında düşüş yaşanırken zamanla başlangıç değerlerine dönüş söz konusudur.

Çizelge 3.12. Marjinal beklenen değerlerin farklarına ilişkin sonuçlar.

Differences of Least Squares Means								Adj=Bonferroni	
Effect	grup	zaman2	_grup	_zaman2	Estimate	SE	DF	t Value	Pr> t
grup*zaman2	0	0	0	5	44.5556	3.1315	165	14.23	<.0001
grup*zaman2	0	0	0	30	7.5185	3.1375	165	2.40	0.7956
grup*zaman2	0	0	0	90	2.8889	3.1869	165	0.91	1.0000
grup*zaman2	0	0	0	180	1.0370	3.3483	165	0.31	1.0000
grup*zaman2	0	0	1	0	1.7000	4.8042	165	0.35	1.0000
grup*zaman2	0	0	1	5	71.3000	4.7894	165	14.89	<.0001
grup*zaman2	0	0	1	30	43.1000	4.7181	165	9.14	<.0001
grup*zaman2	0	0	1	90	27.4000	4.5641	165	6.00	<.0001
grup*zaman2	0	0	1	180	16.4667	4.3835	165	3.76	0.0107
grup*zaman2	0	5	0	30	-37.0370	3.1356	165	-11.81	<.0001
grup*zaman2	0	5	0	90	-41.6667	3.1810	165	-13.10	<.0001
grup*zaman2	0	5	0	180	-43.5185	3.3368	165	-13.04	<.0001
grup*zaman2	0	5	1	0	-42.8556	4.7878	165	-8.95	<.0001
grup*zaman2	0	5	1	5	26.7444	4.7730	165	5.60	<.0001
grup*zaman2	0	5	1	30	-1.4556	4.7014	165	-0.31	1.0000
grup*zaman2	0	5	1	90	-17.1556	4.5468	165	-3.77	0.0101
grup*zaman2	0	5	1	180	-28.0889	4.3655	165	-6.43	<.0001
grup*zaman2	0	30	0	90	-4.6296	3.1562	165	-1.47	1.0000
grup*zaman2	0	30	0	180	-6.4815	3.2835	165	-1.97	1.0000
grup*zaman2	0	30	1	0	-5.8185	4.7084	165	-1.24	1.0000
grup*zaman2	0	30	1	5	63.7815	4.6934	165	13.59	<.0001
grup*zaman2	0	30	1	30	35.5815	4.6205	165	7.70	<.0001
grup*zaman2	0	30	1	90	19.8815	4.4632	165	4.45	0.0007
grup*zaman2	0	30	1	180	8.9481	4.2783	165	2.09	1.0000
grup*zaman2	0	90	0	180	-1.8519	3.1869	165	-0.58	1.0000
grup*zaman2	0	90	1	0	-1.1889	4.5367	165	-0.26	1.0000
grup*zaman2	0	90	1	5	68.4111	4.5210	165	15.13	<.0001
grup*zaman2	0	90	1	30	40.2111	4.4454	165	9.05	<.0001
grup*zaman2	0	90	1	90	24.5111	4.2816	165	5.72	<.0001
grup*zaman2	0	90	1	180	13.5778	4.0885	165	3.32	0.0497
grup*zaman2	0	180	1	0	0.6630	4.3342	165	0.15	1.0000
grup*zaman2	0	180	1	5	70.2630	4.3178	165	16.27	<.0001
grup*zaman2	0	180	1	30	42.0630	4.2385	165	9.92	<.0001
grup*zaman2	0	180	1	90	26.3630	4.0664	165	6.48	<.0001
grup*zaman2	0	180	1	180	15.4296	3.8626	165	3.99	0.0044
grup*zaman2	1	0	1	5	69.6000	3.7943	165	18.34	<.0001
grup*zaman2	1	0	1	30	41.4000	3.7988	165	10.90	<.0001
grup*zaman2	1	0	1	90	25.7000	3.8356	165	6.70	<.0001
grup*zaman2	1	0	1	180	14.7667	3.9574	165	3.73	0.0118
grup*zaman2	1	5	1	30	-28.2000	3.7973	165	-7.43	<.0001
grup*zaman2	1	5	1	90	-43.9000	3.8311	165	-11.46	<.0001
grup*zaman2	1	5	1	180	-54.8333	3.9486	165	-13.89	<.0001
grup*zaman2	1	30	1	90	-15.7000	3.8126	165	-4.12	0.0027
grup*zaman2	1	30	1	180	-26.6333	3.9082	165	-6.81	<.0001
grup*zaman2	1	90	1	180	-10.9333	3.8356	165	-2.85	0.2215

Çizelge 3.12'deki sonuçlara bakılacak olursa, hastanede tedavi alan bireylerde, ameliyat öncesi (0. gün) ve 5. gün ölçüm değerleri arasında fark bulunmuştur ($p<0,001$). Yine aynı grupta 0. gün ve 30. gün, 0. gün ve 90. gün ile 0. gün ve 180. gün ölçüm değerleri arasındaki fark anlamlı değildir (sırasıyla $p=0,796$; $p=1,00$; $p=1,00$). Ancak ikinci grupta 0. gün ile diğer bütün günler arasındaki anlamlı fark bulunmuştur (sırasıyla $p<0,001$; $p<0,001$; $p<0,001$; $p=0,012$). Bunlara ek olarak ilk ölçüm zamanında gruplar arası fark anlamlı değilken ($p=1,00$), 5. gün, 30. gün, 90. gün ve 180. gün ölçümleri her iki grupta farklıdır (sırasıyla $p<0,001$; $p<0,001$; $p<0,001$; $p<0,001$; $p=0,004$). Diğer bütün karşılaştırmalar yine Çizelge 3.12'de yer almaktadır.

