

**T.C.**  
**Mersin Üniversitesi**  
**Eđitim Bilimleri Enstitüsü**  
**Eđitim Bilimleri Anabilim Dalı**  
**Eđitimde Ölçme ve Deđerlendirme Bilim Dalı**

**KÜÇÜK ÖRNEKLEMLERDE TEST EŐİTLEME YÖNTEMLERİNİN**  
**ÇEŐİTLİ FAKTÖRLERE GÖRE İNCELENMESİ**

**YÜKSEK LİSANS TEZİ**

**Semih AŐİRET**

**Mersin, 2014**



**T.C.**  
**MERSİN ÜNİVERSİTESİ**  
**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**  
**EĞİTİM BİLİMLERİ ANABİLİM DALI**  
**EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI**

**KÜÇÜK ÖRNEKLEMLERDE TEST EŞİTLEME YÖNTEMLERİNİN**  
**ÇEŞİTLİ FAKTÖRLERE GÖRE İNCELENMESİ**

**YÜKSEK LİSANS TEZİ**

**SEMİH AŞİRET**

**TEZ DANIŞMANI**

**YRD. DOÇ. DR. SEÇİL ÖMÜR SÜNBÜL**

**MERSİN, 2014**

## KABUL VE ONAY

Mersin Üniversitesi Eğitim Bilimleri Enstitüsü Müdürlüğü'ne,

Semih AŞİRET tarafından hazırlanan "Küçük Örneklerde Test Eşitleme Yöntemlerinin Çeşitli Faktörlere Göre İncelenmesi" başlıklı bu çalışma, jürimiz tarafından Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan .....

Prof. Dr. Adnan KAN

Üye.....

Doç. Dr. Devrim ALICI

Üye.....

Yrd. Doç. Dr. Seçil ÖMÜR SÜNBÜL (Danışman)

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

21.8/2014  
Prof. Dr. Yüksel Keleş  
Enstitü Müdürü

## TEŞEKKÜR

Çalışmamın her aşamasını büyük bir özen ve titizlikle takip eden, inceleyen, katkı sağlayan, desteğini ve yardımını hiçbir zaman esirgemeyen danışmanım Yrd. Doç. Dr. Seçil ÖMÜR SÜNBÜL'e çok teşekkür ederim.

Yüksek lisans eğitimim boyunca beni yetiştiren ve tez sürecinde büyük yardımlarını gördüğüm, bilgi ve deneyimlerinden yararlandığım Yrd. Doç. Dr. Önder SÜNBÜL'e, yüksek lisans eğitimim boyunca ders aldığım ve tezimi okuyarak görüşleri ve eleştirileriyle çalışmama katkıda bulunan değerli hocalarım Prof. Dr. Adnan ERKUŞ'a, Doç. Dr. Devrim ÖZDEMİR ALICI'ya, tezimde jüri üyesi olarak katkıda bulunan Prof. Dr. Adnan KAN'a teşekkürlerimi sunarım.

Yüksek lisans eğitimine başlamam için beni teşvik eden ve bu süreçte desteğini esirgemeyen değerli dostum Mükremin ACAR'a, özellikle tez sürecinde sürekli yanımda olan ve sürekli desteğini gördüğüm nişanlım Hilal TORUN'a, gerek ders, gerekse tez sürecinde desteklerini esirgemeyen, sürekli motive olmamı sağlayan arkadaşlarım Seyran ERGİN AYDEMİR ve Seha YORMAZ'a çok teşekkür ederim.

Bu süreç boyunca sürekli arkamda olup desteklerini benden esirgemeyen anneme, babama, abim Hasan AŞİRET'e ve kardeşim Mehmet AŞİRET'e, ayrıca isimlerini buraya yazamadığım, bu anlamlı süreçte bana destek olan ve emeği geçen tüm hocalarıma ve arkadaşlarıma sonsuz teşekkür ederim.

## ÖZET

# KÜÇÜK ÖRNEKLEMLERDE TEST EŞİTLEME YÖNTEMLERİNİN ÇEŞİTLİ FAKTÖRLERE GÖRE İNCELENMESİ

**Semih AŞİRET**

Yüksek Lisans Tezi

Eğitim Bilimleri Enstitüsü

Yrd. Doç. Dr. Seçil ÖMÜR SÜNBÜL (Danışman)

MERSİN, 2014

121 Sayfa

Bu araştırmada, küçük örneklerde seçkisiz gruplar desenine dayalı yapılan test eşitleme yöntemlerinin, örneklem büyüklüğüne, formlar arasındaki ortalama güçlük düzeyleri farkına, şans başarısına ve test uzunluğuna göre karşılaştırılması amaçlanmıştır. Bunun yanı sıra, hangi koşullar altında hangi yöntemin daha iyi sonuç verdiği araştırılmıştır.

