

**KAYIP VERİNİN TEST EŐİTLEMESİNE ETKİSİNİN
İNCELENMESİ**

**INVESTIGATING THE EFFECT OF MISSING DATA ON
TEST EQUATING**

Duygu Gizem ERTOPRAK

Hacettepe Üniversitesi

Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalı

Doktora Tezi

olarak hazırlanmıştır.

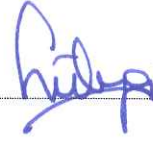
2017

KABUL ve ONAY

Eđitim Bilimleri Enstitüsü M¼d¼rl¼ę¼'ne,

Duygu Gizem ERTOPRAK'ın hazırladıđı “Kayıp Verinin Test Eđitlemeye Etkisinin İncelenmesi” bařlıklı bu alıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölme ve Deđerlendirme Bilim Dalı'nda Doktora Tezi** olarak kabul edilmiřtir.

Bařkan Prof. Dr. H¼lya KELECİOđLU



¼ye (Danıřman) Prof. Dr. Selahattin GELBAL



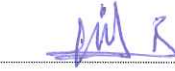
¼ye Do. Dr. Neře G¼LER



¼ye Do. Dr. İsmail KARAKAYA



¼ye Do. Dr. Burcu ATAR



ONAY

Bu tez Hacettepe Üniversitesi Lisans¼st¼ Eđitim-Öđretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından 24 / 03 / 2017 tarihinde uygun gör¼lm¼ř ve Enstit¼ Yönetim Kurulunca / / tarihinde kabul edilmiřtir.

Prof. Dr. Ali Ekber řAHİN
Eđitim Bilimleri Enstitüsü M¼d¼r¼

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Tezimin/Raporumun tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.

(Bu seçenikle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etmeniz ve kütüphane bu talebinizi yerine getirirse bile, teziniz arama motorlarının önbelleklerinde kalmaya devam edebilecektir)

Tezimin/Raporumun 11 / 04 / 2020 tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.

(Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisi alınabilir).

Tezimin/Raporumun tarihine kadar erişime açılmasını istemiyorum ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.

Serbest Seçenek/Yazarın Seçimi:

11 / 04 / 2017

Duygu Gizem ERTOPRAK

ETİK BEYANNAMESİ

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

Duygu Gizem ERTOPRAK

TEŞEKKÜR

Eğer bu teşekkür metnini okuyorsanız sizi temin ederim ki bu noktaya ulaşabilmem aşağıda sıraladığım çok kıymetli insanlar ve kurumlar sayesinde olmuştur...

Benim için "Nasıl bir akademisyen olmak istersiniz?" sorusuna verilecek en güzel yanıt olan çok değerli hocam tez danışmanım Prof. Dr. Selahattin GELBAL'a,

Yüksek lisans öğrenimine başladığım ilk günden bu yana yoluma ışık tutan çok değerli hocalarım Prof. Dr. Hülya KELECİOĞLU ve Prof. Dr. Nuri DOĞAN'a,

Değerli görüş ve önerileri ile tezin bugünkü halini almasını sağlayan Doç. Dr. Neşe GÜLER, Doç. Dr. Burcu ATAR ve Doç. Dr. İsmail KARAKAYA'ya,

Bambaşka bir ülke ve farklı bir akademik ortamı deneyimlememi sağlayan Prof. Dr. Larry H. LUDLOW'a,

En takıldığım anlarda değerli vakitlerini sorularımı yanıtlamaya ayıran Yrd. Doç. Dr. Sevilay KİLMEN, Arş. Gör. Onur TOKA ve Yrd. Doç. Dr. Fatih ORÇAN'a,

Yüzlerini görmekten, seslerini duymaktan her daim güç aldığım çok sevgili arkadaşlarım Didem ÖZDOĞAN, Çiğdem AKIN ARIKAN, Neşe KUTLU ABU, Serpil PEKDOĞAN ve Kayhan BOZGÜN'e,

Bugün bu teze büyük bir gururla adımları yazabilmemi sağlayan canımdan öte ailem babam Zeynel SARAL, annem Serpil SARAL ve biricik kardeşim Damla Hazal SARAL'a,

Zoru kolay, geceyi gündüz, imkansızı mümkün kılan ve varlığı ile hayatımı anlamlandıran yol arkadaşım sevgili eşim İzzet Eray ERTOPRAK'a,

Öğretim Görevlisi olarak yürüttüğüm bu tezi bitirebilmem için beni destekleyen Amasya Üniversitesi'ne,

Doktora öğrenimim boyunca yurt içi ve yurt dışı doktora burs olanağı sağlayan TÜBİTAK'a,

sonsuz teşekkürlerimi sunarım.

KAYIP VERİNİN TEST EŞİTLEMeye ETKİSİNİN İNCELENMESİ

Duygu Gizem ERTOPRAK

ÖZ

Bu arařtırmada, kayıp verilerin test formları ierisindeki yeri, mekanizması, oranı ve bařa ıkma yöntemlerine göre oluřturulan farklı kořullardaki testlerden elde edilen puanların, ortak test deseni kapsamında ve MTK'ya dayalı Stocking-Lord yöntemi kullanılarak eřitlenmesi sonucunda kestirilen madde ve yetenek parametrelerine iliřkin eřitleme hatası (RMSE) ve eřitleme yanlılıđı (BIAS) deđerlerini karřılařtırmak amalanmıřtır.

Bu ama dođrultusunda, temelde kayıp veri probleminin test eřitleme üzerindeki etkilerini incelemek iin üç ařamadan oluřan bir veri türetme süreci takip edilmiřtir. İlk ařamada, eřitlenmesi planlanan her iki test formu iin de 3PL modele uygun ikili puanlanan maddelerden oluřan tam veri setleri türetilmiřtir. İkinci ařamada, test formları iin birinci ařamada türetilen tam veri setleri üzerinde kullanılan veri silme algoritmaları ile üç farklı kayıp veri yeri (eřitlenecek test-ET, her iki test-HİT, ortak test-OT), üç farklı kayıp veri mekanizması (tamamen rastgele kayıp-TRK, rastgele kayıp-RK, rastgele olmayan kayıp-ROK) ve üç farklı kayıp veri oranına (%10, %20, %40) sahip kayıp verili veri setleri yaratılmıřtır. Üüncü ve son ařamada ise, kayıp verili maddelerin bulunduđu veri setleri üzerinde dört farklı bařa ıkma yönteminin (uygulanmamıř gibi davranma-UGD, yanlıř yanıtlanmış gibi davranma-YYGD, lojistik regresyona dayalı oklu deđer atama-LRDA, diskriminant fonksiyonuna dayalı oklu deđer atama-DFDA) kullanılması ile kayıp veri sorunu özölmüş veri setlerine ulařılmıřtır.

Verilerin türetilmesinin ardından eřitleme sürecine geilmiřtir. Bu kapsamda, tam veri setlerine sahip test formlarının eřitlendiđi referans kořul ile tam veri setlerinin manipöle edilmesiyle elde edilen test formlarının eřitlendiđi 108 (3x3x3x4) farklı simölasyon kořulu da dahil olmak üzere toplam 109 eřitleme iřlemi yürütölmüřtür. R paket programı üzerinden gerekleřtirilen tüm analiz iřlemleri iin 50'řer tekrar yapılmıřtır. Analiz sonuçlarının deđerlendirilmesi ařamasında ise öncelikle her bir simölasyon kořulu kapsamında yürütölen eřitlemeler ile elde edilen ortalama hata (RMSE) ve ortalama yanlılık (BIAS) deđerleri raporlanmış, ardından bu deđerlerin

referans koşul kapsamında yürütülen eşitlemeler ile elde edilen değerlere olan yakınlığına bakılmıştır.

Araştırmadan elde edilen sonuçlar, kayıp veri yeri faktörüne ait koşullar için en az hatalı ve en az yanlı bulguların HİT, en çok hatalı ve en çok yanlı bulguların ET; kayıp veri mekanizması faktörüne ait koşullar için en az hatalı ve en az yanlı bulguların TRK, en çok hatalı ve en çok yanlı bulguların ROK; kayıp veri oranı faktörüne ait koşullar için en az hatalı ve en az yanlı bulguların %10, en çok hatalı ve en çok yanlı bulguların ise %40 koşulları için elde edildiğini göstermiştir. Kayıp veri başa çıkma yöntemlerine ilişkin olarak elde edilen sonuçlar ise referans koşul eşitlemeleri sonunda kestirilen özellikle yetenek parametrelerine ait ortalama RMSE ve BIAS değerlerine en yakın ve en düşük hata değerlerini üreten yöntemin DFÇDA olduğunu ve ikili puanlanan maddelerden oluşan veri setlerindeki kayıp veri probleminin çözülmesi için en sık başvurulan UGD ve YYGD yöntemlerinin, test eşitleme üzerinde hatalı ve yanlı kestirimler ürettiğini kaydetmiştir.

Anahtar sözcükler: Kayıp veri, test eşitleme, eşitleme hatası, eşitleme yanlılığı, kayıp veri başa çıkma yöntemleri

Danışman: Prof. Dr. Selahattin GELBAL, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

INVESTIGATING THE EFFECT OF MISSING DATA ON TEST EQUATING

Duygu Gizem ERTOPRAK

ABSTRACT

The purpose of this study is to compare equating error (RMSE) and equating bias (BIAS) values with regard to estimated item and ability parameters obtained as a result of equating scores of tests in different conditions which created according to missing data location, mechanism and handling methods using IRT-based Stocking-Lord method within anchor test design.

In accordance with this purpose, a data generation process which consisted three stage were followed in order to examine the effects of missing data problem on test equating. In the first stage, complete data sets which composed binary items fit to 3PL model were generated to create each test form. In the second stage, missing data sets which have three different missing data location (new test-NT, both test-BT, anchor test-AT), three different missing data mechanism (missing completely at random-MCAR, missing at random-MAR, missing not at random-MNAR) and three different missing data rate (10%, 20%, 40%) were created with using data deletion algorithms on complete data sets which generated in the first stage. In the third and last stage, data sets which handled missing data problem were reached with using four handling methods (treating as not administered-TNA, treating as incorrect-TI, logistic regression-based multiple imputation-LRMI, discriminant function-based multiple imputation-DFMI) on missing data sets.

After data generation, equating process were conducted. In this context, equatings were carried out separately on reference condition whose test forms that have complete data sets and 108 (3x3x3x4) different simulation conditions whose test forms that were obtained by manipulating of complete data sets. To sum up, total 109 equating process were performed. 50 replication were done one by one to entire analysis process which conducted through R package. At the stage of evaluating of analysis results, average equating error (RMSE) and average equating bias (BIAS) values which obtained by equatings on every simulation condition were reported. And then these values and the values which obtained by equatings on reference condition were compared and looked how close they were.

The results obtained from study showed that BT had the least equating error and the least equating bias values while NT had the most equating error and the most equating bias values in terms of missing data location factor, MCAR had the least equating error and the least equating bias values while MNAR had the most equating error and the most equating bias values in terms of missing data mechanism factor, 10% had the least equating error and the least equating bias values while 40% had the most equating error and the most equating bias values in terms of missing data rate factor. The results about missing data handling methods showed that DFMI had produced the closest and the lowest average RMSE and BIAS values to the average values of reference condition. Results also showed that TNA and TI methods which frequently used to handle missing data problem in binary items produced inaccurate and biased estimations on test equating.

Keywords: missing data, test equating, equating error, equating bias, missing data handling methods

Advisor: Prof. Dr. Selahattin GELBAL, Hacettepe University, Department of Educational Sciences, Division of Measurement and Evaluation in Education

İÇİNDEKİLER

KABUL ve ONAY.....	ii
ETİK BEYANNAMESİ	iv
TEŞEKKÜR.....	v
ÖZ	vi
ABSTRACT	viii
İÇİNDEKİLER.....	x
TABLolar DİZİNİ	xiii
ŞEKİLLER DİZİNİ.....	xiv
SİMGELER VE KISALTMALAR DİZİNİ	xvi
1. GİRİŞ	1
1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı ve Önemi:.....	6
1.3. Problem Cümlesi:	7
1.3.1. Alt Problemler:.....	7
1.4. Sınırlılıklar:.....	8
1.5. Araştırmanın Kuramsal Temeli	9
1.5.1. Madde Tepki Kuramı ile ilgili Kuramsal Çerçeve	9
1.5.1.1. Madde Karakteristik Eğrisi (MKE)	10
1.5.1.2. İkili Puanlanan Madde Yanıtları Üzerinde Kullanılan Madde Tepki Modelleri	12
1.5.2. Test Eşitleme ile İlgili Kuramsal Çerçeve	13
1.5.2.1. Bağlama (Linking).....	13
1.5.2.2. Test Eşitleme	15
1.5.2.3. Eşitleme Desenleri.....	16
1.5.2.4. MTK'ya Dayalı Eşitleme Yöntemleri.....	18
1.5.2.4.1. Ayrı kalibrasyon.....	18
1.5.2.4.2. Eşzamanlı kalibrasyon	20
1.5.2.5. Eşitleme Hatası.....	21
1.5.3. Kayıp Veri ile ilgili Kuramsal Çerçeve	22
1.5.3.1. Veri Matrisi.....	22
1.5.3.2. Tam Olmayan Veri Matrisi ve Kayıp Veri Gösterimi.....	23
1.5.3.3. Kayıp Veri Problemi	24
1.5.3.4. Kayıp Veri Teorisi	27
1.5.3.4.1. Kayıp veri mekanizması	27
1.5.3.4.2. Kayıp veri mekanizmalarına yönelik ihmal edilebilirlik varsayımı	30
1.5.3.4.3. Kayıp veri başa çıkma yöntemleri	33
1.5.4. Test Eşitleme ve Kayıp Veri	42
2. İLGİLİ ARAŞTIRMALAR.....	47
2.1. Kayıp Veri Probleminin Test Eşitlemeye Etkisi ile İlgili Çalışmalar.....	47
2.2. Ortak Test Deseni Kapsamında Kullanılan Farklı Eşitleme Yöntemlerinin Kayıp Veri Varsayımları ile İlgili Araştırmalar	47

2.3. Kayıp Veri Probleminin Geçerlik ve Güvenirliğe Etkisi ile İlgili Çalışmalar	51
2.4. Kayıp Veri Başa Yöntemlerinin Performanslarının Karşılaştırılması ile İlgili Çalışmalar	54
2.5. İlgili Araştırmalar Özet	64
3. YÖNTEM	67
3.1. Araştırmanın Türü	67
3.2. Eşitleme Deseni	67
3.3. Simülasyon Faktörleri ve Koşulları	68
3.4. Verilerin Türetilmesi	70
3.4.1. Tam Veri Setlerinin Türetilmesi	71
3.4.2. Veri Silme Algoritmaları aracılığıyla Kayıp Veriye Sahip Maddeler İçeren Veri Setlerinin Elde Edilmesi	72
3.4.3. Kayıp Veri Başa Çıkma Yöntemlerinin Kullanılması ile Yeni Veri Setlerinin Elde Edilmesi	77
3.5. Eşitleme Süreci	78
3.6. Verilerin Analizi ve Değerlendirme Ölçütleri	79
4. BULGULAR	83
5. TARTIŞMA	104
6. SONUÇ ve ÖNERİLER	108
6.1. Sonuçlar	108
6.2. Öneriler	111
6.2.1. Araştırmaya Dönük Öneriler	111
6.2.2. Yeni Araştırmalara Dönük Öneriler	111
KAYNAKÇA	113
EKLER DİZİNİ	120
EK 1. ETİK KOMİSYONU ONAY BİLDİRİMİ	121
EK 2. ORJİNALLİK RAPORU	122
EK 3. TAM VERİ SETLERİNİN TÜRETİLMESİNDE KULLANILAN KODLAR	124
EK 4. TRK MEKANİZMASINA SAHİP VERİ SETLERİNİN ELDE EDİLMESİNDE KULLANILAN KODLAR	126
EK 5. RK MEKANİZMASINA SAHİP VERİ SETLERİNİN ELDE EDİLMESİNDE KULLANILAN KODLAR	127
EK 6. ROK MEKANİZMASINA SAHİP VERİ SETLERİNİN ELDE EDİLMESİNDE KULLANILAN KODLAR	129
EK 7. TRK MEKANİZMASI KAPSAMINDA ELE ALINAN SİMÜLASYON KOŞULLARINA GÖRE KESTİRİLEN MADDE PARAMETRELERİNE İLİŞKİN RMSE VE BIAS DEĞERLERİ	131
EK 8. RK MEKANİZMASI KAPSAMINDA ELE ALINAN SİMÜLASYON KOŞULLARINA GÖRE KESTİRİLEN MADDE PARAMETRELERİNE İLİŞKİN RMSE VE BIAS DEĞERLERİ	132

EK 9. ROK MEKANİZMASI KAPSAMINDA ELE ALINAN SİMÜLASYON KOŞULLARINA GÖRE KESTİRİLEN MADDE PARAMETRELERİNE İLİŞKİN RMSE VE BIAS DEĞERLERİ	133
EK 10. ARAŞTIRMA KAPSAMINDA ELE ALINAN SİMÜLASYON KOŞULLARINA GÖRE KESTİRİLEN YETENEK PARAMETRELERİNE İLİŞKİN RMSE VE BIAS DEĞERLERİ.....	134
ÖZGEÇMİŞ	135

TABLULAR DİZİNİ

Tablo 1.1: Kayıp Verinin Etkisi.....	4
Tablo 1.2: Tek Grup Deseni	16
Tablo 1.3: Rastgele Grup Deseni	17
Tablo 1.4: Ortak Test Deseni.....	17
Tablo 1.5: Farklı Kayıp Veri Mekanizmalarına Göre Oluşturulmuş İş Performans Ölçümleri	28
Tablo 3.1: Araştırmanın Eşitleme Deseni	68
Tablo 3.2: Araştırmanın Simülasyon Faktörleri ve Koşulları	68
Tablo 3.3: Araştırma Kapsamında Ele Alınan Simülasyon Koşulları	70
Tablo 3.4: RK Mekanizmasına Sahip Kayıp Verilerin Elde Edilmesinde Kullanılan Kayıp Veri Oranları.....	75
Tablo 3.5: ROK Mekanizmasına Sahip Kayıp Verilerin Elde Edilmesinde Kullanılan Kayıp Veri Oranları.....	77
Tablo 4.1: Madde Parametrelerine ait RMSE ve BIAS Değerleri için Anlamlı Bulunan ANOVA Sonuçları	93
Tablo 4.2: Tam Veriye Sahip Test Formlarının Eşitlenmesi ile Kestirilen Madde Parametrelerine ilişkin RMSE ve BIAS Değerleri	94
Tablo 4.3: Yetenek Parametresine ait RMSE ve BIAS Değerleri için Anlamlı Bulunan ANOVA Sonuçları	101
Tablo 4.4: Tam Veriye Sahip Test Formlarının Eşitlenmesi ile Kestirilen Yetenek Parametrelerine ilişkin RMSE ve BIAS Değerleri	102

ŞEKİLLER DİZİNİ

Şekil 1.1. Tipik Bir Madde Karakteristik Eğrisi	10
Şekil 1.2. Bağlama Yöntemlerinin Sınıflandırılması	14
Şekil 1.3. Veri Matrisi Gösterimi	23
Şekil 1.4. Tam Olmayan Veri Matrisi Gösterimi	24
Şekil 1.5. Kayıp Veri Gösterge Matrisi	24
Şekil 1.6. ÇDA İşlem Şeması	38
Şekil 4.1. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Hataları	85
Şekil 4.2. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Hataları	85
Şekil 4.3. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Yanlılıkları	86
Şekil 4.4. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Yanlılıkları	86
Şekil 4.5. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Hataları	88
Şekil 4.6. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Hataları	88
Şekil 4.7. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Yanlılıkları	89
Şekil 4.8. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Yanlılıkları	89
Şekil 4.9. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Hataları	91
Şekil 4.10. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Hataları	91
Şekil 4.11. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Yanlılıkları	92
Şekil 4.12. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Yanlılıkları	92
Şekil 4.13. Araştırma Koşulları Kapsamında Gerçekleştirilen Eşitlemeler Sonucunda Kestirilen a Parametrelerine ait Ortalama Eşitleme Hataları	94
Şekil 4.14. Araştırma Koşulları Kapsamında Gerçekleştirilen Eşitlemeler Sonucunda Kestirilen b Parametrelerine ait Ortalama Eşitleme Hataları	95
Şekil 4.15. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Hataları	98

Şekil 4.16. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Yanlılıkları	98
Şekil 4.17. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Hataları	99
Şekil 4.18. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Yanlılıkları	99
Şekil 4.19. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Hataları	100
Şekil 4.20. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Yanlılıkları	100
Şekil 4.21. Araştırma Koşulları Kapsamında Gerçekleştirilen Eşitlemeler Sonucunda Kestirilen Yetenek Parametrelerine ait Ortalama Eşitleme Hataları.....	103

SİMGELER VE KISALTMALAR DİZİNİ

MTK (IRT): Madde Tepki Kuramı

KTK (CTT): Klasik Test Kuramı

MKE (ICC): Madde Karakteristik Eğrisi

UGD: Uygulanmamış gibi Davranma

YYGD: Yanlış Yanıtlanmış gibi Davranma

ÇDA: Çoklu Değer Atama

LRÇDA: Lojistik Regresyona Dayalı Çoklu Değer Atama

DFÇDA: Diskriminant Fonksiyonuna Dayalı Çoklu Değer Atama

SL: Stocking-Lord

3PLM: Üç Parametrelili Lojistik Model

RMSE: Hata Kareleri Ortalamasının Karekökü

BIAS: Eşitleme Yanlılığı

1. GİRİŞ

Bu bölümde problem durumu, araştırmının amacı ve önemi, problem cümlesi, alt problemler, sayılılar, sınırlılıklar ve araştırmının kuramsal temeli başlıklarına yer verilmiştir.

1.1. Problem Durumu

Bireylere ait fiziksel özellikler, standart araçlar kullanılarak ölçülmek istendiğinde, aynı araçların herhangi bir nitelik kaybına uğramadan defalarca kullanılabilirdiği görülmektedir. Örneğin bir bireyin boy uzunluğunu ölçmek amacıyla kullanılan metre, herhangi bir güvenilirlik kaybına uğramadan başka bireylerin boy uzunluklarını ölçmek için de kullanılabilir. Üstelik, aynı ölçme aracı, kullanılmaya devam edildiği sürece, aynı özelliği aynı güvenilirlikte ölçmeye devam edecektir. Eğer ilgili ölçme aracında herhangi bir işlev kaybı meydana gelirse (örneğin bir parçası silinirse veya kırılırsa), aynı araç ile eşit güvenilirlikte ölçüm yapan bir yenisi daha yapılabilir. Bunun aksine, bireylerin başarı ve yetenek gibi gizil psikolojik özelliklerini ölçmeye yönelik hazırlanmış standart araçlar defalarca kullanılmak istendiğinde, soruların gizlilik ilkesinin kaybolması ve testin güvenliğinin düşmesi gibi ciddi sıkıntılarla karşılaşılabilir. Bu sıkıntıların önüne geçilebilmek için ölçülmek istenen özellik ile benzer özellikleri aynı güvenilirlikte ölçmeyi amaçlayan alternatif test formları kullanılabilir (Bränberg, 2010).

Özellikle bireylere ait başarı ve yetenek düzeylerinin belirlenmesinde işe koşulan ulusal ve uluslararası geniş ölçekli eğitim testi uygulamalarında, alternatif form kullanımlarına sıklıkla başvurulduğu görülmektedir. Ülkemizde yürütülen Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı (ALES) ve Yabancı Dil Bilgisi Seviye Tespit Sınavı (YDS) ile uluslararası arenalarda yürütülen Lisansüstü Eğitime Giriş Sınavları (GRE ve GMAT), Uluslararası İngilizce Dil Yeterlik Sınavları (TOEFL ve IELTS) gibi her biri farklı amaca hizmet eden test uygulamaları bunun en güzel örneklerini oluşturmaktadır. Ancak alternatif test formu uygulamaları hakkında sorgulanması gereken bazı önemli hususlar bulunmaktadır. Örneğin biri ilkbahar diğeri sonbahar dönemi ALES formunu alan A ve B bireylerinin, aynı ilan numaralı akademik kadroya başvuruda bulduklarını düşünelim. Böyle bir durumda, farklı soruları yanıtlayan A ve B bireylerinin ALES puanları karşılaştırılabilir midir?

Bu soruya verilecek yanıtın "evet" olabilmesi için birbirlerine alternatif olarak gösterilen formların aynı kapsama ve mümkün olduğunca aynı güçlük düzeyine sahip olacak biçimde hazırlanmış olması gerekir. Ancak her bir uygulamada farklı sorular sorulduğu sürece, formların güçlük düzeylerinde bir miktar farklılıkların ortaya çıkması da oldukça doğal olacaktır (Kolen & Brennan, 2004). Uygulama açısından beklenen bu durum, diğerine göre daha kolay olan ALES formunu alan bireyi avantajlı, daha zor olan ALES formunu alan bireyi ise dezavantajlı konuma getirir. Böyle durumlarda bireyler arasında daha adil karşılaştırmalar yapılabilmesi için formlar arasındaki güçlük farklılıklarını mümkün olduğunca ortadan kaldırmayı amaçlayan "test eşitleme" süreçlerine başvurulması gerekir (Cook & Eignor, 1991).

İstatistiksel anlamda test eşitleme, birden fazla testin puan dağılımları arasındaki fonksiyonel ilişkiyi tanımlar. Benzer istatistiksel özelliklere sahip test formları arasında kurulan bu fonksiyonel eşitleme ilişkisi, bir formdan alınan puanların diğer form üzerindeki bilinmeyen değerinin kestirilmesi amacına hizmet eder (Holland & Dorans, 2006). Eğer ALES'in ilkbahar ve sonbahar dönemi formlarından alınan puanlar arasındaki eşitleme ilişkisi doğrulukla kurulursa, birbirlerinin sorularını görmeyen A ve B bireyleri, sanki aynı dönem uygulanan test formunu almış gibi olurlar. Böylelikle bu iki birey, aynı akademik kadro ilanına başvuruda bulunmak için farklı formlardan almış oldukları puanları kullanabilir hale gelirler.

Bir testin alternatif formlarından alınan puanların eşitlenmesi sırasında, formlar arasında kurulacak eşitleme ilişkisini etkileyecek bazı faktörler bulunmaktadır. Bu faktörlerden biri, eşitlemesi yapılacak test formlarında bulunan maddelerin bazılarında yanıt alınamadığı durumlarda ortaya çıkan kayıp verilerdir (Little & Rubin, 1987; Allison, 2002). Bireylerin bir maddeyi yanıtsız bırakma davranışına sebep olabilecek durumlar, De Ayala, Plake ve Impara (2001) tarafından üç temel başlık altında özetlenmektedir:

1. Testin uygulama desenine bağlı olarak ortaya çıkan cevap vermeme davranışı; farklı kitapçıkların kullanıldığı test uygulamalarında, bireylerin almadıkları kitapçıklara ait sorular için yanıt üretmedikleri durumlarda ortaya çıkabilir.

2. Erişilemeyen maddelere bağlı olarak ortaya çıkan cevap vermeme davranışı; hız faktörünün etkili olduğu test uygulamalarında, bireylerin bütün sorulara yanıt vermek için yeterli zaman bulamamaları ile testin sonlarında yer alan sorular için ortaya çıkabilir.
3. Kasıtlı olarak ortaya çıkan cevap vermeme davranışı; yeterli uygulama süresine sahip olan test uygulamalarında, bireylerin bazı maddeleri kasıtlı olarak boş bırakmalarına bağlı olarak ortaya çıkabilir.

Bireylerin test maddelerine verdikleri gözlenebilir yanıtlardan gözlenemeyen (gizil) özellikleri hakkında çıkarımlarda bulunmayı amaçlayan test uygulamaları sırasında cevap vermeme davranışına bağlı olarak ortaya çıkan tüm bu kayıplar, bireyler hakkında doğru kestirimlere ulaşılması konusunda ciddi bir engel oluşturur (Hohensinn & Kubinger, 2011). Kestirimlerin gücünün azalması, standart analiz yöntemlerinin tam veriler üzerinden yapılandırılmış olması ve tam olmayan veri setleri üzerinde bu yöntemlerin kullanımının mümkün olmamasından kaynaklanır (Allison, 2002). Farklı özelliklerdeki veri setlerinde ortaya çıkan farklı koşullardaki kayıp verilerin; parametre kestirimlerinde yanlılığın artması, standart hataların büyümesi ve işe koşulan istatistiksel testin gücünün düşmesi gibi olumsuz etkilerini gösteren çalışmalar, bu açıklamaların birer kanıtı niteliğindedir (De Ayala, Plake & Impara, 2001; de Leeuw, Hox & Huisman, 2003; Acock, 2005; Ambler, Omar & Royston, 2007; Hedeker, Mermelstein & Demirtaş, 2007; Van Buuren, 2007; Finch, 2008; Rose, Davier & Xu, 2010; Hohensinn & Kubinger, 2011; Doğanay-Erdoğan, 2012; Toka, 2012; Demir, 2013). Ayrıca kayıp verilerin varlığının, ölçme işlemlerinin geçerlik ve güvenilirliklerini olumsuz yönde etkilediği de birçok çalışma ile ortaya konmuştur (Brown, 1994; Marsh, 1998; Enders & Bandalos, 2001; Bal, 2003; Enders, 2003, 2004; Van Ginkel, 2007; Bernaards & Sijtsma, 1999, 2000; Chen, Wang & Chen, 2011; Çokluk & Kayrı, 2011; Demir & Parlak, 2012; Demir, 2013; Akbaş & Tavşancıl, 2015; Şahin-Kürşad & Nartgün, 2015).

Kayıp verilerin geçerlik ve güvenilirlik üzerindeki etkilerinin incelendiği yukarıdaki çalışmalara ek olarak; kayıp verilerin ölçme işlemlerinin güvenilirlik ve geçerliği, araştırma sonuçlarının iç ve dış geçerliği ile sonuçların genellenebilirliği üzerindeki etkileri Tablo 1.1'deki gibi özetlenmiştir (McKnight, Mcknight, Sidani & Figueredo, 2007):

Tablo 1.1: Kayıp Verinin Etkisi

<i>Etkileme tipi</i>	<i>Ortaya çıkma biçimi</i>	<i>Etkileme biçimi</i>
Ölçme (güvenirlik, yapı geçerliği)	Ölçme (maddeler, tekli ölçümler, kayıp veriye sahip yapının çoklu ölçümleri)	1. ↓ madde havuzu → ↑ hata varyansı → ↓ ölçmenin güvenilirliği 2. ↓ bilgi → kapsamın eksikliği → ↓ ölçmenin geçerliği
	Örneklem seçimi (yanıtlayıcılar ile yanıtlayıcı olmayanlar, testi tamamlayanlar ile tamamlamayanlar, yanıtları tam olan katılımcılar ile yanıtları tam olmayan katılımcılar arasındaki farklılıklar)	Bu grupların özelliklerindeki farklılıklar → seçim yanlılığı → temsil edici olmayan örneklem → ↓ dış geçerlik
Araştırma sonuçlarının güvenilirlik ve geçerliği (iç geçerlik)	Rastgelelik 1. Katılmama 2. Verinin farklı biçimlerde kaybı	1. ↓ örneklem büyüklüğü → başlangıç eşdeğersizliği → ↓ iç geçerlik 2. başlangıç eşdeğersizliği → ↓ iç geçerlik
	Farklı şekillerde ortaya çıkan yanıtlayıcı kaybı	1. başlangıç eşdeğersizliği → ↓ iç geçerlik 2. eşit olmayan grup büyüklüğü → istatistiksel varsayımların ihlali → ↓ istatistiksel sonuçların geçerliği
	Veri analizleri (örneklem büyüklükleri)	↓ istatistiksel güç ve dağılımsal varsayımların ihlali → ↓ istatistiksel sonuçların geçerliği
Sonuçların genellenebilirliği	Yukarıdaki kayıp veri tiplerinden herhangi biri ya da hepsi	Yukarıdaki sorunların herhangi biri ya da hepsi → istatistiksel çıkarımlarda ve bulguların yorumlanmasında zorluklar → doğru olmayan bilgi dayanakları → yanlış bilgilendirilmiş ve yanlış yönlendirilmiş öneriler

Kaynak: McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. New York: Guilford Press.

Not. Sağa doğru ok işaretleri (→) "sebep olur, yol açar, sonucunu doğurur" olarak okunmalıdır; aşağı ve yukarı doğru olan ok işaretleri (↓↑) ise sonuçtaki azalma veya artışı ifade ederler.

Tablo 1.1, kayıp veri probleminin tipik bir ölçme işleminde bulunması istenen ilkeleri büyük oranda zedeleyebileceğini göstermektedir. Kayıp veri problemi ile ölçmenin ilkeleri arasında teorik olarak belirtilen bu olumsuz ilişki, bazı gerçek test uygulamaları üzerinden de örneklendirilebilir. Örneğin testi tamamlamak için yeterli sürenin verilmediği veya bilimsel olarak yanlışlığı kabul edilen soruların yöneltildiği test uygulamalarında karşılaşılan kayıp veri problemi, genellikle, yanıt alınmayan soruların uygulamaya katılan her birey tarafından doğru yanıtlandırılmış olarak kabul edilmesi ile çözümlenir. Ancak bu çözüm yönteminin işe koşulmasıyla ölçme işlemine karışan sabit hatalar, geçerliğin düşmesine neden olurlar. İlgili soruların istatistiksel analizlerden çıkarıldığı bir diğer çözüm yönteminde ise duyarlılık anlamında güvenilirlik ve kapsam geçerliğinin düşme ihtimali ortaya çıkar. Ayrıca bu

çözümün kullanıldığı durumlarda, test tipik bir psikolojik yapıyı ölçmeye yönelik olarak hazırlanmış ve ilgili yapının tüm kritik boyutlarını ölçen maddelerden oluşuyor ise sorulardan birinin dahi analizden çıkarılması ile yapı geçerliği de zedelenmiş olur. Bu durumun bazı bireylerin başarı düzeylerini belirlemede daha az soru ile kestirim yapılmasına ve dolayısıyla da bu bireyler için güvenirliliğin düşmesine sebep olması da karşılaşılabilecek muhtemel başka bir olumsuzluktur.

Başka bir örnek, ölçmecilerin test uygulama sonuçlarını bilgisayara aktardığı veya test kağıtlarını puanladığı aşamalarda dikkatsizlik yaşadıkları veya testi alan bireylerin uygulamalar sırasında yorgun, uykusuz, hasta veya öfkeli oldukları için bazı soruları kazara okumadıkları durumlarda karşılaşılan kayıp veriler için verilebilir. Burada ölçme işlemine karışan tesadüfi hatalar güvenirliliği düşürücü bir etki yapar. Bunun yanı sıra, ölçme araçlarının hazırlanmasının, çoğaltılmasının, uygulanmasının ve puanlanmasının kolay olması anlamını taşıyan kullanışlılık ilkesi de, kayıp verilerin puanlama sırasında yarattıkları zorluklar sebebiyle zarar görebilir. Zhang ve Walker (2008) ölçme işlemlerine ait olan tüm bu ilkelerin zedelenmesinin; test eşitleme, madde yanlılığı ve sınıflandırma gibi her biri farklı amaca hizmet eden istatistiksel işlemlerin etkililiğini azalttığını belirtmiştir.

Ölçme işlemlerinin ilkelerini iyileştirip test eşitleme gibi birçok istatistiksel sürecin etkililiğini arttırmak adına Türkiye'de yapılan büyük ölçekli test uygulamalarında birçok önlem alınmaktadır. Bunlardan bir tanesi de şans başarısından arındırılmış ham puanların hesaplanmasında düzeltme formülünü işe koşturmasıdır. Ancak yanlış yanıtın doğru yanıtı götürdüğü düşüncesi, uygulamalar sırasında birçok bireyin cevabından emin olmadığı soruları yanıtı bırakmasına sebep olmaktadır. Bu noktada da önemli olan yanıtı bırakmanın hangi miktarlarda ve hangi mekanizmalarda meydana geldiği hususudur. Zira kayıp verilerin veri seti içerisindeki oranı veya rastgele dağılıp-dağılmadıkları, kayıp veriler ile başa çıkmada kullanılacak çözüm yöntemlerinin farklılaşmasına sebep olmaktadır. Bu gerekçelerle, herhangi bir eşitleme çalışmasına başlanmadan önce, farklı test formlarına ait veri setlerinin farklı yerlerinde bulunan farklı oran ve mekanizmalardaki kayıp veri probleminin uygun başa çıkma yöntemleri ile çözülmesi, ardından bu testlerden elde edilecek puanların eşitlenmesi aşamasına geçilmesi gerekmektedir.

1.2. Araştırmanın Amacı ve Önemi:

Eğitim sistemi ve öğelerine yönelik alınan kararların çoğu, bireylere uygulanan ve farklı amaçlara hizmet eden geniş ölçekli testlerden elde edilen bulgulara dayalı olarak verilmektedir. Bu testlerden elde edilen sonuçların her geçen gün sayıları artan daha fazla kurum tarafından kabul görmesi neticesinde, daha çok sayıda bireye ulaşabilmek için daha fazla sayıda test formu geliştirilmektedir. Testlerin güvenliğinin ve gizliliğin korunması adına yapılan bu işlemler ile aynı amaca ve aynı kapsama sahip farklı soruları içeren alternatif test formlarının elde edilmesi amaçlanır.

Alternatif test formu uygulamalarının başarısı, testlerden alınan sonuçların geçerli olduğu süreler içerisinde uygulanan formların istatistiksel açıdan birbirlerine denk olma düzeyleri ile ilişkilidir. Test formlarından birinin diğerlerine göre daha zor olması, zor olan formu alan bireyleri, diğer formları alan bireylere göre daha başarısız bir duruma sokar. Test geliştirme süreçleri açısından da karşılaşılmak istenmeyen bu durumun önüne geçilebilmesi ve test sonuçlarına dayalı olarak alınan kararların doğruluk derecelerinin artırılabilmesi için eşitleme çalışmaları yürütülür. Test eşitleme sayesinde, farklı yanıtlayıcı gruplarına ait puanların aynı ölçek üzerinde ifade edilebilmesini sağlayan eşitleme ilişkileri kurularak, farklı formlardan alınan puanların karşılaştırılabilirliği sağlanmış olur. Ancak puanların karşılaştırılabilir olmasını sağlayan bu ilişkiler, eşitlemesi yapılan test formlarına ait veri setlerinde çeşitli sebeplerle ortaya çıkan kayıp verilerden olumsuz olarak etkilenir. Zira tam veri setlerine göre yapılandırılan standart analiz yöntemlerinin tam olmayan veri setleri üzerinde kullanılması durumunda, ne test eşitleme ne de başka istatistiksel süreçlerden doğru sonuçlar üretmesi beklenir. Bu gerekçeyle, her iki konunun da ortak bir başlık altında ele alındığı çalışmaların yürütülmesine ihtiyaç duyulmaktadır.

Alanyazında test eşitleme ve kayıp veri konusunu ayrı ayrı ele alan oldukça fazla çalışma bulunmasına rağmen, her iki konunun ortak olarak ele alındığı çalışma sayısının az olduğu görülmüştür (Liou & Cheng, 1995; Holland, Sinharay, von Davier & Han, 2008; Miyazaki, Hoshino, Mayekawa & Shigemasa, 2009; Shin, 2009; Sinharay & Holland, 2009; Puhan, 2010). Bu çalışmaların oldukça büyük bir kısmı da, testin uygulanma desenine bağlı olarak ortaya çıkan kayıp veri problemi ile ilgilenmektedir. Başka bir ifadeyle Shin (2009) dışındaki çalışmaların hiçbiri, bu

araştırmanın yürütülme amacı ile aynı amaca sahip değildir. Shin (2009) ise ortak test deseni kapsamında Rasch modeline dayalı olarak yürüttüğü gerçek puan eşitleme çalışmasında, eşitlemesini yaptığı test formlarında bulunan kayıp veriler ile başa çıkmada kullanılan yöntemlerin performanslarını karşılaştırmıştır. Bu amaçla kullandığı test formlarını, 20 ortak maddeye sahip ikili puanlanan 50 soruluk gerçek bir standart başarı testinin (English Language Arts-ELA) 2005 ve 2006 yılı uygulamalarına ait madde parametrelerinden türetmiştir. Bu şekilde elde ettiği test formları üzerinde farklı başa çıkma yöntemlerini uygulamasının ardından performans değerlendirmeleri yapmıştır. Çalışmadan elde edilen sonuçlar, büyük örneklemeler üzerinde kayıp veriye sahip olan maddeleri yanlış olarak kodlamak yerine hiç uygulamamış gibi davranmanın daha doğru eşitleme sonuçları ürettiğini göstermiştir.

Bu çalışmada, kayıp verilerin test formları içerisindeki yeri, mekanizması, oranı ve başa çıkma yöntemlerine göre oluşturulan farklı koşullardaki testlerden elde edilen puanların, ortak test deseni kapsamında ve Madde Tepki Kuramına dayalı Stocking-Lord yöntemi kullanılarak eşitlenmesi sonucunda kestirilen madde ve yetenek parametrelerine ilişkin eşitleme hata (RMSE) ve eşitleme yanlılık (BIAS) değerlerini karşılaştırmak amaçlanmıştır. Bu amaç doğrultusunda araştırmanın, Madde Tepki Kuramı çerçevesinde oluşturulan bir araştırma deseni kapsamında, kayıp veri ile test eşitleme konusunu doğrudan ilişkilendirmesi ve bunu yaparken de farklı simülasyon koşullarının performanslarını değerlendirip en hatasız ve en yansız sonuçları üreten kayıp veri başa çıkma yöntemini belirlemesi ve önermesi açısından ilgili literatüre önemli katkılar sağlayacağı düşünülmüştür.

1.3. Problem Cümlesi:

Kayıp veriye sahip maddelerin bulunduğu test formları eşitlendiğinde kestirilen parametrelere ilişkin eşitleme hatası ve eşitleme yanlılığı; kayıp verinin yerine, mekanizmasına ve oranına göre nasıl değişmektedir?

1.3.1. Alt Problemler:

1. Kayıp veriye sahip maddelerin bulunduğu test formları, kayıp veri başa çıkma yöntemleri içerisinde değer atama içermeyen yöntemler (UGD ve YYG) ve çoklu değer atamaya dayalı yöntemler (LRÇDA ve DFÇDA) ile çözümlenip eşitlendiğinde kestirilen **madde parametrelerine** ilişkin eşitleme hatası ve

eşitleme yanlılığı; kayıp verinin yerine, mekanizmasına ve oranına göre nasıl değişmektedir?

2. Tam veriye sahip test formları ile kayıp veri başa çıkma yöntemleri kullanılarak elde edilen test formları eşitlendiğinde kestirilen **madde parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı nasıl değişmektedir?

3. Kayıp veriye sahip maddelerin bulunduğu test formları, kayıp veri başa çıkma yöntemleri içerisinde değer atama içermeyen yöntemler (UGD ve YYGD) ve çoklu değer atamaya dayalı yöntemler (LRÇDA ve DFÇDA) ile çözümlenip eşitlendiğinde kestirilen **yetenek parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı; kayıp verinin yerine, mekanizmasına ve oranına göre nasıl değişmektedir?

4. Tam veriye sahip test formları ile kayıp veri başa çıkma yöntemleri kullanılarak elde edilen test formları eşitlendiğinde kestirilen **yetenek parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı nasıl değişmektedir?

1.4. Sınırlılıklar:

Araştırma;

- eşitleme deseni olarak ortak test deseni,
- simülasyon koşulları olarak kayıp verinin test formları içerisindeki yeri, kayıp verinin mekanizması, kayıp verinin oranı ve kayıp veri ile başa çıkma yöntemleri,
- verilerin türetilmesinde işe koşulan MTK modeli olarak 3 PL model ve
- değerlendirme ölçütleri olarak da RMSE ve BIAS ile sınırlıdır.

1.5. Araştırmanın Kuramsal Temeli

Bu araştırmanın kuramsal temeli, "Madde Tepki Kuramı", "Test Eşitleme" ve "Kayıp Veri" olmak üzere üç farklı konu başlığı çerçevesinde yapılandırılmıştır. Her bir konu içeriği, araştırma probleminin çözümlenmesinde kullanılan bilgilerin genel hatlarıyla verilmesi amacıyla oluşturulmuştur.

1.5.1. Madde Tepki Kuramı ile ilgili Kuramsal Çerçeve

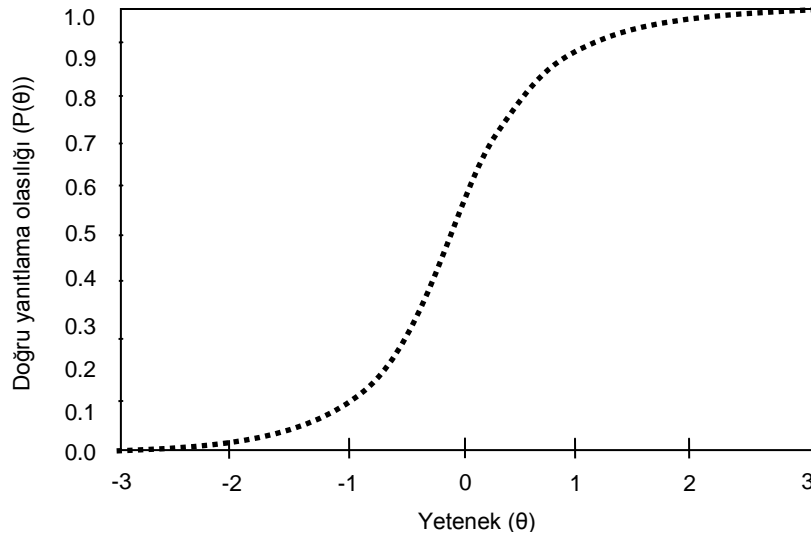
Eğitim ve psikoloji alanında gerçekleştirilen ölçme-değerlendirme faaliyetlerinin temel amacı; bireylerin başarı, yetenek, zeka, tutum gibi doğrudan gözlenemeyen örtük özelliklerine ilişkin güvenilir ve geçerli ölçümler elde etmek, bu ölçümlerden yararlanarak da bireyler hakkında doğru kararlar verilmesini sağlamaktır. Bu amaç doğrultusunda öncelikli olarak yapılması gereken, ölçülmesi istenen örtük özelliği tüm yönleriyle açıklamaktır.

Bir testi yanıtlayan bireyin performansının altında yatan örtük özelliği açıklamak için geliştirilen başlıca ölçme kuramı *Klasik Test Kuramı (KTK)*'dir. KTK; gözlenen puanı ifade eden toplam test puanını, ölçülen özelliğe ilişkin gerçek puan ve ölçme işlemine karışan hata puanının toplamı olarak açıklayan bir modeldir. Bu basit modelde elde edilmesi hedeflenen puan, bireylerin ölçülen değişkenlerine ilişkin gerçek puanlarıdır. Ancak ölçme işlemine karışan çeşitli hatalar sebebiyle, gerçek puan doğrudan elde edilemeyip, gözlenen puanlar yoluyla kestirilir. KTK'nın gerçek puan kestirim sürecinde, madde ve test istatistikleri yanıtlayıcı grubuna, grup istatistikleri de madde ve test istatistiklerine bağımlıdır. Daha işevuruk bir ifadeyle, uygulanan teste ait madde ve test istatistikleri, testi yanıtlayan bireylerin yetenek düzeylerine; bireylerin yetenek düzeyleri de madde ve test istatistiklerine bağlı olarak değişmektedir. Örneğin zor bir testi alan bir birey düşük yetenek düzeyine, kolay bir testi alan birey yüksek yetenek düzeyine sahipmiş gibi; ya da, yüksek yetenek düzeyindeki bir birey tarafından yanıtlanan bir madde kolay, düşük yetenek düzeyindeki bir birey tarafından yanıtlanan bir madde de zormuş gibi görünecektir (Hambleton, Swaminathan & Rogers, 1991). Tüm bunların yanı sıra, gözlenen puanların testin tümüne yönelik olacak şekilde hesaplanmasından dolayı da, testin her bir maddesi üzerinde gösterilecek performansa ilişkin kestirim olanağı sağlanamamaktadır (Hambleton & Swaminathan, 1985). KTK'ya ilişkin tüm bu olumsuzluklar, alternatif kuram arayışlarına yol açan ve çözümlenmesi gereken ciddi sınırlılıklar olmuşlardır.