Çizelge 3.13. Her birey için rasgele kesim noktası ve eğim önkestirimleri.

The Mixed Procedure Solution for Random Effects						
Effect	id	Estimate	Std Err Pred	DF	t Value	Pr > t
Intercept	1	-3.6417	5.8465	165	-0.62	0.5342
zaman	1	0.009980	0.01602	165	0.62	0.5342
.....						
Intercept	23	-31.0971	5.8465	165	-5.32	<.0001
zaman	23	0.08522	0.01602	165	5.32	<.0001
Intercept	24	-0.2805	5.8465	165	-0.05	0.9618
zaman	24	0.000769	0.01602	165	0.05	0.9618
Intercept	25	-6.3735	5.8465	165	-1.09	0.2772
zaman	25	0.01747	0.01602	165	1.09	0.2772
Intercept	26	-0.7327	5.8465	165	-0.13	0.9004
zaman	26	0.002008	0.01602	165	0.13	0.9004
Intercept	27	-3.0983	5.8465	165	-0.53	0.5969
zaman	27	0.008491	0.01602	165	0.53	0.5969
Intercept	28	-9.0106	6.8377	165	-1.32	0.1894
zaman	28	0.02469	0.01874	165	1.32	0.1894
Intercept	29	10.9879	6.8377	165	1.61	0.1100
zaman	29	-0.03011	0.01874	165	-1.61	0.1100
Intercept	30	-2.1285	6.8377	165	-0.31	0.7560
zaman	30	0.005833	0.01874	165	0.31	0.7560
Intercept	31	-35.4446	6.8377	165	-5.18	<.0001
zaman	31	0.09713	0.01874	165	5.18	<.0001
.....						
Intercept	51	25.7137	6.8377	165	3.76	0.0002
zaman	51	-0.07047	0.01874	165	-3.76	0.0002
Intercept	52	-11.4157	6.8377	165	-1.67	0.0969
zaman	52	0.03128	0.01874	165	1.67	0.0969
Intercept	53	-6.1884	6.8377	165	-0.91	0.3668
zaman	53	0.01696	0.01874	165	0.91	0.3668
Intercept	54	-10.2537	6.8377	165	-1.50	0.1356
zaman	54	0.02810	0.01874	165	1.50	0.1356
Intercept	55	-11.8955	6.8377	165	-1.74	0.0838
zaman	55	0.03260	0.01874	165	1.74	0.0838
Intercept	56	5.2102	6.8377	165	0.76	0.4472
zaman	56	-0.01428	0.01874	165	-0.76	0.4472
Intercept	57	10.0914	6.8377	165	1.48	0.1419
zaman	57	-0.02766	0.01874	165	-1.48	0.1419

Çizelge 3.13’de her birey için rasgele kesim noktası (v_{0i}) ve rasgele eğim (v_{1i}) parametrelerinin önkestirimleri ve önemlilikleri yer almaktadır. Çizelge 3.14’de bütün veri setinin yanında, her bireyin her zaman noktasındaki koşullu beklenen yanıt $E(y_i | v_i)$ önkestirimleri ve %95 güven sınırları yer almaktadır. Hasta bazında çıkarsamalar bu tablolar yardımıyla yapılabilir. Çizelge 3.13’e bakılacak olursa; örneğin, 23. bireyin rasgele kesim noktası parametresi $-31,0971 \pm 5,8465$, rasgele eğim parametresi ise $0,08522 \pm 0,01602$ olarak önkestirilmiştir ve her iki parametre de anlamlı bulunmuştur ($p < 0,001$). Çizelge 3.14’e bakılacak olursa, 23. bireyin hastanede tedavi aldığı (grup=0) ve 75 yaşında olduğu görülmektedir. Bu kişinin her zaman noktasındaki önkestirilen yanıt değerleri ve %95 güven sınırları ise yine çizelge 3.14’te görüldüğü gibi; ameliyat öncesi: 146,903 (135,508-158,298), ameliyattan sonra 5. gün: 102,773 (91,512-114,035), 30. gün: 141,941 (131,340-152,542), 90. gün: 151,684 (142,631-160,737), 180.gün: 161,206 (154,310-168,101) olarak önkestirilmiştir. Çizelge 3.14’te yer alan bu sonuçlara Çizelge 3.9 ‘da bulunan sabit etki parametre kestirimleri ve Çizelge 3.13’de bulunan rasgele etki parametre önkestirimleri kullanılarak ulaşılabilir. Örneğin yine 23. birey düşünüldüğünde, bu bireye ait beklenen ortalama önkestirimi, $E(y_i | v_i) = X_i \hat{\beta} + Z_i \hat{v}_i$, 0. gün için

$$y_{23,0} = \beta_1 \text{grup}_0 + \beta_3 \text{grup}_0 \text{zaman} 2_0 + v_{0,23} + v_{1,23} \text{zaman}$$

$$= 176,96 + 1,0371 - 31,0971 + 0,08522 = 146,90$$

şeklinde hesaplanabilir.

Çizelge 3.14. Her bireyin her zaman noktasındaki beklenen fleksiyon yanıt önektirimleri.