Araştırmada, üç parametrelili lojistik modele (3PLM) uygun olarak, seçkisiz gruplar desenine göre, her bir form ( X ve Y) için, 5000'er kişilik, 1-0

şeklinde iki kategorili olarak evren verisi üretilmiştir. Bu araştırmada, bu iki test formunu eşitlemek için seçkisiz gruplar deseni kullanılmıştır. Hem X hem de Y formu için, şans parametre değeri 0.0, 0.1, 0.2, 0.25 olmak üzere dört düzeyde, test uzunluğu 25 ve 50 olmak üzere iki düzeyde değişimlenmiştir. Ayrıca X formu ile Y formu arasındaki güçlük düzey farkı, 0.1, 0.4 ve 0.7 olmak üzere 3 düzeyde değişimlenmiştir. Araştırmada yer alan faktörlere ve değişimlenen düzeylere uygun, X formu için 8 ve Y formu için 24 olmak üzere toplam 32 evren formu oluşturulmuştur. Araştırmada, simülatif olarak üretilen formlar, farklı örneklem büyüklüklerinde (10, 25, 50, 75, 100, 150, 200), özdeş, ortalama, doğrusal, Dairesel yay, C=2 ve C=3 düzeyinde log-linear ön-düzgünleştirme yapılmış eşit yüzdelikli eşitleme yöntemleri kullanılarak 100 replikasyon yapılarak eşitlenmiştir. Verilerin üretiminde ve formların eşitlenmesinde R 3.1 programı kullanılmıştır. Araştırmada yer alan sonuçlar, eşitlemenin standart hatası (SEE), eşitleme yanlılığı (BIAS) ve eşitleme hatası (RMSE) ölçütlerine göre değerlendirilmiştir.

Araştırma sonucunda, 50 örneklem büyüklüğü ve üstünde, formlar arası güçlük düzey farkı 0.4 olduğu durumlarda eşitleme yapmanın faydalı olabileceği sonucuna ulaşılmıştır. Test uzunluğunun artması, tüm yöntemlerin hata değerlerinde doğrusal artışa yol açtığı görülmüştür. Ayrıca, şans parametre değerlerinin değişiminin Dairesel yay yöntemi dışındaki diğer yöntemlerin hata değerlerini önemli ölçüde etkilemediği görülmüştür. Araştırmada yer alan faktörlerin birçok düzeyinde, Dairesel yay ve ortalama eşitleme yöntemlerinin daha az hatalı eşitlemeler yaptığı saptanmıştır.

**Anahtar kelimeler:** Test Eşitleme, Eşitleme Hatası, Küçük Örneklem

## **ABSTRACT**

### **FACTORS AFFECTING THE TEST EQUATING METHOD USING SMALL SAMPLES**

**Semih AŞİRET**

Master thesis

Institute of Education Sciences

Assist. Prof. Dr. Seçil ÖMÜR SÜNBÜL (Advisor)

MERSİN, 2014

121 Pages

In this study, it was aimed to compare the equating methods for random group design using small samples by the variables such as sample size, average difficulty differences between forms, chance success and test length. Beside this, which method gave better results under what conditions was also investigated.



In study, 5000 dichotomous simulated data which was consistent with three parameter logistic model (3PLM) was produced for each form (X and Y). In order to equate two test forms “the random groups design” was used in this study. For both X and Y forms, chance parametre values were altered in four levels (0.00, 0.10, 0.20, 0.25) and test length was altered in two levels (25 and 50). Furthermore, average difficulty difference between X and Y forms were altered in three levels (0.1, 0.4, 0.7). In this study, 8 forms for X and 24 forms for Y, in total 32 forms which were consistent with the variables studied and their levels were simulated. Simulated forms were equated by using identity, mean, linear, 2 and 3 moments pre-smoothing polynomial log-linear equipercentile and Circle-Arc Methods for seven different sample sizes (10, 25, 50, 75, 100, 150 and 200) with 100 replications. R.3.1 programming language was used to produce data and to equate the simulated forms. The results obtained from this simulation study were evaluated based on standard error of equating (SEE), equating bias (BIAS) and equating error (RMSE) criterions.

The findings indicated that in the case when sample size was 50 and more and the level of average difficulty difference between forms was 0.4, it was concluded that equating forms would give better results than not equating. Increasing the test length caused to linear increasing equating error values of all equating methods. Moreover, it was resulted that changing chance parameter (c) didn't effect equating error values of equating methods except for Circle-Arc Method. Circle-Arc and Mean equating methods were found to less equating error than the other equating methods for small sample equating under most of the conditions studied.