İkinci bir ölçme kuramı olan *Madde Tepki Kuramı (MTK)* ise, KTK'dan farklı olarak, bireylerin doğrudan gözlenemeyen örtük özellik düzeyleri ile test maddeleri üzerindeki gözlenen performansları arasındaki ilişkiyi matematiksel olarak ortaya koymayı amaçlayan bir model kurar. Bu model sayesinde, bir bireyin ilgili örtük özelliğe sahip olma düzeyi ile o örtük özelliğe yönelik bir test maddesini doğru yanıtlama olasılığı arasındaki ilişki tanımlanarak; ilgili bireyin her bir test maddesi üzerinde nasıl performans göstereceğine ilişkin kestirim olanağı sağlanır. Kuramın bu kestirimleri sağlamasının temel sebebi, her bir bireye ait yetenek düzeyinin, testi alan yanıtlayıcı grubundan ve yöneltilen test maddelerinden bağımsız olarak hesaplanabilen mutlak (değişmez) bir özellik olduğu varsayımdır. Bu varsayımın doğal bir sonucu olarak, bir birey farklı yanıtlayıcı grupları içerisinde yer alıp farklı testleri alsa dahi, hesaplanacak yetenek düzeyi tüm test uygulamalarında aynı sonucu verecektir (Hambleton & Swaminathan, 1985; Embretson & Reise, 2000).

1.5.1.1. Madde Karakteristik Eğrisi (MKE)

Her bir test maddesine ilişkin gösterilen performansın; bireyin yetenek düzeyi üzerindeki regresyonu *madde karakteristik fonksiyonu (MKF)*, bu fonksiyona ait eğri de *madde karakteristik eğrisi (MKE)* olarak adlandırılır (Lord & Novick, 1968). MKE'nin matematiksel fonksiyonu, belli bir yetenek düzeyindeki (θ) bireylerin belirli bir maddeyi doğru yanıtlama olasılıklarını ($P(\theta)$) verir. $P(\theta)$ değerlerinin, θ 'nın bir fonksiyonu olarak çizilmesi durumunda, Şekil 1.1'deki gibi tipik bir eğri elde edilir:



Şekil 1.1. Tipik Bir Madde Karakteristik Eğrisi (Lord & Novick, 1968)

Şekil 1.1'deki eğriden yola çıkılarak kolayca yorumlanacağı gibi, tipik bir madde için düşük yetenek düzeyindeki bireylerin maddeyi doğru yanıtlama olasılıklarının düşük, yüksek yetenek düzeyindeki bireylerin ise maddeyi doğru yanıtlama olasılıklarının yüksek olması beklenir. Böylelikle, bireyin testle ölçülmek istenen özelliğe sahip olma miktarı arttıkça, ilgili maddeyi doğru yanıtlama konusundaki başarısı da artmaktadır. Kuramın MKE üzerinde gözlemlenebilen başka bir özelliği de, bir bireyin bir test maddesini doğru yanıtlama olasılığının, testi yanıtlayan diğer bireylerin yetenek düzeylerinden bağımsız olarak sadece ilgili maddenin MKE'sinin biçimine bağlı olmasıdır. Benzer şekilde ilgili maddeye ait MKE'nin biçimi de, o maddeyi yanıtlayan bireylerin yetenek düzeylerine bağlı değildir. MTK'nın KTK'ya göre en önemli üstünlüğü olarak kabul edilen bu değişmezlik özelliği sayesinde, test maddelerine ait farklı özellikleri tanımlayan madde parametrelerinin testi alan gruptan, testi alan bireylerin örtük özelliklerini tanımlayan yetenek parametrelerinin ise madde örnekleminden bağımsız olarak kestirilebilmesi sağlanır (Lord & Novick, 1968; Crocker & Algina, 1986).

MKE'yi tanımlamak için maddenin farklı özelliklerini ifade eden üç farklı parametre kullanılmaktadır. Bunlar; sırasıyla madde ayırıcılık (item discrimination), madde güçlük (item difficulty) ve şans (guessing) parametrelerine karşılık gelen a , b ve c parametrelerdir. MKE'yi teknik açıdan tanımlayan ilk parametre olan b parametresi, maddenin yetenek ölçeği üzerinde bulunduğu noktayla ilgili bir indistir ve bir maddenin P olasılıkla doğru yanıtlanması için gerekli olan yetenek düzeyini gösterir (Lord & Novick, 1968; Hambleton & Swaminathan, 1985). b parametresi kuramsal olarak $-\infty$ ile $+\infty$ arasında değerler alabilse de, genellikle -3 ile $+3$ aralığında tanımlanır (Baker, 2001). Maddeye ait b değeri arttıkça, maddeyi doğru yanıtlamak için gerekli olan yetenek düzeyi artar; azaldıkça da azalır. Pozitif b değerleri maddenin zor, negatif b değerleri maddenin kolay, 0 'a yakın olduğu değerler ise maddenin orta güçlükte olduğunu gösterir.

Maddenin ayırıcılık özelliğini tanımlayan a parametresi, bir madde için, bireyin yetenek düzeyinin maddenin güçlük düzeyine ($\theta=b$) eşit olduğu noktadaki eğime karşılık gelmektedir. Bu nedenle madde eğim parametresi olarak da adlandırılan a , madde ile ölçülmek istenen örtük özelliğin ne derece doğru ölçülebildiğinin bir göstergesi olup lojistik modeller için 0 ile 2 aralığında tanımlanır (Hambleton & Swaminathan, 1985; Crocker & Algina, 1986; Baker, 2001). Düşük ayırıcılığa

sahip maddelerde maddenin doğru yanıtlanma olasılığı düşük ve yüksek yetenek düzeyindeki bireyler için hemen hemen aynı olmakla birlikte; yüksek ayırtıcılığa sahip maddelerde maddenin doğru yanıtlanma olasılığı düşük ve yüksek yetenek düzeyindeki bireyler için farklılıklar göstermektedir. Başka bir ifadeyle, maddeye ait a değerinin artması, ilgili maddenin bilen öğrenciler ile bilmeyen öğrencileri iyi ayırt ettiği anlamına gelir.

MKE'yi tanımlayan bir diğer parametre olan c parametresi ise bir maddenin sadece tahminle doğru yanıtlanma olasılığını ifade eder. c parametresi 0 ile 1 değerleri aralığında tanımlansa da, uygulamada .35'den daha yüksek olan c değerleri kabul görmemektedir (Hambleton & Swaminathan, 1985; Baker, 2001). c parametresinin değeri arttıkça ilgili maddenin şansa doğru yanıtlanma olasılığı artarken, değer azaldıkça bu olasılık da azalır.

1.5.1.2. İkili Puanlanan Madde Yanıtları Üzerinde Kullanılan Madde Tepki Modelleri

MTK kapsamında ele alınan modeller test maddelerinin puanlanma biçimine bağlı olarak değişiklik gösterirler. İkili maddeler (binary items) olarak da nitelendirilen maddelerin 1 (doğru) ya da 0 (yanlış) şeklinde iki kategorili (dichotomous) olarak puanlandığı durumlarda kullanılan MTK modelleri; *bir parametrelili lojistik (1 PL)*, *iki parametrelili lojistik (2 PL)* ve *üç parametrelili lojistik (3 PL)* modellerdir.

1PL modelde, bireyin bir maddeyi doğru yanıtlanma olasılığı ile yetenek düzeyi arasındaki ilişki b parametresi üzerinden tanımlanmaktadır. Bu model kapsamında a parametreleri eşit ve çoğunlukla 1, c parametreleri ise 0 olarak kabul edilir (Crocker & Algina, 1986; Embretson & Reise, 2000). Bu durumda, θ yetenek düzeyindeki bir bireyin maddeyi doğru yanıtlanma olasılığı şu şekilde ifade edilir:

$$P_i(\theta) = \frac{e^{D_i(\theta-b_i)}}{1 + e^{D_i(\theta-b_i)}}$$

$P_i(\theta)$: θ yeteneğindeki bir bireyin i maddesini doğru yanıtlanma olasılığı

b_i : i maddesinin güçlük indeksi

D : ölçek sabiti (1,7)

2PL modelde, bireyin bir maddeyi doğru yanıtlanma olasılığı ile yetenek düzeyi arasındaki ilişki b ve a parametresi üzerinden tanımlanmaktadır (Crocker & Algina,

1986; Embretson & Reise, 2000). Bu durumda, θ yetenek düzeyindeki bir bireyin maddeyi doğru yanıtlama olasılığı şu şekilde ifade edilir:

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}$$

$P_i(\theta)$: θ yeteneğindeki bir bireyin i maddesini doğru cevaplama olasılığı

b_i : i maddesinin güçlük indeksi

a_i : i maddesinin ayırıcılık gücü indeksi

D : ölçek sabiti (1,7)

3PL modelde, bireyin bir maddeyi doğru yanıtlama olasılığı ile yetenek düzeyi arasındaki ilişki b , a ve c parametresi üzerinden tanımlanmaktadır (Crocker & Algina, 1986; Embretson & Reise, 2000). Bu durumda, θ yetenek düzeyindeki bir bireyin maddeyi doğru yanıtlama olasılığı şu şekilde ifade edilir:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}$$

$P_i(\theta)$: θ yeteneğindeki bir bireyin i maddesini doğru cevaplama olasılığı

b_i : i maddesinin güçlük indeksi

a_i : i maddesinin ayırıcılık gücü indeksi

c_i : i maddesinin şans parametresi

D : ölçek sabiti (1,7)

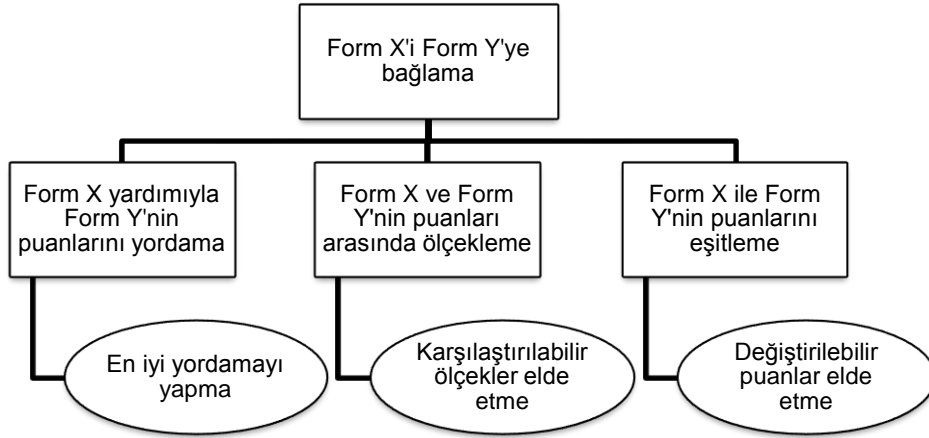
MTK'nın, güçlü matematiksel altyapıya sahip modellerini işe koşması sayesinde özellikle test geliştirme, soru bankası oluşturma, bireye uyarlanmış ölçme aracı hazırlama, madde yanlılığı belirleme, madde seçeneklerini ağırlıklandırma ve test eşitleme konularında karşılaşılan sorunlara çözüm getirdiği iddia edilmektedir (Hambelton & Swaminathan, 1985).

1.5.2. Test Eşitleme ile İlgili Kuramsal Çerçeve

1.5.2.1. Bağlama (Linking)

Angoff (1971), herhangi iki testten elde edilen puanlar arasında bir bağ oluşturmak için birinci testteki bir puanın ikinci testteki bir puana dönüştürülmesi gerektiğini belirtmiştir. Holland ve Dorans (2006), bir test ile başka bir testten elde edilen

puanlar arasındaki dönüşümleri ifade eden tüm işlemlerin en genel sınıfını bağlama (linking) olarak adlandırmış ve farklı test formları arasında bağ kurmak için kullanılan yöntemleri; yordama (predicting), ölçekleme (scaling) ve eşitleme (equating) olmak üzere üç temel kategoriye ayırmışlardır. Bağlama yöntemlerine ilişkin bu sınıflandırma biçimi ve yöntemlere ilişkin genel amaçlar Şekil 1.2'deki gibidir.



Şekil 1.2. Bağlama Yöntemlerinin Sınıflandırılması (Holland & Dorans, 2006)

Yordama yönteminin, bir test veya başka bir ölçme işleminden elde edilen puanları tahmin etmek için kullanımı, puanları bağlarken dikkate alınan muhtemelen en eski yaklaşımdır. Yordama çalışmaları, testi alan bireyler hakkında var olan bilgilerden yola çıkarak test formlarından elde edilecek puanlar hakkında tahminde bulunmayı amaçlar. Tahmin sürecinde kullanılan bilgi kaynakları, bireylerin başka testlerden aldıkları puanlar veya demografik bilgiler olabilir (Holland, 2007).

Holland ve Dorans'ın (2006) sınıflamasında ikinci kategoriye oluşturan ölçekleme yöntemi, karşılaştırılabilir puanlar elde edilmesi amacıyla iki farklı test formundan elde edilen puanların ortak bir ölçeğe dönüştürülmesini ifade eder. Test formlarını ortak bir ölçek üzerinde ifade ederek testler arasında yaratılan dolaylı bağlantılar, X ve Y'den elde edilen puanların karşılaştırılabilir hale getirilmelerini sağlar. Ölçekleme yöntemleri; test bataryası ölçekleme (battery scaling), ortak madde ölçekleme (anchor scaling), dikey ölçekleme (vertical scaling), kalibrasyon (calibration) ve uyum (concordance) isimleriyle adlandırılan birçok alt kategoriden oluşmaktadır (Holland, 2007). Bu kategoriler içerisinde uygulamalarda çok sık karşılaşılan dikey ölçekleme, farklı gruplar üzerinde uygulanan aynı yapıyı farklı

içeriklerle ve aynı güvenilirlikte ölçmeyi amaçlayan iki test formundan elde edilen puanların aynı ölçek üzerine yerleştirilmesini ifade eder (Kolen & Brennan, 2004).

Üçüncü bağlama yöntemi olan eşitleme ise, yordama ve ölçeklemeden daha fazla anlam taşıyarak, iki test formundan elde edilen puanlar arasında bağlantı kurmanın en güçlü hali olarak ifade edilir (Holland & Dorans, 2006). Eşitleme yönteminin işe koşulması ile iki test formundan elde edilen puanlar arasında öyle bir bağlantı üretilir ki, işlem sonucunda her bir test formundan elde edilen puanlar sanki aynı teste aitlermiş gibi işlem görürler (Dorans, Moses & Eignor, 2010).

1.5.2.2. Test Eşitleme

Angoff (1971) test eşitlemenin, bir testin birim sisteminin diğer test formunun birim sistemine dönüştürülmesi anlamına geldiğini ve bu dönüştürme işleminden sonra her iki formdan elde edilen puanların eşdeğer olacağını belirtmiştir. Kolen (1988) test eşitlemenin, benzer içerik ve benzer güçlük düzeyinde geliştirilen test formları arasındaki farklılıkları düzenleyerek, bu test formlarından elde edilen puanların birbiri yerine kullanılabilmesini sağlayan istatistiksel bir süreç olduğunu belirtmiştir. Woldbeck (1998) için test eşitleme, test geliştiricilere ve uygulayıcılara bir testin iki farklı formundan alınan puanları karşılaştırma olanağı sağlayan bir grup yöntemi tanımlar. Felan (2002) ise test eşitlemenin, ancak bir testin birden fazla formunun var olduğu durumlarda kullanılabileceğini ve eşitleme işlemi sonunda farklı formları alan kişilerin birbiriyle karşılaştırılabileceğini belirtmiştir. Tanımlar incelendiğinde, test formları arasındaki güçlük farklılıklarının düzenlenerek, elde edilen puanların birbirleri yerine kullanılabilmesini sağlayan istatistiksel sürecin test eşitleme olarak adlandırıldığı görülmektedir.

Eşitleme çalışmalarının yürütülebilmesi için karşılanması gereken bazı koşullar bulunmaktadır. KTK kapsamında gerçekleştirilmesi planlanan eşitleme çalışmaları için karşılanması gereken koşullar; eşitlik, simetri, gruplar arası değişmezlik, aynı yapıyı ölçme ve tek boyutluluktur (Hambleton & Swaminathan, 1985). MTK kapsamında gerçekleştirilmesi planlanan eşitleme çalışmalarında ise model-veri uyumunun sağlanması, eşitleme işlemlerine başlamak için yeterli görülmektedir (Kolen, 1981). Tercih edilen kurama göre gerekli koşullar karşılandıktan ve test formlarından elde edilen puanların eşitlenebileceği kararı alındıktan sonra eşitleme deseninin seçilmesi aşamasına geçilebilir (Kolen & Brennan, 2004).

1.5.2.3. Eşitleme Desenleri

Eşitlemesi yapılacak test formlarına ait verileri toplama şeklini ifade eden eşitleme desenleri; *tek grup deseni (single group design)*, *rastgele grup deseni (random group design)* ve *ortak test deseni (anchor test design)* olmak üzere üç temel başlık altında gruplandırılmaktadır.

En basit eşitleme deseni olan *tek grup deseninde*, eşitlemesi yapılmak istenen iki test formu aynı yanıtlayıcı grubuna uygulanır. Aynı yanıtlayıcı grubunun her iki test formunu da alması nedeniyle, yanıtlayıcıların yetenek düzeylerinden kaynaklanan bir hata ortaya çıkmamaktadır. Ancak aynı gruba iki farklı uygulama yapmak, yanıtlayıcıların ikinci forma ait uygulama sırasında yorgun düşmelerine neden olabilir. Başka bir ifadeyle, test formlarının veriliş sırasından kaynaklanan bir hata eşitleme sürecine karışabilir (Kolen & Brennan, 2004). Tek grup desenin uygulama şekli Tablo 1.2'deki gibidir.

Tablo 1.2: Tek Grup Deseni

<i>Evren</i>	<i>Örneklem</i>	<i>Form X</i>	<i>Form Y</i>
P	1	✓	✓

Test formlarının veriliş sırasından kaynaklanabilecek olası hataları engellemek için sıra dengeleme işlemi önerilmiştir (Kolen, 1988). Bireylerin yarısına önce Form X sonra Form Y, diğer yarısına ise önce Form Y sonra Form X uygulayarak yapılan bu dengeleme işlemi sayesinde, aynı yanıtlayıcı grubuna iki uygulama yaptıktan kaynaklanan etkiler ortadan kalkmaktadır (Kolen & Brennan, 2004).

Rastgele grup deseninde, yanıtlayıcı grubu rastgele olarak ikiye ayrılır ve gruplardan birine Form X diğerine ise Form Y uygulanır. Uygulama sonrasında iki formdan elde edilen performans düzeyleri arasındaki farklılık, doğrudan formların güçlük düzeyleri arasındaki farklılığı verir. Bu farklar kullanılarak kurulan eşitleme ilişkileri sayesinde formlardan elde edilen puanlar birbirine dönüştürülür. Bu desen, yanıtlayıcıların sadece tek bir test formunu alması nedeniyle yorgunluk etkisinden kaynaklanabilecek hataları ortadan kaldırmaktadır. Bunun yanı sıra, zamandan tasarruf sağlaması sebebiyle de tek grup desenine göre daha fazla tercih edilmektedir (Kolen, 1988). Ancak, ikiye ayrılan gruplardaki yanıtlayıcıların yetenek dağılımlarındaki farklılıklar, eşitleme sürecinde belirlenemeyen bir yanlılık açığa

çıkarabilir (Kolen & Brennan, 2004). Rastgele grup deseninin uygulama şekli Tablo 1.3'deki gibidir.

Tablo 1.3: Rastgele Grup Deseni

<i>Evren</i>	<i>Örneklem</i>	<i>Form X</i>	<i>Form Y</i>
P	1	✓	
P	2		✓

Ortak test deseni, aynı test formunun tekrar uygulanması gerektiği ancak güvenlik nedenlerinden dolayı uygulanamadığı koşullarda kullanılır (Kolen & Brennan, 2004). Desenin uygulaması, ortak maddelere sahip iki farklı test formunun iki farklı yanıtlayıcı grubuna verilmesi şeklinde gerçekleştirilir. Bu test deseninde kullanılan ortak maddelerin işlevi, farklı formları alan farklı yanıtlayıcı grupları arasındaki farklılıkları kontrol ederek eşitleme ilişkilerini ortaya çıkarmaktır. Başka bir ifadeyle, ortak maddeler yoluyla edinilen bilgiler kullanılarak test formlarından elde edilen puanlar birbirlerine dönüştürülür (Crocker & Algina, 1986). Ortak test deseninin uygulama şekli Tablo 1.4'deki gibidir.

Tablo 1.4: Ortak Test Deseni

<i>Evren</i>	<i>Örneklem</i>	<i>Form X</i>	<i>Ortak Test</i>	<i>Form Y</i>
P	1	✓	✓	
Q	2		✓	✓

Ortak test deseni uygulamalarında, test formları arasındaki eşitleme ilişkisinin belirlenmesinde önemli rol oynayan ortak maddelerin belli başlı bazı özelliklere sahip olması gerekir. Bu özelliklerden ilki, ortak maddelerin eşitlenmek istenen test formlarının içeriğini ve özelliklerini doğrulukla yansıtmasıdır. Başka bir ifadeyle, test formlarındaki maddeler ile ortak test içerisindeki maddeler olabildiğince özdeş olmalıdır. Angoff (1971) ortak maddelerin; kapsam, yapı, madde türü, güçlük düzeyi gibi özellikler açısından tüm testin iyi birer temsilcisi olması gerektiğini savunmuştur. Kolen ve Brennan'da (2004) benzer bir şekilde, ortak maddelerin özellikle güçlük ve içerik bakımından tüm testi iyi bir şekilde yansıtması gerektiğini belirtmiştir. Ayrıca, ortak maddelerin her iki test formunda da herhangi bir kelime veya ifade değişikliği yapılmadan ve aynı soru numaralarına denk gelecek şekilde yerleştirilmeleri gerektiğini vurgulamışlardır.

Ortak maddelerin sahip olması gereken bir diğer özellik, maddelerin tüm test formu içerisindeki miktarının yeterli olmasıdır. Angoff (1971) ortak testin uzunluğunun,

tüm testteki madde sayısının %20'si kadar olması gerektiğini savunmuştur. Hambleton, Swaminathan ve Rogers (1991) ile Woldbeck (1998), ortak maddeler için gereksinim duyulan sayının, testteki madde sayısının yaklaşık olarak %20 ile %25'i arasında olması gerektiğini ifade etmişlerdir. Kolen ve Brennan (2004) ise, ortak test deseni kullanılarak yürütülen eşitleme çalışmalarındaki ortak madde miktarının artmasının eşitleme hatasını azalttığını raporlamışlardır.

1.5.2.4. MTK'ya Dayalı Eşitleme Yöntemleri

Yürütülecek araştırmaya ait verilerin seçilen desene uygun şekilde toplanmasının ardından hangi istatistiksel yöntem ile eşitleme ilişkilerinin kestirileceğine karar verilir. MTK'ya dayalı olarak yapılacak eşitleme çalışmalarındaki yöntem seçimi, veri toplama aşamasında kullanılan eşitleme desenine bağlı olarak değişiklik gösterir. Ortak test deseni dışındaki tüm desenlerde, yanıtlayıcı grupları aynı veya eş yetenek düzeyinde oldukları için kestirilen parametreler aynı ölçek üzerinde yer alırlar. Ortak test deseninde ise farklı çalışma gruplarından gelen farklı yetenek düzeylerindeki yanıtlayıcı gruplarına uygulanan test formlarından elde edilen parametrelerin aynı ölçek üzerine yerleştirilmesi gerekir. Parametrelerin ortak bir ölçeğe yerleştirilmesi için yapılacak ek ölçek dönüştürme işleminin *ayrı kalibrasyon (separate calibration)* ve *eşzamanlı kalibrasyon (concurrent calibration)* olmak üzere iki yolu vardır (Kolen & Brennan, 2004).

1.5.2.4.1. Ayır kalibrasyon

Ayır kalibrasyon kullanıldığında, farklı test formlarından elde edilen madde ve yetenek parametrelerini aynı ölçeğe yerleştirmek için aşağıdaki adımlar takip edilir (Kolen & Brennan, 2004):

1. Eşitlemesi yapılacak her iki test formuna ait parametreler ayrı ayrı kestirilir.

Bu adımda, X ve Y test formlarını alan yanıtlayıcılara ait yetenek parametre değerleri (θ) ve test formlarındaki maddelere ait madde parametre değerleri (a , b , c) uygun istatistik programları aracılığıyla kestirilir.

2. Moment yöntemleri veya karakteristik eğri yöntemleri kullanılarak A (eğim) ve B (kesişim) olarak ifade edilen eşitleme katsayıları hesaplanır.

Moment yöntemleri, ortak maddelerin a ve b parametre değerleri üzerinden ölçek dönüştürme işlemleri yapan yöntemlerdir. Bunlar, ortalama-ortalama ve ortalama-

standart sapma yöntemleridir. Ortalama-ortalama yönteminde, A ve B katsayılarının hesaplanması için ortak maddelerin a ve b parametrelerine ilişkin ortalamalar kullanılırken; ortalama-standart sapma yönteminde, ortak maddelerin sadece b parametrelerine ilişkin ortalama ve standart sapma değerleri kullanılır. Karakteristik eğri yöntemleri ise, ortak maddelerin tüm madde parametrelerini hesaba katarak, madde karakteristik eğrileri arasındaki farkları azaltmaya yönelik olarak geliştirilen ölçek dönüştürme yöntemleridir. Bunlar, Haebara ve Stocking-Lord yöntemleridir.

Haebara yönteminde, belli bir yetenek düzeyindeki yanıtlayıcılar için madde karakteristik eğrileri arasındaki fark, her bir maddeye ait madde karakteristik eğrileri arasındaki farkın karelerinin toplamıdır. Bu yöntemle ait fonksiyonun matematiksel olarak ifadesi aşağıdaki gibidir.

$$L(\theta_i) = \sum_{j=1}^m [p_{ij}(\theta_i, a_{XLj}, b_{XLj}, c_{XLj}) - p_{ij}(\theta_i, \frac{a_{YLj}}{A}, Ab_{YLj} + B, c_{YLj})]^2$$

Bu denklemde, j . bireyin i maddesini doğru yanıtlama olasılığı p_{ij} , X formundaki j . ortak madde için sırasıyla madde ayırıcılık, madde güçlük ve şans parametreleri $a_{XLj}, b_{XLj}, c_{XLj}$, Y formundaki j . ortak madde için sırasıyla madde ayırıcılık, madde güçlük ve şans parametreleri $a_{YLj}, b_{YLj}, c_{YLj}$ şeklinde ifade edilmektedir. Denklemde, Y formundaki ortak maddelere ilişkin dönüştürülmüş parametreler kullanılmaktadır.

Stocking-Lord yönteminde, belli bir yetenek düzeyindeki yanıtlayıcılar için madde karakteristik eğrileri arasındaki fark, her bir maddeye ait madde karakteristik eğrileri arasındaki farkın toplamının karesidir. Bu yöntemle ait fonksiyonun matematiksel olarak ifadesi aşağıdaki gibidir.

$$L(\theta_i) = [\sum_{j=1}^m p_{ij}(\theta_i, a_{XLj}, b_{XLj}, c_{XLj}) - \sum_{j=1}^m p_{ij}(\theta_i, \frac{a_{YLj}}{A}, Ab_{YLj} + B, c_{YLj})]^2$$

Ayrı kalibrasyon yöntemleri konusunda yapılan çalışmalar, karakteristik eğri yöntemlerinin ve özellikle de karakteristik eğri yöntemleri kapsamındaki Stocking-Lord yönteminin, ortalama-ortalama ve ortalama-standart sapma yöntemlerinden daha kararlı sonuçlar üretme eğiliminde olduğunu ortaya koymaktadır (Baker & Al-

Karni, 1991, Hanson & Beguin, 2002, Kolen & Brennan, 2004, Kilmen, 2010, Gök, 2012).

3. Ayrı kalibrasyon yöntemlerinden biri kullanılarak hesaplanan A ve B katsayıları ile ölçek dönüştürme işlemleri gerçekleştirilir.

Bu adımda, X formunu alan belli yetenek düzeyindeki yanıtlayıcıların almadıkları Y formu için yetenek düzeyleri karşılıkları ve benzer şekilde X formundaki madde parametrelerinin de Y formundaki karşılıkları elde edilir. Dönüştürme işlemleri ile X formuna ait parametrelerin aynı ölçek üzerine yerleştirilmiş olan diğer formdaki karşılıkları, dönüştürülmüş parametreler olarak adlandırılır.

Ayrı kalibrasyonda, moment veya karakteristik eğri yöntemleri kullanılarak elde edilen A ve B katsayılarının, yetenek ve madde parametrelerine ait ölçek dönüştürme formülleri içerisinde kullanımı aşağıdaki gibidir.

$$\theta_{Yi} = A\theta_{Xi} + B$$

Bu denklemde, X formunu (referans form) alan *i* bireyine ait yetenek düzeyi θ_{Xi} , aynı bireyin girmedığı Y formundan alacağı yetenek düzeyi ise θ_{Yi} ile ifade edilmektedir.

$$a_{Yj} = \frac{a_{Xj}}{A}$$

$$b_{Yj} = Ab_{Xj} + B$$

$$c_{Yj} = c_{Xj}$$

X formundaki *j* maddesine ait madde parametreleri a_{Xj} , b_{Xj} , c_{Xj} ; Y formundaki aynı maddeye ait dönüştürülmüş madde parametreleri ise a_{Yj} , b_{Yj} , c_{Yj} ile ifade edilmektedir.

1.5.2.4.2. Eşzamanlı kalibrasyon

Eşzamanlı kalibrasyon sırasında, eşitlemesi yapılacak test formlarında bulunan maddelere ait parametreler birlikte (eşzamanlı olarak) kestirilir. Kestirim işlemleri için ortak maddelerin her iki test formunda da aynı madde parametrelerine sahip olduğu varsayımı bulunmaktadır. Bu varsayımdan hareketle, parametre kestirim işlemlerini yapacak istatistik program yalnızca bir kere çalıştırılmakta ve aynı ölçek üzerinde yer alan parametre değerleri üretilmektedir. Dolayısıyla eşzamanlı

kalibrasyonda, ayrı kalibrasyon yöntemlerinde olduğu gibi eşitleme katsayılarının hesaplanmasına ve ek bir dönüştürme işlemi yapılmasına ihtiyaç yoktur (Kim, 2007). Kullanılan madde tepki modeline ait varsayımların tam anlamıyla karşılandığı durumlarda, eşzamanlı kalibrasyon yönteminin ayrı kalibrasyon yöntemine göre daha doğru sonuçlar ürettiği bilinmektedir (Hanson & Beguin, 2002).

1.5.2.5. Eşitleme Hatası

Seçilen eşitleme yöntemine uygun olacak biçimde yürütülen eşitleme işlemleri sonrasında elde edilen eşitleme sonuçlarının değerlendirilebilmesi için eşitlemede hata kavramının anlaşılması gereklidir (Kolen & Brennan, 2004). Eşitlemede hata kavramı, bireyin aldığı test için kestirilen yetenek düzeyi ile almadığı test için kestirilen yetenek düzeyi arasındaki farkla açıklanmaktadır (Cook & Eignor, 1991). Bu açıklamaya göre, bireylerin kendilerine uygulanan test formuna ilişkin gerçek yetenek düzeyleri ile almadıkları test formuna ilişkin dönüştürülmüş yetenek düzeylerinin birbirine eşit çıkmaması, eşitleme işlemine karışan hatalardan kaynaklanmaktadır.

Eşitleme işleminin toplam hatası, *rastgele (random)* ve *sistemik (bias)* olmak üzere iki tür eşitleme hatasının toplamıyla ifade edilmektedir. Eşitlemesi yapılacak test formlarının uygulandığı örneklem özellikleriyle ilgili bir hata türü olan rastgele eşitleme hatası, eşitlemenin standart hatası kavramıyla da anılmaktadır. Rastgele hatanın miktarı, eşitleme sürecine katılan örneklem büyüklüğü arttıkça azalacak, örneklem büyüklüğü azaldıkça da artacaktır. Çok büyük örneklerde ise elde edilen hata miktarı önemsiz hale gelecektir. Bu nedenle, test formlarının daha büyük örneklerde uygulanması ile rastgele eşitleme hatası azaltılabilir (Kolen, 1988; Felan, 2002; Kolen & Brennan, 2004).

Eşitleme yanlılığı olarak da ifade edilen sistemik hata ise, uygun eşitleme deseni ve eşitleme yönteminin kullanılmadığı veya bunların varsayımlarının ihlal edildiği durumlarda ortaya çıkan hata türüdür. Örneğin, tek grup deseni kullanılarak yürütülen eşitleme çalışmalarında bireylerin ikinci test formunu yanıtlamaları sırasında yorgun düşmelerinden kaynaklanabilecek etkiler veya ortak test deseni uygulamalarında yanıtlayıcı gruplarının yetenek düzeylerinin büyük ölçüde farklılık göstermesinden veya ortak maddelerin toplam test formunu iyi bir şekilde temsil

etmemesinden kaynaklanabilecek etkiler sistematik hatayı arttıracaktır. Bu nedenle sistematik hata; dikkatli bir test geliştirme, eşitleme desenlerine ait dezavantajların kontrol edilerek uygulanması ve uygun eşitleme yöntemlerinin kullanılmasıyla en aza indirgenebilir (Kolen, 1988; Felan, 2002; Kolen & Brennan, 2004).

Sonuç olarak, yürütülen herhangi bir eşitleme çalışması için hem rastgele hem de sistematik hatanın olabildiğince küçük çıkması, eşitleme işlemlerinin doğrulukla gerçekleştirildiğinin bir göstergesidir. Tam tersi durumlarda, bireylere ilişkin dönüştürülen yetenek düzeyleri, eşitleme işlemine karışan hatalar sebebiyle gerçek değerlerinden büyük ölçüde farklılık göstereceğinden, eşitlenmiş puanlara ve bu puanlar sonucunda alınan kararlara duyulan güven azalır. Daha hatasız ve daha doğru sonuçlar elde edilebilmesi için eşitleme çalışmalarının tasarlanma ve yürütülme aşamalarında, eşitleme hatalarının artmasına yol açacak unsurların mümkün olduğunca azaltılmasına çalışılmalıdır. Bu çabanın doğal bir sonucu olarak da, ölçme ve değerlendirme faaliyetlerinin kalitesinin artması beklenir.

1.5.3. Kayıp Veri ile ilgili Kuramsal Çerçeve

1.5.3.1. Veri Matrisi

Ölçme işlemleri, çoğunlukla sayısal olarak ifade edilen, üst seviyede matematiksel işlem ve yorumlama yapılmasına olanak sağlayan ölçme sonuçları açığa çıkarırlar. İlk elde edildiklerinde üzerlerinde herhangi bir istatistiksel işlem yapılmadığı için ham veri olarak adlandırılan bu ölçme sonuçları, satırlarında gözlemler (bireyler) ve sütunlarında değişkenler bulunan iki boyutlu veri matrisleri üzerinde gösterilirler. Veri matrisleri dikdörtgen biçimli bir görünüme sahiptir ve standart istatistiksel yöntemler bu biçimdeki veri setlerini analiz etmek üzere geliştirilmişlerdir (Little & Rubin, 1987, Schafer & Graham, 2002).

Çok değişkenli istatistikte *veri matrisi*, araştırmaya katılan gözlem sayısını ifade eden n tane satırdan ve tüm gözlemler için ölçülen değişkenlerin sayısını ifade eden p tane sütundan oluşan $n \times p$ boyutlu değerler topluluğu olarak ifade edilir ve en genel haliyle Şekil 1.3'deki gibi gösterilir:

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2p} \\ y_{31} & y_{32} & \dots & y_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{np} \end{bmatrix}$$

Şekil 1.3. Veri Matrisi Gösterimi

Veri matrisleri alfabenin herhangi bir harfi ile temsil edilebilirken, matrisin herhangi bir gözlemi i , herhangi bir değişkeni ise j indisiyle gösterilir. Burada Y ile temsil edilen veri matrisinin y_{ij} . elemanı, i . gözlemin j . değişken için aldığı değeri ifade eder. Başka bir ifadeyle, tipik bir veri matrisine ait her bir hücre, ilgili gözlemlerin ilgili değişkenlere ait ölçme sonuçlarını barındırır.

1.5.3.2. Tam Olmayan Veri Matrisi ve Kayıp Veri Gösterimi

Rubin'in (1976) kayıp veri teorisine göre, tipik bir veri matrisindeki her bir hücre için "gözlendi" veya "gözlenemedi" olmak üzere iki farklı ölçme sonucu bulunmaktadır. Şekil 1.3'deki veri matrisinde bulunan y_{ij} değerlerinin tamamının gözlenmiş olması durumunda Y matrisi *tam veri matrisi* olarak adlandırılır. Neredeyse tüm veri toplama süreçlerinde, araştırma kapsamında ele alınan tüm değişken bilgilerini içeren tam veri matrislerine ulaşmak istenir. Ancak ölçme işlemine karışan çeşitli sebeplerle bunun gerçekleştirilmesi oldukça güç olmaktadır. Bunun yerine sıklıkla elde edilen, veri matrisine ait bir veya birden fazla y_{ij} değerinin gözlenemediği ve gözlenemeyen y_{ij} değerlerinin yer aldığı hücrelerin boş kaldığı durumlarda ortaya çıkan *tam olmayan veri matrisleridir*. Bu matrislerde, gözlenen ile gözlenemeyen değerler bir arada yer alırlar.

Y veri matrisindeki gözlenemeyen değerlerin z_{ij} ile temsil edildiği düşünülduğünde, z gösterge değişkeni için i . gözlemin j . değişkene ait değeri basitçe *kayıp* anlamına gelir. Buna göre tam olmayan veri vektörü $y=Y_{gözlenen}$ ve kayıp veri vektörü $z=Y_{kayıp}$ olmaktadır. Dolayısıyla tam veri $Y=(Y_{gözlenen}, Y_{kayıp})$ biçimini almaktadır (Peng, Harwell, Liou & Ehman, 2007). Bu durumda, kayıp değerleri içeren yeni Y matrisi Şekil 1.4'deki gibi gösterilir:

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2p} \\ z_{31} & y_{32} & \dots & z_{3p} \\ y_{41} & z_{42} & \dots & z_{4p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & z_{n2} & \dots & y_{np} \end{bmatrix}$$

Şekil 1.4. Tam Olmayan Veri Matrisi Gösterimi

Uygulamalar sırasında kayıp veriler ile çalışmada kolaylık sağlaması amacıyla kayıp veri gösterge matrislerinden yararlanılmaktadır. Tipik bir veri matrisindeki değerlerin hangilerinin gözlenen, hangilerinin kayıp olduklarının belirtilmesinde R gösterge değişkeninden faydalanılır. Burada ifade edilen R değişkeni, kayıplılık (missingness) olarak adlandırılmakta ve araştırma kapsamında ölçülen değişken bilgisinin elde edildiği durumlar için 1 ($r=1$), kayıp olduğu durumlar için 0 ($r=0$) değerini almaktadır (Schafer & Graham, 2002; Graham, 2009; Enders, 2010). Y veri matrisine uygulanan bu basit matematiksel dönüşüm ile Şekil 1.5'deki kayıp veri gösterge matrisi elde edilir:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ r_{31} & r_{32} & \dots & r_{3p} \\ r_{41} & r_{42} & \dots & r_{4p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{np} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 1 \end{bmatrix}$$

Şekil 1.5. Kayıp Veri Gösterge Matrisi

Şekil 1.5'de de görüldüğü gibi tam olmayan veri matrislerinde bulunan kayıp veriler, Little ve Rubin'in (1987) belirttiği dikdörtgensel biçimin bozulmasına sebep olurlar. Bu bilgiden hareketle, dikdörtgen biçimindeki veri setlerini analiz etmek için geliştirilen standart istatistiksel yöntemlerin, kayıp veriye sahip veri setleri için kullanılabilir olmadığı söylenebilir. İstatistiksel işlemler için engel teşkil eden bu önemli durum, başlı başına *kayıp veri* probleminin anlaşılması ve çözümlenmesi gerektiğine işaret eder.

1.5.3.3. Kayıp Veri Problemi

Tipik bir veri setinde, bazı gözlemler için bazı değişken bilgilerine ulaşılamadığı durumlarda kayıp veri problemi ile karşılaşılır (Little & Rubin, 1987; Allison, 2002). Kayıp veri basitçe, elde edilmesi planlanan veri seti ile elde edilen veri seti

arasındaki farklılıkları ifade eder (Enders, 2010). Bu tanımlamada bahsi geçen *elde edilmesi planlanan veri seti*, istatistiksel olarak tam bilgi içeren tam veri setlerini; *elde edilen veri seti* ise, gözlenen değerlerin yanı sıra gözlenemeyen değerleri de içeren tam olmayan veri setlerini ifade eder. Böyle herhangi iki veri seti arasındaki farklılık, araştırmacıları doğrudan kayıp verilere götürür.

Kayıp veri problemi farklı araştırmacılar tarafından farklı şekillerde dile getirilmiştir. Rubin (1976, 1987) uygulamaya dönük araştırmalarda kayıp veri ile karşılaşmanın kaçınılmaz olduğunu; Allison (2002) istatistiksel analiz yapan herkesin er ya da geç kayıp veri problemi ile başa çıkmak durumunda kaldıklarını; Graham (2009) araştırma tarihinin başlangıcından beri araştırmacıların bu problem ile uğraşmak zorunda kaldıklarını; Finch (2010) ise istatistikçiler ve veri analistlerinin sık sık kayıp veri problemi ile yüz yüze geldiklerini belirtmiştir.

Rubin (1987) kayıp veriye bağlı olarak ortaya çıkabilecek en önemli sorunları; bilgi eksikliğinden dolayı veri setinin daralmasına bağlı olarak yapılacak kestirimlerin gücünün azalması, gözlenen ile kayıp veriler arasındaki sıklıkla sistematik olan farklılıktan dolayı olası bir yanlılığın ölçme işlemine karışma riski ve kayıp değerler içeren tam olmayan veri setleri üzerinde, tam veri setlerine göre yapılandırılmış standart istatistiksel analiz yöntemlerinin kullanılamaması olarak belirtmektedir. Peng, Harwell, Liou ve Ehman (2007) ise, kayıp verinin gerçekte neden bir sorun teşkil ettiğini, kayıp veriye bağlı olarak ortaya çıkabilecek dört önemli problem üzerinden açıklamaktadır:

- i. Kayıp veri konusu ile ilgili en ciddi endişe, istatistiksel bir model kullanılarak elde edilecek kestirimlerde, kayıp veriden kaynaklanabilecek yanlılıkların ortaya çıkmasıdır. Örneğin gözlenemeyen yanıtlara sahip olan yanıtlayıcıların yanıtlama profillerinin yanıtları tam olanların profillerinden farklı olması, yanlılığın bir nedeni olabilmektedir. Bu örnekte, örneklemin yanıtlayanlardan oluşan kısmı "seçkisizlik (randomization)" özelliğini taşımamaktadır. Başka bir ifadeyle, tüm soruları yanıtlayan bireylerden oluşan bir örneklemin, asıl örnekleme ve evreni temsil etme düzeyi düşüktür. Bu durumda araştırmacı sadece yanıtlayanlardan oluşan örneklemden elde ettiği verileri kullanarak istatistiksel işlemler yaparsa, sonuçların yanlı olma ihtimali oldukça yüksek olur.

- ii. Kayıp veri, bilgi eksikliğine ve buna bağlı olarak istatistiksel analizlerin gücünün azalmasına sebep olur. Bir veya birkaç değişkene ait bilgisi gözlenemeyen bazı yanıtlayıcıların analizden çıkarılması veya kayıp değer içeren değişken veya değişkenlerin analiz dışı bırakılması, kullanılan istatistiğin gücünün azalması ve daha yüksek standart hata değerlerinin elde edilmesi ile sonuçlanmaktadır. Çünkü bu tip durumlarda, analiz sonuçlarının evrene genellenmesini engelleyecek şekilde, yanıtlayıcıdan yanıtlayıcıya farklılaşan bir yanıtlanma mekanizması ve değişkenden değişkene farklılaşan bir örneklem oluşumu vardır.
- iii. Yaygın olarak kullanılan istatistiksel yöntemlerin, kayıp bilgi içeren veri setleri üzerinde kullanımları uygun değildir. Örneğin, kayıp verili veri setleri üzerinde faktör analizi uygulamalarının gerçekleştirilmesi, analiz deseninde dengesizliğe yol açmaktadır. Bununla birlikte, çok değişkenli istatistiksel yöntemlerin analiz edilmesi için geliştirilmiş olan ticari istatistiksel yazılımlar da tam veri setlerini analiz etmek üzere yapılandırılmışlardır.
- iv. Kayıp veri problemi, daha önemli konularda değerlendirilebilecek kaynakların boşa harcanmasına sebep olurlar. Örneğin, boylamsal araştırmalarda veya ön test-son teste dayalı deneysel araştırmalarda ilk uygulamada bilgileri toplanan bireylere diğer uygulamalarda da ulaşılabilmesi için ekstra zaman, çaba ve para harcanmaktadır. Kaybolan bilgilerin kazanılması ve daha yüksek yanıtlanma oranlarının elde edilebilmesi için harcanan tüm bu kaynaklar, çoğu zaman, harcanılanları amorti bile etmemektedir.

Acock (2012) ise kayıp veriye bağlı olarak ortaya çıkabilecek sorunları farklı bir bakış açısı ile ele almış ve diğer araştırmacılar gibi kayıp verinin yaratacağı sorunlara değil, kayıp verilerin varlığında, araştırmacıların amaçlaması gereken üç temel noktaya odaklanmıştır. Bunlar; (1) istatistiksel analizlerde kullanılan bilginin maksimuma çıkarılması, (2) model parametrelerinin kestirimi sırasında ortaya çıkacak yanlılığın minimuma indirgenmesi ve (3) parametrelerin kestirimi ile ilişkili olan standart hataların kestirimi sırasında ortaya çıkacak yanlılığın minimuma indirgenmesidir.

1.5.3.4. Kayıp Veri Teorisi

1.5.3.4.1. Kayıp veri mekanizması

Rubin'in (1976) kayıp veri teorisindeki anahtar fikir, kayıplılığın olasılık dağılımına sahip bir değişken olmasıdır. Veri matrislerindeki gözlenen ve kayıp değerlerin kolaylıkla görülebilmelerini sağlamak amacıyla R gösterge değişkeni ile ifade edilen kayıplılık, araştırma kapsamında ele alınan herhangi bir değişkene ait ölçme sonuçlarının kayıp olma olasılığını belirtir. İstatistik literatüründe bu olasılığın dağılımı, bazen *yanıtlama mekanizması* bazen de *kayıplılık mekanizması* olarak adlandırılmaktadır ve kayıplılık ile veri seti arasındaki ilişkinin doğası, farklı kayıp veri mekanizmalarının ortaya çıkmasına sebep olmuştur (Schafer & Graham, 2002; Enders, 2010).

Kayıp veri için araştırmacılar tarafından en kabul görmüş sınıflama Little ve Rubin (1987) tarafından yapılmıştır. Bu sınıflandırmada, *tamamen rastgele kayıp-TRK* (*missing completely at random*), *rastgele kayıp-RK* (*missing at random*) ve *rastgele olmayan kayıp-ROK* (*missing not at random*) olmak üzere üç farklı kayıp veri mekanizması tanımlanmıştır.