Model V: Zaman Hem Sürekli Hem de Kategorik, Rasgele Kesim Noktası & Eğim Modeli														16	
10:01 Tuesday, June 29, 2007															
Obs	id	grup	yas	vki	zaman	fleks	zaman2	Pred	StdErr	Pred	DF	Alpha	Lower	Upper	Resid
1	1	0	54	26.44	0	180	0	174.358	5.77109	165	0.05	162.964	185.753	5.6417	
2	1	0	54	26.44	5	110	5	129.853	5.70373	165	0.05	118.591	141.114	-19.8526	
3	1	0	54	26.44	30	170	30	167.139	5.36910	165	0.05	156.538	177.740	2.8608	
4	1	0	54	26.44	90	180	90	172.368	4.58505	165	0.05	163.315	181.421	7.6324	
5	1	0	54	26.44	180	180	180	175.118	3.49219	165	0.05	168.223	182.013	4.8824	
.....															
99	20	0	63	28.92	90	165	90	160.316	4.58505	165	0.05	151.263	169.369	4.6840	
100	20	0	63	28.92	180	172	180	167.012	3.49219	165	0.05	160.116	173.907	4.9884	
101	21	0	54	27.72	0	180	0	176.556	5.77109	165	0.05	165.161	187.951	3.4440	
102	21	0	54	27.72	5	120	5	132.020	5.70373	165	0.05	120.759	143.282	-12.0203	
103	21	0	54	27.72	30	170	30	169.156	5.36910	165	0.05	158.555	179.757	0.8438	
104	21	0	54	27.72	90	180	90	174.023	4.58505	165	0.05	164.970	183.076	5.9767	
105	21	0	54	27.72	180	180	180	176.231	3.49219	165	0.05	169.336	183.126	3.7687	
106	22	0	46	26.00	0	180	0	165.674	5.77109	165	0.05	154.280	177.069	14.3255	
107	22	0	46	26.00	5	90	5	121.288	5.70373	165	0.05	110.026	132.550	-31.2878	
108	22	0	46	26.00	30	160	30	159.169	5.36910	165	0.05	148.568	169.770	0.8307	
109	22	0	46	26.00	90	170	90	165.826	4.58505	165	0.05	156.773	174.878	4.1744	
110	22	0	46	26.00	180	175	180	170.717	3.49219	165	0.05	163.822	177.613	4.2826	
111	23	0	75	32.24	0	155	0	146.903	5.77109	165	0.05	135.508	158.298	8.0971	
112	23	0	75	32.24	5	80	5	102.773	5.70373	165	0.05	91.512	114.035	-22.7735	
113	23	0	75	32.24	30	135	30	141.941	5.36910	165	0.05	131.340	152.542	-6.9410	
114	23	0	75	32.24	90	145	90	151.684	4.58505	165	0.05	142.631	160.737	-6.6838	
115	23	0	75	32.24	180	160	180	161.206	3.49219	165	0.05	154.310	168.101	-1.2055	
116	24	0	42	31.54	0	180	0	177.719	5.77109	165	0.05	166.325	189.114	2.2805	
117	24	0	42	31.54	5	145	5	133.168	5.70373	165	0.05	121.906	144.429	11.8323	
118	24	0	42	31.54	30	165	30	170.224	5.36910	165	0.05	159.623	180.825	-5.2240	
119	24	0	42	31.54	90	165	90	174.900	4.58505	165	0.05	165.847	183.953	-9.8998	
120	24	0	42	31.54	180	173	180	176.821	3.49219	165	0.05	169.926	183.716	-3.8208	
121	25	0	68	29.12	0	170	0	171.627	5.77109	165	0.05	160.232	183.021	-1.6265	
122	25	0	68	29.12	5	120	5	127.158	5.70373	165	0.05	115.897	138.420	-7.1583	
123	25	0	68	29.12	30	165	30	164.632	5.36910	165	0.05	154.031	175.233	0.3680	
.....															
.....															
.....															

3.2. Tekrarlı Ölçümlerde Varyans Analizi Sonuçları

Çizelge 3.15. Küresellik testi.

Sphericity Tests				
Variables	Mauchly's		Chi-Square	Pr > ChiSq
	DF	Criterion		
Transformed Variates	9	0.1210951	112.77215	<.0001
Orthogonal Components	9	0.1210951	112.77215	<.0001

Küresellik varsayımı reddedildiği için ($p < 0,001$), zaman ve zaman*grup etkileşimi için çok değişkenli varyans analizi sonuçları verilmiştir (çizelge 3.16 ve çizelge 3.17).

Çizelge 3.16. Zaman etkisi için çok değişkenli varyans analizi sonuçları.

The GLM Procedure					
Repeated Measures Analysis of Variance					
MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no zaman Effect					
H = Type III SSCP Matrix for zaman					
E = Error SSCP Matrix					
Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.15975841	68.37	4	52	<.0001
Pillai's Trace	0.84024159	68.37	4	52	<.0001
Hotelling-Lawley Trace	5.25945149	68.37	4	52	<.0001
Roy's Greatest Root	5.25945149	68.37	4	52	<.0001

Çizelge 3.17. Zaman*grup etkileşimi için çok değişkenli varyans analizi sonuçları.

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no zaman*grup Effect					
H = Type III SSCP Matrix for zaman*grup					
E = Error SSCP Matrix					
Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.48315582	13.91	4	52	<.0001
Pillai's Trace	0.51684418	13.91	4	52	<.0001
Hotelling-Lawley Trace	1.06972568	13.91	4	52	<.0001
Roy's Greatest Root	1.06972568	13.91	4	52	<.0001

Çizelge 3.16 ve 3.17'deki sonuçlara göre hem zaman etkisi hem de zaman*grup etkileşimi anlamlı bulunmuştur ($p < 0,001$). Grup etkisinin testi Çizelge 3.18'de yer almaktadır. Sonuçlara göre grup etkisi de anlamlı bulunmuştur ($p < 0,001$).

Çizelge 3.18. Grup etkisi için varyans analizi sonuçları.