Rubin (1976) oldukça teknik ifadelerle, TRK türü kayıpta, kayıplılık dağılımının ne araştırma kapsamında ölçülen değişken değerlerine ne de kayıp verinin ortaya çıktığı değişken değerine bağlı olmadığını; RK türü kayıpta, kayıplılık dağılımının başka bir değişken değerine bağlı olduğunu; ROK türü kayıpta ise, kayıplılık dağılımının kayıp verinin ortaya çıktığı değişken değerine bağlı olduğunu belirtmiştir. Little ve Rubin (1987) daha anlaşılır ifadelerle, bir verinin kayıp olma olasılığının araştırma kapsamında ölçülen veya ölçülmeyen bütün özelliklerden bağımsız olma durumunu TRK; bir verinin kayıp olma olasılığının kayıp verinin yer aldığı değişken dışında, araştırma kapsamında ölçülen veya ölçülmeyen diğer bazı özelliklere veya değişkenlere bağlı olma durumunu RK; bir verinin kayıp olma olasılığının kayıp verinin yer aldığı değişkene bağlı olma durumunu ise ROK olarak tanımlamıştır. Finch (2008) ise, TRK türünde, yanıtızlık davranışı ile ilişkili olan herhangi bir sistematik mekanizma bulunmadığını; RK türünde, yanıtızlık davranışı ile maddenin ölçmeyi amaçladığı gizil özelliğin ilişkili olduğunu, ancak o maddeye verilmiş olan gerçek yanıtın ilişkili olmadığını; ROK türünde ise, yanıtızlık davranışı ile maddeye verilmiş olan gerçek yanıtın direk olarak ilişkili olduğunu vurgulamıştır.

Yukarıdaki tanımlamalara göre TRK türü kayıpta, kayıp veriye sahip olan değişken ile ilişkili olan veya kayıp verinin ortaya çıkmasına neden olan herhangi bir değişken bulunmamaktadır. Örneğin, herhangi bir test maddesinin gözden kaçırılması sebebiyle yanıtlanamaması veya sonradan yanıtlanmak üzere bırakılıp geri dönüş yapılmaması durumlarında ortaya çıkar. RK türü kayıpta, kayıp verinin ortaya çıkma nedeni ile kayıp veriye sahip olan değişken ilişkilidir. Bu tür kayıp, herhangi bir test maddesinin cinsiyet, eğitim seviyesi, gelir düzeyi, etnik köken gibi değişkenlerle ilişkili olarak yanlılık içermesi sebebiyle bazı gruplar tarafından kasıtlı olarak yanıtlanmaması durumlarında ortaya çıkar. Örneğin, kadın yanıtlayıcılar erkeklere hitap eden bazı test maddelerini anlayamadıkları için yanıtlanmadıklarında RK türü kayıp meydana gelir. ROK türü kayıpta, kayıp verinin ortaya çıkma nedeni ile kayıp veriye sahip olan değişken aynıdır. Örneğin, herhangi bir test maddesinin açık-anlaşılır ifade edilmemesinden dolayı anlaşılamadığı için yanıt alınamaması veya maddenin çok zor olmasından dolayı yanıtlanamaması durumlarında ortaya çıkar.

Enders (2010) farklı kayıp veri mekanizmalarını, 20 bireye ait zeka puanları (IQ) ve iş performans (IP) ölçümlerini kullandığı bir örnek üzerinde açıklamıştır. Bu örnekte, bireylerin kayıp veri içermeyen performans ölçümlerinden TRK, RK ve ROK mekanizmalarına uygun olacak şekilde veri silme işlemleri yapılmış ve yorumlamalar Tablo 1.5'de verilen tüm bu bilgiler ışığında yapılmıştır.

Tablo 1.5: Farklı Kayıp Veri Mekanizmalarına Göre Oluşturulmuş İş Performans Ölçümleri

Sıra	Zeka puanı (IQ)	İş performans ölçümleri (IP)			
		Tam veri	TRK	RK	ROK
1	78	9	—	—	9
2	84	13	13	—	13
3	84	10	—	—	10
4	85	8	8	—	—
5	87	7	7	—	—
6	91	7	7	7	—
7	92	9	9	9	9
8	94	9	9	9	9
9	94	11	11	11	11
10	96	7	—	7	—
11	99	7	7	7	—
12	105	10	10	10	10
13	105	11	11	11	11
14	106	15	15	15	15
15	108	10	10	10	10
16	112	10	—	10	10
17	113	12	12	12	12
18	115	14	14	14	14
19	118	16	16	16	16
20	134	12	—	12	12

İş performansı ölçümlerinde yer alan kayıp verilere ait mekanizmanın TRK olarak tanımlanabilmesi için, kayıp verilerin hem zeka puanı hem de iş performans ölçümleri değişkeni ile ilişkili olmaması gerekmektedir. Tablo 1.5'in TRK başlıklı dördüncü sütunu incelendiğinde, performans ölçümlerinde ortaya çıkan kayıp verilerin, zeka puanı ve performans ölçümlerinin düşük ya da yüksek olmasıyla ilişkili olmadığı görülmektedir. Buradaki TRK durumu Rubin'in (1976) teorisindeki olasılık terminolojisine göre aşağıdaki gibi ifade edilebilir:

$$P(IP \text{ kayıp veri} | IP, IQ) = P(IP \text{ kayıp veri})$$

Buna göre IP'deki kayıplılık olasılığı ne IP'ye ne de IQ'ya bağlı değilse, IP değişkenindeki kayıp veri mekanizması TRK'dır (Allison, 2002; Schafer & Graham, 2002; Allison, 2003).

İş performansı ölçümlerinde yer alan kayıp verilere ait mekanizmanın RK olarak tanımlanabilmesi için, kayıp verilerin zeka puanı değişkeni ile ilişkili olması gerekmektedir. Tablo 1.5'in RK başlıklı beşinci sütunu incelendiğinde, performans ölçümlerinde ortaya çıkan kayıp verilerin, zeka puanı ile ilişkili olduğu; ancak, iş performansı bakımından yüksek ya da düşük ölçümlere sahip olmakla ilişkili olmadığı görülmektedir. Buna göre, zeka puanı düşük olan katılımcıların iş performansı ölçümlerinde, zeka puanı yüksek olan katılımcılara göre daha fazla kayıp veri meydana gelmiştir. Bununla birlikte, düşük zeka puanına sahip katılımcıların performans ölçümlerinde gözlenen kayıp veriler, aynı katılımcıların iş performansı ölçümlerinin düşük ya da yüksek olmasından bağımsızdır. Buradaki RK durumu Rubin'in (1976) teorisindeki olasılık terminolojisine göre aşağıdaki gibi ifade edilebilir:

$$P(IP \text{ kayıp veri} | IP, IQ) = P(IP \text{ kayıp veri} | IQ)$$

Buna göre IP'deki kayıplılık olasılığı IQ'ya bağlı ancak IP'ye bağlı değilse, IP değişkenindeki kayıp veri mekanizması RK'dır (Allison, 2002; Schafer & Graham, 2002; Allison, 2003).

İş performansı ölçümlerinde yer alan kayıp verilere ait mekanizmanın ROK olarak tanımlanabilmesi için, kayıp verilerin yine performans ölçümleri değişkeni ile ilişkili olması gerekmektedir. Tablo 1.5'in ROK başlıklı altıncı sütunu incelendiğinde, performans ölçümlerinde ortaya çıkan kayıp verilerin, iş performansı değişkenindeki ölçümlerle ilişkili olduğu görülmektedir. Buna göre, iş performansı

ölçümleri düşük olan katılımcıların performans ölçümlerinde, iş performansı ölçümleri yüksek olan katılımcılara göre daha fazla kayıp veri meydana gelmiştir. Buradaki ROK durumu Rubin'in (1976) teorisindeki olasılık terminolojisine göre aşağıdaki gibi ifade edilebilir:

$$P(IP \text{ kayıp veri} | IP, IQ) = P(IP \text{ kayıp veri} | IP)$$

Buna göre IP'deki kayıplılık olasılığı IP'ye bağlı ise, IP değişkenindeki kayıp veri mekanizması ROK'tur (Allison, 2003). Schafer ve Graham (2002), Enders (2010), Doğanay-Erdoğan (2012) ve Toka (2012) ise ROK mekanizmasına sahip kayıp verilerin, değişkenin kendi değerlerinin yanı sıra araştırma kapsamında incelenen diğer değişken değerlerine de bağlı olabileceğini belirtmişlerdir. Bu ifade, düşük performans ölçümlerine sahip katılımcıların performans ölçümlerinde ortaya çıkan kayıp verilerin, aynı zamanda araştırma kapsamında ele alınan zeka puanı değişkeninden de etkilenebileceği anlamını taşır. Tablo 1.5 kontrol edildiğinde, iş performansı değişkenindeki kayıp verilerin, dolaylı olarak zeka puanının düşük ya da yüksek olması ile de ilişkili olarak da meydana geldikleri görülmektedir.

1.5.3.4.2. Kayıp veri mekanizmalarına yönelik ihmal edilebilirlik varsayımı

Kayıp veri mekanizmasının RK olması ve kayıp veri sürecine ait parametreler ile kestirilmek istenen parametreler arasında herhangi bir ilişkinin olmaması durumlarında, kayıp veriler ihmal edilebilir. *İhmal edilebilirlik (ignorability)*, temelde, kestirim sürecinin bir parçası olarak herhangi bir kayıp veri modellenmesine ihtiyaç olmadığı anlamına gelir. *İhmal edilemezlik (nonignorability)* ise ihmal edilebilirliğin zıttı olarak, veri türünün RK olmadığı durumlarda, kayıp veri mekanizmasının göz ardı edilemez olduğu ve daha etkili parametre kestirimleri için kayıp veri mekanizmasının modellenmesine ihtiyaç duyulduğu anlamına gelir (Allison, 2002). Schafer ve Graham'da (2002), RK veri türünü aynı zamanda ihmal edilebilir yanıtızlık (ignorable nonresponse), ROK veri türünü ihmal edilemez yanıtızlık (nonignorable nonresponse) olarak adlandırır.

Araştırmacılar genellikle, herhangi bir değişkene ait bilgisi kayıp olan bireylerin, aynı değişken için bilgileri elde edilen diğer bireylerden herhangi bir farklarının olmadığını kanıtlama eğilimindedirler. Başka bir ifadeyle, araştırmalar sonucunda elde edilen veri setlerinde kayıplar olmasının veya olmamasının, bulgularda anlamlı bir farklılığa yol açmayacağı gösterilmeye çalışılmaktadır. Örneğin, gelir

durumlarını beyan eden ve etmeyen bireyler arasında, araştırma kapsamında incelenen diğer değişkenler açısından anlamlı farklılıklar oluşmadığına yönelik kanıtlar toplanmaya çabalanır. Böylelikle, herhangi bir kuramsal temele dayanmaksızın, çalışılan veri setlerinde bulunan kayıpların *RK türünde veya ihmal edilebilir* olduğu varsayımında bulunulup, bu varsayıma bağlı olarak kayıp veriler analiz dışında bırakılırlar (Allison, 2002).

Benzer şekilde Rubin'de (1976), çoğu araştırmada kayıp verilerin analiz dışı bırakılmasının temelinde, bu verilerin bir takım varsayımlara dayanarak ihmal edilebilir olarak görülmesi görüşünün bulunduğunu belirtmiştir. Örneğin, kayıp değerlerin veri seti içerisinde rastgele bir şekilde ortaya çıkması nedeniyle veri dağılımında bir çarpıklık oluşmayacağı, çok değişkenli normal dağılım varsayımının karşılanması durumunda her bir değişken için kayıp verinin meydana gelme olasılığının eşit olacağı ya da yürütülen araştırmanın bağımlı değişkeninde meydana gelen kayıpların, gözlenen değişken veya değişkenlerle ilgisinin olmadığı gibi varsayımlara bağlı olarak kayıp veriler analiz dışı bırakılabilmektedir. Ancak kayıp verileri analiz dışında bırakma kararı öylesine ve çabucak alınabilecek bir karar değildir. Böylesine önemli bir karar için kayıp veri dağılımına ilişkin rastgeleliğin kontrol edilmesi ve kayıp veri mekanizmalarının doğrulukla tespit edilmesi gereklidir. Allison (2002), en güçlü varsayımlara sahip olan TRK mekanizmasının kanıtlanması sürecinde, bir değişkende meydana gelen kayıp verilerin olasılığının, bu değişkenin kendi değerleri ya da başka değişkenlerin değeriyle ilişkili olmadığını ispatlanması gerektiğini belirtmiştir. Bu ispatın sağlanması durumunda, kayıp veriler dışında kalan ve gözlenen verilerden oluşan tam veri seti, orjinal gözlemler kümesinin basit rastgele bir örnekleme olarak kabul edilebilmektedir.

TRK mekanizmasının kontrolü için Little (1988), veri setindeki her bir değişkene ait ortalamalar arasındaki farklılıkları eşzamanlı olarak değerlendiren ve *t*-testinin çok değişkenli uzantısı olan bir yöntem önermiştir. Little'ın TRK testi (Little's MCAR test) olarak adlandırılan bu yöntem, tek değişkenli *t*-testinin aksine tüm veri setine uygulanmakta ve aynı kayıp veri örüntüsünü paylaşan gözlemlerin alt gruplarının ortalama farklılıklarının karşılaştırılması şeklinde gerçekleştirilmektedir. Little'ın test istatistiği, alt grupların ortalamaları ve genel ortalama arasındaki standartlaştırılmış

farkların ağırlıklı toplamı olarak aşağıdaki şekilde ifade edilmekte ve anlamlı bir d^2 istatistiği TRK aleyhinde kanıt sağlamaktadır (Enders, 2010).

$$d^2 = \sum_{j=1}^J n_j (\hat{\mu}_j - \hat{\mu}_j^{(ML)})^T \hat{\Sigma}_j^{-1} (\hat{\mu}_j - \hat{\mu}_j^{(ML)})$$

Kayıp veri dağılımlarının TRK mekanizmasına uygunluğuna ilişkin istatistiksel kanıtlar sağlanabilmesine rağmen, RK ve ROK mekanizmalarını test etmek için yeterli kanıt elde edilememektedir. RK mekanizması için bir değişkende meydana gelen kayıp verilerin olasılığının, analizdeki diğer değişkenler kontrol altına alındığında, bu değişkenin kendi değerleri ile ilişkisiz olduğunu kanıtlamak gerekir. Başka bir ifadeyle, herhangi bir değişkende kayıp veri oluşma olasılığının, bu değişkenin kendisi dışındaki diğer bazı değişkenler ile ilişkili olması durumunda, söz konusu kayıp verilerin RK mekanizmasına uygun olduğu ileri sürülebilir. Örneğin performans ölçümlerinde, belli bir zeka puanının altında kayıp verilerin meydana gelmesi RK mekanizmasına işaret eder. Ancak mekanizmanın gerçekten de RK olduğundan emin olunması için performans değişkenindeki kayıp verilerin, performans değişkenindeki değerler ile ilişkisiz olduğunun ispatlanması gerekir. Kayıp verilere ait gerçek değerlerin bilinmemesi nedeniyle de bu ilişkisizlik ispatlanamamaktadır. Aynı sebepten dolayı ROK mekanizmasının da manidarlığını test etmek mümkün değildir (Allison, 2002; Enders, 2010). Sonuç olarak kayıplar, RK ve ROK mekanizmalarına uygun dağılımlar sergileseler de, bu durumun kesin olarak ortaya konulabilmesi için ilgili gözlemlere ulaşıp gözlenememiş bilgilerin tamamlanması gerekmektedir. Bilgiler tamamlanmadığı sürece, RK ve ROK mekanizmalarının doğruluğunu kanıtlayacak testler de kullanılamamaktadır (Schafer & Graham, 2002).

RK ve ROK mekanizmalarını kanıtlayacak testler bulunmasa da, TRK ve ROK'un tanımlarına uygun olmayacak şekilde herhangi bir araştırma kapsamında ölçülen diğer değişkenlerle ilişkili olarak ortaya çıkan kayıp veriler, RK mekanizmasının varlığına ve rastgeleliğe işaret ederler. Bu güçlü işaretler, RK mekanizmasına sahip olduğu düşünülen kayıp verilerin ihmal edilebilir olarak kabul edilmesine sebep olurlar. Ancak Finch'e (2008) göre, kayıp verilerin ihmal edilebilir olup olmadığına ilişkin kesin kararın verilebilmesi için kayıp veri mekanizmalarının yanı sıra kayıp verilerin miktarına veya veri seti içerisindeki oranına da bakmak

gereklidir. Büyük bir veri setinde, verinin %5'i veya daha azı rastgele olarak kayıpsa (TRK veya RK) çok ciddi problemlerle karşılaşılmaz ve kayıp veriler ihmal edilebilir. Ancak, küçük veya orta büyüklükteki bir veri setinde çok sayıda ve rastgele olmayan kayıp (ROK) veri mekanizmasında bir kayıp varsa, bu veri seti için uygun kayıp veri başa çıkma yöntemlerinden birine başvurulmalıdır.

Yukarıdaki bilgiler ışığında, kayıp verilerin ihmal edilebilir olup olmaması kararının, kayıp veri mekanizmalarına ve kayıp veri miktarına bağlı olarak değişiklik gösterdiği görülmektedir. Buna göre büyük bir veri setinde rastgele olarak ve az miktarda ortaya çıkan kayıp veriler dışındaki tüm durumlardaki kayıp verilerle uygun yöntemler kullanılarak baş edilmesi gerekir.

1.5.3.4.3. Kayıp veri başa çıkma yöntemleri

Kayıp verilerin hemen her araştırmada karşılaşılan bir problem olması, verilerin analizi aşamasında bazı istatistiksel çözüm yöntemlerinin gelişmesine zemin oluşturmuştur. Allison (2009), araştırma sırasında karşılaşılan kayıp veri sorunu ile başa çıkmada kullanılacak iyi bir yöntemin (i) kayıp verilerden kaynaklanabilecek yanlılığı minimuma indirilmesi, (ii) elverişli bilgilerin kullanımını maksimuma çıkarması ve (iii) standart hata, güven aralığı ve olasılık değerlerine yönelik titiz kestirimler sağlaması olduğunu belirtmiştir. Literatürde kayıp veri ile başa çıkmada kullanılan birçok yöntem ve bu yöntemlere ilişkin farklı sınıflandırmalar yer almakla birlikte, bu çalışmada Finch'in (2008) sınıflandırma biçimi kullanılmıştır. Bu sınıflandırma biçimine göre kayıp veri ile başa çıkma yöntemleri, *değer atama içermeyen yöntemler (nonimputation-based methods)* ve *değer atamaya dayalı yöntemler (imputation-based methods)* olmak üzere iki temel başlık altında toplanmaktadır.

Kayıp verilerle başa çıkmak için geliştirilen yöntemlere ait uygulamalar, bazıları kayıp veri yerine değer atama yöntemlerini de içeren; R, SAS, SPSS ya da BILOG-MG gibi standart istatistik yazılımlarında gerçekleştirilmektedir (Finch, 2008). Bu yazılımlar ile birlikte; S-Plus, STATA, AMOS, Mplus, HLM, WINSTEPS, BIGSTEPS, ConQuest, LOGIST, MULTILOG, PARSCALE ve TESTFACT programlarında da kayıp veriler için basit düzeyde çözüm önerileri bulunmaktadır (du Toit, 2003; Acock, 2005). Ancak Amelia II, IVEware, SOLAS ve R yazılımına

ait Amelia, missForest, Hmisc, mi ve mice paketleri özellikle kayıp veri problemi ile başa çıkmak için geliştirilmiş programlardır.

- **Değer atama içermeyen yöntemler**

Kayıp veri problemi ile başa çıkmada kullanılan değer atama içermeyen yöntemler; *parametre kestirim sürecine dayalı yöntemler (parameter estimation-based methods)* ile *silmeye dayalı yöntemler (deletion-based methods)* olmak üzere iki başlık altında incelenmektedir. Parametre kestirim sürecine dayalı yöntemler; kayıp verileri, madde parametrelerinin kestirim sürecinde çözülebilen bir sorun olarak ele almaktadır. Bu başlık kapsamında ele alınan yöntemlerden herhangi birisi kullanıldığında, kayıp veri problemi ile başa çıkmak için, parametre kestirimi için kullanılan istatistik programından başka bir programın daha kullanılmasına veya ön bir istatistiksel işlem yapılmasına gerek kalmamaktadır.

Bu yöntemlerden ilki, uygulamalar sonucunda kayıp veri ile sonuçlanan maddelere, bireylere hiç uygulanmamış (not-administered) veya hiç sunulmamış (not-presented) gibi davranmaktadır. Bu yöntemin kullanıldığı durumlarda parametre kestirim işlemleri gerçekleştirilirken, testi alan herhangi bir bireye ait kayıp bilgi içeren maddeler -sanki o birey ilgili maddeyi hiç okumamış veya görmemiş de bu sebeple yanıt alma fırsatı bulamamış gibi- işleme alınır ve parametre kestirim sürecine dahil edilmezler. Başka bir ifadeyle, kayıp veriye sahip olan ilgili maddeler istatistiksel süreçlerde görmezden gelinirler (Finch, 2008). Görmezden gelme (ignoring missing data) olarak da adlandırılan bu yöntem, kayıp veri mekanizması hususunda herhangi bir varsayım gerektirmemesi ve kestirim sürecinin sadece gözlenebilen veriler üzerinden yürütülmesi sebepleriyle, neredeyse hiçbir yöntemin işe koşulmaması anlamını taşımaktadır (Delucci, 1994).

Parametre kestirim sürecine dayalı ikinci yöntem, kayıp bilgiye sahip olan maddelere, sanki yanıtlayıcılar tarafından yanlış yanıtlanmışlar (incorrect) gibi davranmaktadır. Başka bir ifadeyle bu yöntem kullanılarak testte yer alan maddelerin puanlanması ve parametrelerinin kestirilmesi sürecinde, kayıp veriye sahip olan maddelerin her biri yanlış yanıtlandırılmış gibi işleme alınır. Parametre kestirim sürecine dayalı olan son yöntem ise, kayıp bilgi içeren madde yanıtlarına aşamalı olarak doğru yanıtlanmışlar (fractionally correct) gibi davranmaktadır. Örneğin beş seçenekli maddelerden oluşan bir çoktan seçmeli testte, yanıtı kayıp olan bir

madde için kaybın yerine gelebilecek beş farklı alternatif yanıt mevcuttur. Bu durumda, kayıp olan yanıt seçenek sayısına bağlı olarak 1/5 oranında aşamalı olarak doğru yanıtlandırılmış gibi puanlandırılabilir (Finch, 2008).

Parametre kestirim sürecine dayanan tüm kayıp veri baş etme yöntemleri, ikili puanlanan madde yanıtları üzerinde kullanılan MTK modellerine göre parametre kestirim işlemleri yapan bir yazılım olan BILOG-MG ile şu adımlar takip edilerek gerçekleştirilir: BILOG-MG kayıp veriye sahip veri setleri üzerinde madde ve yetenek parametre kestirimlerini yapmadan önce, veri setinde 1 ya da 0 puanlama biçimine göre herhangi bir karşılığı bulunmayan kayıp verilere nasıl davranılmasının istendiğini sorar. Bu kısımda araştırmacı bu tip maddeler için seçeneği sunulan *uygulanmamış gibi davranmak (UGD)*, *yanlış yanıtlanmış gibi davranmak (YYGD)* veya *aşamalı olarak doğru yanıtlanmış gibi davranmak (ADYGD)* yöntemlerinden birini seçer ve parametre kestirim işlemleri bu seçime göre yapılır. BILOG-MG, aksi belirtilmediği veya kod dosyasına elle girilmediği sürece, varsayılan yöntem olarak YYGD'yi kullanır (Zimowski, Muraki, Mislevy & Bock, 2003). Sık kullanılan diğer parametre kestirim yazılımları olan MULTILOG ve PARSCALE'de, kayıp veri problemini ele alırken BILOG-MG'dekine benzer bir kodlama süreci kullanıp aynı çözüm önerilerini sunarlar ancak varsayılan yöntem olarak UGD'ye başvururlar (du Toit, 2003). TESTFACT programında ise, BILOG-MG'dekine benzer biçimde, kayıp veri problemi ile başa çıkmada varsayılan yöntem olarak YYGD kullanılır (Wood, Wilson, Gibbons, Schilling, Muraki & Bock, 2003).

Birçok MTK modeline uygun verinin analiz edilmesine imkan sağlayan R yazılımı ltm (Latent Trait Models under IRT) ve TAM (Test Analysis Modules) paket programları da, parametre kestirim süreçlerinde kayıp veri problemini çözmeye yönelik olarak hazırlanmış özel kodlamalara sahiptirler. Bu paket programlar, veri setinde kayıpların bulunması durumunda, veri setinin sadece gözlenen kısmını dikkate alan ve kayıp verileri görmezden gelen UGD'yi varsayılan yöntem olarak kullanırlar (Rizopoulos, 2015; Kiefer, Robitzsch & Wu, 2017). ltm paket programında madde parametre kestirim işlemini gerçekleştiren *tpm* fonksiyonuna ait açıklama kısmında kayıp veri problemi ile başa çıkma süreci şu ifadelerle belirtilir: Madde parametre kestirimi yapılmak istenen veri setinde kayıp verilerin olması durumunda, parametre kestirim sürecine veri setinin sadece gözlenebilir

olan kısmı dahil edilir. Yani model sadece var olan veriyi kullanarak parametre kestirim işlemlerini gerçekleştirir (Rizopoulos, 2015). TAM paket programında da, veri setindeki tüm kayıpların hesaplama işlemlerinden çıkarıldığı ve parametre kestirimlerinin veri setindeki gözlenen kısımdan elde edilen bilgilerle gerçekleştirildiği belirtilir (Kiefer, Robitzsch & Wu, 2017). Başka bir ifadeyle, her iki paket programında da varsayılan yöntem olarak UGD kullanılır.

MTK'ya ait uygulamaların yürütüldüğü tüm bu programlardaki kayıp veri başa çıkma seçeneklerine bakıldığında, kayıp verilerin mekanizmasına bakılmaksızın, varsayılan yöntemler olarak UGD veya YYGD'ye başvurulduğu görülmektedir. Parametre kestirim sürecine dayalı yöntemlerin yanı sıra, değer atama içermeyen yöntemler kapsamında ele alınan silmeye dayalı yöntemler için süreç biraz daha farklı ve katı kurallarla işlemektedir. Silmeye dayalı yöntemler kapsamında sıklıkla başvurulan iki yöntemden ilki, gözlem silme (casewise deletion) veya tam gözlem analizi (complete case analysis) olarak da adlandırılan liste bazında silme-LBS (listwise deletion) yöntemi; ikincisi ise, uygun gözlem analizi (available case analysis) olarak da adlandırılan çiftler bazında silme-ÇBS (pairwise deletion) yöntemidir. Bu yöntemlerin geneli, yeterli büyüklükte örnekleme uygulanan veri setlerinde meydana gelen TRK mekanizmasına sahip kayıp veri probleminin çözümünde başarılı sonuçlar vermektedirler. Ancak kayıp veri mekanizması TRK değilse, kestirimler ciddi oranda yanlılık içerdiği için bu koşullarda silmeye dayalı yöntemlerin kullanımı önerilmemektedir (Allison, 2002; Acock, 2005).

Değer atama içermeyen yöntemler kapsamında kayıp veri içeren maddelere hiç uygulanmamışlar gibi davranıp görmezden gelmek veya yanlış yanıtlanmışlar gibi davranmak, istatistiksel testlerin gücünü azaltır. Benzer şekilde, kayıp veriye sahip yanıtlayıcıları ya tüm çalışmadan ya da bazı analizlerden çıkarmak, kayıp verinin miktarı ölçüsünde potansiyel olarak hem analizlerin gücünü hem de doğruluğunu azaltır (Madow, Olkin & Rubin, 1983; Roth & Switzer, 1995). Bu tip olumsuzluklar, birçok istatistik programında kayıp veriler ile baş etmede varsayılan olarak kullanılan bu yöntemlerin aksine, veri setindeki var olan bilgilerden yola çıkarak kayıp veriler yerine başka değerler koyan değer atamaya dayalı yöntemlerin kullanımını yaygınlaştırmıştır.

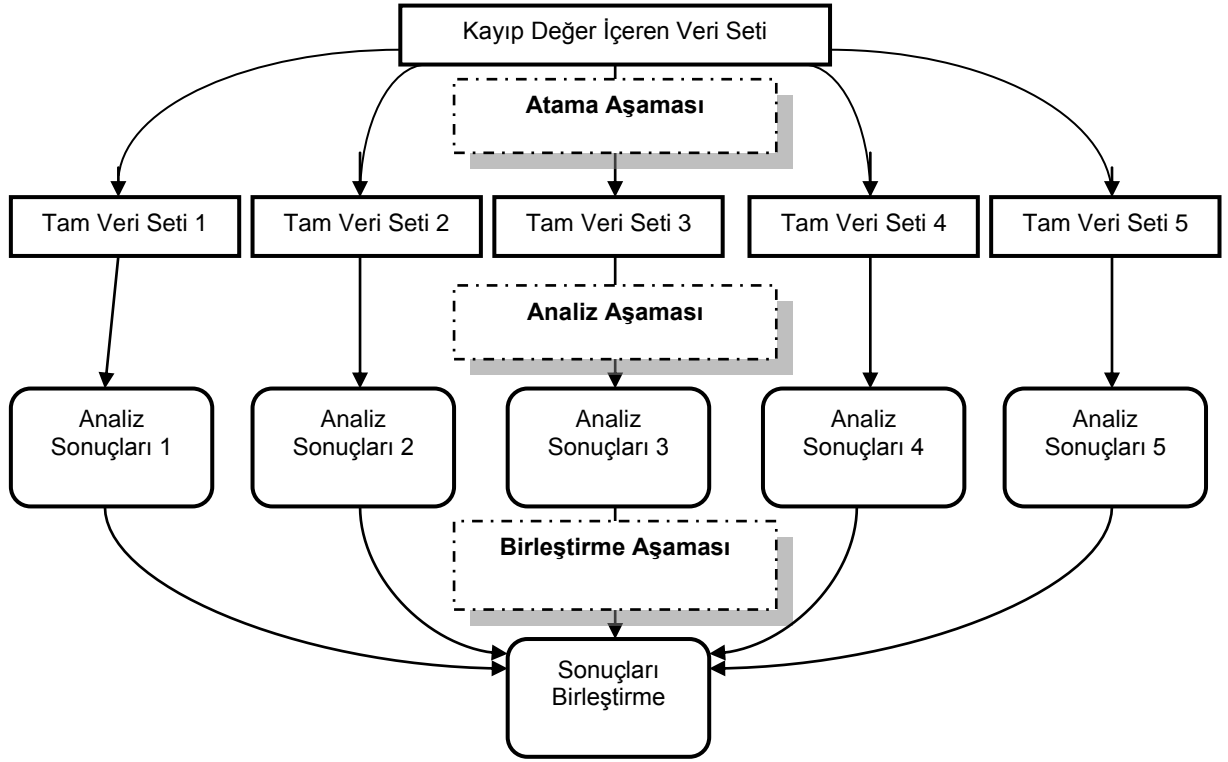
- **Değer atamaya dayalı yöntemler**

Schafer ve Graham (2002) ile Enders (2010), kayıp veriler yerine değer atama işlemlerinin yapılması halinde tam bilgi içeren dikdörtgen biçimli veri setlerine ulaşılabileceğini ve bu durumun özellikle silmeye dayalı yöntemlerin örneklem büyüklüğü üzerindeki olumsuz etkilerine karşı oldukça büyük bir avantaj olduğunu belirtmişlerdir. Değer atamaya dayalı yöntemler; *basit atamaya dayalı yöntemler (single imputation methods)* ile *çoklu değer atamaya dayalı yöntemler (multiple imputation methods)* olmak üzere iki temel başlık altında incelenmektedir.

Basit atamaya dayalı yöntemler kapsamında, çok sayıda yöntem bulunmaktadır. Bunlardan bazıları; ortalama atama-OA (mean substitution), ortanca atama-ORA (median imputation), mod atama-MA (mode imputation), rastgele atama-RA (random imputation), regresyon atama-REA (regression imputation), koşullu ortalama atama-KOA (conditional mean imputation), deste yardımıyla atama-DYA (hot-deck ve cold-deck imputation), en yakın komşu-EYK (nearest neighbor) yöntemi, yanıt fonksiyonu ile atama-YFA (response function imputation), son gözlemi ileri taşıma-SGİT (last observation carried forward), sonraki gözlemi geri taşıma-SGGT (next observation carried backward) olarak adlandırılmaktadır. Bu yöntemler oldukça sık kullanılmalarına rağmen, özel durumlar dışındaki kayıp veriler ile başa çıkmada ideal çözüm önerileri sunmamaktadırlar. Ayrıca yöntemlerin kullanımı, çoğunlukla, pozitif ya da negatif yönde ciddi yanlılıklar, artan Tip II hataları ve olduğundan daha az olarak kestirilen korelasyon değerleri ile sonuçlanmaktadır. Bununla birlikte bu yaklaşımların çoğunun kullanımı, kayıp veri mekanizmasının rastgelelik özelliğini karşıladığı durumlar için geçerlidir (Acock, 2005). Bu ciddi dezavantajlarından dolayı, basit atamaya dayalı yöntemlerin çoğu, evren parametreleri ve standart hata kestirimlerinin doğrulukla elde edilmesi bakımından yetersiz kalmaktadır (Schafer & Graham, 2002).

Basit atamaya dayalı yöntemlerin aksine, her biri standart yollarla analiz edilebilen ve birden fazla atanmış veri seti meydana getiren çoklu değer atamaya dayalı yöntemler, kayıp veriye sahip olan değişken dışındaki diğer değişken bilgilerini de, daha doğru değerler elde edebilme adına atama süreci içerisine dahil ederler (Finch & Margraf, 2008). İlk kez Rubin (1987) tarafından temelleri atılan ÇDA yöntemi, temelde, $m > 1$ sayıda eksiksiz veri seti elde edilecek şekilde atama işleminin yapılması, elde edilen m farklı tamamlanmış veri setinin standart

tekniklerle analiz edilmesi ve ulaşılan sonuçların birleştirilmesinden oluşan üç aşamalı bir süreci ifade eder (Rubin, 1987; Schafer & Graham, 2002). Atama sayısının 5 ($m=5$) olduğu durumda, ÇDA sürecine ilişkin şematik gösterim Şekil 1.6'daki gibidir:



Şekil 1.6. ÇDA İşlem Şeması (Takahashi & Ito, 2013)

ÇDA yöntemi uygulamalarında takip edilen süreç hemen hemen aynı olsa da, değer atama aşamasında farklı hesaplama algoritmaları kullanılmaktadır. Algoritma seçimini etkileyen en önemli faktör, veri setlerinin normal model yaklaşımına uygun olarak çok değişkenli normal-ÇDN (multivariate normality) dağılım varsayımını karşılayıp karşılamaması ile ilgilidir. Değer atama işlemini ÇDN dağılım kullanarak yapan ve bu sebeple sürekli veri setlerindeki kayıp değerlere ait atama işlemlerinin gerçekleştirilmesinde kullanılan algoritmalar *Markov zinciri Monte Carlo-MZMC (Markov chain Monte Carlo)* (Rubin, 1987) ve *önyüklemeli beklenti maksimizasyonu-ÖBM (expectation-maximization with bootstrapping)* (Honaker & King, 2010) olarak adlandırılırlar. Atama yapmak için ÇDN dağılımlar yerine her bir değişkene ait farklı çokterimli (multinomial) dağılımları kullanan ve bu sebeple farklı tip değişkenleri içeren veri setlerindeki kayıp değerlere ait atama işlemlerinin gerçekleştirilmesinde kullanılabilen

algoritma ise *zincirleme denklemler ile çok değişkenli atama-ZDÇDA (multivariate imputation by chained equations)* olarak adlandırılır (Schafer, 1997).

MZMC, R yazılımı Norm paketi ile SAS PROC MI programlarında uygulamaları gerçekleştirilen ve ÇDN dağılım varsayımını karşılayan sürekli değişkenlerin bulunduğu veri setlerindeki kayıp veri probleminin çözümlenmesinde kullanılan bir veri artırma algoritmasıdır. ÖBM, R yazılımı Amelia II paketinde uygulamaları gerçekleştirilen ve beklenti maksimizasyonu-BM (expectation-maximization) ile parametrik olmayan önyükeme-Ö (non-parametric bootstrapping) süreçlerini birleştiren bir hesaplama algoritmasıdır (Takahashi & Ito, 2013).

ZDÇDA algoritması; sürekli (continuous), ikili (binary), çok kategorili (polytomous) ve bunların karışık olduğu farklı tip değişkenlerdeki kayıp verilerle baş edebilmektedir. Algoritmanın bu özelliği, her bir değişkenin kendi tipine uygun olan atama modellerini kullanmasından kaynaklanır. Kayıp veri içeren hedef değişken sürekli olduğunda, doğrusal regresyon model-DRM (linear regression model), tahmine dayalı eşleştirme-TDE (predictive mean matching); ikili olduğunda lojistik regresyon-LR (logistic regression) ve diskriminant fonksiyonu-DF (discriminant function); çok kategorili olduğunda çokterimli lojistik regresyon-ÇLR (multinomial logistic regression) ve oranlı olasılık modeli-OOM (proportional odds model) işe koşulmaktadır (van Buuren & Groothuis-Oudshoorn, 2011; White, Royston & Wood, 2011). ZDÇDA algoritmasının uygulamaları; R yazılımı mice paket programı, IVEware ve SOLAS programlarında mevcuttur (Takahashi & Ito, 2013; Wu, Jia & Enders, 2015).

Kayıp veri içeren değişken ikili olarak puanlandığında, değer atama aşamasında kullanılan yöntemlerden ilki ZDÇDA algoritması kapsamında ele alınan LR'dir (Allison, 2005; Ambler, Omar & Royston, 2007; Hedeker, Mermelstein & Demirtaş, 2007; Van Buuren, 2007; Sulis & Porcu, 2008; Finch, 2010; Smolkowski, Danaher, Seeley, Kosty & Severson, 2010; Ma, Akhtar-Danesh, Dolovich & Thabane, 2011; White, Royston & Wood, 2011; Kropko, Goodrich, Gelman & Hill, 2014; Wu, Jia & Enders, 2015). LR analizinin temel amacı, bağımsız değişkenler ile bağımlı değişken arasındaki ilişkilerin incelenmesi sonucunda kategorik bağımlı değişkene ait değerlerin tahmin edilmesidir. Bu tahmin işlemi aslında iki ya da daha fazla gruba ilişkin "üyelik" kestirimi yapılmasını sağlar (Çokluk, 2010).

LR analizinin kayıp veri problemini çözümlemedeki kullanım işlevi, kayıp veriye sahip olan maddelerin hangi puanlama kategorisine sahip olduğunu kestirmede bir regresyon denklemi oluşturmaktır. Bu denklemde, araştırma kapsamında ele alınan değişkenler bağımsız olarak kabul edilerek bağımsız değişkenlerden edinilen bilgiler ile kayıp verilerin gerçekte hangi kategorinin üyesi olduğu kestirilir. Başka bir ifadeyle, kayıp verileri içeren bağımlı değişkenin kategori sayısına göre farklılaşan LR modelleri işe koşularak, kayıp veriler yerine değer atama işlemleri gerçekleştirilir. LR analizine ait modeller, bağımlı değişkenin ölçüldüğü ölçek türü ve kategori sayısına göre farklılaşmaktadır. Bağımlı değişkenin kategori sayısı $k=2$ olduğunda kullanılan model ikili (binary) LR modeli, kategori sayısı $k \geq 3$ olduğunda kullanılan modeller ise çokterimli (multinomial) LR ve oranlı olasılık (proportional odds) LR modelleridir.

1 (doğru) ya da 0 (yanlış) şeklinde iki kategorili olarak puanlanan maddeler içerisinde bulunan kayıp veriler yerine, değişkenin tipine uygun olarak 1 ya da 0 değerlerinin atanmasında ikili LR modelinden yararlanılır. Model parametreleri, kesişim (intercept) ve kestiricilerin her biri ile ilişkili olan regresyon katsayılarını içerirler ve kestiricileri doğrudan iki kategoriden birine dahil olma olasılığının lojit (β) değerlerine bağlarlar (Wu, Jia & Enders, 2015). Bağlama işlemi sonrasında, kayıp verilerin yerine geçecek 1-0 değerlerinin atanması için veri setindeki gözlenen değerlere ait parametrelerin sonsal dağılımları kullanılır. Bu dağılımlardan türetilen LR eşitliğindeki olasılık kurallarına göre de değer atama işlemleri gerçekleştirilir (Allison, 2005). Değer atama işlemleri için kullanılan ikili ve oranlı olasılık modelleri SAS PROC MI, R yazılımı mice paketi ve IVEware programında; çokterimli model ise R yazılımı mice paket programı ve IVEware'de kullanılmaktadır (Wu, Jia & Enders, 2015).

Kayıp veri içeren değişken iki kategorili olarak puanlandığında, değer atama aşamasında kullanılması önerilen yöntemlerden ikincisi ise yine ZDÇA algoritması kapsamında uygulaması gerçekleştirilen DF'dir (Allison, 2005; van Buuren, 2007; van Buuren & Groothuis-Oudshoorn, 2011, Wu, Jia & Enders, 2015). Diskriminant analizi olarak da adlandırılan DF analizi LR analizine oldukça benzer olarak, grup üyeliklerini yordamaya yönelik bir model kurma amacına hizmet eden çok değişkenli bir istatistiksel tekniktir. Model, gruplar arasında en iyi ayrımı yapmayı sağlayan yordayıcı değişkenlerin doğrusal bir kombinasyonuna dayalı olarak

ortaya çıkan diskriminant fonksiyonlarından oluşur. Bu diskriminant fonksiyonları kullanılarak bireylerin ya da birimlerin sınıflandırılması amaçlanır (Büyüköztürk & Çokluk-Bökeoğlu, 2008).

DF analizinin kayıp veri problemini çözümlemedeki kullanım işlevi, kayıp veri içeren maddelerin hangi puanlama kategorisine sahip olduğunu kestirmede bir diskriminant modeli oluşturmaktır. Burada kullanılan model LR'dekinden farklı olarak, puanlama kategorisi üyelik kestiricilerinin, ÇDN dağılıma uygunluğu varsayılan kategori ortalamaları ile kategoriler boyunca değişmeyen bir kovaryans matrisine sahip olduğu sayılına sahiptir. Kestiricilerin doğrusal kombinasyonlarını türetmek ve grup farklılıklarını maksimuma çıkarmak için model içerisinde kullanılacak diskriminant fonksiyonları, grup ortalamaları ve ortaklanmış kovaryans matrisi kestirimlerinin sonsal dağılımlarından çekilir. Elde edilen diskriminant fonksiyonları ile her bir kategorinin örnekleme oranını ifade eden önsel dağılımlar birlikte kullanılarak, kayıp verilerin yerine geçecek değerlere ait atama işlemleri gerçekleştirilir (Allison, 2005; van Buuren & Groothuis-Oudshoorn, 2011). DF modeli ile ÇDA; SAS PROC MI, R yazılımı mice paketi ve IVEware programlarında kullanılmaktadır (Wu, Jia & Enders, 2015).

ÇDA yöntemi kapsamında hangi hesaplama algoritması ve hangi değer atama modeli kullanılırsa kullanılsın, sonuçların genellenebilirliğini arttırmak amacıyla, birden fazla sayıda ($m > 1$) atama işlemi yapılmaktadır. Böylelikle, her biri standart istatistiksel yöntemler ile analiz edilebilecek m sayıda bağımsız tam veri seti elde edilmektedir. Her bir tam veri seti üzerinde teker teker yürütülen aynı analize ait işlem sonuçları ise Rubin'in (1987) basit kuralına göre birleştirilmektedir. Örneğin kayıp veri içeren bir veri setindeki maddelere ait parametre kestirimleri elde edilmek istendiğinde ilk yapılması gereken veri setindeki kayıp veri problemi ile başa çıkmak, ardından madde parametre kestirimlerini elde etmektir. Kayıp veri problemi ile başa çıkmak üzere ÇDA yönteminin kullanımına karar verildikten sonra uygun bir istatistik programı seçilir. Yöntemin uygulanma sürecine ilişkin gerekli işlem adımları takip edildikten sonra m sayıda ataması yapılmış veri seti elde edilir. Ardından, her bir veri seti üzerinde başlangıçta yapılması amaçlanan madde parametre kestirimleri yapılır. Son aşamada, her bir veri seti üzerinde yürütülen analizlerden elde edilen sonuçların ortalaması alınarak tek bir ortak sonuç elde edilir (Schafer & Olsen, 1998). ÇDA'nın ortaklama (pooling) süreci

olarak adlandırılan tüm bu işlem adımlarının son aşamasında kullanılan matematiksel formül şu şekildedir:

$$\bar{Q} = \frac{1}{m} \sum_{i=1}^m \hat{Q}_i$$

m atama sayısı olmak üzere, $i=1$ 'den m 'e kadar elde edilen ve \hat{Q}_i olarak gösterilen parametre kestirimlerinin ($\hat{Q}_1, \hat{Q}_2, \dots, \hat{Q}_m$) aritmetik ortalaması \bar{Q} olarak ifade edilir (Rubin, 1987).

1.5.4. Test Eşitleme ve Kayıp Veri

MTK modelleri, test maddelerine verilen gözlenebilir yanıtlar aracılığıyla, bireylerin testle ölçülmesi hedeflenen örtük özellik düzeylerine ilişkin kestirimler yapılmasına olanak sağlamaktadır. Buna göre, bireylerin ölçülmek istenen örtük özellik düzeyleri hakkında daha doğru ve etkili kestirime ulaşılabilmesi için veri setindeki bilgilerin gözlenebilir olması gerekmektedir. Bu gereklilik, standart istatistiksel yöntemlerde olduğu gibi MTK modellerinin de tam veri varsayımı altında geliştirilmelerine sebep olmuştur. Böylesi bir varsayımın varlığı, tam veri setlerinin elde edilmesine engel olan kayıp verilerin, hangi test uygulamalarında daha sıklıkla ortaya çıktığının tespit edilmesini gerekli kılar (Zhang & Walker, 2008).

Mislevy ve Wu (1996) MTK'ya dayalı test uygulamalarında sıklıkla karşılaşılan kayıp veri biçimlerini aşağıdaki başlıklar altında toplamıştır:

Alternatif test formları (Alternate test forms). Soruların gizliliğinin ve güvenliğinin korunması amacıyla, benzer içeriğe fakat farklı sorulara sahip iki ya da daha fazla test formuna ait uygulamaların genel adıdır. Bu uygulama biçimi ile test formlarından birini alan yanıtlayıcılar için almadıkları diğer test formu için kayıp veriler meydana gelir.

Hedefe yönelik testler (Targeted testing). Daha etkili ölçme işlemleri yapabilme gerekçesiyle, bireylerin yeteneğiyle ilişkili olduğu düşünülen sınıf düzeyi, ders notu gibi başka değişken bilgilerinin dikkate alınması üzerine testlerin güçlük düzeylerinin ayarlandığı uygulamaların genel adıdır. Bu uygulama biçimi ile yanıtlayıcıların düzeylerine uygun olmadığı gerekçesiyle yöneltilmeyen sorular yanıtız kalır.

Uyarlamalı testler (Adaptive testing). Bireysel farklılıkları dikkate alan ölçme işlemleri yapabilme gerekçesiyle, bireylerin ölçülmek istenen yetenek düzeylerine uygun güçlükteki soruların yöneltildiği uygulamaların genel adıdır. Bu uygulama biçimi ile yanıtlayıcıların düzeylerine uygun olmadığı gerekçesiyle seçim dışı kalan sorular yanıtı bırakılır.

Erişilemeyen maddeler (Not-reached items). Tipik test koşulları altında, zaman yetersizliği sebebiyle testin son kısımlarındaki sorulara erişilemeyen durumlarda ortaya çıkar. Bu tip durumlarda, erişilemeyen sorular için kayıp veriler ortaya çıkar.

Boş bırakılan maddeler (Omitted items). Yanıtlayıcıların soruların tümü üzerine düşünmek ve yanıtlamak için zamanı olduğu halde, dikkatsizlik, yanıtlamak istememe veya doğru yanıtı bilmeme gibi sebeplerle boş yanıtlar bıraktıkları durumlarda ortaya çıkar. Böyle durumlarda, boş yanıtlar kayıp verileri ifade eder.

Yanıtlayıcı seçimi (Examinee choice). "Aşağıdaki dört sorudan herhangi iki tanesini seçip yanıtlayınız" gibi ifadelerin yer aldığı bazı test uygulamalarında, yanıtlayıcılardan belli soruları seçmelerinin istenmesi ile ortaya çıkar. Bu tip durumlarda, yanıtlayıcılar tarafından seçilmeyen sorular için kayıp yanıtlar meydana gelir.

Test uygulamaları sırasında boş bırakılan maddeler, araştırmacılar tarafından başa çıkılması en zor olan kayıp veri biçimini ifade ederler. Bu durum, boş bırakılan madde formundaki kayıpların araştırmacıların kontrolü dışında, çok çeşitli sebeplerle ve çoğunlukla engellenemeyecek şekilde ortaya çıkmasından kaynaklanır. Boş bırakılan maddelerden farklı olarak, erişilemeyen madde biçimindeki kayıpların ortaya çıkma olasılığını azaltmak için testin uygulama süresini uzatma, verilen süre içerisinde yanıtlanabilecek kadar soru sorma gibi birtakım düzenlemeler yapılabilir. Alternatif, hedefe yönelik, uyarlamalı ve yanıtlayıcı seçimine yönelik test uygulamalarında ise, araştırma deseninden kaynaklı olarak araştırmacıların bilgi ve beklentisi dahilinde olan kayıp veri biçimi ile karşılaşılır (Mislevy & Wu, 1996).

Test eşitleme ile aynı anlamına gelen alternatif test formu uygulamalarında, eşitleme deseninden kaynaklanan kayıp veri problemi ile karşılaşılmaktadır. Örneğin en çok kullanılan eşitleme deseni olan ortak test deseninde, farklı yanıtlayıcı gruplarında yer alan bireylerin, her iki test formunda da ortak olan

maddeler dışındaki diğer test maddelerini yanıtlama fırsatı bulamaması, yanıtlanamayan test formundaki maddeler için kayıp verilerin ortaya çıkmasına sebep olmaktadır. Bu sebeple ortaya çıkan kayıp veriler ile başa çıkılması sürecinde, farklı test formlarında bulunan ortak maddeler üzerinden kurulan eşitleme ilişkileri kullanılmaktadır.

Bu süreç şöyle işler: Farklı (P ve Q) gruplardan çekilen ve biri Form X 'i diğeri Form Y 'yi alan iki farklı yanıtlayıcı örneklemini, her iki formda da bulunan ortak maddeleri yanıtlarlar. Form X 'i alan 1.yanıtlayıcı grubundaki bireylerin ortak maddelere verdiği yanıtlar üzerinden Form Y 'nin ortak madde dışındaki diğer maddelerine (Y - ortak maddeler), Form Y 'yi alan 2.yanıtlayıcı grubunun ortak maddelere verdikleri yanıtlar üzerinden ise Form X 'in ortak madde dışındaki diğer maddelerine (X - ortak maddeler) ait kestirimler yapılır. Başka bir ifadeyle, farklı gruplarda yer alan bireylerin almadıkları test formu için kayıp olan yanıtları, ortak maddelere bağlı olarak yürütülen kestirim işlemleri sonucunda doldurulur. Bu işlemin ardından, eşitleme deseninden kaynaklanan kayıp verilerin kestirimlerde ciddi hatalara ve yanlışlıklara sebep olup olmadığının araştırılması için parametre kestirim farklılıklarına dayanan farklı değerlendirme ölçütleri hesaplanır (Liou & Cheng, 1995; Holland, Sinharay, von Davier & Han, 2008; Miyazaki, Hoshino, Mayekawa & Shigemasu, 2009; Sinharay & Holland, 2009; Puhan, 2010).

Araştırma deseni sebebiyle ortaya çıkan kayıp veriler ile eşitlemesi yapılacak test formlarından her biri normal ve kendi başına bir test olarak ele alındığında -kasıtlı olarak veya çeşitli sebeplerle yanıtlanmayan maddeler şeklinde- ortaya çıkan kayıp veriler farklı anlamlar taşırlar (Sinharay & Holland, 2009). Bu nedenle, araştırma deseninden tamamen farklı olarak, her biri kendi başına bir test olan alternatif formların uygulanması sonucunda elde edilen kayıp verili veri setlerinin, eşitleme süreçlerinde kullanılması ile elde edilecek çıktılarının incelenmesine de ihtiyaç vardır. Bunun için test formlarının her biri için normalde yürütülmesi gereken kayıp veri başa çıkma süreci takip edilmeli, ardından bu test formlarından elde edilecek puanların eşitlenmesine geçilmelidir.

Kayıp veri başa çıkma sürecinde kullanılacak başa çıkma yöntemi, eşitlenecek test maddelerine ait puanlama biçimine göre değişiklik gösterir. Eğitim alanında yürütülen eşitleme çalışmalarında sıklıkla kullanılan çoktan seçmeli testler, doğru veya yanlış yanıtı sahip olan maddeler içermesi bakımından ikili olarak

puanlanırlar. İkili puanlanan maddelerin yanıtlanmama davranışına bağlı olarak ortaya çıkan kayıp veri problemi ile başa çıkmada kullanılan en yaygın yöntemlerden ikisi, kayıp verilerin hiç uygulanmamış ve yanlış yanıtlanmış olarak kabul edilmeleridir (Lord, 1974; Culbertson, 2011; Hohensinn & Kubinger, 2011). Kayıp verilerin hiç uygulanmamış gibi ele alındığı ilk yöntemde, veri setinde sanki hiç kayıp veri yokmuş gibi sadece gözlenen bilgiler dikkate alınır. Ancak kayıp verilerin hiç yoklarmış gibi istatistiksel işlemlere dahil edilmemesi, veri seti ile ilgili bilgi eksikliğine bağlı olarak hatalı kestirimlerin üretilmesine yol açar (Delucci, 1994; Finch, 2008).

Öte yandan, doğru veya yanlış yanıtı olan test uygulamalarında kayıp verinin ortaya çıkma sebebi, ilgili maddenin ölçtüğü bilginin veya yeteneğin yokluğu ile ilişkilendirildiği için gözlenemeyen yanıtlar çoğunlukla yanlış yanıtlandırılmış gibi işlem görürler (Ludlow & O'Leary, 1999). Yanıt alınamayan maddelere bu şekilde davranılması, yanıtızlığın sebeplerine ilişkin oldukça güçlü varsayımlarda bulunulması anlamına gelir. Örneğin, yanıtızlık davranışının bilgi eksikliğinden kaynaklandığı durumlarda, sıfır değerlerinin kayıp verilerin yerini alması mantıklı olacaktır. Ancak, zaman yetersizliğinden dolayı ortaya çıkan kayıp verilere, eğer yanıtılsalarmış kesinlikle yanlış yanıtlanacaklarmış gibi davranılması hiç mantıklı değildir (Lord, 1974; Emenogu, Falenchuk & Childs, 2010).

Yanlı ve hatalı kestirimlere yol açtıkları birçok araştırmacı tarafından ortaya konan UGD ve YYGD yöntemleri, sanki en etkili çözümlermiş gibi tam veri varsayımı altında geliştirilen birçok istatistik programı tarafından varsayılan yöntemler olarak kullanılmaktadır. Hatalı kestirimler ürettiği bilinen bu yöntemlerin kullanımının bu denli yaygın olması, ikili puanlanan maddelerdeki kayıp veri problemine yönelik olarak geliştirilmiş daha karmaşık yöntemlerin değerlendirilmesini gerekli kılmaktadır (Culbertson, 2011; Hohensinn & Kubinger, 2011). Özellikle de ikili puanlanan maddelerden oluşan test uygulamalarında, yanıtlayıcıların kasıtlı olarak boş bıraktığı maddelere bağlı olarak oluşan kayıp verilerin ihmal edilebilir olmadığı konusunda uzlaşıldığı için bu tür kayıp verilerin varlığında, elde edilen kestirimlerin manidar düzeyde yanlılık içerme olasılığı yüksektir. Bu olasılığı düşürmede ve ihmal edilebilir olmayan kayıp veriler ile başa çıkmada en yüksek performans gösteren kayıp veri başa çıkma yöntemlerinden biri ÇDA'dır (Allison, 2002; Allison, 2005; Finch & Margraf, 2008; Demir, 2013).

ÇDA, kayıp veri içeren veri setleri üzerinde yapılan istatistiksel işlemlerin güvenilirliğini ve dolayısıyla genellenebilirliğini arttırmak amacıyla birden fazla sayıda değer atama işlemi yapan ve bu sayede standart hataların yansız olarak kestirilebilmesini sağlayan bir kayıp veri başa çıkma yöntemidir (Acock, 2005). Kayıp veri içeren değişken ikili olarak puanlandığında, ÇDA yönteminin değer atama aşamasında kullanılması önerilen yöntemlerden ilki LR, ikincisi ise DF'dir (Allison, 2005). Her iki analiz yöntemi de, grup üyeliklerini yordamaya yönelik modeller kurma amacına hizmet eden çok değişkenli istatistiksel tekniklerdir. LR ve DF'nin kayıp veri sorununu çözümlemedeki kullanım işlevleri, kayıp veri içeren maddelerin hangi puanlama kategorisine (1 ya da 0) sahip olduğunu kestirmede sırasıyla lojistik ve diskriminant modellerini işe koşmaktır.

İkili puanlanan maddelerin bulunduğu test formlarında ortaya çıkan kayıp veriler ile yukarıda bahsedilen yöntemler kullanılarak başa çıkıldıktan sonra kayıp veri problemi çözülen test formlarından elde edilen puanlar eşitlenmeye hazır hale gelir. Ardından, araştırma deseni kaynaklı kayıp veri probleminin çözümünde olduğu gibi, eşitlemesi yapılacak test formlarında ortaya çıkan kayıp veri probleminin çözümlenmesinin eşitleme süreçleri üzerindeki etkileri belirlenebilir.

2. İLGİLİ ARAŞTIRMALAR

Bu bölümde, ilgili literatür kapsamında yürütülmüş olan araştırma özetlerine yer verilmiştir.

2.1. Kayıp Veri Probleminin Test Eşitlemeye Etkisi ile İlgili Çalışmalar

Shin (2009) ortak test deseni kapsamında Rasch modeline dayalı olarak yürüttüğü gerçek puan eşitleme çalışmasında, eşitlemesi yapılan test formlarında bulunan kayıp veriler ile başa çıkmada sıklıkla işe koşulan iki farklı yöntemle ait performans karşılaştırması yapmayı amaçlamıştır. Bu amaç doğrultusunda, kayıp veri başa çıkma yöntemlerinin performansları, eşitlenen test formlarına ait ortak test içerisinde parametre kaymasına sahip olan maddelerin (drift items) doğrulukla tespit edilme düzeyleri açısından değerlendirilmiştir. Bu kapsamda, gerçek bir standart başarı testinin (English Language Arts-ELA test) 2005 ve 2006 yılı uygulamalarına ait madde parametrelerinden türetilen yeni test formları üzerinde bir Monte Carlo simülasyon çalışması yürütülmüştür. 50'şer sorudan oluşan test formlarının 20 ortak maddesi içinden seçilen 5 maddenin b parametrelerinde 1.2'lik logit artışlarla kaymalar yaratılmıştır. Bu maddeler ortak test içerisinde yaratılan drift maddeleri oluşturmaktadır. Bu maddelere verilen yanıtlar üzerinde, sırasıyla yüksek, orta ve düşük yetenek düzeyine sahip bireyler için %7, %10 ve %20'lik kayıplar yaratılmıştır. Bu işlemin ardından, kayıp verilere hiç uygulanmamışlar ve yanlış yanıtlanmışlar gibi davranılarak drift maddelerin doğrulukla tespit edilme düzeyleri değerlendirilmiştir. Değerlendirme ölçütü olarak, kalibrasyon sırasında eski forma ait kestirilen parametrelerin sabitlendiği durumlarda drift maddelerin tespit edilmesinde kullanılan iki istatistik (displacement index, robust-z statistic) işe koşulmuştur. Çalışmanın sonuçları, kayıp veriye sahip maddeleri yanlış olarak kodlamak yerine hiç uygulamamış gibi davranmanın, drift maddelerin tespit edilme doğruluğunu arttırdığını göstermiştir.

2.2. Ortak Test Deseni Kapsamında Kullanılan Farklı Eşitleme Yöntemlerinin Kayıp Veri Varsayımları ile İlgili Araştırmalar

Liou ve Cheng (1995) ortak test deseni kapsamında gerçekleştirilen eşitleme işlemlerinde, iki yanıtlayıcı grubuna iki farklı test formu uygulanmasından dolayı sadece yanıtlanan test formu için "gözlenen" yanıtlanmayan test formu için ise "kayıp" puanların ortaya çıktığını ve bu sebeple araştırma deseni kaynaklı kayıp

veri problemi ile karşı karşıya kalındığını belirtmişlerdir. İstatistiksel veri atama yöntemlerinin kullanılması sayesinde kayıp verilerin uygun kestirimler ile yer değiştirebilmesi ve sonuçta test formlarının sanki her iki yanıtlayıcı grubuna uygulanmış gibi doğrudan eşitlenebilmesi fikrinden hareketle yürüttükleri araştırmalarında, eşit yüzdelikli eşitleme (equipercentile equating) yöntemi işe koşulduğunda kullanılabilir farklı veri atama yöntemlerini yapay veri setleri üzerinde değerlendirmeyi amaçlamışlardır. Bu amaçla, ortak test deseninin kullanımı ile ortaya çıkan kayıp verilere ait mekanizmanın ihmal edilemez olarak kabul edilmesi durumunda işe koşulan geleneksel eşitleme yöntemlerinin daha büyük hata değerleri üretebileceği ve böyle durumlarda kayıp puanların, Bayeşçi önsel ve sonsal dağılımlar ile eş zamanlı olarak kullanılan veri atama yöntemleri ile kestirilebileceği belirtilmiştir. Çalışmanın sonucunda, ortak test desenine dayalı eşit yüzdelikli eşitlemeler için kayıp veri atama yöntemlerinin kullanımı önerilmiştir. Veri atama yöntemlerinin eşlik ettiği bir eşitleme deseninin, daha düşük eşitleme hata (conditional mean squared differences-MSD) değerleri ürettiği ve dolayısıyla daha etkili eşitleme sonuçlarının elde edilmesine sebep olduğu bulgusuna ulaşılmıştır.

Holland, Sinharay, von Davier ve Han (2008) birlikte yürüttükleri çalışmada, eşdeğer olmayan gruplarda ortak test deseni (non-equivalent groups with anchor test design-NEAT) kapsamında en çok kullanılan iki gözlenen puan eşitleme yöntemi olan zincir eşitleme (chain equating-CE) ve sonradan tabakalandırma eşitleme (poststratification equating method-PSE) yöntemlerine ait kayıp veri varsayımlarını KTK kapsamında incelemeyi amaçlamışlardır. Tek grup veya eşdeğer gruplar deseni kullanıldığında veri setindeki gözlenen bilgiler yoluyla doğrudan kestirimler yapılabilirken, NEAT deseni kullanıldığında veri setinin ulaşılabilir olan kısmıyla kestirimler yapılmaktadır. Bu nedenle NEAT deseni kapsamında geliştirilen gözlenen puan eşitleme yönteminin en temel amacı, yeterince güçlü kayıp veri varsayımlarının üretilmesini sağlamaktır. CE ve PSE yöntemleri, gözlenen puanlara ait eşitleme fonksiyonlarını hesaplamak için ortak test tarafından sağlanan bilgiyi kullanan tamamen farklı yolları ifade ettikleri için eşdeğer olmayan iki yanıtlayıcı grubunun ortak test üzerinde önemli ölçüde farklılaştığı test uygulamalarında farklı sonuçlar verme eğilimindedirler. Bununla birlikte her iki yöntem de, NEAT deseninde ortaya çıkan kayıp veriler hakkında

farklı varsayımlara sahiptirler. Bu varsayımlara dayanarak, hangi yöntemin spesifik bir durum için daha doğru sonuçlar üreteceğini söylemek oldukça zordur. Bu çalışmada, başka bir çalışmada NEAT deseni kaynaklı olarak ortaya çıkan kayıp verileri ulaşılabılır olan gerçek bir veri seti kullanılarak, yöntemlerin temelinde yatan kayıp veri varsayımlarının değerlendirilmesine imkan sağlayan kestirimlerin karşılaştırılması sağlanmıştır. Çalışmanın sonuçlarına göre, her iki yöneme ait kestirimler oldukça benzer çıkmıştır, ancak CE'ye ait kestirimler nispeten daha doğrudur.

Miyazaki, Hoshino, Mayekawa ve Shigemasu (2009) yaptıkları çalışmada, MTK'ya dayalı NEAT deseni kapsamında eşitlemesi yapılan test formlarından kestirilen madde parametrelerini bağlamak için Monte Carlo beklenti maksimizasyonu (Monte Carlo expectation maximization) algoritmasını işe koşan yeni bir eşzamanlı kalibrasyon yöntemi önermişlerdir. Böyle bir önerinin geliştirilmesinde tetikleyici olan ana düşünce, halihazırda kullanılan eşzamanlı kalibrasyon yönteminin bireylerin almadıkları test formları üzerindeki puanlarını kayıp veri olarak kabul etmesi ve bu kayıp verileri de görmezden gelmesi düşüncesidir. Ancak bireylerin yanıtlayacakları test formlarını kendi yetenek düzeylerine göre seçmeleri, seçilmeyen test formları için meydana gelen kayıpların bireylerin yetenekleri hakkında bilgi taşıdıkları anlamına gelir ve bu nedenle bu şekilde ortaya çıkan kayıp veriler görmezden gelinemezler. Bu çalışmada yapılandırılan yöntem sayesinde, MTK'ya dayalı eşitlemelerde NEAT deseni kapsamında toplanan verilerde ortaya çıkan kayıp veri mekanizmasını madde yerine test formu boyutunda ele alan yeni bir model ortaya konmuştur. Bu model ile kestirilen parametrelerin daha düşük eşitleme hata ve yanlışlık değerlerine sahip olduğu saptanmıştır.

Sinharay ve Holland (2009) yaptıkları çalışmada, araştırma deseninden kaynaklanan kayıp veri probleminin test eşitlemeye etkisini incelemeyi amaçlamışlardır. Bu amaç doğrultusunda, eşitleme çalışmalarında en çok kullanılan veri toplama deseni olan NEAT deseni kapsamında kullanılabilen üç popüler eşitleme yöntemini (poststratification equating method-PSE, chain equipercentile equating method-CE, IRT-observed score equating method-IRT-
OSE) kayıp veri varsayımları ve bu varsayımlara ait uygulama performansları açısından karşılaştırmışlardır. Çalışmanın gerçekleştirilme gerekçesinin, farklı

eşitleme yöntemlerinin karşılaştırılmasında kullanılan daha doğru sonuçların üretilmesini sağlamak için veri setinde bulunan kayıp verilerin ya da kayıp verilerin dağılımlarının bir şekilde doldurulması gerekliliği düşüncesi olduğunu belirtmişlerdir. Araştırmada, üç farklı gerçek veri seti üzerinde üç farklı eşitleme yönteminin kayıp veri varsayımları kullanılarak araştırma deseninden kaynaklanan kayıp veriler doldurulmuştur. Araştırma sonuçlarına göre, diğer yöntemlere göre daha tatmin edici sonuçlar üretmesi sebebiyle, NEAT deseninin kullanıldığı durumlarda araştırma deseninden kaynaklanan kayıp veri problemi ile başa çıkılmasında "CE" yönteminin kullanımı önerilmiştir. Buna ek olarak, araştırma deseni kaynaklı kayıp veri problemi sebebiyle ortaya çıkabilecek olumsuzlukların önüne geçilebilmesi için çoklu eşitleme yöntemlerini aynı anda işe koşabilen programların kullanılması ve elde edilen sonuçlar arasındaki farklılıkların karşılaştırılması gerekliliği vurgulanmıştır.

Sinharay ve Holland (2009) çalışmalarında, NEAT deseni kapsamında kullanılan doğrusal olmayan üç farklı gözlenen puan eşitleme yöntemine (frequency estimation equipercentile equating-FEEE, chain equipercentile equating-CEE, IRT observed score equating-IRT-OSE) ait kayıp veri varsayımlarını tanımlamayı ve sonrasında her bir yöntemin varsayımları ile ilişkili olarak kayıp verilerin nasıl doldurulabileceğini örnek uygulamalar üzerinden göstermeyi amaçlamışlardır. Eşitleme yöntemlerinin her biri NEAT deseni kullanımı kaynaklı olarak ortaya çıkan kayıp veriler hakkında farklı varsayımlar yaparlar. FEEE yöntemi, her iki yanıtlayıcı grubunda da ortak test puanları ile toplam test puanlarının koşullu dağılımının aynı olduğunu varsayar. CEE yöntemi, her iki yanıtlayıcı grubunda da ortak test ile toplam test puanlarının eşit yüzdelikli eşitleme fonksiyonlarının aynı olduğunu varsayar. IRT-OSE yöntemi ise, ortak testteki ve tüm testteki maddelerin her iki yanıtlayıcı grubu boyunca değişen madde fonksiyonlaşması göstermediğini ve MTK modelinin veri setine yeterli derecede uyum sağladığını varsayar. Bu çalışmada, NEAT deseninin kullanımı ile ortaya çıkan kayıp verilerin doldurulması için üç yol gösterilmiştir. Başka bir çalışmada kullanılan gerçek veri setinden sözde testler (pseudo-tests) oluşturularak yürütülen bu eşitleme çalışmasında, en düşük bias (yanlılık) değerlerini veren en iyi yöntem IRT-OSE, ikinci en iyi yöntem ise CEE çıkmıştır.

Puhan (2010) yürüttüğü eşitleme çalışmasında, eski ve yeni test formlarını alan örneklemelere ait yetenek dağılımlarının benzer ve farklı olduğu durumlarda üç farklı eşitleme yöntemine (chained linear equating method-CLE, Tucker equating method-TE, Levine observed score equating method-LOSE) ait kestirimlerden elde edilen eşitleme hatası (equating error), yanlılık (bias) ve RMSE (root mean squared error) değerlerini karşılaştırmıştır. Sonuçlar incelendiğinde, çalışmada ele alınan birçok koşula göre en düşük yanlılık ve hata değerlerini üreten eşitleme yöntemi "CLE" çıkmıştır. CLE yönteminin ardından en düşük yanlılık ve hata sonuçları veren diğer yöntemin ise LOSE olduğu bulgusuna ulaşılmıştır. TE yöntemi ise her zaman en düşük rastgele eşitleme hatalarını üretmesine rağmen, en yüksek yanlılık ve RMSE değerlerini üretmiştir. Eşitleme yöntemlerine ilişkin bu bulgulara ulaşılmasının ardından, her bir yönteme ait kayıp veri varsayımları test edilerek, neden belirli bir koşul altında bir eşitleme yönteminin daha doğru veya daha hatalı sonuçlar ürettiği deneysel olarak gösterilmiştir.

2.3. Kayıp Veri Probleminin Geçerlik ve Güvenirliğe Etkisi ile İlgili Çalışmalar

Enders (2003), farklı kayıp veri koşulları altında farklı başa çıkma yöntemlerinin performanslarını Cronbach- α kestirimleri temelinde incelemeyi amaçlayan bir çalışma yürütmüştür. Çalışma kapsamında; örneklem büyüklüğü (100, 300, 500), madde sayısı (10, 20), madde kategori sayısı (3, 5, 7) ve kayıp veri oranı (%15, %30) değişkenleri manipüle edilmiştir. Farklı kayıp veri yöntemleri (BM, LBS, ÇBS, OA ve BOA-bireysel ortalamaları atama) temelinde elde edilen Cronbach- α kestirimleri BIAS, RMSE ve güven aralığı (confidence interval coverage) ölçütleri kapsamında ele alınmıştır. Çalışmanın bulguları, BM ve ÇBS yöntemlerinin diğerlerine göre daha yüksek kestirimler üretmesine rağmen, tüm yöntemlerin negatif yanlı kestirimler ürettiğini göstermiştir. TRK mekanizmasına sahip kayıp verilerin bulunduğu koşullar için bazı durumlarda LBS yönteminin negatif yanlı kestirimler ürettiği, BM ile elde edilen sonuçların ise örneklem büyüklüğü ve madde sayısının artmasına paralel olarak daha iyi sonuçlar ürettiği saptanmıştır. ROK mekanizmasına sahip kayıp verilerin bulunduğu koşullar için elde edilen tüm kestirimlerin negatif yanlı olduğu ve en iyi sonuçların BOA yöntemi ile elde edildiği görülmüştür.

Enders (2004) başka bir araştırmasında, TRK ve RK koşulları altında farklı kayıp veri başa çıkma yöntemlerinin Cronbach- α kestirimleri üzerindeki etkisini incelemeyi amaçlamıştır. 1000 farklı yapay veri seti üzerinde gerçekleştirilen araştırmada, farklı kayıp veri yöntemleri (LBS, ÇBS, OA ve BM) temelinde elde edilen Cronbach- α kestirimleri BIAS, RMSE ve güven aralığı çerçevesinde ele alınmıştır. Çalışmada TRK koşulu altında BM, LBS ve ÇBS yöntemlerinin yansız, OA yönteminin ise negatif yanlı güvenilirlik kestirimleri ürettiği; RK koşulu altında ise sadece BM yönteminin yansız, diğer yöntemlerin negatif yanlı kestirimler ürettiği sonucuna varılmıştır. Bunun yanı sıra hem RMSE hem de güven aralığı açısından en iyi yöntemin BM, en kötü yöntemin ise OA olduğu ortaya çıkmıştır.

Çokluk ve Kayrı (2011) tarafından gerçekleştirilen çalışmanın genel amacı, kayıp verilerin olmadığı tam veri koşulu ile farklı oranlarda kayıp verilerin olduğu koşullarda, kayıp veriler yerine yaklaşık değer atamada kullanılan farklı yöntemler sonucu elde edilen faktör yapılarının, düzeltilmiş madde-toplam korelasyonlarının ve Cronbach- α iç tutarlık katsayılarının karşılaştırmalı olarak incelenmesidir. Çalışmada, 10 madde içeren gerçek bir ölçeğin 200 birey üzerinde uygulanması ile elde edilen tam veri seti kullanılmış, verilerin analizinde de açıklayıcı faktör analizi işe koşulmuştur. Kayıp veri içermeyen tam veri setindeki değerler rastgele silinerek, her bir değişkendeki kayıp veri oranının yaklaşık olarak %15 ile %20 ve 0 ile %50 aralığında değiştiği iki farklı veri seti elde edilmiştir. Çalışmanın bulguları, kayıp verilerin olmadığı orijinal veri seti için elde edilen "tek faktörlü" yapının, kayıp verilere farklı yöntemlerle (seri ortalaması, yakın noktaların ortalaması, yakın noktaların medyanı, doğrusal değer kestirimi ve noktanın doğrusal eğimi) yaklaşık değerlerin atandığı koşullar için de "tek faktörlü" olarak elde edildiğini ortaya koymuştur. Bununla birlikte, farklı değer atama yöntemlerinin kullanılması ile elde edilen veri setlerinin açıklanan varyans oranları, özdeğerleri ve Cronbach- α iç tutarlılık katsayıları orijinal veri setine göre daha düşük çıkmıştır.

Demir (2013), kayıp verilerin varlığında ikili puanlanan maddelerden oluşan testlerin psikometrik özelliklerini farklı kayıp veri yöntemlerine göre incelemeyi ve bu tür testlerin psikometrik özelliklerinin kestiriminde uygun olan ve olmayan kayıp veri yöntemlerini tespit etmeyi amaçladığı bir çalışma yürütmüştür. Çalışmada, 2011 yılında 527517 öğrenci üzerinde uygulanan ve 20 maddesinin her birinde %2 ile %34,4 arasında değişen kayıp verilerin bulunduğu tespit edilen SBS matematik

testine ait veriler kullanılmıştır. Kayıp verilerin analizinde dizin silme, 0 atama, seri ortalamaları atama, gözlem birimi ortalaması atama, yakın noktaların ortalamasını atama, yakın noktaların ortancasını atama, doğrusal interpolasyon, doğrusal eğilim noktası, regresyonla atama, beklenti-maksimizasyon algoritması, veri çoğaltma ve ÇDA olmak üzere 12 farklı yöntemin uygulaması yapılmıştır. Bu yöntemlerin işe koşulmasıyla elde edilen yeni veri setlerinin yapı geçerliğinin tespit edilmesi adına açımlayıcı ve doğrulayıcı faktör analizleri gerçekleştirilmiş, bunun yanı sıra madde parametreleri (güçlük, ayırıcılık, çift serili korelasyon, nokta çift serili korelasyon, madde güvenilirliği), test parametreleri (ortalama, standart sapma, standart hata, basıklık, çarpıklık) ve iç tutarlılık katsayıları (Cronbach- α) da kestirilmiştir. Elde edilen bulgular, özellikle ihmal edilebilir olmayan kayıp verilerin varlığında ikili puanlanan maddelerden oluşan testlerin psikometrik özelliklerine yönelik analizlerde, uygun bir kayıp veri başa çıkma yönteminin kullanılmasının gerekli olduğunu göstermiştir. Silmeye dayalı yöntemler ve 0 atama yöntemi, bu tür veriler için uygun değilken, basit atama yöntemlerinin de yanlı kestirimler üretme olasılığı yüksek çıkmıştır. Bunun yanı sıra en uygun yöntemlerin de EÇO ve ÇDA oldukları saptanmıştır.

Akbaş ve Tavşancıl'ın (2015) yürüttükleri çalışmanın genel amacı, farklı örneklem büyüklüğü ve farklı kayıp veri mekanizmalarına sahip olan ölçeklerin psikometrik özelliklerini beş farklı kayıp veri başa çıkma yöntemine göre incelemektir. Bu amaç doğrultusunda, örneklem büyüklüğü ve madde sayısının farklı düzeylerde manipüle edildiği 100 farklı yapay tam veri seti türetilmiştir. Ardından, bu veri setleri üzerinde TRK, RK ve ROK mekanizmasına sahip kayıp verili veri setlerini yaratmak için veri silme işlemleri gerçekleştirilmiştir. Bu veri setleri üzerinde liste bazında silme (LBS), Öklid uzaklığı üzerinden benzer tepki örüntüsüne dayalı atama (BTÖDA), stokastik regresyonla değer atama (SRA), beklenti-maksimizasyon algoritması (BM) ve çoklu değer atama (ÇDA) yöntemleri ile değer atama işlemleri yapılmıştır. Atama işlemleri sonrasında ölçeklerin geçerlik ve güvenilirliğine ilişkin incelemeler; Cronbach- α , McDonald ω ve Ω_w kestirimleri, açıklanan toplam varyans oranları, D^2 istatistiği ve model-veri uyumuna ilişkin indeks değerleri üzerinden gerçekleştirilmiştir. Çalışmanın bulgularına göre, sıklıkla kullanılan LBS yöntemi ciddi sorunlara yol açarken, BM ve ÇDA yöntemleri genel olarak yüksek performans göstermiştir. Bununla birlikte, kesin olarak en iyi

sonucu veren tek bir yöntemin olmadığı saptanmıştır. Hem geçerlik hem de güvenilirlik açısından elde edilen sonuçlar örneklem büyüklüğü ve madde sayısı açısından incelendiğinde, elde edilen kestirimlerin büyük ölçüde birbirine benzer çıktığı; kayıp verinin mekanizması ve oranı açısından yapılan incelemelerde ise, yöntemlerin tamamının ROK mekanizmasına sahip veri setleri içeren koşullarda negatif yanlı kestirimler ürettiği görülmüştür. Bunun yanı sıra, ROK kapsamında elde edilen güvenilirlik kestirimlerinin, kayıp veri oranı ve kayıp veri baş etme yöntemi değişkenlerinin tüm düzeyleri için tam veri setlerine göre daha düşük değerlere sahip olduğu tespit edilmiştir.

Çüm ve Gelbal (2015) yürüttükleri çalışmada, farklı oran ve mekanizmalara sahip kayıp verili veri setleri üzerinde, farklı başa çıkma yöntemlerinin kullanılması ile yapılan yaklaşık değer atamalarının model-veri uyum değerlerini nasıl etkilediğini incelenmişlerdir. 2012 yılında uygulanan PISA testine katılan 15 yaş grubundaki 4848 Türk öğrenci arasından seçilen 1578 öğrencinin puanlarının tam veri setini oluşturduğu çalışmada, tam veri seti üzerinden eksiltmeler yapılarak yaklaşık %20 ve %30 oranlarında TRK mekanizmasına sahip ve yaklaşık %20 oranında TRK mekanizmasına sahip olmayan veri setleri yaratılmıştır. Ardından 10 farklı yaklaşık değer atama yönteminin bu veri setleri üzerinde kullanılması ile elde edilen model veri uyum değerlerine bakılmıştır. Çalışmada, kayıp veriler yerine yaklaşık değer atama yöntemlerinin veri dağılımlarını önemli ölçüde değiştirdiği belirlenmiştir. Ayrıca bilinçsizce yapılan atamaların daha sonra yapılacak analizlerin sonuçlarını olumsuz yönde etkileyebileceği sonucuna da ulaşılmıştır.

2.4. Kayıp Veri Başa Yöntemlerinin Performanslarının Karşılaştırılması ile İlgili Çalışmalar

De Ayala, Plake ve Impara (2001) yaptıkları çalışmada, boş bırakılan madde (omitted response) formunda ortaya çıkan kayıp verilerin, bireylerin yetenek kestirimleri üzerindeki etkilerini incelemeyi amaçlamışlardır. Bu amaç kapsamında çalışmada, boş yanıtların MTK kapsamında kullanılan üç farklı yetenek kestirim yöntemi (biweight estimation-BE, expected a posteriori-EAP, maximum likelihood estimation-MLE) üzerindeki etkileri karşılaştırılmıştır. Bunun yanı sıra, her üç yöntem için de, boş yanıtlara aynı zamanda "yanlış yanıtlanmışlar" gibi ya da "hiç uygulanmamışlar" gibi davranılmıştır. UGD yönteminin özellikle ele alınmak istenmesinin sebebinin, ihmal edilemez olan ROK mekanizmasına sahip kayıp

verilerin varlığında, yetenek kestirimlerinin nasıl etkileneceğini araştırmak olduğu belirtilmiştir. Tüm bu koşulların gerçekleştirilmesi için farklı miktarlardaki kayıp verilerin meydana getirildiği bir Monte Carlo çalışması yürütülmüş ve sonuçlar RMSE ve BIAS cinsinden ifade edilmiştir. Çalışmanın sonucunda, en kötü kestirimler, boş bırakılan maddelere yanlış olarak davranıldığı durumlarda elde edilmiştir. Bu durumun aksine, boş bırakılan maddelerin 0.5 değeri ile değiştirildiği durumlarda EAP kestirim yöntemine ait elde edilen sonuçların, tam veriye sahip maddelerin kullanıldığı durumlarda elde edilen yetenek kestirimleri kadar doğru olduğu gözlenmiştir. Elde edilen başka bir sonuç da, boş yanıtların miktarı arttıkça yetenek kestirimlerine ait hata ve yanlışlık değerlerinin arttığıdır. Bunun yanı sıra, ihmal edilemez türde olan kayıp verileri görmezden gelmek (başka bir ifadeyle uygulanmamış gibi davranmak) tüm yetenek kestirim sonuçları üzerinde olumsuz etkiye sahip çıkmıştır. Yine de UGD yöntemine ait sonuçlar YYGD'ye göre daha az hatalı ve daha az yanlış kestirimler üretmiştir.

Acock (2005), kayıp veri problemiyle başa çıkmak için kullanılan geleneksel yöntemler ile çağdaş yöntemleri ayrıntılı bir şekilde tanıtır, bu yöntemlerin kullanılabileceği istatistiksel paket programlara işaret ettikten sonra, yöntem bazında karşılaştırmalar yapmıştır. Bu kapsamda çalışmada, 2002 yılında 818 birey üzerinde uygulanan sosyal bir araştırmaya ait ölçekten elde edilen tam veri seti kullanılmıştır. Tam veri setine ait değişkenler üzerinde, TRK mekanizmasına sahip %50 oranında kayıp veri içeren bir veri seti ve RK mekanizmasına sahip %40 oranında kayıp veri içeren ikinci bir veri seti oluşturulmuştur. Hem tam veri seti hem de kayıp veri içeren diğer iki veri seti üzerinde farklı kayıp veri yöntemlerinin (LBS, ÇBS, OA, BM, ÇDA) kullanılmasıyla elde edilen tamamlanmış veri setleri üzerinde regresyon kestirimleri ve kestirimlere ilişkin hata oranları elde edilmiştir. Araştırmanın bulguları incelendiğinde, OA'nın hem TRK hem de RK varsayımı altında en kötü yöntem olarak tanımlandığı, dizin silmeye dayalı LBS ve ÇBS yöntemlerinin TRK varsayımı altında kullanılabilir olduğu ve son olarak da çağdaş yöntemlerin geleneksel yöntemlere kıyasla daha iyi sonuçlar ürettiği görülmüştür.

Allison (2005) yaptığı çalışmada, ikili puanlanan veri setlerinde bulunan kayıp veri problemi için ÇDN dağılım varsayımına sahip MZMC algoritmasının kullanımının uygun olmadığını göstermeyi amaçlamakla birlikte, LR ve DF modellerine dayalı

ÇDA yöntemlerini tanıtmayı hedeflemiştir. Bu kapsamda, örneklem büyüklüğü 500 olan ikili puanlanan yapay bir veri seti üzerinde TRK ve RK mekanizmasına sahip kayıp verili veri setleri yaratılmıştır. Bu veri setleri üzerinde; LBS, yuvarlamalı MZMC (MCMC-with rounding), yuvarlamasız MZMC (MCMC-without rounding), LR'ye dayalı ÇDA, DF'ye dayalı ÇDA yöntemlerini uygulamış ve yöntemlerin performanslarını model katsayıları, standart hataların ortalaması, kestirimlerin standart sapması ve güven aralıkları açısından karşılaştırmıştır. Çalışmanın sonuçları; yuvarlamalı MZMC yönteminin asla kullanılmaması gerektiğini ve bu yöntemin hiçbir koşul altında yuvarlamasız MZMC yöntemine üstün olmadığını göstermiştir. Bununla birlikte hem TRK hem de RK mekanizması koşulundaki katsayıların kestirilmesinde kullanılan LR ve DF yöntemlerinin, yuvarlamasız MZMC ve LBS yöntemlerine göre daha üstün sonuçlar ürettiği saptanmıştır.

Ambler, Omar ve Royston (2007) yaptıkları çalışmada, ikili puanlanan veri setlerinde bulunan kayıp veriler ile başa çıkmada kullanılan atama yöntemlerinin etkililiklerini ve atama işlemleri sonrasında elde edilen kestirimlerin güvenilirliklerini değerlendirmiştir. Çalışma kapsamında kullanılan yapay veri setlerini oluşturmak için büyük bir ulusal kalp ameliyatı veritabanı kullanılmıştır. Farklı örneklem büyüklüğü ve farklı kayıp veri oranlarına göre oluşturulan veri setleri üzerinde TRK ve RK mekanizmasına sahip kayıp verileri içerecek şekilde kayıp verili veri setleri yaratılmıştır. Ardından bu veri setleri üzerinde farklı kayıp veri atama yöntemleri (LBS, OA, MA, KOA, hot-deck atama, ZDÇDA'ya dayalı ÇDA) kullanılmıştır. ZDÇDA algoritması ile değer ataması işlemleri için ikili puanlanan maddelerin doğasına uygun LR modeli işe koşulmuştur. Çalışma sonucunda, kayıp verileri görmezden gelmenin (UGD) ve LBS yöntemini işe koşmanın, tedavi yöntemini ve alınacak diğer kararları önemli ölçüde etkileyen yüksek yanlılık ve zayıf kestirim değerleri ürettiği; LR modeline dayalı ZDÇDA yönteminin ise en düşük yanlılık ve en güçlü kestirimleri üreten yöntem olduğu belirlenmiştir. ZDÇDA'dan sonra en iyi sonuçları üreten yöntemin ise KOA olduğu saptanmıştır.

Hedeker, Mermelstein ve Demirtaş (2007) yürüttükleri çalışmalarında, ikili puanlanan maddelerde ihmal edilemez türde olan kayıp verilerin varlığında sıklıkla kullanılan farklı başa çıkma yöntemlerinin (missing=smoking, SGIT, LR'ye dayalı ÇDA) performanslarını karşılaştırmayı ve ÇDA yönteminin sadece ÇDN dağılım varsayımına sahip veri setleri için değil aynı zamanda ikili puanlanan veri setleri

için de kullanılabilceğini göstermeyi amaçlamışlardır. Bu amaçla çalışmada, 489 bireyin bir yıl boyunca sigara kullanma programına katılım durumlarına ilişkin veriler kullanılmıştır. Araştırmada, bireye yönelik tedavinin etkililiğini arttırmak için kullanılan kayıp veri başa çıkma yönteminin etkililiğinin oldukça önemli olması sebebiyle, bu çalışmaya ait sonuçların da önemli olacağına vurgu yapılmıştır. Çalışmanın bulgularında, regresyon kestirimlerine ait en yüksek hata miktarlarının en tutucu yöntemlerden olan missing=smoking yönteminin kullanımıyla; en düşük hata miktarlarının ise LR'ye dayalı ÇDA yöntemine ait olduğu gözlenmiştir.

Van Buuren (2007), kayıp veri sorunu ile başa çıkmada kullanılan ÇDA yönteminin amacının, kayıp veri içeren veri setlerine ait istatistiksel kestirimlerin elde edilmesinde geçerli çıkarımların üretilmesini sağlamak olduğunu belirtmiştir. Bu amaçla da, çok değişkenli veri setlerinde bulunan kayıp veriler yerine değer atamada kullanılan iki farklı algoritma bulunduğunu; bunlardan ilkinin ÇDN dağılım varsayımını karşılayan ve parametrik istatistik teorisine dayalı veri setleri için uygun olan MZMC algoritması, ikincisinin ise her bir değişkene ait koşullu dağılımları kullanan ve yarı-parametrik istatistik teorisine dayalı veri setleri için uygun olan ZDÇDA algoritması olduğunu ifade etmiştir. Bu çalışmada, 3801 Hollandalı kız öğrenciye uygulanan ve farklı kategori sayılarına ($k=2, 5$ ve 6) sahip değişkenleri üzerinde kayıp değer bulunduran ergenlik gelişimi verileri üzerinde hem MZMC hem de ZDÇDA kullanılmıştır. MZMC algoritmasına göre yürütülen değer atama işlemleri sonrasında, kayıp veri içeren değişkenin doğasına uygun olmayacak biçimde atanmış olan değerler en yakın kategori sayısına yuvarlanmıştır. ZDÇDA algoritmasına göre yürütülen değer atama işlemleri sırasında, kayıp veri içeren değişkenlerin her birinin kategori sayısına uygun olan ($k=2$ için LR, $k>2$ için çoklu LR) modeller kullanılmıştır. Çalışmanın bulgularına göre, en yanlı sonuçlar MZMC algoritmasının kullanıldığı koşullarda ortaya çıkmıştır. Buna göre, bu tip kategorik değişkenlerden oluşan veri setlerindeki kayıp değerler yerine çoklu değerlerin atanması aşamasında, MZMC algoritmasına oldukça uygun bir alternatif olan ZDÇDA algoritmasının kullanımı önerilmiştir.

Finch'in (2008) yürüttüğü araştırmanın temel amacı, kayıp veriler ile başa çıkmada kullanılan yöntemlerin MTK çerçevesinde kestirilen madde parametreleri üzerindeki etkilerini belirlemektir. Bunun için araştırmanın referans koşulunu oluşturacak şekilde 3PL modele uygun biçimde türetilen ikili puanlanan veri setleri

üzerinde %5, %15 ve %30 oranlarında RK ve ROK mekanizmasına sahip kayıp veriler içeren veri setleri yaratılmıştır. Ardından, bu veri setlerinin üzerinde farklı kayıp veri başa çıkma yöntemleri kullanılmış ve elde edilen her bir veri seti için madde ayırıcılık, güçlük ve şans parametreleri kestirilmiştir. Elde edilen kestirimler ile referans koşula (tam veri setine) ait kestirimler arasındaki farkın miktarına göre yöntemlerin performansı karşılaştırılmıştır. Çalışmada ele alınan kayıp veri başa çıkma yöntemleri; değer atama içermeyen yöntemler kapsamındaki UGD, YYGD, ADYGD ve değer atamaya dayalı yöntemler kapsamındaki CM-corrected item mean substitution, BM, ÇDA ve YFA yöntemleridir. Araştırmada, RK varsayımı karşılanırsa bile, kayıp veriye sahip ikili puanlanan maddelerin bulunduğu testlerdeki kayıpların YYGD ve UGD olarak kodlanmasının madde parametre kestirimlerinde yüksek yanlılık değerlerine yol açtığı sonucuna ulaşılmıştır. YYGD yöntemi araştırma kapsamında ele alınan diğer başa çıkma yöntemlerine göre, hem ayırıcılık hem de güçlük parametreleri açısından en yanlı sonuçları içeren yöntemdir. Bunun yanı sıra, ÇDA'nın diğer yöntemlere göre en düşük yanlılık içeren yöntem olduğu saptanmıştır. Sonuçlar kayıp veri mekanizması açısından incelendiğinde ise, en yüksek yanlılık değerlerine sahip olan kestirilen madde parametrelerinin, ROK mekanizmasındaki kayıp verilerin bulunduğu veri setlerine ait olduğu tespit edilmiştir.

Sulis ve Porcu (2008), kayıp veri sorunu ile başa çıkmada kullanılan ÇDA yöntemi kapsamında her bir değişkene uygun regresyon modellerini işe koşan ZDÇDA algoritmasını tanıtmayı ve ardından da yapay bir veri seti üzerinde uygulamasını göstermeyi amaçlamışlardır. Bu amaçla, R paket programı aracılığıyla, Monte Carlo yaklaşımı ile türetilen Likert tipi tam bir veri seti üzerinde %5, %10, %15, %20 ve %25 oranlarında TRK ve RK mekanizmasına sahip olan kayıp verili veri setleri yaratılmıştır. Bu işlemin ardından, hem tam hem de kayıp verili setler üzerinde LR modeli ile eğim kestirimleri elde edilmiştir. Çalışmanın bulguları; kayıp veri oranı %15'den az olup LR'ye dayalı ÇDA yöntemi işe koşulduğunda elde edilen kestirimlerin, tam veri setinden elde edilen kestirime oldukça yakın olduğunu ve kayıp veri oranı arttıkça elde edilen kestirimlerin tam veri setine ait kestirimlerden gittikçe uzaklaştığını göstermiştir. Bununla birlikte çalışmada, özellikle TRK mekanizmasına sahip düşük kayıp veri oranlarında daha etkili ve daha yansız kestirimlerin elde edildiği vurgulanmıştır.

Rose, Davier ve Xu (2010) yürüttükleri çalışmada, eğitim araştırmalarında kullanılan geniş ölçekli test uygulamaları sonucunda meydana gelen kayıp verilerin etkilerini incelemişlerdir. Bu kapsamda araştırma kapsamında, kayıp verilerin ortaya çıkış sebepleri araştırılmadan başvuru en sık yöntemler olan UGD ve YYGD'nin her ikisinin de aslında problemlili yaklaşımlar olduğunu göstermeyi amaçlamışlardır. Çalışmanın sonucunda, ihmal edilemez türdeki kayıp verilerin varlığında, hem KTK'ya hem de MTK'ya dayalı test sonuçlarının geçerlik ve güvenilirliğinin tehdit altında olduğu belirlenmiştir. Örneğin, hem yapay hem de gerçek veri setinde (PISA-2006) bulunan ROK mekanizmasına sahip kayıp verilerin, MTK kapsamında yanlış madde ve yetenek parametre kestirimlerine yol açtıkları saptanmıştır. Bu yanlış kestirimler, KTK için de maddelerin doğru yanıtlanma oranları üzerinde gözlenmiştir. Ayrıca, ROK mekanizmasına sahip kayıp verilerin veri seti içerisindeki oranı %30'dan %50'ye yükselince, ciddi oranda yanlış kestirimler kaydedilmiştir.