The GLM Procedure					
Repeated Measures Analysis of Variance					
Tests of Hypotheses for Between Subjects Effects					
Source	DF	Type III SS	Mean Square	F Value	Pr > F
grup	1	30720.50842	30720.50842	41.28	<.0001
Error	55	40935.39333	744.27988		

Bunların yanı sıra, zamanın karesel etkisi anlamlı bulunmamıştır ($p=0,896$). Buna ilişkin sonuçlara burada yer verilmemiştir. Çizelge 3.19.'de ise, her zaman noktasında alınan ölçümler bakımından gruplar arası farkın anlamlılığı test edilmiştir.

Çizelge 3.19. Her zaman noktasında gruplar arası farkın testi.

The GLM Procedure				
Least Squares Means				
Adjustment for Multiple Comparisons: Bonferroni				
grup	fleks1 LSMEAN	Standard Error	H0:LSMEAN=0 Pr > t	H0:LSMean1= LSMean2 Pr > t
0	178.000000	1.158282	<.0001	0.2916
1	176.300000	1.098842	<.0001	
grup	fleks2 LSMEAN	Standard Error	H0:LSMEAN=0 Pr > t	H0:LSMean1= LSMean2 Pr > t
0	133.444444	5.159214	<.0001	0.0004
1	106.700000	4.894460	<.0001	
grup	fleks3 LSMEAN	Standard Error	H0:LSMEAN=0 Pr > t	H0:LSMean1= LSMean2 Pr > t
0	170.481481	3.427502	<.0001	<.0001
1	134.900000	3.251614	<.0001	
grup	fleks4 LSMEAN	Standard Error	H0:LSMEAN=0 Pr > t	H0:LSMean1= LSMean2 Pr > t
0	175.111111	3.346828	<.0001	<.0001
1	150.600000	3.175080	<.0001	
grup	fleks5 LSMEAN	Standard Error	H0:LSMEAN=0 Pr > t	H0:LSMean1= LSMean2 Pr > t
0	176.962963	1.776470	<.0001	<.0001
1	161.533333	1.685307	<.0001	

Çizelge 3.19'deki sonuçlara göre ilk ölçüm zamanında gruplar arası fark anlamsız ($p=0,292$), 5. gün, 30. gün, 90. gün ve 180. günlerde gruplar arası fark anlamlı bulunmuştur ($p<0,001$).

Her grup için zaman noktalarının birbirleri ile karşılaştırmaları Bonferroni düzeltilmeli bağımlı gruplarda t-testi kullanılarak yapılmıştır. Bulunan p değerleri, PROC MIXED işlemi sonucunda bulunan p değerleri (Çizelge 3.12) ile birlikte Çizelge 3.20'de verilmiştir.

Çizelge 3.20. Gruplar için zamanlara ait çoklu karşılaştırmalar.

Grup=0 GLM							
0-5	$p<0,001$	5-30	$p<0,001$	30-90	$p=0,007$	90-180	$p=0,045$
0-30	$p=0,002$	5-90	$p<0,001$	30-180	$p=0,004$		
0-90	$p=0,008$	5-180	$p<0,001$				
0-180	$p=0,257$						
Grup=0 MIXED							
0-5	$p<0,001$	5-30	$p<0,001$	30-90	$p=1,000$	90-180	$p=1,000$
0-30	$p=0,795$	5-90	$p<0,001$	30-180	$p=1,000$		
0-90	$p=1,000$	5-180	$p<0,001$				
0-180	$p=1,000$						
Grup=1 GLM							
0-5	$p<0,001$	5-30	$p<0,001$	30-90	$p<0,001$	90-180	$p<0,001$
0-30	$p<0,001$	5-90	$p<0,001$	30-180	$p<0,001$		
0-90	$p<0,001$	5-180	$p<0,001$				
0-180	$p<0,001$						
Grup=1 MIXED							
0-5	$p<0,001$	5-30	$p<0,001$	30-90	$p=0,002$	90-180	$p=0,221$
0-30	$p<0,001$	5-90	$p<0,001$	30-180	$p<0,001$		
0-90	$p<0,001$	5-180	$p<0,001$				
0-180	$p=0,012$						

Çizelge 3.20'de yer alan sonuçlara bakıldığında, her iki yöntemle bulunan anlamlılıkların farklı olduğu görülmektedir.

4. TARTIŞMA

Uzunlamasına veriler, tekrarlı ölçümler biçiminde toplanan verilerde, tekrarların zaman içinde alındığı özel bir durumdur. Doğrusal karma etki modelleri, ilk olarak Laird ve Ware (1982) tarafından karma etki modelleri iki aşamalı model yaklaşımı yaklaşımı kullanılarak uygulamalı olarak anlatılmıştır. Bundan sonra doğrusal karma etki modelleri uzunlamasına verilerin analizinde esnek çözümler sağlaması ile son zamanda bu alanda bilgisayar yazılımlarının da gelişmesi ile oldukça popüler bir yöntem haline gelmiştir.

Bu çalışmada, fleksiyon ölçümlerinin gruplara göre zaman içindeki değişimlerinin analizi için beş farklı doğrusal karma etki modeli kurulmuştur. Bunlar arasından model seçimi AIC ve BIC bilgi kriterlerine göre değerlendirilmiştir. Bütün modellerde varyans kovaryans kestirimlerinde SEÇOK yöntemi kullanılmıştır. Kurulan üç modelde zaman sürekli değişken olarak alınmış ve rasgele kesim noktası ile eğim tanımlanmıştır.

İlk modele grup, zamanın birinci dereceden ve zamanın ikinci dereceden ana etkileri ile grup ve zaman arasındaki etkileşimler alınmıştır. Ayrıca yaş ve vki ortak değişkenleri de modele dahil edilmiştir. İlk modelde zamanın birinci dereceden ve karesel ana etkisinin anlamlı olmadığı ($p=0,206$; $p=0,948$), ayrıca yaş ile vki'nin de fleksiyon ölçümleri üzerinde anlamlı bir değişim yapmadıkları bulunmuştur (sırasıyla $p=0,369$; $p=0,146$). İlk modelde grup ana etkisi ile grup*zaman etkileşimi ve grubun zamanın karesi ile olan etkileşimi (grup*zaman*zaman) anlamlı olarak bulunmuştur (sırasıyla $p<0,001$; $p=0,038$; $p=0,029$).

İkinci modelde ise grup ana etkisi ile grup*zaman etkileşimi ve grubun zamanın karesel etkisi ile etkileşimi (grup*zaman*zaman) alınmıştır. Ancak bu modelde, grup ve zamanın karesel etkisi arasındaki etkileşim anlamsız bulunmuştur ($p=0,077$). Grup ana etkisi ile grup*zaman etkileşimi anlamlı bulunmuştur (sırasıyla $p<0,001$; $p=0,018$).

Üçüncü modelde yalnızca grup ana etkisi ile grup*zaman etkileşimi alınmıştır ve her iki etki de anlamlı bulunmuştur ($p<0,001$).

Dördüncü model rasgele kesim noktası modelidir ve zaman değişkeni modele kategorik olarak alınmıştır. Bu modelde grup ve zaman ana etkileri ile grup*zaman etkileşimi yer almaktadır. Her üç etki de anlamlı olarak bulunmuştur. ($p<0,001$).

Beşinci model bir rasgele kesim noktası ve eğim modelidir. Bu modelde zaman hem sürekli hem de kategorik olarak modele alınmıştır. Modelde grup ve zaman ana etkisi ile bunların etkileşimleri yer almaktadır. Her üç etki de anlamlı olarak bulunmuştur ($p<0,001$). Veriye en uygun doğrusal karma etki modelinin, AIC ve BIC bilgi kriterlerine göre beşinci model olduğu söylenebilir.

Seçilen doğrusal karma etki modeli sonucuna göre, hastanede tedavi alanların ameliyat öncesi ortalama fleksiyon değeri $178,00\pm 3,27$ ile ev programı alan hastaların ameliyat öncesi ortalama fleksiyon değeri $176,30\pm 3,52$ farklılık bulunmamıştır ($p=1,000$). Ameliyat sonrası beşinci gün değerlerinde her iki grupta da önemli bir düşüş olmuştur ve her iki grupta beşinci gün ortalama fleksiyon değerleri arasında anlamlı fark vardır (sırasıyla $133,44\pm 3,24$; $106,70\pm 3,50$; $p<0,001$). Hastanede tedavi gören grup ameliyat öncesi değerlerine daha çabuk yaklaşırken ev programı alan grupta iyileşme daha geç olmuştur ve her iki grubun 30., 90. ve 180. günlerdeki ortalama ölçüm değerleri arasında önemli fark bulunmuştur (hastanede tedavi görenler ve ev programı alanlar için ortalama±shata sırasıyla, 30. gün $170,48\pm 3,12$; $134,90\pm 3,40$ $p<0,001$, 90. gün $175,11\pm 2,86$; $150,60\pm 3,19$ $p<0,001$, 180. gün $176,86\pm 2,52$; $161,53\pm 2,92$ $p=0,004$).

Veri setine klasik tekrarlı ölçümlerde varyans analizi uygulandığında da, grup ve zaman etkisi ile grup*zaman etkileşimi anlamlı bulunmuştur ($p<0,001$). Zamanın karesel etkisinin anlamlı olmadığı saptanmıştır ($p=0,896$). Bulunan sonuçlar doğrusal karma etki modeli ile benzerdir.

Doğrusal karma etki modelleri popülasyon hakkında çıkarsamaların yanı sıra bireysel bazda da çıkarsamalara olanak sunmaktadır. Örneğin; hastanede tedavi görenlerin marjinal ortalama kestirimleri alınan ölçüm sırasına göre; 178; 133,44; 170,48; 175,11 ve 176,96'dır. Rasgele etkilerin testlerine göre 23. bireyin (hastanede tedavi almaktadır) rasgele kesim noktası ($v_{0,23}$) $-31,097 \pm 5,846$; rasgele eğimi ($v_{1,23}$) $0,085 \pm 0,016$ olduğu ve bu değerlerin anlamlı oldukları saptanmıştır ($p < 0,001$). Yani bu bireyin popülasyondan önemli olarak saptığı görülmüştür. Bireye ait ölçüm değerleri sırasıyla 155; 80; 135; 145 ve 160'tır. Önkestirilen koşullu beklenen değerleri, $E(y_{23} | v_{23})$, ise sırasıyla 146,903; 102,773; 141,941; 151,684 ve 161,206'dır.

Quene ve Bergh (2004) yaptıkları simülasyon çalışmasında, doğrusal karma etki modellerinin verideki sabit etkileri ve kontrastları bulmada klasik varyans analizi yaklaşımından daha güçlü sonuçlar verdiğini vurgulamışlardır. Sullivan ve ark. (1999) tarafından yapılan dengeli veri yapıları için yaptıkları simülasyon çalışmasında ise, klasik tekrarlı ölçümlerde varyans analizi sonuçlarında, sabit etkiler için bulunan standart hataların karma etki modeli sonuçlarına göre daha düşük olduğunu bulmuşlar ve bunun Tip I hata miktarını arttıracakını belirtmişlerdir. Benzer biçimde, bu tez kapsamında yapılan uygulama sonuçlarına göre, standart hataların çoğunun klasik varyans analizi sonuçlarında daha düşük olduğu görülmüştür. Krauger ve Tian (2004), dengesiz bir veri seti üzerinde yaptıkları çalışmada doğrusal karma etki modellerinin klasik varyans analizi yaklaşımına göre daha iyi sonuç verdiğini tartışmışlardır ve yanıt üzerinde etkili faktörleri belirlemede daha etkili olduğunu vurgulamışlardır.