Dural (2010) yaptığı çalışmada, örneklem büyüklüğü ve kayıp veri mekanizmasına bağlı olarak, kayıp veri ile başa çıkmada kullanılan yöntemlerin çok göstergeli örtük gelişme modelleri üzerindeki etkisini incelemeyi amaçlamıştır. Çalışmada; kayıp veri yöntemi (LBS, ÇBS, REA, BM, ÇDA), kayıp veri mekanizması (TRK, RK, ROK) ve örneklem büyüklüğü (100, 250, 500) olmak üzere üç bağımsız değişkene ait faktörler ele alınmıştır. Bu faktörlerin model üzerindeki etkileri; yakınsama oranı, model uyumu, parametre tahmin yanlışlığı ve standart hata yanlışlığı olmak üzere dört bağımlı ölçüm üzerinden değerlendirilmiştir. Yakınsama oranı, model uyumu ve parametre tahmin yanlışlığına ilişkin bulgular, çok göstergeli örtük gelişme modellerinde ÇBS'nin diğerlerine kıyasla daha iyi bir yöntem olduğunu; standart hata yanlışlığına ilişkin bulgular ise, söz konusu yöntemin parametre tahminleri için yapılan anlamlılık testlerinin gücünü azalttığına ve kullanılırken daha dikkatli olunması gerektiğine işaret etmiştir. Bununla birlikte, REA yönteminin kullanımı hiç önerilmezken; BM ve ÇDA yöntemlerinin standart hata yanlışlıklarında daha iyi performans gösterdikleri, model uyumlarının kabul edilebilir sınırlar içerisinde kaldığı ve örneklem sayısının büyük olduğu durumlarda yakınsama problemlili yol açmadıkları için çok göstergeli örtük gelişme modelleri kapsamında kullanımları önerilmiştir.

Finch (2010), kayıp veri ile başa çıkmak için geliştirilen yöntemlerin sürekli değişkenler içerisindeki kayıp veri problemi üzerine yoğunlaştığını, kategorik değişkenler için uygun olan yöntemlerin hangileri olduğunun ise yeterince açık ve anlaşılır olmadığını ifade etmiştir. Bu nedenle yürüttüğü çalışmada, kategorik değişkenlerden oluşan veri setlerinde kayıpların bulunduğu durumlarda, normal model ve kategorik model yaklaşımlarının performanslarını karşılaştırmıştır. Bu amaçla kullanılan simülatif veri seti, 2000 birey üzerinde uygulanan ve 22 tutum ifadesinden oluşan gerçek bir 5'li Likert tipi doyum ölçeğine aittir. Ölçeğe ait verilerden rastgele olarak farklı örneklem büyüklüklerinde (N=200, 500 ve 1000) veri setleri oluşturulmuştur. R paket programı aracılığıyla, her bir örneklem üzerinde, seçilen beş değişken için hem TRK hem de RK mekanizmasına sahip %25 oranında kayıp veriler yaratılmıştır. Her bir veri seti için 100 tekrar yapılmıştır. Her bir araştırma koşulu için; LR'ye ait eğitim kestirimleri, standart hatalar ve hipotez testi sonuçları kaydedilmiştir. Çalışmanın bulguları, kategorik değişkenlerin kullanıldığı durumlarda kayıp veri problemi ile başa çıkmak için LR modellerine dayanan kategorik model yaklaşımlarını kullanmanın daha uygun olduğunu göstermiştir.

Smolkowski, Danaher, Seeley, Kosty ve Severson (2010), internet-tabanlı olarak yürütülen tütün ve tütün ürünlerini bırakma müdahale programına ait veritabanından elde edilen ikili puanlanan araştırma sonuçlarının içerisindeki kayıp veri problemi ile başa çıkmak için farklı atama yöntemlerinin performanslarını değerlendirmişlerdir. Bu sebeple veritabanından rastgele seçimle oluşturulan 2523 kişilik veri seti üzerinde, çalışma kapsamında ele alınan yöntemler (BM, SGİT, LR'ye dayalı ÇDA) kullanılmıştır. Araştırmada, birbirlerinden oldukça farklı standart hata değerlerinin ve kestirimlerin üretilmesine sebep olan yöntem seçiminin, kullanılacak tedavinin istatistiksel olarak önemliliğini etkileyebileceğine vurgu yapılmıştır. Çalışmanın sonucunda, en hassas kestirimlerin üretilmesine sebep olduğu için LR'ye dayalı ÇDA yönteminin kullanımı önerilmiştir.

Hohensinn ve Kubinger (2011) gerçekleştirdikleri çalışmada, bir testin geliştirilmesi ve kalibrasyonu aşamalarında dikkat edilmesi gereken en kritik konulardan birinin kayıp veri olduğunu ve bu konu ile ilgili yürütülen daha önceki çalışmalarda farklı kayıp veri başa çıkma yöntemlerinin -özellikle yetenek- parametre kestirimlerinin üzerindeki etkilerini incelemeyi amaçladıklarını belirtmişlerdir. Bu çalışmada ise

farklı olarak, ikili (1-0) puanlanan veri setlerindeki kayıp verilerle başa çıkmada en sık kullanılan iki farklı yöntemin (UGD ve YYGD) spesifik madde uyum istatistiklerinin yanı sıra Rasch modeli uyum geçerliği üzerindeki etkilerini incelemek amaçlanmıştır. UGD ve YYGD yöntemlerine ait etkiler hem gerçek bir deneysel hem de yapay bir veri seti üzerinden incelenmiştir. Yapay veriler üzerinden gerçekleştirilen analizlerde üç ayrı veri seti oluşturulmuştur. İlk veri seti hiç kayıp veri içermeyen orijinal veri setini; ikincisi rastgele şekilde veri silme işlemi gerçekleştirilen ve silinen verileri yanlış olarak değerlendirilen veri setini; üçüncüsü ise yine rastgele şekilde veri silme işlemi gerçekleştirilen ve silinen bu verileri uygulanmamış olarak kodlanarak analiz dışı bırakılan veri setini ifade etmiştir. Çalışmanın sonucunda, kayıp verilerin UGD ya da YYGD olarak kodlanmasının, madde ve yetenek parametre kestirimlerinin yanı sıra model uyum geçerliği üzerinde de anlamlı farklılıklar ortaya çıkardığı saptanmıştır. Ancak YYGD'nin UGD'ye göre daha ciddi oranda yanlış sonuçlar içerdiği de vurgulanmıştır.

Ma, Akhtar-Danesh, Dolovich ve Thabane (2011), ikili değişkenlerde bulunan kayıp veri sorununun çözümlenmesinde kullanılan farklı yöntemlerin (LBS, MZMC, ZDÇDA, generalized estimation equation method-GEE, Random-effects logistic regression-RE) performanslarını karşılaştırmayı amaçlamışlardır. Bu amaç doğrultusunda, 1540 bireyin katıldığı bir klinik araştırmasına ait tam veri seti üzerinde; TRK varsayımına sahip %5, %10, %15, %20, %30 ve %50 oranında kayıp veri içeren veri setleri yaratılmıştır. Bu işlemin ardından, her bir veri seti üzerinde gerçekleştirilen 100 tekrar ile farklı kayıp veri başa çıkma yöntemlerine ait uygulamalar yapılmıştır. Çalışmanın bulguları; kestirilen Kappa istatistikleri, olabilirlik oranları (odds ratio) ve güven aralıkları (confidence interval) kapsamında değerlendirilmiştir. Elde edilen sonuçlar; düşük kayıp veri oranlarında yöntemlere ilişkin sonuçların benzerlik göstermekle birlikte, kayıp veri oranı yükseldikçe MZMC'nin en düşük kestirimleri üretme eğiliminde olan yöntem olduğunu göstermiştir.

White, Royston ve Wood (2011), "Zincirleme Denklemler kullanarak Çoklu Değer Atama (ZDÇDA): Uygulama için bir Kılavuz" başlığı altında yürüttükleri araştırmada; ZDÇDA'nın teorik altyapısına ve kayıp veri içeren nicel değişkenler üzerinde nasıl uygulanması gerektiğine değinerek, aslında ne kadar esnek ve pratik bir yaklaşım olduğunu ortaya koymayı amaçlamışlardır. ZDÇDA'nın farklı tip

değişkenler içeren veri setlerindeki kayıp veri problemi ile başa çıkmada kullanılabilir olmasından hareketle, bu çalışmanın kapsamında hem sürekli hem de farklı kategori sayılarına sahip kategorik değişkenleri içeren karışık bir veri seti kullanılmıştır. Veri seti, Birleşik Devletler sınırları içerisinde akıl sağlığı ile ilgili yürütülen bir araştırmaya ait verilerden çekilen 708 kişinin sekiz farklı değişkene ait değerlerini içermektedir. Stata programı aracılığıyla gerçekleştirilen kayıp verilerin analizinde; sürekli değişkenler için doğrusal LR modeli, iki kategorili değişkenler için ikili LR modeli, sıralı olmayan kategorik (unordered categorical) diğer adıyla nominal değişkenler için multinomial LR modeli, sıralı kategorik (ordered categorical) değişkenler için ise ikili LR modelinin uzantısı olan proportional odds LR modeli kullanılmıştır. Araştırmanın sonunda, ZDÇDA'nın oldukça esnek ve iyi sonuç veren bir yaklaşım olmasına karşın oldukça dikkatli kullanılması gerektiğine vurgu yapılmıştır.

Acock (2012), kayıp verinin ne anlama geldiğini, hangi sebeplerle ortaya çıktığını, hangi tip kayıp veri için hangi çözüm yöntemine başvurulması gerektiğini detaylıca açıklayan oldukça önemli bir çalışma yürütmüştür. Geleneksel yöntemlerin yanı sıra özellikle ÇDA yöntemi üzerinde yoğunlaşan araştırmacı, örnek bir veri seti üzerinde farklı yöntemlerin performanslarını karşılaştırdığı bir süreç yürütmüştür. Kayıp verili veri setlerini yaratmak için kullandığı veri seti, daha önce başka bir çalışma için toplanan ve üzerinde LBS yöntemi işe koşulduğu için hiç kayıp veri içermeyen 249 çocuğa ait farklı tip değişkenleri içeren bir veri setidir. Çalışmada, karşılaştırılan tüm yöntemler açısından, ÇDA için elde edilen model kestirimlerinin tam veri setine ait kestirime oldukça yakın olduğu bulgusuna ulaşılmış ve ÇDA yönteminin kullanımı önerilmiştir.

Doğanay Erdoğan'ın (2012) yürüttüğü çalışmanın amacı, kayıp veriler için ÇDA'ya dayalı farklı yöntemlerin kullanımının Rasch modeli kapsamındaki birey ve madde parametre kestirimleri üzerindeki etkisini incelemektir. Bu amaçla yapılan simülasyon çalışmasında; Rasch modeli kapsamında türetilen tam verilerden TRK, RK ile ROK mekanizmalarında ve %10, %30 ile %50 oranlarında veri silinerek kayıp verili yeni veri setleri oluşturulmuştur. Bu işlemin ardından elde edilen veri setleri üzerinde; YFA, CMI-corrected mean imputation, MZMC ve ZDÇDA kullanılarak tamamlanmış veri setleri elde edilmiştir. Çalışmanın sonucunda, Rasch modeli kapsamındaki kayıp verilerin varlığında en iyi değer atama

performansına sahip yöntemin YFA olduğu ve bu yöntemin Likert tipi ölçekler için kayıp veri oranının %50'den az olduğu ve standart hataların daha doğru kestirilmek istendiği durumlarda kullanılabileceği belirtilmiştir.

Kropko, Goodrich, Gelman ve Hill (2014), kayıp veri içeren kategorik değişkenler için kullanılan MZMC ile ZDÇDA yaklaşımlarını karşılaştırmayı amaçlamışlardır. Her iki yaklaşımda da, kayıp veriler yerine ataması yapılmış olan değerlerin ve model katsayılarının doğruluğu değerlendirilmiştir. Bu kapsamda; içerisinde sürekli, ikili, sıralı (ordinal) ve sıralı olmayan kategorik (unordered categorical) değişkenler bulunan 2008 yılı A.B.D ulusal seçim çalışmasında kullanılan bir veri setinden RK mekanizmasına sahip %25 oranında kayıp veri içeren bir veri seti oluşturulmuştur. Çalışmanın bulgularına göre, farklı tip değişkenlerde bulunan kayıp veriler ile başa çıkmada farklı regresyon modellerini işe koşan ZDÇDA'nın sürekli değişkenler için bile MZMC'den daha doğru sonuçlar ürettiği gözlenmiştir.

Wu, Jia ve Enders (2015), farklı kategori sayılarına ($k=2, 3, 5$ ve 7) sahip sıralı değişkenlerde bulunan kayıp veriler için kullanılan farklı değer atama yöntemlerini karşılaştırmayı amaçlamışlardır. Bu amaçla, toplamda 14 değişken içeren ve yapay olarak türetilen veri seti; farklı örneklem büyüklüğü ($N=125$ ve 500), farklı kayıp veri oranları (%30 ve %50) ve RK mekanizmasına sahip olacak şekilde oluşturulmuştur. Kategorik model yaklaşımına uygun olarak kayıp verilerin bulunduğu iki kategorili değişkenler için ikili LR ve DF modelleri, çok kategorili değişkenler için multinomial ve proportional odds LR modelleri; normal model yaklaşımına uygun olarak ise kayıp verilerin bulunduğu değişkenlerin tümü için ataması yapılan değerlerin en yakın kategoriye yuvarlandığı ve yuvarlanmadığı MZMC algoritması kullanılmıştır. Araştırma kapsamında kullanılan tüm yöntemler; güvenilirlik katsayıları, ortalama ölçek puanları ve regresyon katsayıları çerçevesinde değerlendirilmiştir. Araştırmanın sonuçları, örneklem büyüklüğü ve kayıp veri oranlarına bakılmaksızın, normal model yaklaşımının hem ikili hem de çoklu değişkenler üzerinde iyi işlediğini; DF'ye dayalı yaklaşımın ikili puanlanan değişkenler için en iyi olduğunu; yuvarlamalı MZMC algoritması veya LR'ye dayalı modelin kullanımının ise potansiyel olarak önemli ölçüde yanlılık üretmesi açısından sıralı değişkenlere sahip veri setleri üzerinde önerilmediğini göstermiştir.

2.5. İlgili Araştırmalar Özet

Herhangi bir ölçme aracının işe koşulduğu özellikle sosyal bilimler alanında yürütülen araştırmalarda, kayıp veri problemi ile karşılaşmanın neredeyse kaçınılmaz olduğu birçok araştırmacı tarafından defalarca dile getirilmiştir (Little & Rubin, 1987; Schafer, 1997; Allison, 2002; Finch, 2010; Acock, 2012). İlgili literatür kapsamında daha önce yürütülmüş olan ve yukarıda özetlerine yer verilmiş her bir çalışma da bu durumun en somut kanıtlarını teşkil etmektedir. Çalışmalar genel hatlarıyla incelendiğinde, kayıp veri probleminin geçerlik ve güvenirlik çerçevesinde ele alınması ve farklı kayıp veri başa çıkma yöntemlerinin performanslarının karşılaştırılması üzerine yoğunlaşıldığı gözlenmiştir.

Kayıp veri problemini geçerlik ve güvenirlik çerçevesinde ele alan araştırmalar; farklı özelliklere sahip biçimlerde ortaya çıkan kayıp verilerin, ölçme işlemlerinin güvenirlik ve geçerliği, araştırma sonuçlarının iç ve dış geçerliği, sonuçların genellenebilirliği üzerindeki etkilerini incelemişler ve çoğunlukla olumsuz çıktılar kaydetmişlerdir (Brown, 1994; Marsh, 1998; Enders & Bandalos, 2001; Bal, 2003; Enders, 2003, 2004; Van Ginkel, 2007; Bernaards & Sijtsma, 1999, 2000; McKnight, Mcknight, Sidani & Figueredo, 2007; Chen, Wang & Chen, 2011; Çokluk & Kayrı, 2011; Demir & Parlak, 2012; Demir, 2013; Akbaş & Tavşancıl, 2015; Şahin-Kürşad & Nartgün, 2015). Yürütülen ölçme işlemlerinin geçerlik ve güvenirliğinin azalmasının ise; parametre kestirimlerinde yanlılık, şişirilmiş standart hatalar ve işe koşulan istatistiksel testin gücünün düşmesi gibi oldukça ciddi sorunları da beraberinde getirdiği vurgulanmıştır (de Leeuw, Hox & Huisman, 2003).

Kayıp veri yöntemlerinin performanslarının karşılaştırıldığı araştırmalar; kayıp veri problemi ile karşılaşıldığında sıklıkla başvuru ve kullanımı nispeten kolay olan geleneksel yaklaşımlarla, bilgisayar tabanlı istatistik programlarının hızla gelişmesi sonucunda yaygınlaşan ve kullanımı bilgi ile uzmanlık gerektiren çağdaş yaklaşımların etkililiklerini değerlendirmişlerdir (Delucchi, 1994; De Ayala, Plake & Impara, 2001; Acock, 2005; Allison, 2005; Ambler, Omar & Royston, 2007; Hedeker, Mermelstein & Demirtaş, 2007; Van Buuren, 2007; Finch, 2008; Sulis & Porcu, 2008; Rose, Davier & Xu, 2010; Dural, 2010; Finch, 2010; Smolkowski, Danaher, Seeley, Kosty & Severson, 2010; Hohensinn & Kubinger, 2011; Ma, Akhtar-Danesh, Dolovich & Thabane, 2011; White, Royston & Wood, 2011; Acock,

2012; Dođanay Erdoğan, 2012; Kropko, Goodrich, Gelman & Hill, 2014; Wu, Jia & Enders, 2015).

Genellikle simülatif veri setleri kullanılarak yürütölen bu alıřmalar ile yapılan deđerlendirmeler sonucunda; kayıp verilerin yanlış olarak kodlanmasının dođru bir özüm řekli olmadığı, özellikle de hız testlerinde erişilemeyen maddeler olarak ortaya ıkan kayıpların yanlış veya uygulanmamış olarak kodlanmasının ciddi problemlere sebep olduđu belirtilmiştir (Lord, 1974). Bununla birlikte, bazı alıřmalar UGD'nin YYGD'den iyi (De Ayala, Plake & Impara, 2001; Finch, 2008; Hohensinn & Kubinger, 2011), bazıları YYGD'nin UGD'den iyi (Shin, 2009) sonuçlar ürettiđini; Rose, Davier ve Xu'nun, (2010) alıřması ise her ikisinin de sıklıkla başvurulmasına rađmen iyi yöntemler olmadıklarını göstermiştir.

Bu yöntemlerin dođru kestirimler ve düşük standart hatalar üretme açısından yetersiz bulunmaları sebebiyle, Rubin (1987) tarafından temelleri atılan DA yönteminin ve uygulamalarının tanıtıldıđı oldukça açıklayıcı kaynaklar mevcuttur (Schafer, 1997; Allison, 2002; Schafer & Graham, 2002; Enders, 2010; White, Royston & Wood, 2011; Acock, 2012). Diđer başa ıkma yöntemlerinin aksine DA yöntemi, istenilen istatistiksel analizleri gerçekleřtirmek üzere m sayıda ataması yapılmış veri seti meydana getirerek, bir evreni örneklemenin dođasında var olan belirsizlik için bir açıklama getirmektedir. Aynı zamanda, kayıp veriler yerine daha dođru deđerlerin atanmasını sağlamak için kayıp içermeyen diđer deđerşkenlerden elde ettiđi bilgileri de deđer atama sürecine dahil etmektedir (Finch, 2010). Bu açıklamalar ile tutarlı bir biçimde, geleneksel yöntemlerle DA yöntemini karşılařtırıp da DA'nın üstün olmadığını kanıtlayan herhangi bir alıřmaya rastlanmamıştır.

İkili olarak puanlanan deđerşkenler içerisinde kayıp veri problemi ile karşı karşıya kalındıđında, oklu deđer atama aşamasında kategorik model yaklaşımına uygun hesaplama algoritmalarını işe kořan modellerin kullanılması birçok arařtırmacı tarafından önerilmiştir (Allison, 2005; Ambler, Omar & Royston, 2007; Hedeker, Mermelstein & Demirtaş, 2007; Van Buuren, 2007; Sulis & Porcu, 2008; Finch, 2010; Smolkowski, Danaher, Seeley, Kosty & Severson, 2010; Ma, Akhtar-Danesh, Dolovich & Thabane, 2011; White, Royston & Wood, 2011; Kropko, Goodrich, Gelman & Hill, 2014). Özellikle LR (Allison, 2005) ve DF modellerinin

kullanımını öneren çalışmalar (Allison, 2005; Wu, Jia & Enders, 2015) ışığında bu çalışmada da bu iki modele başvurulmuştur.

Araştırmanın iki temel konusunu oluşturan "kayıp veri" ile "test eşitleme"yi ayrı ayrı ele alan oldukça fazla sayıda çalışma olmasına rağmen, kayıp veri probleminin test eşitleme üzerindeki olası etkilerini incelemeyi amaçlayan çok az çalışmaya rastlanmıştır (Shin, 2009). Her iki konuyu, ortak test deseni kapsamında kullanılan farklı eşitleme yöntemlerine ait kayıp veri varsayımlarının ele alınması açısından birleştiren diğer çalışmalar ise, bu araştırmada ele alınan kayıp veri biçiminden farklı bir kayıp veri biçiminin etkilerini incemeleyi amaçlamışlardır (Liou & Cheng, 1995; Holland, Sinharay, von Davier & Han, 2008; Miyazaki, Hoshino, Mayekawa & Shigemasu, 2009; Sinharay & Holland, 2009; Puhan, 2010). Bu araştırmanın, boş bırakılan madde kaynaklı olarak ortaya çıkan kayıp veri ile test eşitleme konusunu doğrudan ilişkilendirmesi ve bunu yaparken de farklı simülasyon koşullarının performanslarını değerlendirip en hatasız ve en yansız sonuçları üreten kayıp veri başa çıkma yöntemini belirleme çabası açısından ilgili literatüre katkı sağlayacağı düşünülmüştür.

3. YÖNTEM

Bu bölümde araştırmanın türü, eşitleme deseni, simülasyon faktörleri ve koşulları, verilerin türetilmesi, eşitleme süreci, verilerin analizi ve değerlendirme ölçütleri başlıklarına yer verilmiştir.

3.1. Araştırmanın Türü

Araştırmada, kayıp veri probleminin test eşitleme üzerindeki etkisini incelemek amaçlanmıştır. Bu amaçla, tam veri setleri türetilmiş ve bu veri setlerinin eşitlenmesi ile tüm araştırma koşulları için referans (karşılaştırma) olan eşitleme sonuçları elde edilmiştir. İşlemin ardından, tam veri setleri üzerinde araştırma koşullarına göre manipüle edilen farklı veri silme algoritmalarının kullanılmasıyla, kayıp veriye sahip maddelerin bulunduğu veri setleri meydana getirilmiş ve bu veri setleri üzerinde farklı kayıp veri başa çıkma yöntemlerinin uygulanması ile elde edilen veri setleri üzerinde eşitleme işlemleri gerçekleştirilmiştir. Böylelikle referans koşul için elde edilen hata ve yanlılık değerlerine en yakın olan ve en düşük değerleri veren araştırma koşulunun ve kayıp veri başa çıkma yönteminin tespiti planlanmıştır. Bu plan doğrultusunda bu araştırma, var olan yöntemlerin simülatif veri setleri üzerinde yeni durumlarda uygulanmasına yönelik bir çalışma olması nedeniyle betimsel, belirli koşullar açısından farklılaşan gruplardan elde edilen bilgilerin birbirleriyle karşılaştırılarak incelenmesi açısından da ilişkisel araştırma niteliği taşımaktadır (Karasar, 2009).

3.2. Eşitleme Deseni

Araştırmanın veri toplama deseni olarak, ortak test deseni kullanılmıştır. Bu eşitleme deseninin kullanılabilmesi için ortak maddeler içeren iki farklı test formuna ve bu test formlarını alan iki farklı yanıtlayıcı grubuna ihtiyaç vardır.

Angoff'un (1971) ortak madde sayısının tüm testteki madde sayısının en az %20'si olması gerektiğini belirten önerisi üzerine, son 20 maddeleri ortak olmak üzere her biri 100'er maddeden oluşan Form X ve Form Y adlı iki test formu oluşturulmuştur. Bu araştırmada Form X eşitlenecek (yeni) test formunu, Form Y eski (base) test formunu ifade etmektedir. Form X'i alan yanıtlayıcı grubu Grup 1, Form Y'yi alan yanıtlayıcı grubu Grup 2 ile temsil edilmektedir ve her iki grup da standart normal

dağılım gösteren 2000'er bireyden meydana gelmektedir. Araştırmada kullanılan eşitleme deseni ile ilgili verilen bu bilgiler Tablo 3.1'de özetlenmiştir.

Tablo 3.1: Araştırmanın Eşitleme Deseni

Örnekleme	Form X (Yeni/Eşitlenecek Test Formu)		
	Form X'in Ortak Maddeleri İçermeyen Kısmı	Ortak Maddeler	Form Y'nin Ortak Maddeleri İçermeyen Kısmı
Grup 1 (2000 birey)	✓ (80 madde)	✓ (20 madde)	
Grup 2 (2000 birey)		✓ (20 madde)	✓ (80 madde)

Form Y (Eski/Base Test Formu)

Not. Test maddelerinin ve bireylerin türetilmesine ilişkin ayrıntılı bilgiler "3.4. Verilerin Türetilmesi" başlığı altında yer almaktadır.

3.3. Simülasyon Faktörleri ve Koşulları

Araştırmada ele alınan simülasyon faktörleri; kayıp verinin yeri, mekanizması, oranı ve başa çıkma yöntemleridir. Bu faktörlere ait koşullara ilişkin bilgiler Tablo 3.2'de verilmiştir.

Tablo 3.2: Araştırmanın Simülasyon Faktörleri ve Koşulları

Kayıp Veri Faktörleri	Koşullar	Koşul Sayısı
Yer	ET, HİT, OT	3
Mekanizma	TRK, RK, ROK	3
Oran	%10, %20, %40	3
Başta çıkma yöntemleri	UGD, YYGD, LRÇDA, DFÇDA	4
Toplam koşul sayısı		108+1*=109

* Araştırmanın referans koşulunu temsil eden ve hiç kayıp veriye sahip olmayan bu tam veri koşulunda (S1), kayıplılık durumu söz konusu olmadığından yer veya mekanizma tanımlanmamış ve bu nedenle herhangi bir başta çıkma yöntemi de kullanılmamıştır. Bu sebeple S1 koşulu, toplam koşul sayısına +1 olarak eklenmiştir.

** ET: eşitlenecek test, HİT: her iki test, OT: ortak test, TRK: tamamen rastgele kayıp, RK: rastgele kayıp, ROK: rastgele olmayan kayıp, UGD: uygulanmamış gibi davranma, YYGD: yanlış yanıtlanmış gibi davranma, LRÇDA: lojistik regresyona dayalı çoklu değer atama, DFÇDA: diskriminant fonksiyonuna dayalı çoklu değer atama.

Kayıp veri faktörlerine ilişkin daha ayrıntılı ifadelerle aşağıda yer verilmiştir.

a. Yer: Araştırma kapsamında ele alınan yer faktörüne ilişkin koşullardan ET ifadesi sadece Form X'i, HİT ifadesi hem Form X hem de Form Y'yi, OT ifadesi ise hem Form X'in hem de Form Y'nin ortak maddelerini ifade etmektedir. Bu ifadeler ışığında, eşitlemesi yapılan test formlarında bulunan kayıp veriye sahip maddeler; ilk koşulda sadece eşitlenecek test-ET formu (Form X) içerisinde, ikinci koşulda her iki test-HİT formu (hem Form X hem de Form Y) içerisinde, üçüncü koşulda ise hem Form X hem de Form Y'nin sadece ortak test-OT formlarının içerisinde olacak şekilde yerleştirilmiştir. Başka bir ifadeyle, ET koşulundaki kayıplar ortak maddeleri

de içeren 100 maddelik Form X, HİT koşulundaki kayıplar ortak maddeleri de içeren 100'er maddelik Form X ve Form Y, OT koşulundaki kayıplar ise sadece her iki test formunun da 20'şer maddelik ortak maddeleri içerisinde yaratılmıştır.

b. Mekanizma: Eşitlemesi yapılan test formlarında bulunan kayıp veriye sahip maddeler; ilk koşulda TRK mekanizması, ikinci koşulda RK mekanizması, üçüncü koşulda ise ROK mekanizmasına sahip olacak şekilde ayarlanmıştır.

c. Oran: Eşitlemesi yapılan test formlarındaki kayıp veriye sahip maddelerin oranı, kayıplılığın yaratılmak istendiği yere (ET, HİT veya OT) bağlı olarak, *maddexbirey* sayısının %10, %20 ve %40'ını oluşturacak şekilde düzenlenmiştir. Bu değerler, kayıp veri oranının %1 ile %40 arasında değiştiği daha önceki çalışmalara göre belirlenmiştir (Ake, 2005; Enders, 2004; Sitjsma & van der Ark, 2003; İnan, 2009; Akbaş, 2014). %10 kayıp veri oranı düşük, %20 kayıp veri oranı orta, %40 kayıp veri oranı ise kayıp verilerin yüksek miktarda olduğu durumu temsil etmektedir. %40 oranının üst sınır olarak alınmasının temel sebebi, kayıp verilerin sayısı fazla olduğunda, elde edilen eşitleme hatası ve yanlılığının gözle görülür ölçüde artırtmayacağına bakmaktır.

d. Başa çıkma yöntemleri: Eşitlemesi yapılan test formlarındaki kayıp veri problemini çözmek üzere farklı başa çıkma yöntemlerine başvurulmuştur. Bu yöntemlerden ilk ikisi araştırmacılar tarafından sıklıkla tercih edilen ve çoğu istatistik programında varsayılan yöntemler olarak kullanılan UGD ve YYGD (Finch, 2008), sonraki ikisi ise alternatif çözüm yolları deneyen araştırmacılar tarafından tercih edilen ve son yıllarda birçok istatistik programında uygulama menüsüne ulaşılabilmesi sebebiyle kullanımı yaygınlaşan ÇDA'ya dayalı LRÇDA ve DFÇDA (Allison, 2005) yöntemleridir.

Aşağıda araştırma kapsamında ele alınan 4 farklı faktöre ait 108 koşulun simülasyon numaralarına Tablo 3.3'te yer verilmiştir.

Tablo 3.3: Araştırma Kapsamında Ele Alınan Simülasyon Koşulları

<i>Kayıp Veri Faktörleri</i>						
<i>Mekanizma</i>	<i>Oran</i>	<i>Başa çıkma yöntemleri</i>	<i>ET</i>	<i>Yer HİT</i>	<i>OT</i>	
TRK	10%	UGD	S2	S3	S4	
		YYGD	S29	S30	S31	
		LRÇDA	S56	S57	S58	
	20%	DFÇDA	S83	S84	S85	
		UGD	S5	S6	S7	
		YYGD	S32	S33	S34	
	40%	LRÇDA	S59	S60	S61	
		DFÇDA	S86	S87	S88	
		UGD	S8	S9	S10	
	RK	10%	YYGD	S35	S36	S37
			LRÇDA	S62	S63	S64
			DFÇDA	S89	S90	S91
20%		UGD	S11	S12	S13	
		YYGD	S38	S39	S40	
		LRÇDA	S65	S66	S67	
40%		DFÇDA	S92	S93	S94	
		UGD	S14	S15	S16	
		YYGD	S41	S42	S43	
ROK		10%	LRÇDA	S68	S69	S70
			DFÇDA	S95	S96	S97
			UGD	S17	S18	S19
	20%	YYGD	S44	S45	S46	
		LRÇDA	S71	S72	S73	
		DFÇDA	S98	S99	S100	
	40%	UGD	S20	S21	S22	
		YYGD	S47	S48	S49	
		LRÇDA	S74	S75	S76	
	ROK	10%	DFÇDA	S101	S102	S103
			UGD	S23	S24	S25
			YYGD	S50	S51	S52
20%		LRÇDA	S77	S78	S79	
		DFÇDA	S104	S105	S106	
		UGD	S26	S27	S28	
40%		YYGD	S53	S54	S55	
		LRÇDA	S80	S81	S82	
		DFÇDA	S107	S108	S109	

* S: Simülasyon

** S1: hiç kayıp verinin olmadığı tam veri setlerinde eşitleme yapılan referans koşul.

Tablo 3.3 incelendiğinde, UGD yöntemine ait koşulların S2-S28, YYGD yöntemine ait koşulların S29-S55, LRÇDA yöntemine ait koşulların S56-S82, DFÇDA yöntemine ait koşulların ise S83-S109 aralığında olduğu görülmektedir.

3.4. Verilerin Türetilmesi

Bu araştırmanın verilerinin türetilmesi üç aşamada gerçekleştirilmiştir. İlk aşamada eşitlemesi yapılacak her iki test formuna ait tam veri setleri türetilmiş, ikinci aşamada farklı veri silme algoritmaları hazırlanarak bir önceki aşamada türetilen tam veri setlerinden kayıp veriye sahip veri setleri elde edilmiş, üçüncü aşamada ise farklı kayıp veri başa çıkma yöntemlerinin kullanılması ile bir önceki aşamada

elde edilen kayıp veriye sahip veri setleri yeni veri setlerine dönüştürülmüştür. Yapılan tüm bu işlemler aşağıda ayrıntıları ile açıklanmaktadır.

3.4.1. Tam Veri Setlerinin Türetilmesi

a. Madde ve yetenek parametrelerinin belirlenmesi: Tam veri setlerinin türetilmesi için test formlarındaki maddelerin parametre değerlerinin ve bu test formlarını alan bireylerin yetenek parametre değerlerinin belirlenmiş olması gerekir. Bu gerekçeyle öncelikle hangi parametre değerlerinin veya değer aralıklarının araştırma konusu kapsamındaki gerçek uygulamaları daha iyi yansıtacağına karar verilmiş, sonrasında ise alınan kararlara uygun biçimde parametre türetme işlemleri gerçekleştirilmiştir.

Bu araştırmada eşitlemesi yapılan her iki test formuna ait parametreler, güçlük (b) parametresi dışında aynı değer aralıkları içerisinde olacak şekilde belirlenmiştir. Formlar arasındaki güçlük farklılığı ise, eşitlemenin amacı olan test formları arasındaki güçlük farklılıklarını dengeleme işlevine hizmet etmesi amacıyla oluşturulmuştur (Gök, 2012). Bu doğrultuda, 20 tanesi ortak olan 100'er maddelik her iki test formundaki ayırıcılık parametreleri (a_i) 0.5 ile 2, güçlük parametreleri Form X için (b_i) -2.5 ile 2.5, Form Y için (b_i) -2 ile 3, şans parametreleri (c_i) ise 0.2 ile 0.3 değerleri arasında olacak şekilde türetilmiştir. Ortak maddelerin ayırıcılık ve şans parametrelerinin türetilmesinde kullanılan değerler Form X ve Form Y'nin ortak maddeleri haricindeki maddeler için kullanılan değerler ile aynı bırakılmış, güçlük parametrelerinin türetilmesinde ise orta güçlükteki maddelere karşılık gelen -1 ile 1 değerleri kullanılmıştır. Test formlarını alan bireylerin yetenek dağılımları ise standart normal dağılım ($\theta \sim N(0, 1)$) gösterecek şekilde örneklenmiştir.

b. Madde yanıtlarının türetilmesi: Test formlarına ait maddelerin ve bu formları alan bireylerin yetenek parametre değerleri türetildikten sonra R yazılımı *irtos* (Partchev, 2016) paket programı kullanılarak, MTK'nın 3 PL modeline uygun biçimdeki iki kategorili (1-0) madde yanıtları türetilmiştir. Başka bir ifadeyle eşitleme işleminde kullanılmak üzere iki farklı test formu türetilmiştir. Form X kendine özgü 80 madde ve Form Y ile ortak olan 20 maddeden oluşmaktadır. Form Y ise kendine özgü 80 madde ve Form X ile ortak olan 20 maddeden oluşmaktadır. Dolayısıyla Form X'e özgü maddeler, Form Y'ye özgü maddeler ve

ortak maddeler olmak üzere üç farklı veri setine ait türetme işlemleri ayrı ayrı gerçekleştirilmiştir.

Sonuçların genellenebilirliğini arttırmak amacıyla, bu aşamada türetilen her iki test formuna ait 2000x100'lük tam veri seti için 50 tekrar yapılmıştır. Böylelikle eşitlemesi yapılan her iki test formu için de sınırları belirlenen madde ve yetenek parametreleri doğrultusunda 50'şer tam veri seti elde edilmiştir. Bu amaçla yazılan R fonksiyonu örneği EK-3'te verilmiştir.

c. Tam veri setlerine ait parametrelerin kestirilmesi: Her iki test formuna ait tüm tekrarlar için R yazılımı ltm (Rizopoulos, 2015) paket programı kullanılarak madde ve yetenek parametre kestirimleri yapılmıştır. Böylelikle araştırmanın referans parametrelerinin elde edilmesi sağlanmıştır.

Bu araştırmada kullanılan tüm veri setlerine ait parametrelerin kestirilmesinde, ltm tarafından madde parametrelerinin kestirimi aşamasında varsayılan olarak kullanılan Marjinal En Çok Olabilirlik (Marginal Maximum Likelihood), yetenek parametrelerinin kestirimi için ise Beklenen Sonsal Dağılım (Expected a Posteriori) yöntemleri kullanılmıştır (Rizopoulos, 2015).

3.4.2. Veri Silme Algoritmaları aracılığıyla Kayıp Veriye Sahip Maddeler İçeren Veri Setlerinin Elde Edilmesi

Verilerin türetilmesinin ikinci aşamasında, tam veri içeren 50 tane test formunun her biri üzerinde, yer \times mekanizma \times oran faktörlerine ait koşulları karşılayan kayıp verili veri setlerinin elde edilmesi amaçlanmıştır. Bu doğrultuda De Ayala, Plake ve Impara (2001), Enders (2004), İnan (2009), Doğanay Erdoğan (2012) ve Toka (2012) tarafından yapılan simülasyon çalışmalarında, farklı kayıp veri mekanizmalarına sahip veri setlerini oluşturmak üzere kullanılan veri silme işlem adımları baz alınmış ve R yazılımı aracılığıyla üç farklı kayıp veri mekanizması için üç farklı veri silme algoritması hazırlanmıştır. Bu işlemin ardından her bir algoritmadaki yer (ET, HİT, OT) ve oran (%10, %20, %40) faktörlerine ait koşulların ifadeleri değiştirilerek tüm araştırma koşullarının oluşturulması sağlanmıştır.

Örneğin sadece ET içerisinde %10 oranında TRK mekanizmasına sahip kayıp veri bulunduran bir veri seti yaratılmak istendiğinde, TRK teorisine uygun olan algoritma seçilip, sadece Form X'e ait maddexbirey sayısının (100x2000=200000)

%10'una denk gelen miktarda (20000) verinin silme işlemi gerçekleştirilir. Aynı mekanizmaya ve aynı orana sahip kayıp HİT içerisinde yaratılmak istendiğinde, veri silme işlemine Form Y'de dahil edilir. OT içerisinde kayıp veri oluşturulmak istendiğinde ise dikkat edilmesi gereken önemli bir nokta bulunur. Örneğin OT içerisinde yer alan %20 oranında RK mekanizmasına sahip bir veri seti elde edilmek istendiğinde RK teorisine uygun olan algoritma seçilip, Form X ve Form Y'nin ortak maddelerine ait maddexbirey sayısının ($20 \times 2000 = 40000$ 'er) %20'sine denk gelen miktarda (8000'er) verinin silme işlemi gerçekleştirilir. Ancak bu silme işleminin hemen ardından kayıp veriye sahip ortak test maddeleri (20 madde) ile Form X ve Form Y'nin geri kalan maddelerinin (80'er madde) birleştirilmesi gerekir. Böylelikle sadece ortak maddeleri içerisinde kayıp veri bulunduran yeni test formları elde edilmiş olur.

TRK mekanizmasına sahip kayıp verilerin elde edilmesinin amaçlandığı ilk algoritma sürecinde, her iki test formu için türetilen tam veri setleri içerisinde hangi değerlerin silineceğinin tamamen rastgele olarak belirlendiği bir silme işlemi yapılmıştır. Buna göre kayıp verilerin ortaya çıkmasının, araştırma kapsamındaki tüm değişken değerlerinden bağımsız olması sağlanmıştır. Algoritmanın bu mantığa göre yürütülmesi, TRK mekanizmasına sahip kayıp verilerin araştırma kapsamında ölçülen veya ölçülmeyen bütün değişken değerlerinden bağımsız olarak ortaya çıkmasından kaynaklanır.

TRK mekanizmasına sahip kayıp verilerin elde edilmesinin amaçlandığı algoritma sürecinde, test formlarına ait tam veri setlerinin her bir tekrarı üzerinde aşağıdaki işlem adımları takip edilmiştir:

- Veri seti içerisinde silinecek verilerin yerine göre, kayıp veri oranlarına karşılık gelen miktarlar belirlenmiştir. Örneğin, Form X için türetilen tam veri seti içinde %10 oranında kayıp veri yaratılmak istendiğinde öncelikle 2000×100 'lük madde yanıt matrisinde silinmesi gereken hücre sayısına karar verilmiştir. Ardından bu işlemin sonucuna göre 20000'e karşılık gelen sayıdaki hücre kadar bilginin silinmesi için bir sonraki adıma geçilmiştir.
- Her defasında birbirinden farklı ve rastgele olacak şekilde, satır için 1 ile 2000 ve sütun için 1 ile 100 arasında rakam üretilmesini sağlayan kodlamalar, belirli satır ve sütun numaralarının seçilmesine sebep olmuştur.

Seçilen numaraların kesiştikleri hücrede bulunan değerler silinmiş ve aynı işleme yaratılması istenen kayıp veri miktarına ulaşıncaya kadar devam edilmiştir. Bu amaçla yazılan R fonksiyonu örneği EK-4'te verilmiştir.

RK mekanizmasına sahip kayıp verilerin elde edilmesinin amaçlandığı algoritma sürecinde, her iki test formu için türetilen tam veri setleri içerisinde hangi değerlerin silineceğinin bireylerin yetenek düzeylerine göre belirlendiği bir silme işlemi yapılmıştır. Buna göre bireylerin yetenek düzeyleri azaldıkça en çok, bireylerin yetenek düzeyleri arttıkça en az miktarda kayıp verinin ortaya çıkması sağlanmıştır. Algoritmanın bu mantığa göre yürütülmesi, RK mekanizmasına sahip kayıp verilerin araştırma kapsamında incelenen değişkenlerin değerlerine bağlı olarak ortaya çıkmasından kaynaklanır.

RK mekanizmasına sahip kayıp verilerin elde edilmesinin amaçlandığı algoritma sürecinde, test formlarına ait tam veri setlerinin her bir tekrarı üzerinde aşağıdaki işlem adımları takip edilmiştir:

- Veri setinin en sonuna her bir bireyin yetenek düzeyini gösteren değerler (θ) 101.sütun olarak eklenmiştir. Ardından bu sütundaki değerler küçükten büyüğe doğru sıralanmıştır. Böylelikle bireylerin matris içindeki yerleri de (sıra numaraları da) değişmiştir. Daha sonra bireyler dört gruba ayrılmıştır. Veri matrisinin en üstünde yer alan 500 birey en düşük yetenek düzeyine sahip yanıtlayıcı grubunu, sonra gelen 500 birey düşük yetenek düzeyine sahip yanıtlayıcı grubunu, hemen altındaki 500 birey yüksek yetenek düzeyine sahip yanıtlayıcı grubunu, en altta yer alan 500 birey ise en yüksek yetenek düzeyine sahip yanıtlayıcı grubunu temsil eder.
- Yanıtlayıcı grupları oluşturulduktan sonra en düşük yetenek düzeyine sahip 500 bireyin madde yanıtları içerisinde rastgele seçilen %40, düşük yetenek düzeyine sahip 500 bireyin madde yanıtları içerisinde rastgele seçilen %30, yüksek yetenek düzeyine sahip 500 bireyin madde yanıtları içerisinde rastgele seçilen %20 ve en yüksek yetenek düzeyine sahip 500 bireyin madde yanıtları içerisinde rastgele seçilen %10 oranına karşılık gelen sayıdaki veri silinmiştir. Böylece düşük yetenek düzeyindeki bireylerin maddelere verdikleri yanıtlar içerisinde toplam %70, yüksek yetenek düzeyindeki bireylerin maddelere verdikleri yanıtlar içerisinde toplam %30

oranında hücre bilgisi silinmiştir. Yapılan tüm işlemler RK mekanizmasının tanımına uygun olacak şekilde, kayıp verilerin araştırma kapsamında incelenen başka bir değişkene bağlanmasını ve bu değişkendeki değerlere göre yaratılmasını sağlamıştır. Bu amaçla yazılan R fonksiyonu örneği EK-5'te verilmiştir. Tablo 3.4 ise RK mekanizmasına sahip kayıp verilerin elde edilmesinde kullanılan kayıp veri oranlarını gösterecek şekilde oluşturulmuştur.

Tablo 3.4: RK Mekanizmasına Sahip Kayıp Verilerin Elde Edilmesinde Kullanılan Kayıp Veri Oranları

Yanıtlayıcı no	Madde no						Yetenek düzeyi (θ)
	1	2	100	
1							min.
2							
...							
...							
500							
501							
502							
...							
...							
1000							
1001							
1002							
...							
...							
1500							
1501							
...							
...							
2000							max.





ROK mekanizmasına sahip kayıp verilerin elde edilmesinin amaçlandığı son algoritma sürecinde, her iki test formu için türetilen tam veri setleri içerisinde hangi değerlerin silineceğinin RK koşulunda olduğu gibi yine düşük yetenek grubunda daha fazla olacak şekilde, ancak bu sefer sadece yetenek düzeyi değişkenine göre değil, kayıp verinin yer aldığı değişkenin madde güçlük değerlerine göre de belirlendiği bir silme işlemi yapılmıştır. Buna göre maddeler zorlaştıkça ve bireylerin yetenek düzeyleri azaldıkça en çok, maddeler kolaylaştıkça ve bireylerin yetenek düzeyleri arttıkça en az miktarda kayıp verinin ortaya çıkması sağlanmıştır. Algoritmanın bu mantığa göre yürütülmesi, ROK mekanizmasına sahip kayıpların, kayıp verilerin meydana geldiği değişkenin kendi değerlerinin yanı sıra, araştırma kapsamında incelenen diğer değişkenlerin değerlerine de bağlı olarak ortaya çıkabilmesinden kaynaklanır.

ROK mekanizmasına sahip kayıp verilerin elde edilmesinin amaçlandığı algoritma sürecinde, test formlarına ait tam veri setlerinin her bir tekrarı üzerinde aşağıdaki işlem adımları takip edilmiştir:

- RK koşulunda yaratılmış ve veri matrisine 101.sütun olarak eklenmiş yetenek düzeylerine göre sıralı olan veri matrisinin 2001.satırına maddelere ait güçlük değerleri eklenmiştir. Ardından veri matrisi madde güçlük değerlerine göre küçükten büyüğe olacak şekilde sıralanmıştır.
- Veri matrisinde yer alan yüksek yetenek düzeyine sahip yanıtlayıcı grubundaki 1000 bireyin kolay olan 50 maddeye verdikleri yanıtlar içerisinde 0, düşük yetenek düzeyine sahip yanıtlayıcı grubundaki 1000 bireyin kolay olan 50 maddeye verdikleri yanıtlar içerisinde rastgele seçilen %20, yüksek yetenek düzeyine sahip yanıtlayıcı grubundaki 1000 bireyin zor olan 50 maddeye verdikleri yanıtlar içerisinde rastgele seçilen %30, düşük yetenek düzeyine sahip yanıtlayıcı grubundaki 1000 bireyin zor olan 50 maddeye verdikleri yanıtlar içerisinde ise rastgele seçilen %50 oranına karşılık gelen sayıda veri silinmiştir. Yapılan tüm işlemler ROK mekanizmasının tanımına uygun olacak şekilde, kayıp verilerin hem bireylerin yetenek düzeylerine hem de maddenin kendisi ile alakalı bir durum olan madde güçlük düzeylerine bağlanmasını ve bu değerlere göre yaratılmasını sağlamıştır. Bu amaçla yazılan R fonksiyonu örneği EK-6'da verilmiştir. Tablo 3.5 ise ROK mekanizmasına sahip kayıp verilerin elde edilmesinde kullanılan kayıp veri oranlarını gösterecek şekilde oluşturulmuştur.