5. SONUÇ VE ÖNERİLER

Çalışmadaki bireylerin fleksiyon değerlerindeki zaman içindeki değişiminin aldıkları tedaviye göre farklılık gösterdiği hem doğrusal karma etki modeli hem de klasik tekrarlı ölçümlerde varyans analizi ile saptanmıştır. Hastanede tedavi gören bireylerde ameliyat sonrası fleksiyon değerlerinin ameliyat öncesi fleksiyon değerine daha hızlı yaklaştığı bulunmuştur. Doğrusal karma etki modelinde her iki grup için farklı denek içi değişim belirlenmiştir. Buna göre hastanede tedavi görenlerin denek içi değişimi ev programı alanlara göre daha düşüktür.

Marjinal ortalama parametre kestirimleri için kurulan doğrusal karma etki modeli ile tekrarlı ölçümlerde varyans analizi sonuçları benzer çıkmıştır. Bunun nedeninin denekler arası değişimin az olması nedeniyle G 'nin bileşenlerinin 0'a yakın çıkmasından kaynaklandığı düşünülmektedir. Denek içi kovaryans yapısı da $\sigma^2 I_{n_i}$ biçiminde tanımlandığından, tekrarlı ölçümler için ortaya çıkacak kovaryans yapısı da, $Z_i G Z_i' + R$, diyagonal olacaktır. Bu durumda en küçük kareler aynı kovaryans parametrelerine bölüneceğinden, doğrusal karma etki modellerinde kullanılan genelleştirilmiş en küçük kareler kestirimleri, tekrarlı ölçümlerde varyans analizinde kullanılan standart en küçük kareler kestirimlerine eşit olacaktır. Denekler arası değişimin fazla olduğu durumlarda karma etki modellerinin kullanılması daha da etkili olabilir. Karma etki modellerinde, hem ortalama hem de varyans-kovaryans yapısının daha esnek biçimde modellenmesi avantajlı bir durumdur. Ayrıca, karma etki modeli yardımı ile veri hakkında daha fazla çıkarsamada bulunulmuştur. Her bireyin, marjinal ortalama yanıtın nasıl saptığı ve bireylere özel çıkarsamalar rasgele etkiler kullanılarak incelenmiştir. Bunların yanı sıra, karma etki modelleri yardımıyla zamanın sürekli bir değişken olarak etkisi de incelenmiştir. Ayrıca karma etki modellerinde ortalama yapısı ve varyans yapısı daha esnek biçimde modellenebilir. Karma etki modelleri ile denek içi ve denekler arası değişim ayrı ayrı belirlenebilir ve parametre kestirimlerinin standart hataları ile önemlilik testleri kestirilen kovaryans parametrelerine dayanır (Littell ve ark., 2006). Denek içi değişim ve denekler arası değişim gruplar için ayrı olarak belirlenebilir. Bu çalışmada yalnızca denek içi değişim gruplar arasında farklı olarak belirlenmiş ve

denek ii hataların yalnızca lm hataları olduėu varsayılmıřtır. Bundan sonra, veri yapısının dengesiz olduėu, denek ii hataların birbirleri ile iliřkili olduėu durumlarda karma etki modelleri incelenebilir. Bu alıřmada yalnızca iki ařamalı model incelenmiřtir. Bu durum daha da geniřletilebilir. rneėin farklı hastanelerden alınan hastalarda yapılan uzunlamasına alıřmaların analizleri karma etki modelleri yardımı ile yapılabilir.

ÖZET

Uzunlamasına Çalışmaların Analizinde Karma Etki Modelleri

Bu çalışmanın amacı, doğrusal karma etki modellerinin uzunlamasına veri analizinde kullanımını incelemek ve sağlık alanında bir veri seti üzerinde uygulamasını yapmaktır.

Uzunlamasına çalışmaların en iyi biçimde çözümlenmesi ve yorumlanması ile ilgili çalışmalar geçtiğimiz yıllarda büyük bir ilerleme göstermiş ve bu anlamda güçlü birçok yöntem geliştirilmiştir. Karma etki modelleri, bireyden elde edilen gözlemlerin bağımsız varsayılmayacağı, verinin yapısında doğal bir hiyerarşik ya da kümelenmiş yapının bulunduğu durumlar için kullanılır. Her bir bireyden belli bir izlem süresi boyunca toplanan ölçümlerin ilişkili olduğu tekrarlı ölçüm tasarımları, bu tür veriler için en uygun tasarımıdır. Özellikle sağlık alanında, veriler genellikle tekrarlı ölçümler biçiminde toplanır. Doğrusal karma etki modelleri uzunlamasına veriler için esnek çözümler yapılabilmesini sağlar.

Çalışmanın ilk bölümünde, tekrarlı ölçümlerin analizinde yaygın olarak kullanılan tekrarlı ölçümlerde varyans analizi ve kısıtlarına değinilmiş, doğrusal karma etki modellerinin avantajlarından bahsedilmiş, rasgele kesim noktası modeli ile rasgele kesim noktası ve rasgele eğim modeli gibi karma etki modelleri tanıtılmıştır. Ayrıca kestirim yöntemleri hakkında bilgi verilmiştir. İkinci kısımda, sağlık alanında toplanan veri seti üzerinde hem karma etki modellerinin hem de varyans analizinin uygulaması SAS paket programı ile gerçekleştirilmiştir.