Tablo 3.5: ROK Mekanizmasına Sahip Kayıp Verilerin Elde Edilmesinde Kullanılan Kayıp Veri Oranları

Yanıtlayıcı no	Madde no										Yetenek düzeyi (θ)		
	1	2	50	51	52	100			
1												min.	
2		%20 (düşük yetenek düzeyi & kolay maddeler)					%50 (düşük yetenek düzeyi & zor maddeler)						
...													
...													
1000													
1001													
1002		0 (yüksek yetenek düzeyi & kolay maddeler)					%30 (yüksek yetenek düzeyi & zor maddeler)						
...													
...													
2000													max.
Madde güçlük düzeyi (b_i)	min.												max.

3.4.3. Kayıp Veri Başa Çıkma Yöntemlerinin Kullanılması ile Yeni Veri Setlerinin Elde Edilmesi

Verilerin türetilmesinin üçüncü ve son aşamasında, farklı başa çıkma yöntemlerinin kullanılması ile bir önceki aşamada oluşturulan kayıp verili veri setlerinin yeni veri setlerine dönüştürülmesi amaçlanmıştır.

UGD yönteminin kullanımında kayıp veriye sahip maddeler sanki hiç uygulanmamış gibi işlem gördükleri için, veri seti üzerinde herhangi bir değişiklik yapılmamış ve direk olarak parametre kestirimine geçilmiştir. Bunun için parametre kestirim işlemlerinin gerçekleştirildiği R yazılımı ltm paket programında, kayıpların görmezden gelinmesini ve kestirim işlemlerinin sadece gözlenen veriler üzerinden yürütülmesini sağlayan özel bir kodlama yapılmıştır (Rizopoulos, 2015). YYGD yönteminin kullanımında ise kayıp veriye sahip maddeler sanki yanlış yanıtlanmış gibi işlem gördükleri için, öncelikle kayıp veriler sıfır değerleri ile yer değiştirilmiş ve böylelikle eksiksiz hale getirilen yeni tam veri setleri üzerinde aynı paket program kullanılarak parametre kestirim işlemleri gerçekleştirilmiştir.

R yazılımı mice paket programı (van Buuren & Groothuis-Oudshoorn, 2011) kullanılarak gerçekleştirilen ÇDA'ya dayalı başa çıkma yöntemlerinden LRÇDA ve DFÇDA ile kayıp veriye sahip maddeler yerine başka değerlerin çoklu atamaları yapılmıştır. Çoklu değer atama (m) sayısının 5 olarak seçilmesinin doğal bir sonucu olarak, her bir kayıp verili set için ataması yapılmış ve tamamlanmış 5 adet tam veri seti elde edilmiştir. Ardından ltm paket programı kullanılarak elde edilen

tam veri setlerinin her biri için parametre kestirim işlemleri gerçekleştirilmiştir. En sonunda Rubin'in (1987) ortaklama (pooling) süreci ile tüm parametre kestirimlerinin ortalaması alınmış ve her bir koşul için tek bir tane parametre kestirim setinin (single set of results) elde edilmesi sağlanmıştır.

3.5. Eşitleme Süreci

Araştırmanın bu sürecinde her bir simülasyon koşulu altında eşitleme işlemlerinin yapılması sağlanmıştır. Belirtilen işlemlerin yürütülmesinde R yazılımı plink paket programı kullanılmıştır. Bu paket kapsamında herhangi bir eşitleme işleminin gerçekleştirilmesi için gerekli olan argümanlar; eşitlemesi yapılacak test formlarına ait madde ve yetenek parametre setleri, maddelere ait puanlama biçimi (madde kategori sayısı), kullanılan MTK modeli, test formlarındaki ortak maddelerin yerlerini belirten sıra numaraları, test formlarının hangisinin eşitlenecek/yeni form hangisinin alınan/eski form olduğunu gösteren ifadeler ve kullanılması istenen ölçek dönüştürme (rescaling) yönteminin adıdır (Weeks, 2010).

Bu araştırmada, eşitlemesi yapılan farklı test formlarından elde edilen yetenek ve madde parametrelerini aynı ölçeğe yerleştirmek için MTK'ya dayalı ayrı kalibrasyon yöntemleri kapsamında ele alınan Stocking-Lord'un kullanımı tercih edilmiştir. Bu tercihte, ilgili ölçek dönüştürme yönteminin diğer yöntemlere göre daha kararlı sonuçlar üretme eğiliminde olması etkili olmuştur (Baker & Al-Karni, 1991; Hanson & Beguin, 2002; Kolen & Brennan, 2004; Kilmen, 2010; Gök, 2012).

Tüm simülasyon koşulları altında türetilen Form X ve Form Y'nin her bir tekrarı için yukarıda bahsi geçen argümanlar elde edildikten veya tanımlandıktan sonra eşitleme işlemlerine başlanmıştır. Örneğin S1 koşuluna ait eşitleme işleminin yürütülebilmesi için öncelikle Form X ve Form Y'nin tam veri setlerine ait madde ve yetenek parametreleri kestirilmiştir. Bu parametre değerlerinin Stocking-Lord yöntemine ait denklem içerisinde kullanılmasıyla, eski test formuna ait gerçek (true) parametrelerinin yeni test formuna yeniden ölçeklenmesini (rescaling) ve böylece yeniden ölçeklenmiş (rescaled) olan dönüştürülmüş parametrelerin elde edilmesini sağlayan A ve B eşitleme katsayılarına ulaşılmıştır. $\theta^* = A\theta + B$ formülü kullanılarak, bireylerin almadıkları test formu için yetenek düzeylerinin (θ^*) elde edilmesi; $a^* = (1/A).a$, $b^* = Ab + B$, $c^* = c$ formülleri kullanılarak da alınmayan

test formunun maddelerine ilişkin madde özelliklerinin (a^*, b^*, c^*) elde edilmesi sağlanmıştır.

Araştırmada ele alınan tekrar sayısının 50 olması sebebiyle, her bir koşul altında yukarıdaki gibi eşitlemesi yapılan 50'şer test formu bulunmaktadır. Bu nedenle her bir simülasyon koşulu altında 50 olmak üzere, toplam $109 \times 50 = 5450$ eşitleme işlemi yürütülmüştür.

3.6. Verilerin Analizi ve Değerlendirme Ölçütleri

Araştırma kapsamında ele alınan tüm koşullar için elde edilen eşitleme sonuçlarının karşılaştırılmasında, her koşul için ayrı ayrı hesaplanan hata ve yanlışlık değerlerine başvurulmuştur. Finch (2008), Sulis ve Porcu (2008) ile Acock (2012) bu tip simülasyon araştırmalarında, tam veri koşulunu ifade eden referans koşula ait kestirimlerin, yansız ve önemsiz denecek seviyede Tip I hata oranları içermesinden dolayı diğer koşullar için referans olabilen ölçüt değerler ürettiklerini belirtmiştir. Böylece referans koşul için elde edilen hata ve yanlışlık değerleri, diğer araştırma koşulları için elde edilen değerlerin karşılaştırılıp yorumlanmasında ölçüt olarak alınır. Bu nedenle araştırmada, referans koşul (S1) için hesaplanan eşitleme hataları ve eşitleme yanlışlıkları, diğer 108 koşula (S2-S108) ait eşitleme sonuçlarının değerlendirilmesinde kullanılmıştır.

Eşitleme işleminde yapılan toplam hatayı ifade eden RMSE (root mean square error), test formlarını alan örneklemelerin bazı özellikleriyle ilgili bir hata türü olan rastgele eşitleme hatası ile eşitleme varsayımlarının karşılanabilirliğiyle ilgili bir hata türü olan sistematik eşitleme hatasının toplamına karşılık gelir. RMSE, dönüştürülmüş parametre değerleri ile gerçek parametre değerleri arasındaki farkın kareleri toplamının tekrar sayısına oranının karekökünün alınmasıyla elde edilir. Aşağıdaki eşitlikle de hesaplanır:

$$RMSE(\tau_j) = \sqrt{\frac{\sum_{r=1}^R (\hat{\tau}_{jr} - \tau_j)^2}{R}}$$

τ_j : j parametresinin gerçek değeri

$\hat{\tau}_{jr}$: tekrar edilen veri seti ($r = 1 \dots R$) için j parametresinin dönüştürülen değeri

R : tekrar sayısı

Yukarıdaki j parametresi ifadesi, yetenek parametresi (θ) ile a ve b madde parametrelerine karşılık gelmektedir. c parametresinin j parametresi kapsamında ele alınmamasının sebebi, dönüştürülmüş parametrelerin elde edilmesini sağlayan ölçek dönüştürme (kalibrasyon) denklemlerinden de görüleceği üzere, a ve b parametrelerine ait değerlerin dönüştürme işlemi sırasında değişikliğe uğramaları, ancak, c parametresine ait değerlerin dönüştürme işleminden etkilenmemelerinden kaynaklanmaktadır. Bu nedenle, dönüştürülmüş ile gerçek parametre değerlerinin farkına dayanan herhangi bir değerlendirme ölçütünün c parametresi üzerinde hesaplanmasına gerek kalmamıştır. Dolayısıyla bu çalışmanın her bir koşulu için elde edilen RMSE değeri; yetenek parametreleri, madde ayırıcılık parametreleri ve madde güçlük parametreleri için ayrı ayrı hesaplanmıştır. Örneğin, S1 koşulunun her bir tekrarı için yetenek parametrelerine ilişkin RMSE hesaplanırken öncelikle Form X'in dönüştürülmüş yetenek parametreleri ile Form Y'nin gerçek yetenek parametreleri arasındaki farklar alınır. Böylelikle 2000 bireyin her biri için $(\theta^* - \theta)$ değerlerine ulaşılır. Ardından bu değerlerin kareleri alınır. Bu aşamaya kadar tekrar sayısının 50 olması sebebiyle, her birey için 50 tane $(\theta^* - \theta)^2$ değeri elde edilmiştir. Sonraki adımda her bir tekrar için elde edilen bu değerler toplanıp $(\theta^* - \theta)_1^2 + (\theta^* - \theta)_2^2 + \dots + (\theta^* - \theta)_{50}^2$, tekrar sayısına bölündükten sonra karekökü alınır. Böylece her bir birey için bir tane değer elde edilmiş olur. Ancak her koşul için bir tane "yetenek parametrelerine ait ortalama RMSE değeri" elde edilmesi gerektiğinden, aşağıdaki eşitlik işleme konulur. Bu eşitliğe göre bireylerin her birine ait değerler toplanıp, birey sayısı olan 2000'e bölünmüş ve S1 koşulu için bir tane yetenek parametrelerine ilişkin ortalama RMSE değeri elde edilmiştir.

$$\sum_{i=1}^I RMSE(\tau_j)/I$$

I : rapor edilen parametreye bağlı olarak birey ya da madde sayısı

Eşitleme işleminde yapılan sistematik hatayı ifade eden BIAS (yanlılık) ise, dönüştürülmüş parametre değerleri ile gerçek parametre değerleri arasındaki farkın toplamının tekrar sayısına bölünmesiyle elde edilir. Aşağıdaki eşitlikle de hesaplanır:

$$BIAS(\tau_j) = \frac{1}{R} \sum_{r=1}^R (\hat{t}_{jr} - \tau_j)$$

τ_j : j parametresinin gerçek değeri

$\hat{\tau}_{jr}$: tekrar edilen veri seti ($r = 1 \dots R$) için j parametresinin dönüştürülen değeri

R : tekrar sayısı

Aynı RMSE'de olduğu gibi, bu çalışmanın her bir koşulu için elde edilen BIAS değeri; yetenek parametreleri, madde ayırıcılık parametreleri ve madde güçlük parametreleri için ayrı ayrı hesaplanmıştır. Örneğin, S1 koşulunun her bir tekrarı için kestirilen madde güçlük parametrelerine ilişkin BIAS hesaplanırken öncelikle Form X'in dönüştürülmüş madde güçlük parametreleri ile Form Y'nin gerçek madde güçlük parametreleri arasındaki farklar alınır. Böylelikle 100 maddenin her biri için $(b^* - b)$ değerlerine ulaşılır. Bu aşamaya kadar tekrar sayısının 50 olması sebebiyle, her madde için 50 tane $(b^* - b)_1 + (b^* - b)_2 + \dots + (b^* - b)_{50}$ değeri elde edilmiştir. Sonraki adımda her bir tekrar için elde edilen bu değerler toplanıp, tekrar sayısına bölünür. Böylece her bir madde için bir tane değer elde edilmiş olur. Ancak her bir koşul için bir tane "madde güçlük parametrelerine ait ortalama BIAS değeri" elde edilmesi gerektiğinden, aşağıdaki eşitlik işleme konulur. Bu eşitliğe göre, maddelerin her birine ait değerler toplanıp, madde sayısı olan 100'e bölünmüş ve S1 koşulu için bir tane madde güçlük parametrelerine ilişkin ortalama BIAS değeri elde edilmiştir.

$$\sum_{i=1}^I BIAS(\tau_j)/I$$

I : rapor edilen parametreye bağlı olarak birey ya da madde sayısı

S1 koşulu üzerinden yukarıda anlatılan işlem adımlarının hepsi araştırmaya ait diğer koşullar (S2-S108) üzerinde de uygulanmış ve tüm koşullar için yetenek ve madde parametrelerine ilişkin ortalama RMSE ve BIAS değerleri elde edilmiştir. Hem RMSE hem de BIAS değerlerinin olabildiğince küçük çıkması, eşitleme işlemlerinin doğrulukla gerçekleştirildiğinin bir göstergesidir. Tam tersi bir durum, dönüştürülen parametrelerin eşitleme işlemine karışan hatalar sebebiyle gerçek değerlerinden büyük ölçüde farklılık gösterdiği şeklinde yorumlanır. Böyle bir yorumlama aynı zamanda, test eşitleme süreci ile elde edilmesi amaçlanan eşitlenmiş puanlara duyulan güvenin azalmasına sebep olur.

Arařtırmada ele alınan kořulların ana ve ortak etkilerinin, kayıp veri bařa ıkma yntemleri kullanıldıktan sonra eřitlenen test formlarından kestirilen parametrelerin ortalama RMSE ve ortalama BIAS deęerleri zerindeki anlamlılıklarını test etmek amacıyla ANOVA'ya bařvurulmuřtur. Kořullar ve kořullar arası tm etkileřimlerin ele alındıęı bir model kurulduęunda, hataların serbestlik derecesi sıfır ıktıęından analiz bařarı ile gerekleřtirilememiř ve bu sebeple sadece ana etkiler ile iki ynl etkileřimler test edilmiřtir. Analizlere ait sonular tablolařtırılırken, simlasyon faktrlerinin kayıp veri bařa ıkma yntemleri zerindeki performansını gstermek amacıyla ile F deęerlerinin yanı sıra eta-kare deęerleri (η^2) de rapor edilmiřtir.

4. BULGULAR

Bu bölümde alt problemlerin sırasına göre araştırmanın bulgularına yer verilmiştir. Bulgular tamamlandığında ise tartışma başlığı altında araştırmanın bulguları ilgili literatür göz önünde bulundurularak tartışılmıştır.

4.1. Birinci Alt Probleme İlişkin Bulgular

Kayıp veriye sahip maddelerin bulunduğu test formları, kayıp veri başa çıkma yöntemleri içerisinde değer atama içermeyen yöntemler (UGD ve YYGD) ve çoklu değer atamaya dayalı yöntemler (LRÇDA ve DFÇDA) ile çözümlenip eşitlendiğinde kestirilen **madde parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı; kayıp verinin yerine, mekanizmasına ve oranına göre nasıl değişmektedir?

Birinci alt probleme ilişkin bulguların net bir şekilde anlaşılıp yorumlanabilmesi için araştırma kapsamında ele alınan koşullara ait eşitleme sonuçları üç farklı kayıp veri mekanizmasına göre ayrı ayrı tablolaştırılmış ve grafiğe dökülmüştür. Bu işlemin hemen ardından kestirilen madde parametrelerine ilişkin eşitleme hatası ve eşitleme yanlılığı değerlerinin çalışmada ele alınan koşulların ana ve ortak etkilerine göre nasıl değişiklik gösterdiğini belirlemek amacı ile yapılan ANOVA sonuçlarına yer verilmiştir.

TRK, RK ve ROK mekanizmasına sahip kayıp verileri içeren simülasyon koşullarında kestirilen madde parametrelerine ilişkin RMSE ve BIAS değerlerini içeren tablolar sırasıyla EK-7, EK-8 ve EK-9'da verilmiştir. TRK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin RMSE ve BIAS değerlerini gösteren Şekil 4.1-4.4'deki grafikler EK-7'de yer alan tabloya; RK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin RMSE ve BIAS değerlerini gösteren Şekil 4.5-4.8'deki grafikler EK-8'de yer alan tabloya; ROK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin RMSE ve BIAS değerlerini gösteren Şekil 4.9-4.12'deki grafikler EK-9'da yer alan tabloya göre oluşturulmuştur.

TRK mekanizmasına sahip kayıp verileri içeren simülasyon koşulları kapsamında kestirilen madde parametrelerine ait hata ve yanlılık değerlerini gösteren EK-7'de; araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE değerleri

kayıp verinin yerine göre incelendiğinde, genel olarak en küçük değerlere sahip olan en az hata içeren sonuçların her iki testin kestirilen madde parametrelerine ait olduğu, en büyük değerlere sahip olan en hatalı sonuçların ise açık ara farkla eşitlenecek testin kestirilen madde parametrelerine ait olduğu görülmüştür.

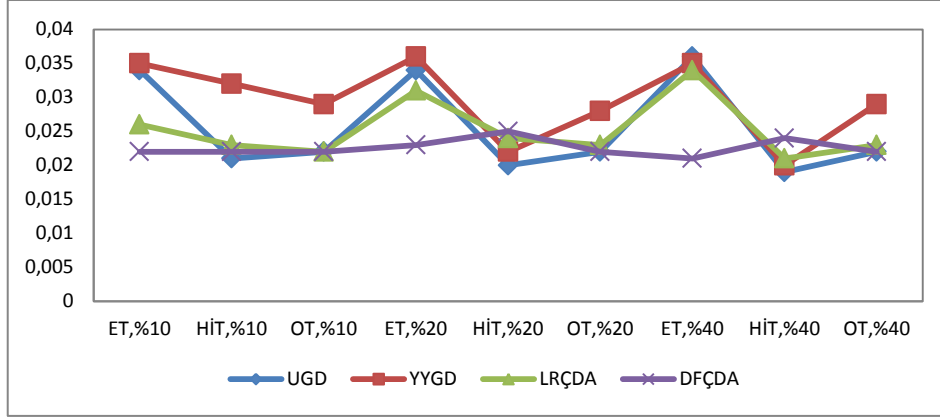
Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen BIAS değerleri kayıp verinin yerine göre incelendiğinde, en küçük değerlere sahip olan en az yanlı sonuçların, ortak testin kestirilen madde parametrelerine ait olduğu görülmektedir. Ancak kayıp veri oranı %40'a çıktığında, ortak test durumunda kestirilen madde parametrelerine ait BIAS değerlerinin her iki teste göre daha yüksek çıktığı durumlarla da karşılaşmıştır. Bu bulgulara göre zaten az sayıda (20) madde içeren ortak testteki kayıp veri sayısının ciddi oranda artması, eşitleme sonuçlarının doğruluğunun olumsuz yönde etkilenmesine sebep olmuş olabilir. Eşitlenecek test ve her iki testin kestirilen madde parametrelerine ait BIAS değerlerine bakıldığında ise, genellikle eşitlenecek testin her iki teste göre daha yüksek yanlılık değerleri ürettikleri bulgusuna ulaşılmıştır.

Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE değerleri kayıp verinin oranına göre incelendiğinde, %10 ve %20 oranları arasında kestirilen madde parametrelerine ait hata değerleri açısından gözle görülür ölçüde farklılık oluşmadığı ancak, %40 oranı için hata miktarlarının arttığı gözlenmiştir. Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen BIAS değerleri kayıp verinin oranına göre incelendiğinde, kayıp veri oranı arttıkça yanlılığın da arttığı görülmüştür. Burada yapılan karşılaştırmalar için de, kayıp veri oranının %40 olduğu simülasyon koşulları en yanlı sonuçları içermektedir.

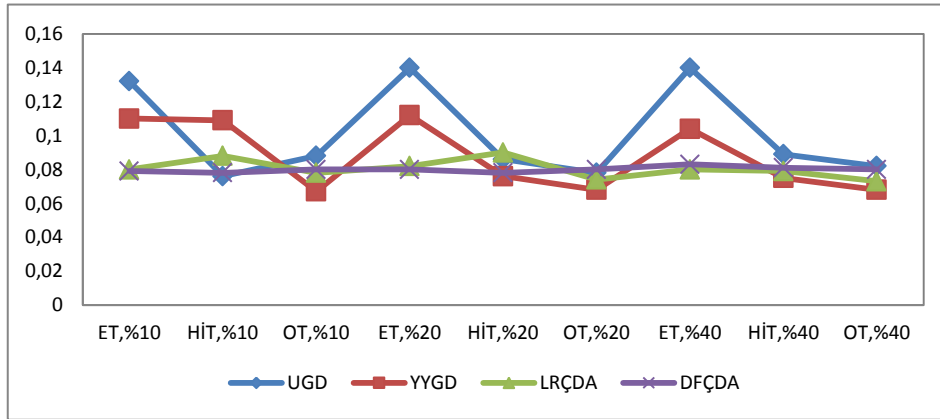
Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri incelendiğinde, genel olarak en düşük hata ve yanlılık değerlerini üreten yöntemlerin LRÇDA ve DFÇDA oldukları, en yüksek hata ve yanlılık değerlerini üreten yöntemin ise UGD olduğu görülmektedir. UGD ile YYYGD yöntemleri arasında çok ciddi farklılıklar olmasa da, genel olarak YYGD'ye ait hata ve yanlılık değerlerinin UGD'ye göre daha düşük çıktığı gözlenmiştir. Bu bulguların yanı sıra, DFÇDA'nın LRÇDA'ya göre özellikle a parametresinde daha düşük hata ve yanlılık değerleri ürettiği gözlenirken, b parametresinin bazı değerleri için de tam tersi bir durum tespit edilmiştir. Ayrıca eşitlenecek teste ait koşullarda ÇDA yöntemlerine doğru gidildikçe hata ve yanlılık değerlerinin düşüşü daha net bir şekilde

gözlemlenebilirken, her iki test ve ortak testin bazı değerleri için ufak miktarlarda artışlar kaydedilmiştir.

TRK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin RMSE değerleri Şekil 4.1 ile 4.2'de verilmiştir.



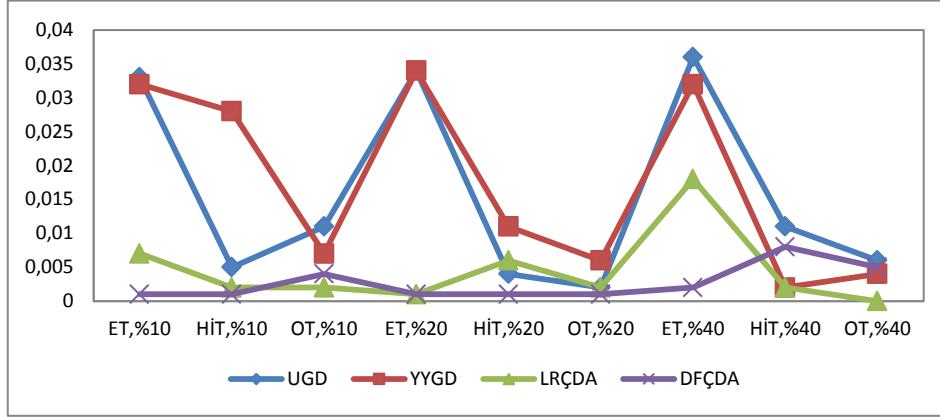
Şekil 4.1. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Hataları



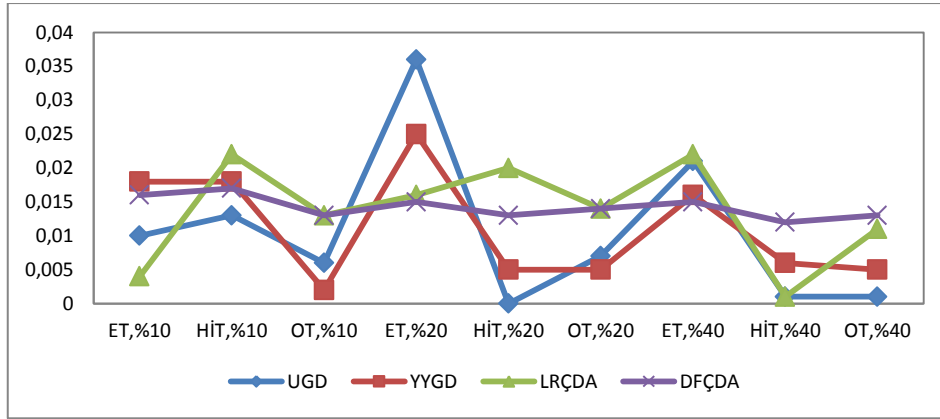
Şekil 4.2. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Hataları

Şekil 4.1 ile 4.2 genel hatlarıyla incelendiğinde, a parametresine ait en düşük eşitleme hatalarının DFÇDA, en yüksek eşitleme hatalarının ise YYGD yöntemi için; b parametresine ait en düşük eşitleme hatalarının DFÇDA ve LRÇDA, en yüksek eşitleme hatalarının ise UGD yöntemi için elde edildiği görülmüştür. Bunun yanı sıra, kayıp verilerin eşitlenecek test formu içerisinde olduğu koşulların gözle görülür şekilde daha yüksek hata değerleri ürettiği gözlenmiştir.

TRK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin BIAS değerleri Şekil 4.3 ile 4.4'te verilmiştir.



Şekil 4.3. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Yanlılıkları



Şekil 4.4. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Yanlılıkları

Şekil 4.3 ile 4.4 genel hatlarıyla incelendiğinde, a parametresine ait en düşük eşitleme yanlılıklarının DFÇDA, en yüksek eşitleme yanlılıklarının ise UGD ve YYGD yöntemleri için; b parametresine ait en yüksek eşitleme yanlılıklarının özellikle eşitlenecek test içerisinde %20 oranında kayıp veri bulunduğu UGD yöntemi için elde edildiği görülmüştür. Bunun yanı sıra kayıp veri başa çıkma yöntemlerine ilişkin BIAS değerlerinin RMSE değerlerine göre birbirlerinden daha çok farklılaştıkları gözlenmiştir.

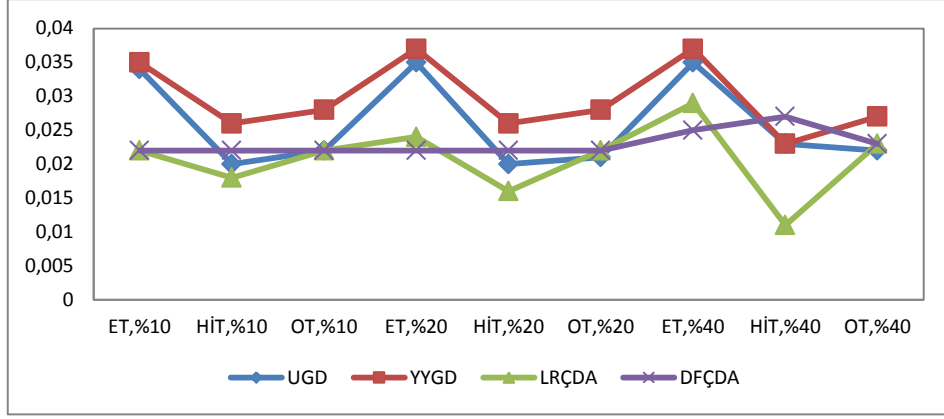
RK mekanizmasına sahip kayıp verileri içeren simülasyon koşulları kapsamında kestirilen madde parametrelerine ait hata ve yanlılık değerlerini gösteren EK-8'de; araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS

değerleri kayıp verinin yerine göre incelendiğinde, genel olarak en küçük değerleri içeren en az hatalı ve en az yanlı sonuçların her iki testin kestirilen madde parametrelerine ait olduğu, en büyük değerlere sahip olan en hatalı ve en yanlı sonuçların ise eşitlenecek testin kestirilen madde parametrelerine ait olduğu görülmüştür.

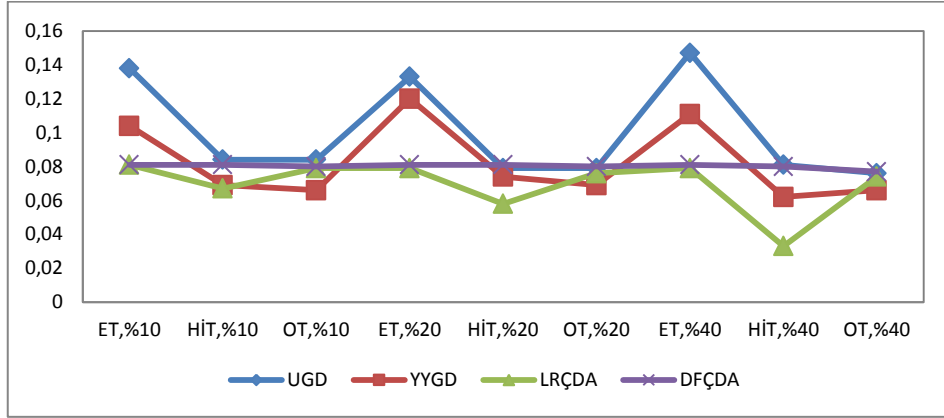
Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE değerleri kayıp verinin oranına göre incelendiğinde, %10 ve %20 arasında kestirilen madde parametrelerine ait hata değerleri açısından gözle görülür farklılıkların oluşmadığı, ancak BIAS değerleri açısından inceleme yapıldığında kayıp veri oranı arttıkça yanlılığın da arttığı görülmüştür. RK kapsamındaki karşılaştırmalar için de aynı TRK'de olduğu gibi, kayıp veri oranının %40 olduğu simülasyon koşulları en yanlı sonuçları üretmektedir.

Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri incelendiğinde, genel olarak en düşük hata ve yanlılık değerlerini üreten yöntemlerin LRÇDA ve DFÇDA oldukları, en yüksek hata ve yanlılık değerlerini üreten yöntemin ise UGD olduğu görülmektedir. UGD ile YYGD yöntemleri arasında çok ciddi farklılıklar olmasa da, genel olarak YYGD'ye ait hata ve yanlılık değerlerinin UGD'ye göre daha düşük çıktığı gözlenmiştir. Bu bulguların yanı sıra DFÇDA'nın da LRÇDA'ya göre daha düşük yanlılık değerleri ürettiği tespit edilmiştir. Ayrıca her iki teste ait koşullarda ÇDA yöntemlerine doğru gidildikçe hata ve yanlılık değerlerinin düşüşü net bir şekilde gözlemlenebilirken, eşitlenecek test ve ortak testin bazı değerleri için ufak miktarlarda artışlar kaydedilmiştir.

RK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin RMSE değerleri Şekil 4.5 ile 4.6'da verilmiştir.



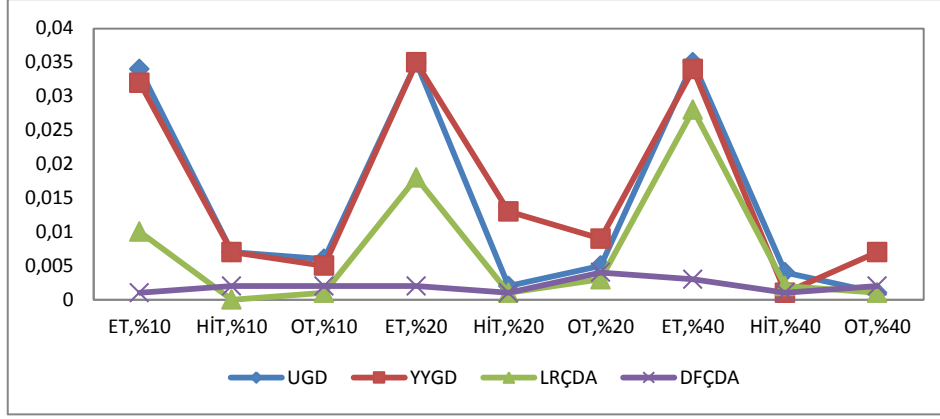
Şekil 4.5. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen α Parametresine İlişkin Eşitleme Hataları



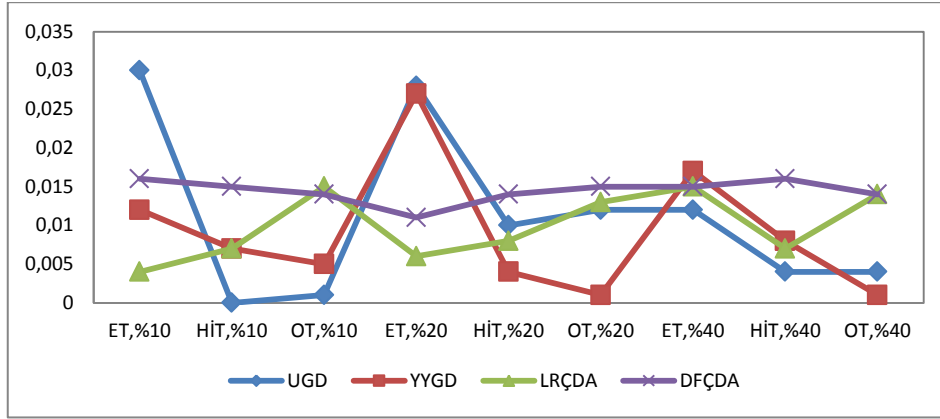
Şekil 4.6. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Hataları

Şekil 4.5 ile 4.6 genel hatlarıyla incelendiğinde, a parametresine ait en düşük eşitleme hatalarının her iki test formunda da %40 oranında kayıp veri olduğu koşullarda LRÇDA yöntemi, en yüksek eşitleme hatalarının ise YYGD yöntemi için; b parametrelerine ait en düşük eşitleme hatalarının yine her iki test formunda da %40 oranında kayıp veri olduğu koşullarda LRÇDA yöntemi, en yüksek eşitleme hatalarının ise eşitlenecek test formunda %40 oranında kayıp veri bulunduğu koşullarda UGD yöntemi için elde edildiği görülmüştür.

RK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin BIAS değerleri Şekil 4.7 ile 4.8'de verilmiştir.



Şekil 4.7. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Yanlılıkları



Şekil 4.8. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Yanlılıkları

Şekil 4.7 ile 4.8 genel hatlarıyla incelendiğinde, DFÇDA yönteminin tüm simülasyon koşulları kapsamında her iki madde parametresi için de birbirine en yakın olan eşitleme yanlılıklarını ürettiği gözlenmiştir. Bunun yanı sıra, a parametresine ait en düşük eşitleme yanlılık değerlerinin DFÇDA yöntemi, genel olarak her iki madde parametresine de ait en yüksek eşitleme yanlılık değerlerinin ise UGD ve YYGD yöntemleri için elde edildiği görülmüştür.

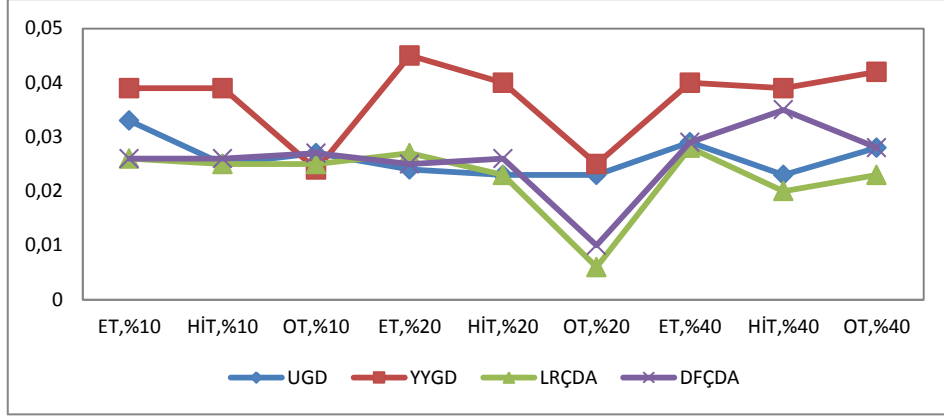
ROK mekanizmasına sahip kayıp verileri içeren simülasyon koşulları kapsamında kestirilen madde parametrelerine ait hata ve yanlılık değerlerini gösteren EK-9'da; araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri kayıp verinin yerine göre incelendiğinde, genel olarak en küçük değerleri içeren en az hatalı ve en az yanlı sonuçların her iki testin kestirilen madde parametrelerine ait olduğu, en büyük değerlere sahip olan en hatalı ve en yanlı sonuçların ise eşitlenecek testin kestirilen madde parametrelerine ait olduğu

görülmüştür. ROK için üretilen RMSE değerleri üzerinde parametre bazında karşılaştırmalar yapıldığında, b parametresi için en az yanlılık içeren kayıp veri yeri koşulunun her iki test olduğu, ancak a parametresi için bu durumun değişken sonuçlar ürettiği tespit edilmiştir.

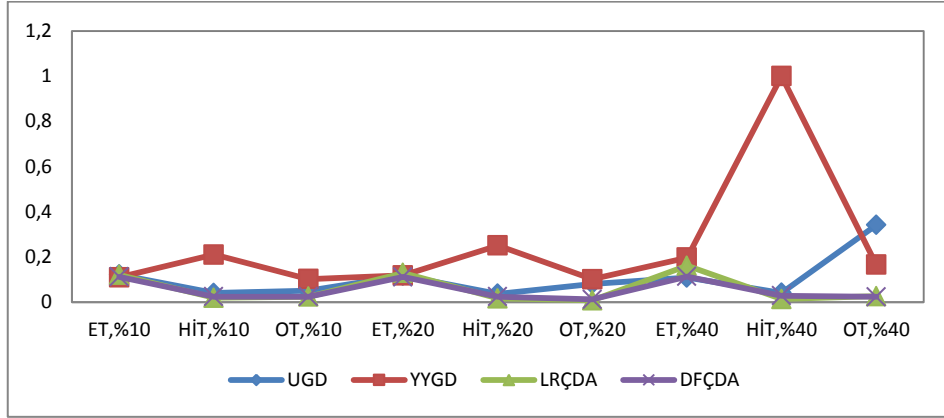
Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE değerleri kayıp verinin oranına göre incelendiğinde, her iki test ve eşitlenecek test koşullarında %10'dan %40'a doğru gidildikçe hata oranının arttığı, ancak ortak test için bu durumun net olarak ifade edilemeyeceği tespit edilmiştir. Araştırmada ele alınan başa çıkma yöntemlerine ilişkin BIAS değerleri kayıp verinin oranına göre incelendiğinde ise karşılaştırmalar arasında kayıp altına alınmaya geçecek tutarlılıklar gözlenememiştir.

Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri incelendiğinde, genel olarak en düşük hata ve yanlılık değerlerini üreten yöntemin LRÇDA olduğu ve DFÇDA'nın da hemen LRÇDA'nın ardından geldiği tespit edilirken, en yüksek hata ve yanlılık değerlerini üreten yöntemin ise YYGD olduğu görülmektedir. UGD ile YYGD yöntemleri arasında çok ciddi farklılıklar gözlenemese de, genel olarak UGD'ye ait hata ve yanlılık değerlerinin YYGD'ye göre daha düşük çıktığı tespit edilmiştir. Raporlanması gereken başka bir bulgu da, ortak test koşulundaki kayıp veriler için özellikle ÇDA yöntemlerinin kullanıldığı durumlarda elde edilen BIAS değerlerinin sıfıra çok yaklaştıkları, başka bir ifadeyle yanlılığının yok denecek kadar az olduğudur.

ROK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin RMSE değerleri Şekil 4.9 ile 4.10'da verilmiştir.



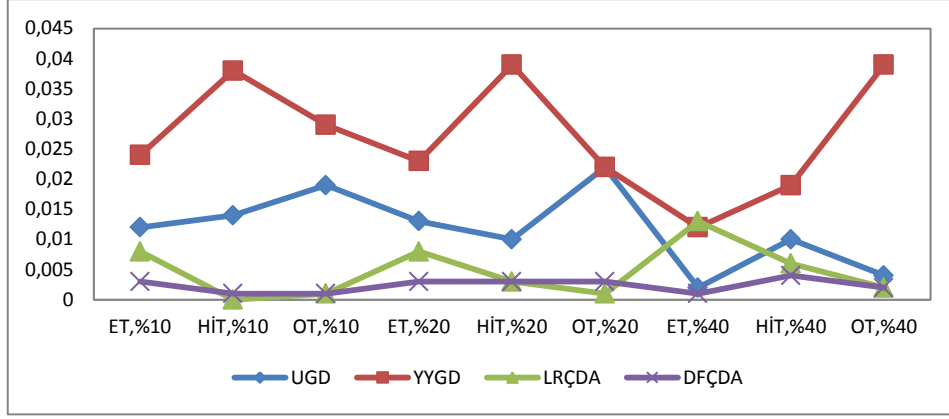
Şekil 4.9. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen α Parametresine İlişkin Eşitleme Hataları



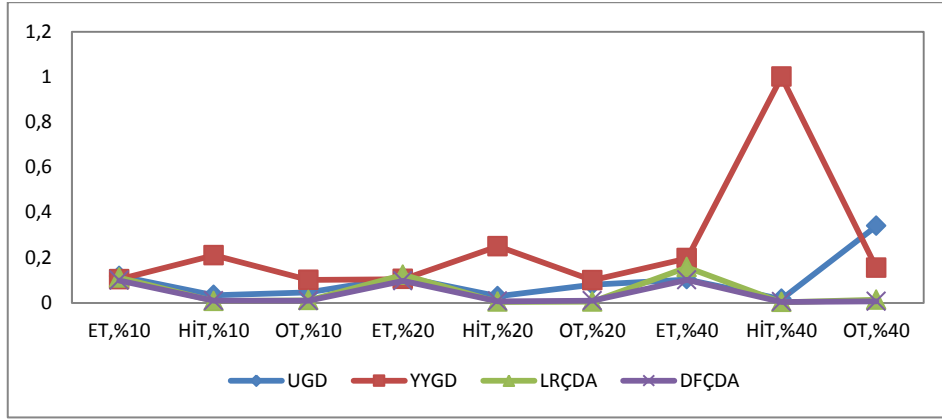
Şekil 4.10. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Hataları

Şekil 4.9 ile 4.10 genel hatlarıyla incelendiğinde, madde parametrelerine ait en düşük eşitleme hatalarının DFÇDA ve LRÇDA'ya, en yüksek eşitleme hatalarını üreten yöntemin ise YYGD'ye ait olduğu gözlenmiştir. Bununla birlikte her iki parametre için de DFÇDA ve LRÇDA'nın birbirlerine oldukça yakın eşitleme hata değerleri ürettiği bulgusuna ulaşılmıştır.

ROK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen madde parametrelerine ilişkin BIAS değerleri Şekil 4.11 ile 4.12'de verilmiştir.



Şekil 4.11. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen a Parametresine İlişkin Eşitleme Yanlılıkları



Şekil 4.12. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen b Parametresine İlişkin Eşitleme Yanlılıkları

Şekil 4.11 ile 4.12 genel hatlarıyla incelendiğinde, her iki madde parametresine de ait olan en düşük eşitleme yanlılıklarının genel olarak DFÇDA'ya, en yüksek eşitleme yanlılıklarını üreten yöntemin ise YYGD'ye ait olduğu gözlenmiştir. Bununla birlikte %40 oranında kayıp veri içeren simülasyon koşulları kapsamında YYGD'nin işe koşulduğu durumlarda kestirilen özellikle b parametresine ait yanlılık değerlerinin, %10 ve %20 oranındakilere göre önemli ölçüde artış gösterdiği saptanmıştır.

Bu alt problem kapsamında, EK-7-8-9'daki kayıp veriye sahip maddeleri içeren testler üzerinde UGD, YYGD, LRÇDA ve DFÇDA yöntemleri kullanılıp eşitleme yapıldığında kestirilen madde parametrelerine ilişkin RMSE ve BIAS değerleri bağımlı, araştırmada incelenen simülasyon faktörleri de (yer, mekanizma, oran) bağımsız değişkenler olarak ele alındığında; bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkilerini belirlemek amacıyla ayrı ayrı ANOVA yapılmıştır.

Madde parametreleri için anlamlı bulunan etkilerin başa çıkma yöntemlerine göre F değerleri ve etki büyüklükleri (η^2) Tablo 4.1'de verilmiştir.

Tablo 4.1: Madde Parametrelerine ait RMSE ve BIAS Değerleri için Anlamlı Bulunan ANOVA Sonuçları

		<i>Kayıp Veri Başa Çıkma Yöntemleri</i>									
				<i>UGD</i>		<i>YYGD</i>		<i>LRÇDA</i>		<i>DFÇDA</i>	
<i>Değerlendirme Ölçütleri</i>	<i>Etkiler</i>	<i>sd</i>	<i>F</i>	η^2	<i>F</i>	η^2	<i>F</i>	η^2	<i>F</i>	η^2	
a	RMSE	Y	2	110.443*	1	-	-	-	-	-	
	BIAS	Y	2	44.821*	.40	-	-	68.325*	.50	-	
		Y*M	4	16.730*	.40	-	-	-	-	-	
b	RMSE	Y	2	-	-	-	-	143.547*	.35	579.247*	.29
		M	2	-	-	-	-	28.327*	.06	473.437*	.24
	BIAS	Y*M	4	-	-	-	-	103.693*	.53	535.764*	.48
		Y	2	-	-	-	-	170.006*	.42	190.508*	.21
	BIAS	M	2	-	-	-	-	93.517*	.23	322.587*	.36
		Y*M	4	-	-	-	-	69.199*	.34	195.921*	.42

*Y: yer, M: mekanizma, Y*M: yer-mekanizma etkileşimi.

** $p < 0.002$

Tablo 4.1'de bulunan madde parametrelerine ilişkin RMSE değerlerine ait ANOVA sonuçları anlamlı bulunan etkiler açısından incelendiğinde, a parametresi için yer ana etkisinin sadece UGD yöntemi için anlamlı olduğu gözlenmiştir. Bununla birlikte DFÇDA için hiçbir faktörün tek başına veya etkileşimlerinin etkisi anlamlı çıkmamıştır. b parametresi için yer ve mekanizma ana etkileri ile yer-mekanizma etkileşim etkisinin hem LRÇDA hem de DFÇDA için anlamlı olduğu ancak, başka etkilerin diğer yöntemlerin hiçbiri üzerinde anlamlı etkilerinin olmadığı kaydedilmiştir.

Tablo 4.1'de bulunan madde parametrelerine ilişkin BIAS değerlerine ait ANOVA sonuçları anlamlı bulunan etkiler açısından incelendiğinde; a parametresi için yer ana etkisinin UGD ve LRÇDA için anlamlı olduğu, yer-mekanizma etkileşim etkisinin sadece UGD için anlamlı olduğu gözlenmiştir. Bu parametre kapsamında YYGD ve DFÇDA için anlamlı sonuç veren herhangi bir etki bulunmamaktadır. b parametresi için yer ve mekanizma ana etkisi ile yer-mekanizma etkileşim etkisinin hem LRÇDA hem de DFÇDA için anlamlı olduğu, UGD ve YYGD yöntemleri için ise hiçbir etkinin anlamlı olmadığı tespit edilmiştir.

4.2. İkinci Alt Probleme İlişkin Bulgular

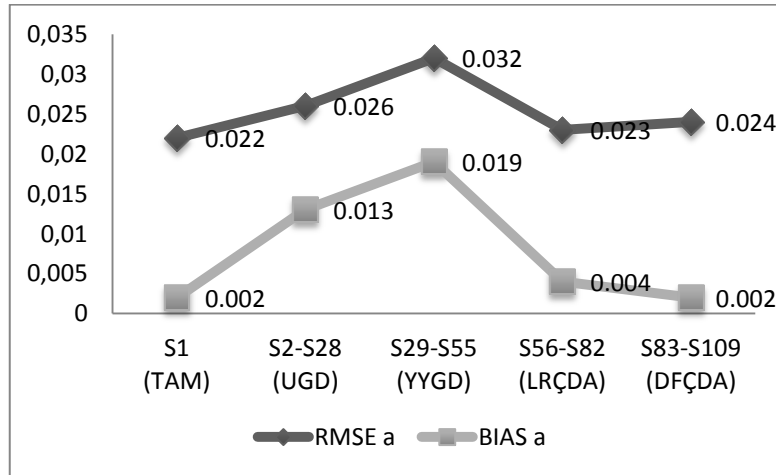
Tam veriye sahip test formları ile kayıp veri başa çıkma yöntemleri kullanılarak elde edilen test formları eşitlendiğinde kestirilen **madde parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı nasıl değişmektedir?

İkinci alt problemin çözümlenmesinde; tam veriye sahip test formlarının eşitlendiği S1 koşulunda kestirilen madde parametrelerine ilişkin RMSE ve BIAS değerleri, kayıp veri başa çıkma yöntemleri kullanılarak elde edilen madde parametre kestirimlerine ilişkin eşitleme hatalarının ve yanlılıklarının nasıl değiştiğinin tespit edilmesi adına referans parametreler olarak kullanılmıştır. Ölçüt amaçlı işe koşulan bu değerler Tablo 4.2'de verilmiştir.

Tablo 4.2: Tam Veriye Sahip Test Formlarının Eşitlenmesi ile Kestirilen Madde Parametrelerine İlişkin RMSE ve BIAS Değerleri

Ölçüt türü	Madde parametreleri	
	a	b
RMSE	0.022	0.080
BIAS	-0.002	0.014

Şekil 4.13 ile 4.14'teki grafiklerde S1 ile ifade edilen koşul tam veriye sahip test formlarının eşitlenmesi ile kestirilen madde parametrelerine ait ortalama hata ve yanlılık değerlerini, S2-S109 aralığında ifade edilen koşullar ise kayıp veri içeren maddelere sahip test formlarının üzerinde farklı başa çıkma yöntemlerinin kullanılmasının ardından eşitlenmesi ile kestirilen madde parametrelerine ait ortalama hata ve yanlılık değerlerini vermektedir. Madde parametrelerine ilişkin ortalama yanlılık değerleri hesaplanırken BIAS değerlerinin mutlak ifadeleri alınmıştır.

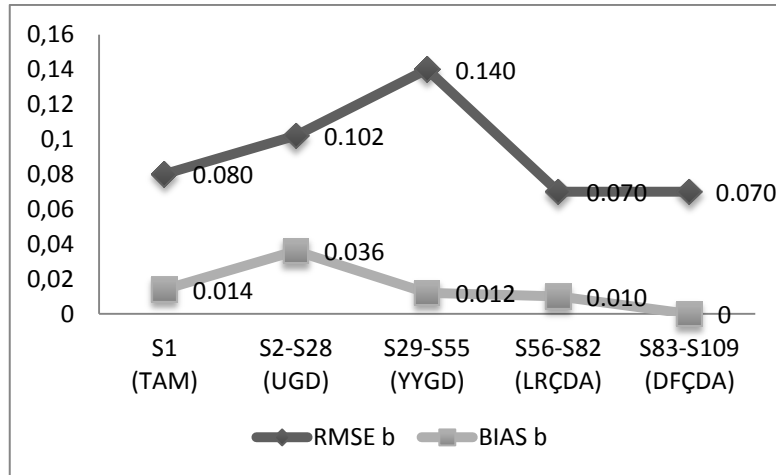


Şekil 4.13. Araştırma Koşulları Kapsamında Gerçekleştirilen Eşitlemeler Sonucunda Kestirilen a Parametrelerine ait Ortalama Eşitleme Hataları

Referans koşula ait kestirilen a parametresi üzerinden hesaplanan ortalama RMSE değerine (0.022) en yakın ve en düşük ortalama $RMSE_a$ değerini (0.023)

veren kayıp veri başa çıkma yöntemi S56 ile S83 koşulları arasında kullanılan LRÇDA'dır. Bu yöntemin hemen ardından gelen DFÇDA (0.024) yöntemi de, tam veriye sahip test formlarından elde edilen eşitleme sonuçlarına oldukça yakın sonuçlar vermiştir. Bu nedenle her iki ÇDA yönteminin de en az hata içeren α parametre kestirimlerinin üretilmesine sebep oldukları söylenebilir. DFÇDA'nın arkasından gelen yöntem 0.026 değeri ile UGD ve en hatalı sonuçlar üreten yöntem ise 0.032 değeri ile YYGD'dir.

Referans koşula ait kestirilen α parametresi üzerinden hesaplanan ortalama BIAS değerine (0.002) eşit ve en düşük ortalama $BIAS_\alpha$ değerini (0.002) veren kayıp veri başa çıkma yöntemi S83 ile S109 koşulları arasında kullanılan DFÇDA'dır. Bu yöntemin hemen ardından gelen LRÇDA (0.004) yöntemi de, tam veriye sahip test formlarından elde edilen eşitleme sonuçlarına oldukça yakın sonuçlar vermiştir. Bu nedenle her iki ÇDA yönteminin de en az yanlışlık içeren α parametre kestirimlerinin üretilmesine sebep oldukları söylenebilir. LRÇDA'nın arkasından gelen yöntem 0.013 değeri ile UGD ve en yanlış sonuçlar üreten yöntem ise 0.019 değeri ile YYGD'dir.



Şekil 4.14. Araştırma Koşulları Kapsamında Gerçekleştirilen Eşitlemeler Sonucunda Kestirilen b Parametrelerine ait Ortalama Eşitleme Hataları

Referans koşula ait kestirilen b parametresi üzerinden hesaplanan ortalama RMSE değerine (0.080) en yakın ve en düşük ortalama $RMSE_b$ değerlerini (0.070) veren kayıp veri başa çıkma yöntemleri S56 ile S83 koşulları arasında kullanılan LRÇDA ve S83 ile S109 koşulları arasında kullanılan DFÇDA'dır. Bu nedenle her iki ÇDA yönteminin de en az hata içeren b parametre kestirimlerinin üretilmesine

sebepler oldukları söylenebilir. ÇDA yöntemlerinin arkasından gelen yöntem 0.102 değeri ile UGD ve en hatalı sonuçlar üreten yöntem ise 0.140 değeri ile YYGD'dir.

Referans koşula ait kestirilen b parametresi üzerinden hesaplanan ortalama BIAS değerine (0.014) en yakın ortalama $BIAS_b$ değerini (0.012) veren kayıp veri başa çıkma yöntemi S29 ile S55 koşulları arasında kullanılan YYGD'dir. Bu yöntemin hemen ardından gelen LRÇDA (0.010) yöntemi de YYGD yöntemine oldukça yakın sonuçlar vermiştir. Ancak S83 ile S109 koşulları arasında kullanılan DFÇDA'nın en düşük ortalama $BIAS_b$ değerini (0) veren kayıp veri başa çıkma yöntemi olduğunun tespitinin ardından, ortalama yanlılık değerleri için başa çıkma yöntemleri arasında DFÇDA<LRÇDA<YYGD<UGD sıralaması yapılmıştır.

4.3. Üçüncü Alt Probleme İlişkin Bulgular

Kayıp veriye sahip maddelerin bulunduğu test formları, kayıp veri başa çıkma yöntemleri içerisinde değer atama içermeyen yöntemler (UGD ve YYGD) ve çoklu değer atamaya dayalı yöntemler (LRÇDA ve DFÇDA) ile çözümlenip eşitlendiğinde kestirilen **yetenek parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı; kayıp verinin yerine, mekanizmasına ve oranına göre nasıl değişmektedir?

Üçüncü alt probleme ilişkin bulguların net bir şekilde anlaşılıp yorumlanabilmesi için araştırma kapsamında ele alınan koşullara ait eşitleme sonuçları üç farklı kayıp veri mekanizmasına göre tablolandırılmış ve grafiğe dökülmüştür. Bu işlemin hemen ardından kestirilen yetenek parametrelerine ilişkin eşitleme hatası ve eşitleme yanlılığı değerlerinin çalışmada ele alınan koşulların ana ve ortak etkilerine göre nasıl değişiklik gösterdiğini belirlemek amacı ile yapılan ANOVA sonuçlarına yer verilmiştir.

Araştırma kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ilişkin RMSE ve BIAS değerlerini içeren tablo EK-10'da verilmiştir. TRK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ilişkin RMSE ve BIAS değerlerini gösteren Şekil 4.15-4.16'daki grafikler, RK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ilişkin RMSE ve BIAS değerlerini gösteren Şekil 4.17-4.18'deki grafikler ve ROK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek

parametrelerine ilişkin RMSE ve BIAS değerlerini gösteren Şekil 4.19-4.20'deki grafikler EK-10'da yer alan tabloya göre oluşturulmuştur.

Araştırma kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ait hata ve yanlılık değerlerini gösteren EK-10'da; araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri kayıp verinin mekanizmasına göre incelendiğinde, en hatalı ve en yanlı değerlerin açık ara farkla ROK mekanizmasına ait olduğu belirlenmiştir. Genel olarak en yansız ve en hatasız olarak tanımlanabilecek kayıp veri mekanizmasının ise TRK olduğu gözlenmiştir. Bu duruma TRK mekanizmasına sahip kayıp verilerin rastgele ve ihmal edilebilir olarak ortaya çıkması sebep olmuş olabilir. Bulguların yanı sıra RMSE ve BIAS değerleri üzerinde kayıp veri mekanizması incelemesi yapıldığında, farklılaşmanın en çok ROK mekanizmasına sahip olan kayıp verilerin YYGD ile çözümlenmeye çalışıldığı durumlarda ortaya çıktığı gözlenmiştir.

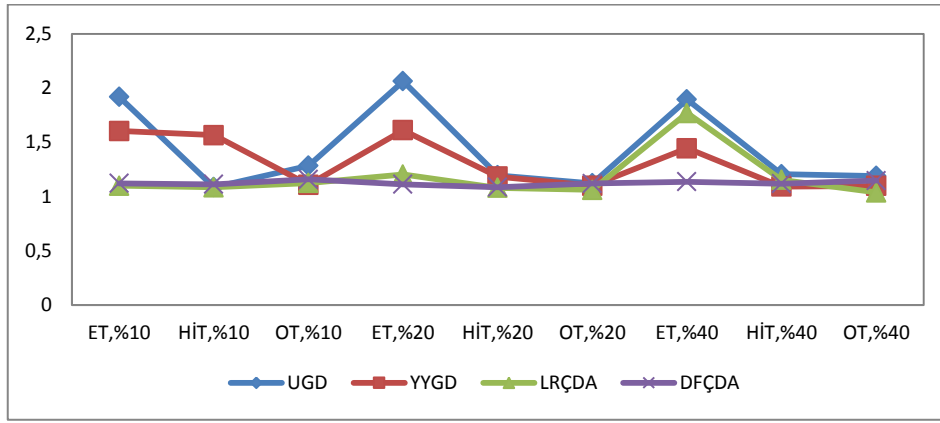
Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri kayıp verinin yerine göre incelendiğinde, genel olarak en küçük değerleri içeren en az hatalı ve en az yanlı sonuçların her iki test ve ortak teste ait olduğu gözlenmiştir. En büyük değerlere sahip olan sonuçların ise eşitlenecek testin kestirilen yetenek parametrelerine ait olduğu tespit edilmiştir. Bununla birlikte ROK mekanizması için en yanlı ve en hatalı sonuçları içeren kayıp veri yeri koşulu her iki test olarak kaydedilmiştir.

Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri kayıp verinin oranına göre incelendiğinde, %10'dan %40'a doğru gidildikçe hata oranlarının ve yanlılıkların arttığı gözlenmiştir. Ancak RMSE değerleri üzerinde apaçık bir şekilde gözlemlenebilen bu bulgular, BIAS değerleri için aynı netlikte değildir. Araştırmada ele alınan başa çıkma yöntemlerine ilişkin BIAS değerleri kayıp verinin oranına göre incelendiğinde ise, karşılaştırmalar arasında kayıt altına alınmaya değer tutarlılıklar gözlenememiştir. Bununla birlikte ROK tipi kayıp verilerdeki oran artışının sebep olduğu artan hata ve yanlılık değerleri, TRK ve RK için aynı netlikte gözlenememiştir.

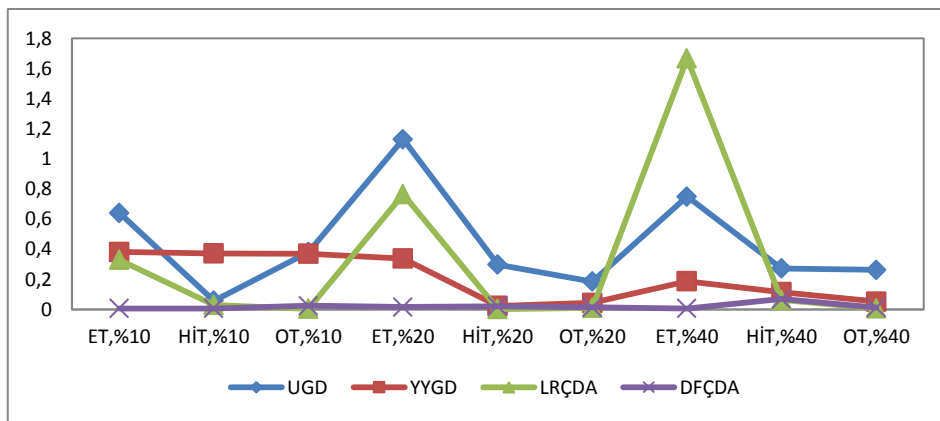
Araştırmada ele alınan başa çıkma yöntemlerine ilişkin elde edilen RMSE ve BIAS değerleri incelendiğinde, genel olarak en düşük hata ve yanlılık değerlerini üreten yöntemin LRÇDA olduğu ve DFÇDA'nın da hemen LRÇDA'nın ardından geldiği

tespit edilirken, en yüksek hata ve yanlılık değerlerini üreten simülasyon koşullarının ise ROK mekanizmasına ait kayıplar içeren veri setleri üzerinde YYGD'nin kullanıldığı koşullar olduğu görülmektedir. Bununla birlikte ROK mekanizmasına sahip kayıp verileri içeren veri setleri üzerinde yürütülen eşitlemeler sonucunda elde edilen RMSE değerlerine ait dizi genişliğinin oldukça büyük olduğu tespit edilmiştir. Raporlanması gereken bir başka bulgu ise, BIAS değerleri için genel olarak en düşük hata ve yanlılık değerlerini veren yöntemlerin LRÇDA ve DFÇDA olduğu ve en yanlı yöntemin ise YYGD çıktığıdır.

TRK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ilişkin RMSE değerleri Şekil 4.15, BIAS değerleri ise Şekil 4.16'da verilmiştir.



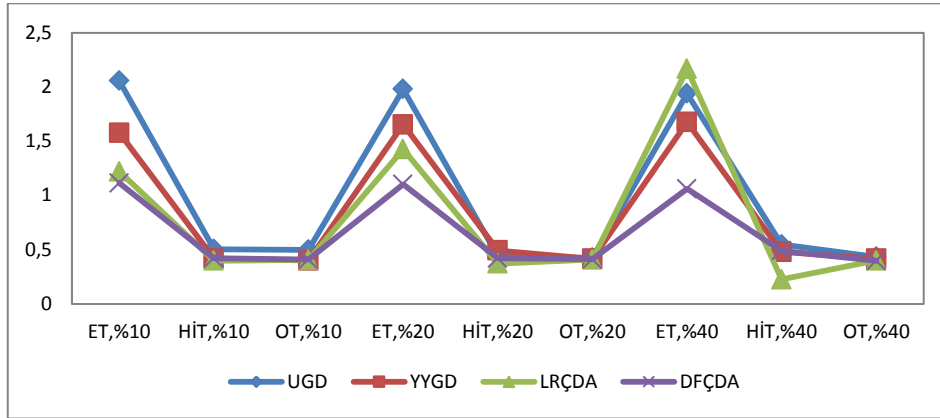
Şekil 4.15. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Hataları



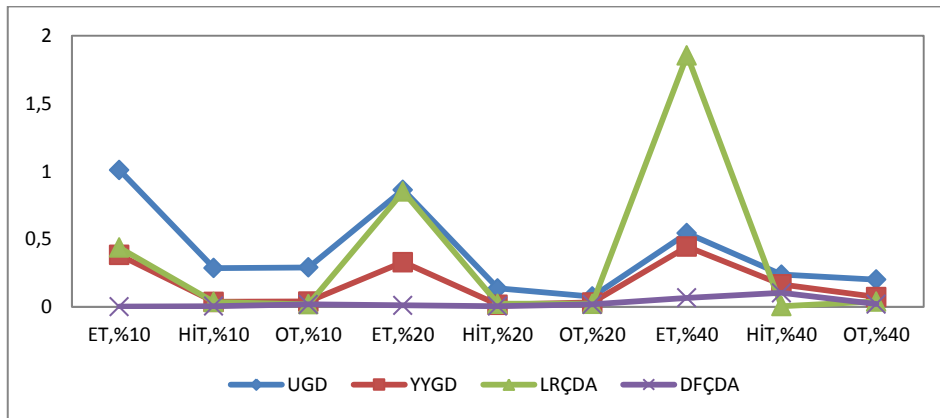
Şekil 4.16. TRK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Yanlılıkları

Şekil 4.15 genel hatlarıyla incelendiğinde, kayıp verilerin özellikle eşitlenecek test formu içerisinde olduğu koşullarda daha yüksek hata değerleri ürettiği görülmüştür. Bununla birlikte en yüksek hata değerlerini üreten kayıp veri başa çıkma yönteminin UGD olduğu saptanmıştır. Şekil 4.16 genel hatlarıyla incelendiğinde ise, LRÇDA yönteminin şaşırtıcı bir şekilde yüksek yanlılık değerleri ürettiği araştırma koşullarına rastlanmıştır. En düşük yanlılık değerlerini üreten başa çıkma yönteminin ise DFÇDA olduğu belirlenmiştir.

RK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ilişkin RMSE değerleri Şekil 4.17, BIAS değerleri ise Şekil 4.18'de verilmiştir.



Şekil 4.17. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Hataları

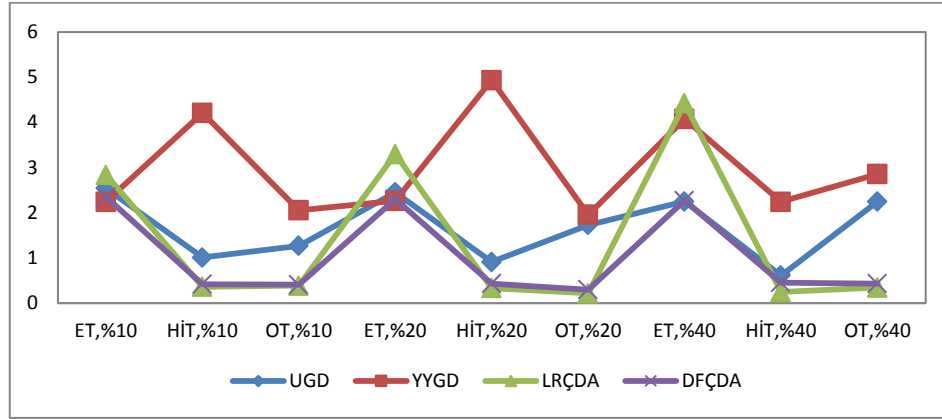


Şekil 4.18. RK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Yanlılıkları

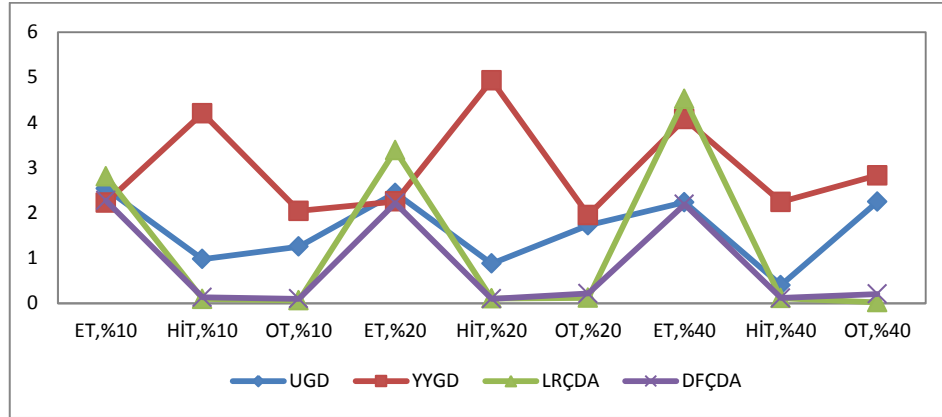
Şekil 4.17 genel hatlarıyla incelendiğinde, kayıp veri başa çıkma yöntemlerinin hemen hemen tüm koşullar kapsamında birbirleriyle tutarlı olacak şekilde hata değerleri ürettiği gözlenmiştir. Özellikle de UGD ve LRÇDA yöntemlerinin

performansları birbirlerine oldukça yakın çıkmıştır. Şekil 4.18 genel hatlarıyla incelendiğinde ise, kayıp verilerin eşitlenecek test içerisinde %40 oranında bulunduğu koşul için en yüksek yanlılık değerlerinin elde edildiği tespit edilmiştir. En düşük yanlılıkları üreten yöntemin ise yine DFÇDA olduğu belirlenmiştir.

ROK mekanizması kapsamında ele alınan simülasyon koşullarına göre kestirilen yetenek parametrelerine ilişkin RMSE değerleri Şekil 4.19, BIAS değerleri ise Şekil 4.20'de verilmiştir.



Şekil 4.19. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Hataları



Şekil 4.20. ROK Mekanizması Kapsamında Ele Alınan Simülasyon Koşullarına Göre Kestirilen θ Parametresine İlişkin Eşitleme Yanlılıkları

Şekil 4.19 genel hatlarıyla incelendiğinde, elde edilen hata değerleri bakımından tüm kayıp veri başa çıkma yöntemleri için tutarsızlıklar gözlenmiştir. Bunun yanı sıra, en düşük hata değerlerini üreten yöntemin DFÇDA olduğu saptanmıştır. Şekil 4.20 genel hatlarıyla incelendiğinde ise, ufak farklılıklar gözlemlense de, Şekil 4.19'da elde edilen hata değerlerine benzer bulguların aynı koşullarda yanlılık değerleri için de gözlemlendiği tespit edilmiştir.

Bu alt problem kapsamında, EK-10'daki *kayıp veriye sahip maddeleri içeren testler üzerinde UGD, YYGD, LRÇDA ve DFÇDA yöntemleri kullanılıp eşitleme yapıldığında kestirilen yetenek parametrelerine ilişkin RMSE ve BIAS değerleri* bağımlı, *araştırmada incelenen simülasyon faktörleri de (yer, mekanizma, oran)* bağımsız değişkenler olarak ele alındığında; bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkilerini belirlemek amacıyla ayrı ayrı ANOVA yapılmıştır. Yetenek parametreleri için anlamlı bulunan etkilerin başa çıkma yöntemlerine göre *F* değerleri ve etki büyüklükleri (η^2) Tablo 4.3'te verilmiştir.

Tablo 4.3: Yetenek Parametresine ait RMSE ve BIAS Değerleri için Anlamlı Bulunan ANOVA Sonuçları

Değerlendirme Ölçütleri	Etkiler	sd	Kayıp Veri Başa Çıkma Yöntemleri							
			UGD		YYGD		LRÇDA		DFÇDA	
			F	η^2	F	η^2	F	η^2	F	η^2
RMSE	Y	2	69.551*	.66	-	-	315.513*	.54	2058.804*	.47
	M	2	18.758*	.18	21.545*	.63	36.013*	.06	545.934*	.13
	Y*M	4	-	-	-	-	95.824*	.33	870.821*	.40
	Y*O	4	-	-	-	-	14.515*	.05	-	-
BIAS	Y	2	26.397*	.27	-	-	2448.515*	.54	1132.049*	.21
	M	2	54.841*	.57	48.156*	.82	624.542*	.14	1722.111*	.33
	O	2	-	-	-	-	138.362*	.03	-	-
	Y*M	4	-	-	-	-	522.382*	.23	1200.770*	.46
	Y*O	4	-	-	-	-	146.322*	.06	-	-

*Y: yer, M: mekanizma, O: oran, Y*M: yer-mekanizma etkileşimi, Y*O: yer-oran etkileşimi.

**p<0.002

Tablo 4.3'te bulunan yetenek parametrelerine ilişkin RMSE değerlerine ait ANOVA sonuçları anlamlı bulunan etkiler açısından incelendiğinde; yer ana etkisinin YYGD haricindeki diğer üç yöntemde de anlamlı olduğu, yer-mekanizma etkileşim etkisinin LRÇDA ve DFÇDA için anlamlı olduğu, mekanizma ana etkisinin her dört yöntem için de anlamlı olduğu, yer-oran etkileşim etkisinin sadece LRÇDA için anlamlı olduğu gözlenmiştir.

Tablo 4.3'te bulunan yetenek parametrelerine ilişkin BIAS değerlerine ait ANOVA sonuçları anlamlı bulunan etkiler açısından incelendiğinde; yer ana etkisinin YYGD haricindeki diğer üç yöntemde de anlamlı olduğu, mekanizma ana etkisinin her dört yöntem için de anlamlı olduğu, oran etkisinin ve yer-oran etkileşim etkisinin sadece LRÇDA için anlamlı olduğu ve yer-mekanizma etkileşim etkisinin ise hem LRÇDA hem de DFÇDA için anlamlı olduğu gözlenmiştir. Sonuç olarak mekanizma-oran etkileşim etkisi hiçbir yöntemin kestirilen yetenek parametrelerine ait ortalama RMSE ve BIAS değerleri üzerinde anlamlı bir etkiye sahip değildir. Burada bahsedilmesi gereken başka bir bulgu da, LRÇDA yöntemine ilişkin elde

edilen hem RMSE hem de BIAS deęerleri üzerinde, mekanizma-oran etkileşim etkisi hariç olmak üzere, tüm ana etkiler ve etkileşimlerin etkileri anlamlıdır.

4.4. Dördüncü Alt Probleme İlişkin Bulgular

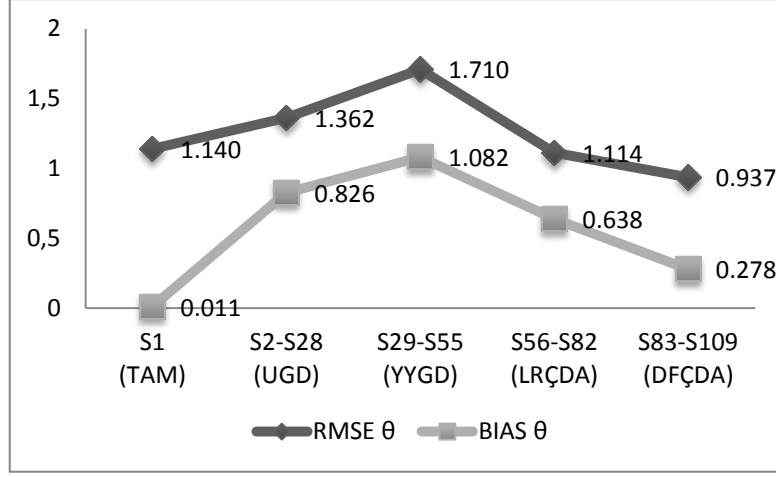
Tam veriye sahip test formları ile kayıp veri başa çıkma yöntemleri kullanılarak elde edilen test formları eşitlendiğinde kestirilen **yetenek parametrelerine** ilişkin eşitleme hatası ve eşitleme yanlılığı nasıl deęişmektedir?

Dördüncü alt problemin çözümlenmesinde; tam veriye sahip test formlarının eşitlendiği S1 koşulunda kestirilen yetenek parametrelerine ilişkin ortalama RMSE ve BIAS deęerleri, kayıp veri başa çıkma yöntemleri kullanılarak elde edilen yetenek parametre kestirimlerine ilişkin eşitleme hatalarının ve yanlılıklarının nasıl deęiştiğinin tespit edilmesi adına referans parametreler olarak kullanılmıştır. Ölçüt amaçlı işe koşulan bu deęerler Tablo 4.4'te verilmiştir.

Tablo 4.4: Tam Veriye Sahip Test Formlarının Eşitlenmesi ile Kestirilen Yetenek Parametrelerine İlişkin RMSE ve BIAS Deęerleri

<i>Ölçüt türü</i>	<i>Yetenek parametresi</i>
	θ
RMSE	1.140
BIAS	0.011

Şekil 4.21'deki grafikte S1 ise ifade edilen koşul tam veriye sahip test formlarının eşitlenmesi ile kestirilen yetenek parametrelerine ait ortalama hata ve yanlılık deęerlerini, S2-S109 aralığında ifade edilen koşullar ise kayıp veri içeren maddelere sahip test formlarının üzerinde farklı başa çıkma yöntemlerinin kullanılmasının ardından eşitlenmesi ile kestirilen yetenek parametrelerine ait ortalama hata ve yanlılık deęerlerini vermektedir. Yetenek parametrelerine ilişkin ortalama yanlılık deęerleri hesaplanırken BIAS deęerlerinin mutlak ifadeleri alınmıştır.



Şekil 4.21. Araştırma Koşulları Kapsamında Gerçekleştirilen Eşitlemeler Sonucunda Kestirilen Yetenek Parametrelerine ait Ortalama Eşitleme Hataları

Referans koşula ait kestirilen θ parametresi üzerinden hesaplanan ortalama RMSE değerine (1.140) en yakın ortalama $RMSE_{\theta}$ değerini (1.114) veren kayıp veri başa çıkma yöntemi S56 ile S82 koşulları arasında kullanılan LRÇDA'dır. Bu yöntemin hemen ardından gelen yöntem 0.937 değeri ile DFÇDA'dır. Ancak S83 ile S109 koşulları arasında kullanılan DFÇDA'nın en düşük ortalama $RMSE_{\theta}$ değerini veren kayıp veri başa çıkma yöntemi olduğunun tespitinin ardından, ortalama hata değerleri için yöntemler arasında $DFÇDA < LRÇDA < UGD < YYGD$ sıralaması yapılmıştır.

Referans koşula ait kestirilen θ parametresi üzerinden hesaplanan ortalama BIAS değerine (0.011) en yakın ve en düşük ortalama $BIAS_{\theta}$ değerini (0.278) veren kayıp veri başa çıkma yöntemi S83 ile S109 koşulları arasında kullanılan DFÇDA'dır. Bu yöntemin hemen ardından 0.638 değeri ile LRÇDA yöntemi gelmektedir. Bu nedenle, her iki yöntemin de en az yanlılık içeren θ parametre kestirimlerinin üretilmesine sebep oldukları söylenebilir. LRÇDA'nın arkasından gelen yöntem 0.826 değeri ile UGD ve en yanlı sonuçlar üreten yöntem ise 1.082 değeri ile YYGD çıkmıştır.

5. TARTIŞMA

Bu bölümde araştırmanın bulguları, hem alt problemler bazında değişkenlere yönelik karşılaştırmalar yapılarak hem de başka araştırmaların sonuçlarından yararlanılarak tartışılmıştır.

de Leeuw, Hox ve Huisman (2003) birlikte yürüttükleri çalışmalarında, veri setlerinde çeşitli sebeplerle ortaya çıkan kayıpların; ölçme işlemlerinin geçerlik ve güvenilirliğinin azalması, parametre kestirimlerinde yanlılık, şişirilmiş standart hatalar ve işe koşulan istatistiksel testlerin gücünün düşmesi gibi oldukça ciddi sorunları da beraberinde getirdiğini belirtmişlerdir. Bu sorunların, kayıp verinin ortaya çıkma sebebinin kaybın meydana geldiği değişken ile doğrudan ilişkili olduğu durumlarda daha da önemli ve göz ardı edilemez olduğunu ve büyük miktarlardaki kayıp verilerin istatistiksel testlerin gücünü önemli derecede tehlikeye düşürdüğünü ifade etmişlerdir. Rubin (1987), Allison (2002), Schafer ve Graham (2002), Peng, Harwell, Liou ve Ehman (2007), Enders (2010) ve Acock'da (2012) bu ifadeleri destekleyici nitelikli açıklamalar yapmışlardır. Aşağıda daha detaylı bir şekilde tartışılan bu araştırmanın bulguları da, özellikle büyük miktarlarda ve ROK mekanizmasına sahip kayıp verilerin varlığında, test eşitleme işlemlerinden elde edilen parametre kestirimlerine ait hata ve yanlılık değerlerinin daha yüksek çıktığını kanıtlamıştır.

Bu araştırmada simülasyon faktörleri (kayıp verinin yeri, mekanizması, oranı, başa çıkma yöntemleri) kapsamında incelenen tüm koşulların kestirilen madde ve yetenek parametrelerine ait ortalama hata (RMSE) ve yanlılık (BIAS) değerleri incelenmiş ve birbirleriyle karşılaştırılmıştır. "Kayıp verinin test içerisinde bulunduğu yer" faktörüne göre kestirilen parametrelere ilişkin en düşük ortalama RMSE ve BIAS değerlerinin, kayıpların her iki test (Form X ve Form Y) içerisinde bulunduğu koşullar; en yüksek ortalama RMSE ve BIAS değerlerinin ise kayıpların sadece eşitlenecek test (Form X) içerisinde bulunduğu koşullar için elde edildiği bulgusuna ulaşılmıştır. Bu faktör kapsamında elde edilen bulguları destekleyen veya çürüten başka araştırmalara ulaşılamasa da, eşitlemesi yapılan test formları içerisindeki kayıp veri miktarının en fazla olduğu (özellikle %40 kayıp veri oranı için) OT koşulunun, en yüksek ortalama hata ve yanlılık değerlerini üretmesi beklenmiştir. Bu beklentinin bir diğer sebebi ise veri toplama deseni olarak işe

koşulan ortak test deseninde, farklı gruplardan gelen farklı yetenek düzeylerindeki bireylere uygulanan farklı test formlarından elde edilen parametrelerin aynı ölçek üzerine yerleştirilmesi işleminin ortak maddeler üzerinden yapılmasıdır.

"Kayıp verinin sahip olduğu mekanizmaya" göre kestirilen parametrelere ilişkin en düşük ortalama RMSE ve BIAS değerlerinin, kayıpların TRK mekanizmasına uygun biçimde ortaya çıktığı koşullar; en yüksek ortalama RMSE ve BIAS değerlerinin ise, kayıp verilerin ROK mekanizmasına uygun biçimde ortaya çıktığı koşullar için elde edildiği bulgusuna ulaşılmıştır. Özellikle de yetenek parametreleri için elde edilen en yüksek RMSE ve BIAS değerlerinin, ROK mekanizmasına sahip kayıp veriler ile baş edilmesi sırasında UGD ve YYGD yöntemlerine başvurulduğu zaman ortaya çıktığı saptanmıştır. Bu bulgular doğrudan Lord (1974), De Ayala, Plake ve Impara (2001), Allison (2002), Rose, Davier ve Xu (2010), Finch (2008) ve Demir'in (2013) yürüttükleri çalışmaların bulguları ile benzerlik göstermektedir. Zira ilgili çalışmalarda da kayıp veri problemi ile başa çıkmada kullanılan en yaygın yöntemler olan UGD ve YYGD'nin, ancak rastgelelik varsayımının karşılandığı kayıp veri mekanizmaları için kullanılabilir olduğu belirtilmiştir.

Enders (2004), Akbaş (2014), Akbaş ve Tavşancıl'ın (2015) farklı kayıp veri mekanizmalarının performanslarını güvenilirlik kapsamında karşılaştırdıkları çalışmalarda da; ROK mekanizmasına sahip kayıp verileri içeren veri setlerine ait güvenilirlik kestirimlerinin, tam veri setlerine ait güvenilirlik kestirimlerine göre daha düşük ve negatif yanlı sonuçlar ürettiği bulgusuna ulaşılmıştır. ROK mekanizması ile ilgili tüm bu olumsuz bulguların yanı sıra, Sulis ve Porcu'da (2008) özellikle TRK mekanizmasına sahip kayıp verilerin bulunduğu veri setlerine ait kestirimlerin daha etkili ve daha yansız çıktığını vurgulamışlardır.

"Kayıp verinin oranına" göre kestirilen parametrelere ilişkin en düşük ortalama RMSE ve BIAS değerlerinin, kayıpların %10 oranında ortaya çıktığı koşullar; en yüksek ortalama RMSE ve BIAS değerlerinin ise, kayıpların %40 oranında ortaya çıktığı koşullar için elde edildiği bulgusuna ulaşılmıştır. Bu bulgudan hareketle veri setleri içerisindeki kayıp verilerin miktarı arttıkça, hata ve yanlılıkların da arttığı gözlenmiştir. Bulgular farklı özelliklere sahip veri setlerinde ortaya çıkan artan miktarlardaki kayıp verilerin; bireylerin yetenek düzeyleri, maddelerin parametre değerleri ve standart hata değerlerinin kestirim gücü üzerindeki olumsuz etkilerini

gösteren çalışmaları destekler niteliktedir (Enders, 2003, 2004; Acock, 2005; Ake, 2005; Van Ginkel, 2007; Finch, 2008; Sulis & Porcu, 2008; Rose, Davier & Xu, 2010; Ma, Akhtar-Danesh, Dolovich & Thabane, 2011; Wu, Jia & Enders, 2015). Tüm bu çalışmalar genel hatlarıyla, daha düşük kayıp veri oranlarında elde edilen kestirim sonuçlarının, daha yüksek kayıp veri oranlarında elde edilenlere göre daha az yanlı ve daha az hatalı olduğunu göstermektedir. Kayıp veri oranı ile ilgili olarak elde edilen ve göze çarpan bir diğer bulgu ise, özellikle OT içerisinde %40 oranında ROK mekanizmasına sahip kayıp veri içeren koşullar için kestirilen parametrelere ait hata ve yanlılık miktarlarının oldukça yüksek olduğudur. Bu bulgu Rose, Davier ve Xu (2010) ile Doğanay Erdoğan'ın (2012) yürüttüğü ve ROK mekanizmasına sahip kayıpların veri seti içerisindeki oranının %30'dan %50'ye yükseldiği durumlarda ciddi derecede yanlı kestirimlerin elde edildiğine yönelik ifadeler ile paralellik göstermektedir.

"Kayıp veri ile başa çıkma yöntemlerine" göre kestirilen parametrelere ilişkin en düşük ortalama RMSE ve BIAS değerlerinin, ÇDA'ya dayalı LR ve DF modellerinin kullanıldıkları koşullar; en yüksek ortalama RMSE ve BIAS değerlerinin ise, UGD ve YYGD yöntemlerinin kullanıldıkları koşullar için elde edildiği bulgusuna ulaşılmıştır. UGD ile YYGD yöntemleri arasında çok ciddi farklılıklar olmasa da, genel olarak YYGD'ye ait hata ve yanlılık değerlerinin UGD yöntemine göre daha yüksek çıktığı gözlenmiştir. Bu iki yöntemle ilişkin bulgular De Ayala, Plake ve Impara (2001), Finch (2008) ile Hohensinn ve Kubinger'in (2011) yürüttüğü çalışmalarda bulgular ile tutarlılık göstermektedir. ÇDA yöntemi ile ilgili olarak elde edilen bulgular ise, kayıp verileri içeren değişken ikili olarak puanlandığında ÇDA yönteminin kullanılmasını ve çoklu değerleri atama aşaması için de LR ya da DF modellerinin işe koşulması önerilerini desteklemektedir (Allison, 2005; van Buuren & Groothuis-Oudshoorn, 2011; Wu, Jia & Enders, 2015). Uygun kayıp veri başa çıkma yönteminin kullanılmadığı durumlarda ortaya çıkan hata ve yanlılık değerlerinin de, yürütülen çalışmanın güvenilirlik ve geçerlik gibi önemli niteliklerini olumsuz yönde etkileyeceği unutulmamalıdır.

Allison (2002) hiçbir kayıp veri yönteminin diğerlerine göre daha iyi veya daha üstün sayılmayacağını ve farklı koşullar altında yürütülen çalışmalarda farklı sonuçların elde edilebileceğini dile getirmiştir. Bu aynı zamanda, bir çalışmanın bulgularına göre en iyi olan X yönteminin, başka bir çalışma için en kötü sonuçları

üreten yöntem olabileceği anlamına gelmektedir. Bu nedenle de, kayıp veriler ile başa çıkmada işe koşulacak en doğru yaklaşım, hiç kaybın olmaması yönünde çabaların sarf edilmesidir. Özellikle de kayıp verinin miktarının fazla olduğu durumlarda, izleme (follow-up) yöntemi ile kayıp olarak elde edilen değerlerin tamamlanması ve bu çabaların yetersiz kaldığı durumlarda ise veri toplama süreçlerinin tekrarlanması önerilmektedir. Benzer şekilde de Leeuw, Hox ve Huisman'da (2003), veri setinde kayıpların olmasının önüne tam anlamıyla geçilemeyeceğini ve bu gerekçeyle kaybın ortaya çıkma nedenine ilişkin sorgulamaların yapılması gerektiğini, aksi takdirde özensiz bir şekilde yürütülen araştırmalar için özensiz sonuçların elde edileceğini ve bu sonuçlar üzerinde yapılacak istatistiksel düzeltmelerin de bir anlamı olmayacağını belirtmişlerdir.

6. SONUÇ ve ÖNERİLER

Bu bölümde, araştırmanın bulgularına dayalı olarak ulaşılan sonuçlara ve bu sonuçlardan yola çıkarak geliştirilen önerilere yer verilmiştir.

6.1. Sonuçlar

Araştırmada kayıp verilerin test formları içerisindeki yeri, mekanizması, oranı ve başa çıkma yöntemlerine göre oluşturulan farklı koşullardaki testlerden elde edilen puanların, ortak test deseni kapsamında ve MTK'ya dayalı SL yöntemi kullanılarak eşitlenmesi sonucunda kestirilen madde ve yetenek parametrelerine ilişkin eşitleme hatası (RMSE) ve eşitleme yanlılığı (BIAS) değerlerini karşılaştırmak amaçlanmıştır. Bu amaç doğrultusunda elde edilen araştırma bulgularına dayalı sonuçlar aşağıda verilmiştir.

Araştırmada ele alınan simülasyon koşulları dikkate alındığında;

- Başa çıkma yöntemlerine ilişkin elde edilen ortalama RMSE ve BIAS değerlerinin incelemesi *kayıp verilerin test içerisinde bulunduğu yere* göre yapıldığında, genel olarak en küçük değerlere sahip en az hata ve yanlılık içeren sonuçların HİT'in kestirilen madde ve yetenek parametrelerine ait olduğu, en büyük değerlere sahip en hatalı ve en yanlı sonuçların ise neredeyse tüm koşullar için ET'nin kestirilen madde ve yetenek parametrelerine ait olduğu saptanmıştır. Kayıpların yeri OT içerisinde olduğunda ise, ET'den daha küçük ancak HİT'den daha büyük hata ve yanlılık değeri üretilmiştir ($\bar{X}_{hata,HİT} < \bar{X}_{hata,OT} < \bar{X}_{hata,ET}$).
- Başa çıkma yöntemlerine ilişkin elde edilen ortalama RMSE ve BIAS değerlerinin incelemesi *kayıp verilerin mekanizmasına* göre yapıldığında, genel olarak en hatalı ve yanlı sonuçların ROK mekanizmasına ait olduğu, en yansız ve hatasız olarak tanımlanabilecek kayıp veri mekanizmasının ise TRK olduğu görülmüştür ($\bar{X}_{hata,TRK} < \bar{X}_{hata,RK} < \bar{X}_{hata,ROK}$). Özellikle de yetenek parametreleri kullanılarak elde edilen en yüksek RMSE ve BIAS değerlerinin, ROK mekanizmasına sahip kayıp verilerin bulunduğu veri setleri üzerinde UGD ve YYGD yöntemlerinin kullanıldığı koşullarda ortaya çıktığı görülmüştür.

- Başa çıkma yöntemlerine ilişkin elde edilen ortalama RMSE ve BIAS değerlerinin incelemesi *kayıp verilerin test içerisinde bulunma oranına* göre yapıldığında, genel olarak kayıp veri miktarının %10'dan %20 ve %40'a doğru arttığı koşullarda hataların ve yanlılıkların da arttığı saptanmıştır ($\bar{X}_{hata,\%10} < \bar{X}_{hata,\%20} < \bar{X}_{hata,\%40}$). Özellikle de OT içerisinde %40 oranında kayıplar içeren koşullara ilişkin kestirilen parametrelere ait hata ve yanlılık değerlerinin oldukça yüksek oldukları sonucuna ulaşılmıştır.
- Başa çıkma yöntemlerine ilişkin elde edilen ortalama RMSE ve BIAS değerlerinin incelemesi *kayıp veri sorununun çözümünde işe koşulan başa çıkma yöntemine* göre yapıldığında, tüm koşullar için en düşük hata ve yanlılık değerlerini üreten yöntemlerin ÇDA'ya dayalı LRÇDA ve DFÇDA oldukları, en yüksek hata ve yanlılık değerlerini üreten yöntemlerin ise UGD ve YYGD oldukları saptanmıştır. UGD ile YYGD yöntemleri arasında çok ciddi farklılıklar olmasa da, genel olarak YYGD'ye ait hata ve yanlılık değerlerinin UGD'ye göre biraz daha yüksek çıktığı sonucuna ulaşılmıştır ($\bar{X}_{hata,LRÇDA} \& \bar{X}_{hata,DFÇDA} < \bar{X}_{hata,UGD} < \bar{X}_{hata,YYGD}$).

Araştırmada ele alınan koşulların ana ve ortak etkileri dikkate alındığında;

- Madde parametrelerine ilişkin ortalama RMSE değerlerine ait ANOVA sonuçları incelendiğinde; genel olarak yer ve mekanizma ana etkilerinin tüm yöntemler için anlamlı bulunduğu, yer-mekanizma etkileşim etkisinin UGD ve YYGD yöntemine ait kestirilen *b* parametreleri için anlamlı bulunduğu saptanmıştır.
- Madde parametrelerine ilişkin ortalama BIAS değerlerine ait ANOVA sonuçları incelendiğinde; genel olarak yer ve mekanizma ana etkilerinin tüm yöntemler için anlamlı bulunduğu, yer-mekanizma etkileşim etkisinin UGD yöntemine ait kestirilen *a* parametreleri ile LRÇDA ve DFÇDA yöntemine ait kestirilen *b* parametreleri için anlamlı bulunduğu saptanmıştır. Bununla birlikte *a* parametresi kapsamında YYGD ve DFÇDA için anlamlı sonuç veren herhangi bir etki bulunmazken, *b* parametresi kapsamında UGD ve YYGD yöntemleri için ise hiçbir etkinin anlamlı olmadığı tespit edilmiştir.
- Yetenek parametrelerine ilişkin ortalama RMSE değerlerine ait ANOVA sonuçları incelendiğinde; yer ana etkisinin YYGD haricindeki diğer üç

yöntemde de anlamlı olduğu, yer-mekanizma etkileşim etkisinin LRÇDA ve DFÇDA için anlamlı olduğu, mekanizma ana etkisinin her dört yöntem için de anlamlı olduğu, yer-oran etkileşim etkisinin sadece LRÇDA için anlamlı olduğu gözlenmiştir.