Meme kanseri olan kadınlardan toplanan veri seti üzerinde hem doğrusal karma etki modeli hem de tekrarlı ölçümlerde varyans analizi uygulanmıştır. Karma etki modelleri, tekrarlı ölçümlerin analizinde etkili ve alternatif bir çözüm yolu olarak önerilmiştir.

Anahtar Sözcükler: Hiyerarşik modeller, karma etki modelleri, sınırlı en çok olabilirlik, tekrarlı ölçümler, uzunlamasına çalışmalar.

SUMMARY

Mixed Effects Models for Analyzing Longitudinal Studies

The purpose of this study was to investigate the use of linear mixed models in longitudinal data analysis and provide a sample application for a data set collected in a health study.

In recent years, studies for analyzing and interpreting longitudinal data have improved significantly and many effective methods have been presented in this frame. Mixed effects models are applied to data that are hierarchical or clustered in nature, in which individual observations cannot be assumed independent. Repeated measures designs, where measurements taken over time on each subject will be correlated, are an important source of such data. Especially in health studies, data are usually collected over time. Linear mixed models provide flexible analysis for longitudinal data.

In the first part of this study, repeated measurements ANOVA, which is widely used in the analysis of repeated measurements and its limitations, were introduced. Then, advantages of mixed models over conventional method were mentioned. Moreover, such mixed models like random intercept, both random intercept and random slope models were described and information about estimation methods was given. In the second part, mixed models and repeated measurements ANOVA were applied on a health study using SAS software.

Data from women with breast cancer were analyzed both using mixed models and ANOVA method. Mixed models analyses were suggested as an alternative and effective analysis tool for repeated measurements data.

Key Words: Hierarchical models, longitudinal data, mixed effects models, repeated measurements, restricted maximum likelihood.

KAYNAKLAR

- ALBERT, P. (1999). Longitudinal data analysis (repeated measures) in clinical trials, *Statistics in Medicine*, **18**: 1707-1732.
- ARMİTAGE, P., COLTON, T. (1998). Encyclopedia of Biostatistics, Wiley, New York , Vol. 4. p.:2657-2662.
- CNAAN, A., LAIRD, N.M., SLASOR, P. (1997). Using the general linear mixed model to analyze unbalanced repeated measures and longitudinal data, *Statistics In Medicine*, **16**: 2349-2380.
- DAVIDIAN, M. (2007). Applied longitudinal data analysis. Erişim: [<http://www.stat.ncsu.edu/people/davidian/st732/>]. Erişim Tarihi :01.03.2007
- FITZMAURICE, G., LAIRD, N., WARE, J. (2004). Applied Longitudinal Analysis, Wiley Series in Probability and Statistics.
- HARTLEY, H.O., RAO J.N.K. (1967). Maximum likelihood estimation for mixed models analysis of variance model, *Biometrika*, **54**: 93-108.
- HEDEKER D., GIBBONS R.D. (2006). Longitudinal Data Analysis, John Wiley & Sons, Inc., Hoboken, New Jersey.
- JUDITHD. S., JOHN B. W. (2003). Applied Longitudinal Data Analysis: Modelling Change And Event Occurrence, Oxford University Pres.
- KAART, T. (2005). The reliability of linear mixed models in genetic studies, Ph.D. Thesis, University of Tartu.
- KHATREE, R. (1999). Nonnegative estimation of variance components: a modification of the Henderson's ANOVA methodology, *The Indian Journal of Statistics*, 61. B. **2**: 261-265.
- KRUEGER, C., TIAN, L. (2004). A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points, *Biological Research for Nursing*, Vol. 6. **2**: 151-157.
- LAIRD, N.M., WARE, J.H. (1982). Random effects models for longitudinal data, *Biometrics*, **38**: 963-974.
- LITTELL, R.C., HENRY P. R., AMMERMAN C. B. (1998). Statistical analysis of repeated measures data using SAS procedures, *J. Anim. Sci.* **76**: 1216-1231.

- LITTELL, R.C., MILLIKEN, G.A., STROUP, W.W., WOLFINGER D.R., SHABENBERGER, O. (2006). SAS for Mixed Models, 2nd Ed., SAS Institute Inc., Cary, NC, USA.
- McCULLOCH, E.C. (2005). Deming short course materials, Lab notes and solutions: Lab 1 notes., 61st Annual Deming Conference, Atlantic City NJ., 5-9 December, Erişim: [<http://www.biostat.ucsf.edu/deming.html>]. Erişim Tarihi: 15.02.2007
- MOLENBERGHS, G., VERBEKE, G. (2001). A review on linear mixed models for longitudinal data, possibly subject to dropout, *Statistical Modelling*, **1**: 235-269.
- ÖZKAN, L. KURT, M., KAHRAMAN, S., KÜÇÜKOĞLU, S., TAŞDELEN, İ., ENGİN, K. (1998). Aksiller diseksiyon ve radyoterapi uygulanan olgularda omuz kısıtlılığı ve lenfödem gelişiminin önlenmesinde rehabilitasyonun rolü, *Türkiye Fiziksel Tıp ve Rehabilitasyon Dergisi [Elektronik Dergi]*, Haziran Sayı:2, Erişim: [<http://www.ftrdergisi.com/>] Erişim tarihi: 12.06.2007
- PATTERSON, H.D., THOMPSON, R. (1971). Recovery of inter-block information when block sizes are unequal, *Biometrika*, Vol. 58. **3**: 545-554.
- QUENE, H., BERGH, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication*, **43**: 103-121.
- RAO, J.N.K. (1998). Marginal Models For Repeated Observations: Inference With Survey Data, Caxleton University School of Mathematics and Statistics Erişim: [http://www.amstat.org/sections/srms/Proceedings/papers/1998_011.pdf] Erişim tarihi: 02.03.2007
- ROBINSON, G.K. (1991). That BLUP is a good thing: the estimation of random effects, *Statistical Science*, **6**: 15-32.
- SERALE, S.R., CASELLA, G., McCULLOCH, C.E. (1992). Variance Components, John Wiley & Sons, New York.
- SULLIVIAN, L. M., DUKES, K. A., LOSINA, E. (1999). An introduction to hierarchical linear modelling, *Statistics in Medicine*, **18**: 855-888.
- SWALLOW, W. H., MONAHAN, J. F. (1984). Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components, *Technometrics*, **26**: 47-57.
- VERBEKE, G. ,MOLENBERGHS, G. (2000). Linear Mixed Models for Longitudinal Data.New York John Wiley Sons.