- Yetenek parametrelerine ilişkin ortalama BIAS değerlerine ait ANOVA sonuçları incelendiğinde; yer ana etkisinin YYGD haricindeki diğer üç yöntemde de anlamlı olduğu, mekanizma ana etkisinin her dört yöntem için de anlamlı olduğu, oran etkisinin ve yer-oran etkileşim etkisinin sadece LRÇDA için anlamlı olduğu ve yer-mekanizma etkileşim etkisinin ise hem LRÇDA hem de DFÇDA için anlamlı olduğu gözlenmiştir. Tüm bunlarla birlikte mekanizma-oran etkileşim etkisi hiçbir yöntemin kestirilen yetenek parametrelerine ait ortalama RMSE ve BIAS değerleri üzerinde anlamlı bir etkiye sahip değildir. En çok sayıda anlamlı etki LRÇDA yöntemine ait kestirilen yetenek parametrelerinin eşitleme hataları üzerinde iken, en az sayıda anlamlı etki ise YYGD yöntemi için elde edilmiştir.

Tam veriler kapsamında gerçekleştirilen eşitlemeler ile simülasyon koşulları kapsamında gerçekleştirilen eşitlemeler karşılaştırıldığında;

- Araştırmanın referans koşulunu oluşturan tam veri setlerine sahip test formlarının eşitlenmesi ile kestirilen madde ve yetenek parametreleri üzerinden hesaplanan ortalama hata ve yanlılık değerlerine en yakın ve en düşük ortalama hata ve yanlılık değerlerini veren kayıp veri başa çıkma yönteminin genel olarak DFÇDA olduğu saptanmıştır. Bu yöntemin hemen ardından gelen LRÇDA yöntemi de hem DFÇDA yöntemi kullanılan test formlarından hem de tam veriye sahip test formlarından elde edilen eşitleme sonuçlarına oldukça yakın sonuçlar vermiştir. Dolayısıyla her iki ÇDA yönteminin de en az hata ve yanlılık içeren parametre kestirimlerinin üretilmesine sebep oldukları söylenebilir.
- Araştırmanın referans koşulunu oluşturan tam veri setlerine sahip test formlarının eşitlenmesi ile kestirilen madde ve yetenek parametreleri üzerinden hesaplanan ortalama hata ve yanlılık değerlerine en uzak ve en yüksek ortalama hata ve yanlılık değerlerini veren kayıp veri başa çıkma

yöntemlerinin ise birbirlerine oldukça benzer sonuçlar veren UGD ve YYGD olduğu sonucuna ulaşılmıştır.

6.2. Öneriler

Bu bölümde araştırma sonuçları doğrultusunda geliştirilen önerilere yer verilmiştir.

6.2.1. Araştırmaya Dönük Öneriler

Araştırma farklı koşullara sahip kayıp verilerin test eşitleme üzerindeki etkilerini inceleyerek aşağıdaki önerilerin üretilmesine sebep olmuştur.

1. Özellikle tamamen rastgele olmayan kayıp veri mekanizmasına ve %20'den yüksek kayıp veri oranlarına sahip test formlarının işe koşulması ile yürütülecek eşitleme işlemlerinden önce, veri setlerinde bulunan kayıplar ile başa çıkılması için birçok istatistik programının öntanımlı olarak önerdiği ve işe koştuğu parametre kestirimine dayalı UGD ve YYGD yöntemlerinin kullanılmaması önerilmektedir.
2. İkili (1/0) puanlanan maddelerden oluşan veri setlerine sahip test formlarının işe koşulması ile yürütülecek eşitleme işlemlerinden önce, kayıp veriler ile başa çıkılması için çoklu değer atamaya dayalı olan LR veya DF yöntemlerine başvurulması önerilmektedir.
3. Herhangi bir eşitleme çalışmasına başlanmadan önce, test formlarına ait veri setlerinin içerdiği kayıp veriler için bir çözüm yöntemi kullanılmadan; kayıpların mekanizmasına, ortaya çıktıkları yere ve oranına ilişkin bilgilerin incelenmesi gerekmektedir. Özellikle rastgele olarak ortaya çıkmayan kayıp veri mekanizmasına sahip ve büyük miktarlardaki kayıp veriler için izleme (follow-up) işlemlerinin yapılıp mümkün olunan ölçüde daha çok veriye ulaşılması, bu çabanın yeterli olmadığı durumlarda ise veri toplama süreçlerinin tekrarlanması önerilmektedir.

6.2.2. Yeni Araştırmalara Dönük Öneriler

1. Farklı şekillerde ortaya çıkan kayıp verilerin test eşitleme üzerindeki etkilerini incelemeyi amaçlayan bu araştırma kapsamında; üç farklı kayıp veri yeri (ET, HİT, OT), üç farklı kayıp veri mekanizması (TRK, RK, ROK), üç farklı kayıp veri oranı (%10, %20, %40) ve dört farklı kayıp veri başa çıkma yöntemi (UGD, YYGD, LRÇDA, DFÇDA) ele alınmıştır. Araştırmanın

referans koşulunu temsil eden ve hiç kayıp veriye sahip olmayan S1 koşulu da tüm simülasyon koşullarına eklenerek toplam 109 farklı koşul ile çalışılmasına sebep olmuştur. Çalışmada sabit tutulanlar ise eşitleme deseni (ortak test), eşitleme yöntemi (SL), eşitlenen test formlarının uzunlukları (100'er madde), ortak madde sayıları (20'er madde), örneklem büyüklükleri (2000'er birey) ve test formlarını alan bireylerin yetenek dağılımları ($\theta \sim N(0,1)$). Bundan sonra yürütülecek çalışmalarda burada sabit bırakılanların her biri birer değişkene çevrilebilir.

2. Araştırmada kayıp verinin yeri faktöründe ele alınan koşullara göre yaratılan tüm kayıplar test formlarındaki ortak maddelerin kayıp veriye sahip olacağı şekilde tasarlanmıştır. Benzer çalışmalar, kayıp veri içerme bakımından ortak maddeleri süreç dışında bırakabilir.
3. Araştırmada 1 ve 0 şeklinde puanlanan iki kategorili veriler ile çalışılmıştır. Benzer koşullar altında, ikiden fazla şekilde puanlanan çok kategorili verilerin de işe katıldığı eşitleme süreçleri gerçekleştirilebilir.
4. Araştırmada RMSE ve BIAS olmak üzere iki farklı değerlendirme ölçütü işe koşulmuştur. Gelecekte yapılacak çalışmalarda; RMSD (root mean square difference), MSE (mean squared error), SEE (standart equating error) adlı değerlendirme ölçütleri de kullanılabilir.
5. Araştırma simülasyon verileri üzerinde yürütülmüştür. Benzer amaçların güdüldüğü çalışmalar, gerçek veri setleri ile simülasyon verilerini birlikte ele alacak şekilde yürütülebilir.

KAYNAKÇA

- Acock, A. (2005). Working with missing values. *Journal of Marriage and Family*, 67(4), 1012-1028.
- Acock, A. (2012). What to do about missing values. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf and K. J. Sher (Eds.). *APA handbook of research methods in psychology: Vol 3. Data analysis and research publication*, 27-50. Washington: American Psychological Association.
- Akbaş, U. ve Tavşancıl, E. (2015). Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 38-57.
- Akbaş, U. (2014). *Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi*. (Yayımlanmamış Doktora Tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ölçme ve Değerlendirme Anabilim Dalı, Ankara.
- Ake, C. F. (2005). Rounding after multiple imputation with non-binary categorical covariates. *SUGI 30 Proceedings*, 112(30), 1-11.
- Allison, P. D. (2002). *Missing data*. California: Sage Publication, Inc.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, 112(4), 545-557.
- Allison, P. D. (2005). Imputation of categorical variables with PROC MI. *SUGI 30 Proceedings*, 113(30), 1-14.
- Allison, P. D. (2009). Missing data. R. E. Millsao and A. Maydeu-Olivares (Eds.). *Quantitative methods in psychology*, 72-89. London: SAGE Publication.
- Ambler, G., Omar, R. Z., & Royston, P. (2007). A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome. *Statistical Methods in Medical Research*, 16(3), 277-298.
- Angoff, W. H. (1971). Scales, norms and equivalent scores. In R. L. Thorndike (Ed.). *Educational measurement*, 508-600. Washington: American Council on Education.
- Baker, F. B., & Al-Karni, A. (1991). A comparison of two procedures for computing IRT equating coefficients. *Journal of Educational Measurement*, 28(2), 147-162.
- Baker, F. B. (2001). *The basics of item response theory*. Washington: ERIC.
- Bal, C. (2003). *Çok gruplu veri setlerinde eksik gözlem sorununun çözümlenmesi ve sağlık alanında bir uygulama*. (Yayımlanmamış Doktora Tezi). Eskişehir Osmangazi Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ABD, Eskişehir.
- Bernaards, C. A., & Sijtsma, K. (1999). Factor analysis of multidimensional polytomous item response data suffering from ignorable item nonresponse. *Multivariate Behavioral Research*, 34(3), 277-313.

- Bernaards, C. A., & Sijtsma, K. (2000). Influence of imputation and em methods on factor analysis when item nonresponse in questionnaire data is nonignorable. *Multivariate Behavioral Research*, 35(3), 321-364.
- Bränberg, K. (2010). *Observed score equating with covariates*. (Unpublished Doctoral Dissertation). University of Umeå.
- Brown, R. L. (1994). Efficacy of the indirect approach for estimating structural equation models with missing data: A comparison of methods. *Structural Equation Modeling: A Multidisciplinary Journal*, 1(4), 287-316.
- Büyüköztürk, Ş., & Çokluk-Bökeoğlu, Ö. (2008). Discriminant function analysis: Concept and application. *Eurasian Journal of Educational Research (EJER)*, 33, 73-92.
- Chen, S. F., Wang, S., & Chen, C. Y. (2011). A simulation study using EFA and CFA programs based on the impact of missing data on test dimensionality. *Expert Systems With Applications*, 39(2012), 4026-4031.
- Cook, L. L., & Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice*, 10(3), 37-45.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston Inc.
- Culbertson, M. J. (2011). *Is it wrong? Handling missing responses in IRT*. 12th Annual Meeting of the National Council on Measurement in Education. Westin New Orleans Canal Place Hotel New Orleans Louisiana, 9-11 April 2011. [Çevrim-içi: <https://www.ideals.illinois.edu/bitstream/handle/2142/18819/ncme-2011-culbertson.pdf?sequence=2&isAllowed=y>, Erişim tarihi: 11 Mayıs 2016.]
- Çüm, S. ve Gelbal, S. (2015). Kayıp veriler yerine yaklaşık değer atamada kullanılan farklı yöntemlerin model veri uyumu üzerindeki etkisi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 1(35), 87-111.
- Çokluk, Ö. ve Kayrı, M. (2011). Kayıp değerlere yaklaşık değer atama yöntemlerinin ölçme araçlarının geçerlik ve güvenilirliği üzerindeki etkisi. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(1), 289-309.
- Çokluk, Ö. (2010). Lojistik regresyon analizi: Kavram ve uygulama. *Kuram ve Uygulamada Eğitim Bilimleri*, 10(3), 1357-1407.
- De Ayala, R. J., Plake, B. S., & Impara, J.C. (2011). The impact of omitted responses on the accuracy of ability estimation in item response theory. *Journal of Educational Measurement*, 38(3), 213-234.
- de Leeuw, E. D., Hox, J., & Huisman, M. (2003). Prevention and treatment of item nonresponse. *Journal of Official Statistics*, 19(2), 153-176.
- Delucchi, K. L. (1994). Methods for the analysis of binary outcome results in the presence of missing data. *Journal of Consulting and Clinical Psychology*, 62(3), 569-575.
- Demir, E. ve Parlak, B. (2012). Türkiye'de eğitim araştırmalarında kayıp veri sorunu. *Eğitimde Ve Psikolojide Ölçme Ve Değerlendirme Dergisi*, 3(1), 230-241.

- Demir, E. (2013). *Kayıp verilerin varlığında iki kategorili puanlanan maddelerden oluşan testlerin psikometrik özelliklerinin incelenmesi*. (Yayımlanmamış Doktora Tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Doğanay-Erdoğan, B. (2012). *Çoklu atama yöntemlerinin Rasch modelleri için performansının benzetim çalışması ile incelenmesi*. (Yayımlanmamış Doktora Tezi). Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.
- Dorans, N. J., Moses, T., & Eignor, D. (2010). Principles and practices of test score equating. *ETS Research Report Series, 2010(2)*, i-41.
- du Toit, M. (2003). *IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, TESTFACT*. Lincolnwood IL: Scientific Software International.
- Dural, S. (2010). *Farklı kayıp veri tekniklerinin çok göstergeli örtük büyüme modelleri üzerindeki etkisi*. (Yayımlanmamış Doktora Tezi). Ege Üniversitesi, Sağlık Bilimleri Enstitüsü, İzmir.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah: Lawrence Erlbaum.
- Emenogu, B. C., Falenchuk, O., & Childs, R. A. (2010). The effect of missing data treatment on Mantel-Haenszel DIF detection. *The Alberta Journal of Educational Research, 56(4)*, 459-469.
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 8(3)*, 430-457.
- Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods, 8(3)*, 322-337.
- Enders, C. K. (2004). The impact of missing data on sample reliability estimates: Implications for reliability reporting practices. *Educational and Psychological Measurement, 64(3)*, 419-436.
- Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Publications, Inc.
- Felan, G. D. (2002). *Test equating: Mean, linear, equipercentile and Item Response Theory*. Annual Meeting of the Southwest Educational Research Association. Austin Texas, 14-16 February 2002. [Çevrim-içi: <http://files.eric.ed.gov/fulltext/ED462436.pdf>, Erişim tarihi: 8 Ağustos 2015.]
- Finch, H. (2008). Estimation of item response theory parameters in the presence of missing data. *Journal of Educational Measurement, 45(3)*, 225-245.
- Finch, H., & Margraf, M. (2008). *Imputation of categorical missing data: A comparison of multivariate normal and multinomial methods*. [Çevrim-içi: www.mwsug.org/proceedings/2008/stats/MWSUG-2008-S05.pdf, Erişim tarihi: 18 Mayıs 2015.]
- Finch, H. (2010). Imputation methods for missing categorical questionnaire data: A comparison of approaches. *Journal of Data Science, 8(2010)*, 361-378.

- Gök, B. (2012). *Denk olmayan gruplarda ortak madde deseni kullanılarak madde tepki kuramına dayalı eşitleme yöntemlerinin karşılaştırılması*. (Yayımlanmamış Doktora Tezi). Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Graham, J. W. (2009). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology*, 60(4), 549-576.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park: Sage.
- Hanson, B. A., & Beguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common-item equating design. *Applied Psychological Measurement*, 26(1), 3-24.
- Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2007). Analysis of binary outcomes with missing data: missing=smoking, last observation carried forward, and a little multiple imputation. *Addiction*, 102(10), 1564-1573.
- Hohensinn, C., & Kubinger K. D. (2011). On the impact of missing values on item fit and the model validness of the Rasch model. *Psychological Test and Assessment Modeling*, 53(3), 380-393.
- Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.). *Educational measurement*, 187-220. Westport: Praeger Publishers.
- Holland, P. W. (2007). A framework and history for score linking. In N. J. Dorans, M. Pommerich and P. W. Holland (Eds.). *Linking and aligning scores and scales*, 5-30. New York: Springer.
- Holland, P. W., Sinharay, S., von Davier, A. A., & Han, N. (2008). An approach to evaluating the missing data assumptions of the chain and post-stratification equating methods for the NEAT design. *Journal of Educational Measurement*, 45(1), 17-43.
- Honaker, J., & King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(2), 561-581.
- İnan, G. (2009). *A simulation study on the comparison of methods for the analysis of longitudinal count data*. (Unpublished Master Thesis). University of Middle East Technical University.
- Karasar, N. (2009). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayınevi.
- Kiefer T., Robitzsch A., & Wu M. (2017). *Package 'TAM'. (Version 1.99993-0)* [Çevrimiçi: <https://cran.r-project.org/web/packages/TAM/TAM.pdf>, Erişim tarihi: 21 Şubat 2017.]
- Kilmen, S. (2010). *Madde tepki kuramına dayalı test eşitleme yöntemlerinden kestirilen eşitleme hatalarının örneklem büyüklüğü ve yetenek dağılımına göre karşılaştırılması*. (Yayımlanmamış Doktora Tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

- Kim, J. (2007). *A comparison of calibration methods and proficiency estimators for creating IRT vertical scales*. (Unpublished Doctoral Dissertation). University of Iowa.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices*. New York: Springer.
- Kolen, M. J. (1981). Comparison of traditional and item response theory methods for equating tests. *Journal of Educational Measurement*, 18(1), 1-11.
- Kolen, M. J. (1988). Traditional equating methodology. *Educational Measurement: Issues and Practice*, 7(4), 29-36.
- Kropko, J., Goodrich, B., Gelman, A., & Hill, J. (2014). Multiple imputation for continuous and categorical data: Comparing joint multivariate normal and conditional approaches. *Political Analysis*, 22(4), 497-519.
- Liou, M., & Cheng, P. E. (1995). Equipercentile equating via data-imputation techniques. *Psychometrika*, 60(1), 119-136.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Lord, F. M. (1974). Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, 39(2), 247-264.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale: Lawrence Erlbaum.
- Lord, F.M., & Novick, M.R. (1968). *Statistical theory of mental test score*. California: Addison-Wesley Publishing.
- Ludlow, L. H., & O'leary, M. (1999). Scoring omitted and not-reached items: Practical data analysis implications. *Educational and Psychological Measurement*, 59(4), 615-630.
- Ma, J., Akhtar-Danesh, N., Dolovich, L., & Thabane, L. (2011). Imputation strategies for missing binary outcomes in cluster randomized trials. *BMC Medical Research Methodology*, 11(18), 1471-2288.
- Madow, W. G., Nisselson, H., Olkin, I., & Rubin, D. B. (1983). *Incomplete data in sample surveys: Theory and bibliographies*. New York: Academic Press.
- Marsh, H. W. (1998). Pairwise deletion for missing data in structural equation models: Nonpositive definite matrices, parameter estimates, goodness of fit, and adjusted sample sizes. *Structural Equation Modeling: A Multidisciplinary Journal*, 5(1), 22-36.
- McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. New York: The Guilford Press.
- Mislevy, R. J., & Wu, P. K. (1996). *Missing responses and IRT ability estimation: omits, choice, time limits, adaptive testing*. (ETS Research Report RR-96-30-ONR). Princeton: Educational Testing Service.

- Miyazaki, K., Hoshino, T., Mayekawa, S. I., & Shigemasu, K. (2009). A new concurrent calibration method for nonequivalent group design under nonrandom assignment. *Psychometrika*, 74(1), 1-19.
- Partchev, I. (2016). *Package 'irtoys'. (Version 0.2.0)* [Çevrim-içi: <https://cran.r-project.org/web/packages/irtoys/irtoys.pdf>, Erişim tarihi: 15 Aralık 2016.]
- Peng, C. Y. J., Harwell, M., Liou, S. M., & Ehman L. H. (2007). Advances in missing data methods and implication for educational research. In S. S. Sawilowski (Ed.). *Real data analysis*, 31-77. USA: Information Age Publishing.
- Puhan, G. (2010). A comparison of chained linear and poststratification linear equating under different testing conditions. *Journal of Educational Measurement*, 47(1), 54-75.
- Rizopoulos, D. (2015). *Package 'ltm'. (Version 1.0-0)* [Çevrim-içi: <https://cran.r-project.org/web/packages/ltm/ltm.pdf>, Erişim tarihi: 17 Kasım 2016.]
- Rose N., von Davier, M., & Xu, X. (2010). Modeling nonignorable missing data with item response theory (IRT). *ETS Research Report Series*, 2010(1), i-53.
- Roth, P. L., & Switzer, F. S. (1995). A Monte Carlo analysis of missing data techniques in a HRM setting. *Journal of Management*, 21(5), 1003-1023.
- Rubin, D. B., (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall/CRC.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177.
- Schafer, J. L., Maren K., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, 33 (4), 545-571.
- Shin, S. H. (2009). How to treat omitted responses in Rasch model-based equating. *Practical Assessment Research & Evaluation*, 14(1), 1-8.
- Sinharay, S., & Holland, P. W. (2009). The missing data assumptions of the nonequivalent groups with anchor test (NEAT) design and their implications for test equating. *ETS Research Report Series*, 2009(1), i-53.
- Sijtsma, K., & Van der Ark, L. A. (2003). Investigation and treatment of missing item scores in test and questionnaire data. *Multivariate Behavioral Research*, 38(4), 505-528.
- Smolkowski, K., Danaher, B. G., Seeley, J. R., Kosty, D. B., & Severson, H. H. (2010). Modeling missing binary outcome data in a successful web-based smokeless tobacco cessation program. *Addiction*, 105(6), 1005-1015.
- Sulis, I., & Porcu, M. (2008). *Assessing the effectiveness of a stochastic regression imputation method for ordered categorical data*. Center for North South Economic

Research, April 2008. [Çevrim-içi: <http://veprints.unica.it/195/1/08-04.pdf>, Erişim tarihi: 1 Eylül 2016.]

- Şahin-Kürşad, M., & Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin ölçeklerin geçerlik ve güvenilirliği bağlamında karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2), 254-267.
- Takahashi, M., & Ito, T. (2013). *Multiple imputation of missing values in economic surveys: comparison of competing algorithms*. Proceedings of the 59th World Statistics Congress of the International Statistical Institute. Hong Kong Convention and Exhibition Centre Hong Kong, 25-30 August 2013. [Çevrim-içi: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.670.5853&rep=rep1&type=pdf>, Erişim tarihi: 9 Mart 2015.]
- Toka, O. (2012). *Kayıp veri durumunda sağlam kestirim*. (Yayımlanmamış Yüksek Lisans Tezi). Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). *Journal of Statistical Software*, 45(3), 1-67.
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16(3), 219-242.
- Van Ginkel, J. R. (2007). *Multiple imputation for incomplete test, questionnaire and survey data*. (Unpublished Doctoral Dissertation). University of Tilburg.
- Weeks, J. P. (2010). Plink: An R package for linking mixed-format tests using IRT-based methods. *Journal of Statistical Software*, 35(12), 1-33.
- White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, 30(4), 377-399.
- Woldbeck, T. (1998). *Basic concepts in modern methods of test equating*. Annual Meeting of the Southwest Psychological Association. New Orleans Los Angeles, 11 April 1998. [Çevrim-içi: <http://files.eric.ed.gov/fulltext/ED417215.pdf>, Erişim tarihi: 17 Temmuz 2015.]
- Wood, R., Wilson, D., Gibbons, R., Schilling, S., Muraki, E., & Bock, R. (2003). *TESTFACT 4 for Windows: Test scoring, item statistics, and full-information item factor analysis [Computer software]*. Lincolnwood: Scientific Software International.
- Wu, W., Jia, F., & Enders, C. (2015). A comparison of imputation strategies for ordinal missing data on Likert scale variables. *Multivariate Behavioral Research*, 50(5), 484-503.
- Zhang, B., & Walker, C. M. (2008). Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, 32(6), 466-479.
- Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). *BILOG-MG 3 for Windows: Multiple-group IRT analysis and test maintenance for binary items [Computer software]*. Skokie: Scientific Software International, Inc.

EKLER DİZİNİ

EK 1. ETİK KOMİSYONU ONAY BİLDİRİMİ

Form: 40

Tez Çalışması Etik Komisyon İzin Muafiyeti Formu

24 / 03 / 2017

Hacettepe Üniversitesi
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Anabilim Dalı Başkanlığı'na

Tez Başlığı / Konusu:	Kayıp Verinin Test Eşitlemeye Etkisinin İncelenmesi
-----------------------	---

Yukarıda başlığı/konusu gösterilen tez çalışmam:

1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır.
2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanımını gerektirmemektedir.
3. Beden bütünlüğüne müdahale içermemektedir.
4. Gözlemsel ve betimsel araştırma (anket, ölçek/skala çalışmaları, dosya taramaları, veri kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.

Hacettepe Üniversitesi Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Komisyondan/Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.


Duygu Gizem ERTOPRAK
(Öğrencinin Adı Soyadı, İmzası)

Öğrenci Bilgileri

Adı Soyadı	Duygu Gizem ERTOPRAK
Öğrenci No	N11246589
Anabilim Dalı	Eğitim Bilimleri
Programı	Eğitimde Ölçme ve Değerlendirme
Statüsü	<input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.

Danışman Görüşü ve Onayı

Yapılan bu tez çalışması için etik kurul izni almaya gerek duyulmamıştır.


Prof. Dr. Selahattin GELBAL
(İmza)
(Danışmanın Unvanı, Adı ve Soyadı)

EK 2. ORJİNALLİK RAPORU



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS/DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLER ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM / BİLİM DALI BAŞKANLIĞI'NA

Tarih: 18/04/2017

Tez Başlığı : Kayıp Verinin Test Eşitlemeye Etkisinin İncelenmesi

Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak **Turnitin** adlı intihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir.

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Endeksi	Gönderim Numarası
13/04/2017	119	223285	24/03/2017	%5	798561655

Uygulanan filtreler:

- 1- Kaynakça hariç
- 2- Alıntılar dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

18/04/2017

Adı Soyadı: Duygu Gizem ERTOPRAK
Öğrenci No: N11246589
Anabilim Dalı: Eğitim Bilimleri
Program: Eğitimde Ölçme ve Değerlendirme
Statüsü: Y.Lisans Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.
Prof. Dr. Selahattin GELBAL



HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
THESIS/DISSERTATION ORIGINALITY REPORT

HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
TO THE DEPARTMENT OF EDUCATIONAL SCIENCES

Date: 18/04/2017

Thesis Title : Investigating The Effect of Missing Data on Test Equating

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defence	Similarity Index	Submission ID
13/04/2017	119	223285	24/03/2017	5%	798561655

Filtering options applied:

1. Bibliography excluded
2. Quotes excluded
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.


18/04/2017

Name Surname: Duygu Gizem ERTOPRAK

Student No: N11246589

Department: Department of Educational Sciences

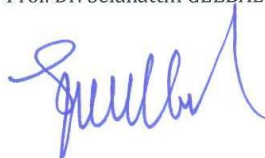
Program: Division of Measurement and Evaluation in Education

Status: Masters Ph.D. Integrated Ph.D.

ADVISOR APPROVAL

APPROVED

Prof. Dr. Selahattin GELBAL



EK 3. TAM VERİ SETLERİNİN TÜRETİLMESİNDE KULLANILAN KODLAR

S1 koşulu kapsamında eşitlenen tam veri setlerine sahip olan total.formx.data[[i]] ve total.formy.data[[i]]'nin elde edilmesi.

```
formx.a <- runif(80, 0.5, 2)
formx.b <- runif(80,-2.5,2.5)
formx.c <- runif(80,0.2,0.3)
```

```
formy.a <- runif(80, 0.5, 2)
formy.b <- runif(80,-2,3)
formy.c <- runif(80,0.2,0.3)
```

```
anchor.a <- runif(20, 0.5, 2)
anchor.b <- runif(20,-1,1)
anchor.c <- runif(20,0.2,0.3)
```

```
Nx <- 2000
nx <- 80
replication <- 50
```

```
formx.person <- vector("list",replication)
formx.item <- vector("list",replication)
formx.data <- vector("list",replication)
```

```
for(i in 1:replication) {
  formx.person[[i]] <- rnorm(Nx,0,1)
  formx.item[[i]] <- cbind(formx.a,formx.b,formx.c)
}
```

```
library(irtoys)
```

```
for(i in 1:replication) {
  formx.data[[i]] <- sim(ip=formx.item[[i]], x=formx.person[[i]])
}
```

```
na <- 20
replication <- 50
```

```
anchor.item <- vector("list",replication)
anchor.formx.data <- vector("list",replication)
```

```
for(i in 1:replication) {
  anchor.item[[i]] <- cbind(anchor.a,anchor.b,anchor.c)
}
```

```
library(irtoys)
```

```
for(i in 1:replication) {
  anchor.formx.data[[i]] <- sim(ip=anchor.item[[i]], x=formx.person[[i]])
}
```

```
total.formx.data <- vector("list",replication)
```

```

for(i in 1:replication) {
total.formx.data[[i]] <- cbind(formx.data[[i]], anchor.formx.data[[i]])
}

Ny <- 2000
ny <- 80
replication <- 50

formy.person <- vector("list",replication)
formy.item <- vector("list",replication)
formy.data <- vector("list",replication)

for(i in 1:replication) {
formy.person[[i]] <- rnorm(Ny,0,1)
formy.item[[i]] <- cbind(formy.a,formy.b,formy.c)
}

library(irtoys)

for(i in 1:replication) {
formy.data[[i]] <- sim(ip=formy.item[[i]], x=formy.person[[i]])
}

anchor.formy.data <- vector("list",replication)

library(irtoys)

for(i in 1:replication) {
anchor.formy.data[[i]] <- sim(ip=anchor.item[[i]], x=formy.person[[i]])
}

total.formy.data <- vector("list",replication)

for(i in 1:replication) {
total.formy.data[[i]] <- cbind(formy.data[[i]], anchor.formy.data[[i]])
}

```

EK 4. TRK MEKANİZMASINA SAHİP VERİ SETLERİNİN ELDE EDİLMESİNDE KULLANILAN KODLAR

S2 koşulunda total.formy.data[[i]] ile eşitlenen S2_3_formx[[i]]'in elde edilmesi.

```
for (i in 1: replication) {  
  S2_3_formx[[i]]<-total.formx.data[[i]]  
  while(sum(is.na(S2_3_formx[[i]]))<20000) {  
    column.sample<-sample(1:100,1,replace=T)  
    row.sample<-sample(1:2000,1,replace=T)  
    S2_3_formx[[i]][row.sample,column.sample]<-NA  
  }  
}
```

EK 5. RK MEKANİZMASINA SAHİP VERİ SETLERİNİN ELDE EDİLMESİNDE KULLANILAN KODLAR

S11 koşulunda total.formy.data[[i]] ile eşitlenen S11_12_formx[[i]]'in elde edilmesi.

```
unsortedformxmar=vector("list", replication)
orderformxmar=vector("list", replication)
sortedformxmar=vector("list", replication)

for (i in 1: replication) {
unsortedformxmar[[i]]=cbind(total.formx.data[[i]],formx.person[[i]])
orderformxmar[[i]]=order(unsortedformxmar[[i]][,101])
sortedformxmar[[i]]=unsortedformxmar[[i]][orderformxmar[[i]],]
}

pureformxmar=vector("list", replication)
for (i in 1: replication) {
pureformxmar[[i]]=sortedformxmar[[i]][,-101]
}

mostlowformxmar=vector("list", replication)
lowformxmar=vector("list", replication)
highformxmar=vector("list", replication)
mosthighformxmar=vector("list", replication)

for (i in 1: replication) {
mostlowformxmar[[i]]<- pureformxmar[[i]][1:500,]
lowformxmar[[i]]<- pureformxmar[[i]][501:1000,]
highformxmar[[i]]<- pureformxmar[[i]][1001:1500,]
mosthighformxmar[[i]]<- pureformxmar[[i]][1501:2000,]
}

mostlowformxmar10=vector("list", replication)
lowformxmar10=vector("list", replication)
highformxmar10=vector("list", replication)
mosthighformxmar10=vector("list", replication)

for (i in 1: replication) {
mostlowformxmar10[[i]]=mostlowformxmar[[i]]
lowformxmar10[[i]]=lowformxmar[[i]]
highformxmar10[[i]]=highformxmar[[i]]
mosthighformxmar10[[i]]=mosthighformxmar[[i]]
}

for (i in 1: replication) {
while(sum(is.na(mostlowformxmar10[[i]]))<8000) {
column.sample<-sample(1:100,1,replace=T)
row.sample<-sample(1:500,1,replace=T)
mostlowformxmar10[[i]][row.sample,column.sample]<-NA
}
}

for (i in 1: replication) {
while(sum(is.na(lowformxmar10[[i]]))<6000) {
```

```

column.sample<-sample(1:100,1,replace=T)
row.sample<-sample(1:500,1,replace=T)
lowformxmar10[[i]][row.sample,column.sample]<-NA
}
}

```

```

for (i in 1: replication) {
while(sum(is.na(highformxmar10[[i]]))<4000) {
column.sample<-sample(1:100,1,replace=T)
row.sample<-sample(1:500,1,replace=T)
highformxmar10[[i]][row.sample,column.sample]<-NA
}
}

```

```

for (i in 1: replication) {
while(sum(is.na(mosthighformxmar10[[i]]))<2000) {
column.sample<-sample(1:100,1,replace=T)
row.sample<-sample(1:500,1,replace=T)
mosthighformxmar10[[i]][row.sample,column.sample]<-NA
}
}

```

```

S11_12_formx =vector("list", replication)
for (i in 1: replication) {
S11_12_formx[[i]]<-
rbind(mostlowformxmar10[[i]],lowformxmar10[[i]],highformxmar10[[i]],mosthighformxmar10
[[i]])
}

```


EK 6. ROK MEKANİZMASINA SAHİP VERİ SETLERİNİN ELDE EDİLMESİNDE KULLANILAN KODLAR

S20 koşulunda total.formy.data[[i]] ile eşitlenen S20_21_formx[[i]]'in elde edilmesi.

```
formx_b_s1<-c(formx.b,anchor.b)
formx_b_s2=matrix(formx_b_s1)
#class(formx_b_s2)
#dim(formx_b_s2)
formx_b=matrix(formx_b_s2,1,100)
#dim(formx_b)

unsortedformxmnr_s1=vector("list",replication)
orderformxmnr_s1=vector("list",replication)
sortedformxmnr_s1=vector("list",replication)

for (i in 1: replication) {
unsortedformxmnr_s1[[i]]=rbind(total.formx.data[[i]],formx_b)
orderformxmnr_s1[[i]]=order(unsortedformxmnr_s1[[i]][2001,])
sortedformxmnr_s1[[i]]=unsortedformxmnr_s1[[i]][,orderformxmnr_s1[[i]]]
}

pureformxmnr_s1=vector("list",replication)
for (i in 1: replication) {
pureformxmnr_s1[[i]]= sortedformxmnr_s1[[i]][-2001,]
}

formx_person_s1=vector("list", replication)
formx_person=vector("list", replication)
for (i in 1: replication) {
formx_person_s1[[i]]<-formx.person[[i]]
formx_person[[i]]=matrix(formx_person_s1[[i]])
}

unsortedformxmnr=vector("list", replication)
orderformxmnr= vector("list", replication)
sortedformxmnr=vector("list", replication)

for (i in 1: replication) {
unsortedformxmnr[[i]]=cbind(pureformxmnr_s1[[i]],formx_person[[i]])
orderformxmnr[[i]]=order(unsortedformxmnr[[i]][,101])
sortedformxmnr[[i]]= unsortedformxmnr[[i]][orderformxmnr[[i]],]
}

pureformxmnr=vector("list", replication)
for (i in 1: replication) {
pureformxmnr[[i]]= sortedformxmnr[[i]][,-101]
}

loweasyformxmnr=vector("list", replication)
lowhardformxmnr=vector("list", replication)
higheasyformxmnr=vector("list", replication)
highhardformxmnr=vector("list", replication)
```

```

for (i in 1: replication) {
loweasyformxmnr[[i]]= pureformxmnr[[i]][1:1000,1:50]
lowhardformxmnr[[i]]= pureformxmnr[[i]][1:1000,51:100]
higheasyformxmnr[[i]]= pureformxmnr[[i]][1001:2000,1:50]
highhardformxmnr[[i]]= pureformxmnr[[i]][1001:2000,51:100]
}

```

```

loweasyformxmnr10=vector("list", replication)
lowhardformxmnr10=vector("list", replication)
higheasyformxmnr10=vector("list", replication)
highhardformxmnr10=vector("list", replication)

```

```

for (i in 1: replication) {
loweasyformxmnr10[[i]]= loweasyformxmnr[[i]]
lowhardformxmnr10[[i]]= lowhardformxmnr[[i]]
higheasyformxmnr10[[i]]= higheasyformxmnr[[i]]
highhardformxmnr10[[i]]= highhardformxmnr[[i]]
}

```

```

for (i in 1: replication) {
while(sum(is.na(loweasyformxmnr10[[i]]))<4000) {
column.sample<-sample(1:50,1,replace=T)
row.sample<-sample(1:1000,1,replace=T)
loweasyformxmnr10[[i]][row.sample,column.sample]<-NA
}
}

```

```

for (i in 1: replication) {
while(sum(is.na(lowhardformxmnr10[[i]]))<10000) {
column.sample<-sample(1:50,1,replace=T)
row.sample<-sample(1:1000,1,replace=T)
lowhardformxmnr10[[i]][row.sample,column.sample]<-NA
}
}

```

```

for (i in 1: replication) {
while(sum(is.na(highhardformxmnr10[[i]]))<6000) {
column.sample<-sample(1:50,1,replace=T)
row.sample<-sample(1:1000,1,replace=T)
highhardformxmnr10[[i]][row.sample,column.sample]<-NA
}
}

```

```

S20_21_formx_s1=vector("list", replication)
S20_21_formx_s2=vector("list", replication)
S20_21_formx=vector("list", replication)

```

```

for (i in 1: replication) {
S20_21_formx_s1[[i]]=cbind(loweasyformxmnr10[[i]], lowhardformxmnr10[[i]])
S20_21_formx_s2[[i]]=cbind(higheasyformxmnr10[[i]], highhardformxmnr10[[i]])
S20_21_formx[[i]]=rbind(S20_21_formx_s1[[i]], S20_21_formx_s2[[i]])
}

```

**EK 7. TRK MEKANİZMASI KAPSAMINDA ELE ALINAN SİMÜLASYON
KOŞULLARINA GÖRE KESTİRİLEN MADDE PARAMETRELERİNE İLİŞKİN
RMSE VE BIAS DEĞERLERİ**

Oran	Yer	Parametre	Başa Çıkma Yöntemleri							
			UGD		YYGD		LRÇDA		DFÇDA	
			RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS
%10	ET	a	0.034	-0.033	0.035	-0.032	0.026	-0.007	0.022	0.001
		b	0.132	-0.010	0.110	-0.018	0.080	-0.004	0.079	0.016
	HiT	a	0.021	0.005	0.032	-0.028	0.023	0.002	0.022	0.001
		b	0.076	0.013	0.109	-0.018	0.088	0.022	0.078	0.017
	OT	a	0.022	-0.011	0.029	-0.007	0.022	-0.002	0.022	-0.004
		b	0.088	-0.006	0.067	0.002	0.078	0.013	0.080	0.013
%20	ET	a	0.034	-0.034	0.036	-0.034	0.031	-0.001	0.023	0.001
		b	0.140	-0.036	0.112	-0.025	0.082	-0.016	0.080	0.015
	HiT	a	0.020	-0.004	0.022	-0.011	0.024	0.006	0.025	0.001
		b	0.086	0.000	0.076	-0.005	0.090	0.020	0.078	0.013
	OT	a	0.022	0.002	0.028	-0.006	0.023	0.002	0.022	-0.001
		b	0.078	0.007	0.068	0.005	0.074	0.014	0.080	0.014
%40	ET	a	0.036	-0.036	0.035	-0.032	0.034	-0.018	0.021	-0.002
		b	0.140	-0.021	0.104	-0.016	0.080	-0.022	0.083	0.015
	HiT	a	0.019	-0.011	0.020	-0.002	0.021	0.002	0.024	-0.008
		b	0.089	0.001	0.075	0.006	0.079	0.001	0.081	0.012
	OT	a	0.022	-0.006	0.029	-0.004	0.023	0.000	0.022	-0.005
		b	0.082	0.001	0.068	0.005	0.073	0.011	0.080	0.013

**EK 8. RK MEKANİZMASI KAPSAMINDA ELE ALINAN SİMÜLASYON
KOŞULLARINA GÖRE KESTİRİLEN MADDE PARAMETRELERİNE İLİŞKİN
RMSE VE BIAS DEĞERLERİ**

Oran	Yer	Parametre	Başa Çıkma Yöntemleri							
			UGD		YYGD		LRÇDA		DFÇDA	
			RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS
%10	ET	a	0.034	-0.034	0.035	-0.032	0.022	-0.010	0.022	-0.001
		b	0.138	-0.030	0.104	-0.012	0.081	0.004	0.081	0.016
	HiT	a	0.020	-0.007	0.026	-0.007	0.018	0.000	0.022	-0.002
		b	0.084	0.000	0.069	0.007	0.067	0.007	0.081	0.015
	OT	a	0.022	-0.006	0.028	-0.005	0.022	-0.001	0.022	-0.002
		b	0.084	-0.001	0.066	0.005	0.079	0.015	0.080	0.014
%20	ET	a	0.035	-0.035	0.037	-0.035	0.024	-0.018	0.022	-0.002
		b	0.133	-0.028	0.120	-0.027	0.079	-0.006	0.081	0.011
	HiT	a	0.020	-0.002	0.026	-0.013	0.016	-0.001	0.022	-0.001
		b	0.079	0.010	0.074	-0.004	0.058	0.008	0.081	0.014
	OT	a	0.021	-0.005	0.028	-0.009	0.022	-0.003	0.022	-0.004
		b	0.079	0.012	0.069	0.001	0.076	0.013	0.080	0.015
%40	ET	a	0.035	-0.035	0.037	-0.034	0.029	-0.028	0.025	0.003
		b	0.147	-0.012	0.111	-0.017	0.079	-0.015	0.081	0.015
	HiT	a	0.023	-0.004	0.023	0.001	0.011	-0.002	0.027	0.001
		b	0.081	0.004	0.062	0.008	0.033	0.007	0.080	0.016
	OT	a	0.022	0.001	0.027	-0.007	0.023	0.001	0.023	0.002
		b	0.076	0.004	0.066	-0.001	0.074	0.014	0.077	0.014

**EK 9. ROK MEKANİZMASI KAPSAMINDA ELE ALINAN SİMÜLASYON
KOŞULLARINA GÖRE KESTİRİLEN MADDE PARAMETRELERİNE İLİŞKİN
RMSE VE BIAS DEĞERLERİ**

Oran	Yer	Parametre	Başa Çıkma Yöntemleri							
			UGD		YYGD		LRÇDA		DFÇDA	
			RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS
%10	ET	a	0.033	-0.012	0.039	-0.024	0.026	-0.008	0.026	-0.003
		b	0.122	-0.118	0.110	-0.104	0.120	-0.111	0.112	-0.099
	HiT	a	0.025	-0.014	0.039	-0.038	0.025	0.000	0.026	-0.001
		b	0.040	-0.034	0.210	-0.210	0.019	0.008	0.024	0.009
	OT	a	0.027	-0.019	0.024	-0.029	0.025	0.001	0.027	-0.001
		b	0.050	-0.046	0.102	-0.101	0.024	0.011	0.024	0.009
%20	ET	a	0.024	-0.013	0.045	-0.023	0.027	-0.008	0.025	-0.003
		b	0.117	-0.114	0.118	-0.105	0.129	-0.124	0.111	-0.097
	HiT	a	0.023	-0.010	0.040	-0.039	0.023	-0.003	0.026	-0.003
		b	0.037	-0.029	0.251	-0.250	0.016	0.004	0.024	0.007
	OT	a	0.023	-0.022	0.025	-0.022	0.006	-0.001	0.010	-0.003
		b	0.081	-0.080	0.101	-0.100	0.008	-0.005	0.012	-0.009
%40	ET	a	0.029	0.002	0.040	-0.012	0.028	-0.013	0.029	-0.001
		b	0.108	-0.103	0.197	-0.197	0.161	-0.157	0.113	-0.101
	HiT	a	0.023	-0.010	0.039	0.019	0.020	-0.006	0.035	-0.004
		b	0.041	-0.017	1.000	1.000	0.013	0.003	0.028	0.004
	OT	a	0.028	0.004	0.042	-0.039	0.023	0.002	0.028	-0.002
		b	0.341	-0.340	0.166	-0.155	0.026	0.013	0.023	0.006

**EK 10. ARAŞTIRMA KAPSAMINDA ELE ALINAN SİMÜLASYON
KOŞULLARINA GÖRE KESTİRİLEN YETENEK PARAMETRELERİNE İLİŞKİN
RMSE VE BIAS DEĞERLERİ**

Tür	Oran	Yer	Başa Çıkma Yöntemleri								
			UGD		YYGD		LRÇDA		DFÇDA		
			RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	
TRK	%10	ET	1.918	-0.641	1.604	-0.381	1.099	-0.329	1.122	-0.006	
		HİT	1.086	-0.057	1.565	-0.372	1.084	-0.030	1.113	-0.006	
		OT	1.283	-0.380	1.108	-0.370	1.124	0.004	1.158	0.024	
	%20	ET	2.063	-1.129	1.612	-0.338	1.202	-0.765	1.113	0.016	
		HİT	1.195	-0.297	1.184	-0.023	1.080	0.002	1.086	0.021	
		OT	1.121	-0.185	1.099	0.043	1.061	0.010	1.120	0.014	
	%40	ET	1.895	-0.749	1.444	-0.187	1.769	-1.666	1.136	0.006	
		HİT	1.205	-0.272	1.092	0.115	1.154	-0.063	1.119	-0.069	
		OT	1.189	-0.262	1.098	0.052	1.037	0.006	1.146	0.012	
	RK	%10	ET	2.057	-1.009	1.578	-0.382	1.220	-0.437	1.114	0.001
			HİT	0.504	-0.285	0.418	0.036	0.398	-0.035	0.420	0.004
			OT	0.499	-0.290	0.400	0.039	0.406	-0.019	0.408	0.018
%20		ET	1.980	-0.863	1.654	-0.329	1.426	-0.850	1.102	0.010	
		HİT	0.436	-0.136	0.492	0.015	0.371	-0.022	0.420	-0.004	
		OT	0.425	-0.075	0.416	-0.032	0.408	0.022	0.412	0.018	
%40		ET	1.939	-0.543	1.677	-0.444	2.166	-1.855	1.061	0.066	
		HİT	0.548	-0.237	0.479	0.165	0.226	-0.005	0.490	-0.103	
		OT	0.433	-0.201	0.416	-0.072	0.400	0.039	0.397	0.020	
ROK		%10	ET	2.541	-2.535	2.241	-2.228	2.837	-2.809	2.346	-2.272
			HİT	1.008	-0.978	4.206	-4.203	0.360	-0.094	0.418	-0.129
			OT	1.268	-1.255	2.053	-2.042	0.384	-0.064	0.408	-0.098
	%20	ET	2.449	-2.441	2.264	-2.252	3.297	-3.386	2.297	-2.218	
		HİT	0.911	-0.878	4.935	-4.935	0.331	-0.104	0.431	-0.097	
		OT	1.728	-1.724	1.953	-1.951	0.217	-0.121	0.298	-0.214	
	%40	ET	2.248	-2.236	4.075	-4.075	4.420	-4.519	2.270	-2.188	
		HİT	0.608	-0.399	2.241	-2.241	0.245	-0.114	0.453	-0.118	
		OT	2.244	-2.244	2.855	-2.830	0.344	-0.020	0.432	-0.205	

ÖZGEÇMİŞ

Kişisel Bilgiler

<i>Adı Soyadı</i>	Duygu Gizem ERTOPRAK
<i>Doğum Yeri</i>	Bolu
<i>Doğum Tarihi</i>	01.09.1988

Eğitim Durumu

<i>Lise</i>	Mersin Dumlupınar (Süper) Lisesi	2006
<i>Lisans</i>	Hacettepe Üniversitesi Eğitim Fakültesi Fen Bilgisi Öğretmenliği	2010
<i>Yüksek Lisans</i>	Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı	2012
<i>Yabancı Dil</i>	İngilizce: Okuma (Çok iyi), Yazma (İyi), Konuşma (İyi)	

İş Deneyimi

<i>Stajlar</i>		
<i>Projeler</i>		
<i>Çalıştığı Kurumlar</i>	Amasya Üniversitesi Eğitim Fakültesi	2012-halen

Akademik Çalışmalar

Yayınlar (Ulusal, uluslararası makale, bildiri, poster vb gibi.)

--

Seminer ve Çalıştaylar

--

Sertifikalar

--

İletişim

<i>e-Posta Adresi</i>	duygugizemertoprak@gmail.com

<i>Jüri Tarihi</i>	24.03.2017
--------------------	------------