VITTINGHOFF, E., SHIBOSKI, S.C., GLIDDEN, D.V., McCULLOCH, C.E. (2005), Regression Methods in Biostatistics: Linear, Logistics, Survival and Repeated Measures Models, Springer Science&Business Media Inc.

WOLFINGER, R., CHANG, M., Comparing the SAS GLM and MIXED procedures for repeated measures, SAS Institute Inc., Cary, NC, Erişim:[<http://www.sas.com/rnd/app/papers/mixedglm.pdf>] Erişim Tarihi: 11.04.2006

WU, H., ZHANG, J. (2006). Nonparametric Regression Methods for Longitudinal Data Analysis, John Wiley & Sons, Inc.

ÖZGEÇMİŞ

I- Bireysel Bilgiler

Adı: Beyza

Soyadı: Doğanay

Doğum yeri ve tarihi: Ankara 1981

Uyruğu: T.C.

Medeni Durumu: Bekar

İletişim adresi ve telefonu: Ankara Üni. Tıp Fak. Morfoloji kampüsü,
Biyostatistik A.B.D., Sıhhiye-Ankara, (0312) 310 30 10 / 326.

II- Eğitimi

Yüksek Lisans: Ankara Üniversitesi Biyoistatistik A.B.D. (2004-)

Lisans: Hacettepe Üniversitesi İstatistik Bölümü (1999-2004)

Yabancı Dili: İngilizce

III- Ünvanları

Araş Gör. - Ankara Üniversitesi Biyoistatistik A.B.D. (2005-)

IV- Mesleki Deneyimi

Araş Gör. - Ankara Üniversitesi Biyoistatistik A.B.D. (2005-)

V- Bilimsel İlgi Alanları

Doğrusal modeller

Karma etki modelleri

Tekrarlı ölçümler

Kayıp gözlem durumları

VI- Bilimsel Etkinlikleri

Özbek U., Doğanay B., Elhan A. H.: Klinik Denemelerde Koşullu Güce Bağlı Olarak Örneklem Büyüklüğünün Hesaplanması. 08-12 Mayıs 4. İstatistik Kongresi sözlü bildiri sunumu,2005.

Dođanay B., Köse S.K., Özbek U., Atakurt Y.: Kategorik Veri Çözümlemesinde Logaritmik Doğrusal Modeller ve Bir Uygulama. 20-22 Eylül 8. Ulusal Biyoistatistik Kongresi sözlü bildiri sunumu, 2005.

Özbek U., Öztuna D., Dođanay B., Köse K.: Tekrarlı Ölçümlerin Eğri Altında Kalan Alan Yardımıyla Analizi. 20-22 Eylül 8. Ulusal Biyoistatistik Kongresi sözlü bildiri sunumu, 2005.

Dođanay B., Köse S. K., Atakurt Y.: Tekrarlı Ölçümlerin Analizinde Karma Etki Modelleri. 5-9 Eylül 9. Ulusal Biyoistatistik Kongresi sözlü bildiri sunumu, 2006.

Dođanay B., Özbek U.: Roc Eğrilerinin Karşılaştırılmasında Doğrusal Regresyon Yaklaşımı. 5-9 Eylül 9. Ulusal Biyoistatistik Kongresi sözlü bildiri sunumu, 2006.

İdilman R., Mızrak D., Soykan I., Çorapçiođlu D., Erden E., Çoban Ş., Bektaş M., Uysal A.R., Emral R., Dođanay B., Özden A.: Insulin-sensitizer agents in the treatment of individuals with non-alcoholic steatohepatitis: A preliminary report. HEPATOLOGY 44 (4): 651A-651A 1240 Suppl. 1 OCT ,2006.

Hançer F., Kunak C.S., Ada O.A., Süzen H.S., Doganay B., Alpar S., Gülhan M., Kurt B., İşcan M.: The Role of GSTP1 ALA114VAL Polyforphism in Drug Resistance in Lung Cancer. Abstract Book of 6th International Congress of Turkish Society of Toxicology, 2-5 November, Antalya, Turkey P110, 2006.

Eđitimci ve Katılımcı Olarak Yer Alınan Kurs ve Çalıştaylar

Klinik Araştırma Metodolojileri Kursu (Uygulamalı İstatistik), National Institute of Health (NIH) "Çocuk Gelişiminin Desteklenmesi: Yale-Ankara İşbirliği" Projesi Gelişimsel Pediatri Ünitesi Eğitim Programları-II, ANKARA, 22 Ekim 2005.