**Keywords:** Test Equating, Equating Error, Small Samples

## İÇİNDEKİLER

TEŞEKKÜR.....	i
ÖZET .....	ii
ABSTRACT .....	iv
İÇİNDEKİLER.....	vi
TABLOLAR DİZİNİ.....	xiv
ŞEKİLLER DİZİNİ.....	xv
GRAFİKLER DİZİNİ.....	xvi
SİMGELER VE KISALTMALAR DİZİNİ .....	xviii
BÖLÜM I: GİRİŞ.....	1
1.TEST EŞİTLEMENİN ÖZELLİKLERİ .....	3
1.1. Simetri Özelliği.....	3
1.2. Aynı Değerler Özelliği.....	3
1.3.Eşitlik Özelliği .....	4
1.4. Gruptan Bağımsızlık Özelliği .....	4
2. VERİ TOPLAMA DESENLERİ .....	5

2.1. Ortak Bireyleri (Common Examinees) Kullanan Desenler .....	5
2.1.1. Tek Grup Deseni (Single Group Design) .....	5
2.1.2. Seçkisiz Gruplar Deseni (Random Groups Design) .....	6
2.1.3. Karşıt dengelenmiş Desen (Counterbalanced Design) .....	7
2.2. Ortak Maddeleri Kullanan Desenler ( Common Anchors Design) .....	8
2.2.1. Denk Gruplarda Ortak Madde Deseni .....	8
2.2.2. Denk Olmayan Gruplarda Ortak Madde Deseni (Nonequivalent Anchor Tests-NEAT) .....	9
2.2.2.1. İç Ankor .....	11
2.2.2.2. Dış Ankor .....	13
<b>3. TEST EŞİTLEME YÖNTEMLERİ.....</b>	<b>15</b>
3.1. Özdeş Eşitleme (Identity Equating) .....	15
3.2. Ortalama Eşitleme (Mean Equating).....	16
3.3. Doğrusal Eşitleme (Linear Equating).....	17
3.4. Eşit Yüzdellikli Eşitleme (Equipercntile Equating) .....	18
3.4.1.Eşit Yüzdellikli Eşitleme Yönteminde Düzgünleştirme (Smoothing in Equipercntile Equating).....	21
3.5.Eşit Yüzdellikli Eşitlemede Polinomal Loglinear Ön-düzgünleştirme Yöntemi (Polynomial Log-Linear Method in Equipercntile Equating).....	22
3.6. Dairesel Yay Yöntemi (Circle-Arc Method).....	24
3.7. Denk Olmayan Gruplarda Ortak Madde (NEAT) Desenlerinde Kullanılan Eşitleme Yöntemleri .....	25
3.7.1. Zincirlenmiş Doğrusal Eşitleme Yöntemi (Chained Linear Equation) .....	26
3.7.2 Zincirlenmiş Eşit Yüzdellikli Eşitleme Yöntemi (Chained Equipercntile Equation).....	26

3.7.3 Tucker Yöntemi .....	27
3.7.4 Levine Yöntemi.....	27
3.7.5 Frekans Kestirim Yöntemi (Frequency Estimation Equipercetile Equating).....	27
<b>4. EŞİTLEME HATASI.....</b>	<b>28</b>
4.1 Önyükleme (Bootstrap) Yöntemi.....	30
<b>5. ÖLÇÜT EŞİTLEME .....</b>	<b>30</b>
<b>6.İLGİLİ ARAŞTIRMALAR .....</b>	<b>31</b>
6.1. Yurtdışında Yapılan Çalışmalar .....	31
6.2. Yurt İçinde Yapılan Çalışmalar .....	37
<b>7. ARAŞTIRMANIN AMACI VE ÖNEMİ.....</b>	<b>40</b>
<b>8. PROBLEM CÜMLESİ .....</b>	<b>42</b>
8. 1. Alt Problemler.....	42
<b>9. SINIRLILIKLAR .....</b>	<b>44</b>
<b>BÖLÜM II: YÖNTEM.....</b>	<b>46</b>
<b>1. Araştırmanın Türü .....</b>	<b>46</b>
<b>2. Araştırma Kapsamında Değişimlenen Faktörler .....</b>	<b>46</b>
<b>3. Veri Üretimi .....</b>	<b>51</b>
<b>4. İşlem ve Verilerin Analizi.....</b>	<b>53</b>

<i>Eşitlemenin Standart Hatası (Standard Error of Equating)</i> .....	54
<i>Eşitleme Yanlılığı (Equating Bias)</i> .....	55
<i>Eşitleme Hatası (Root Mean Square Error)</i> .....	55

## **BÖLÜM III: BULGULAR..... 57**

### **1. Küçük örneklerde, çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen, eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve hata kestirimlerine (RMSE) temel etkileri nasıl değişmektedir? ..... 57**

1.1.1. Küçük örneklerde, örneklem büyüklüğünün, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	57
1.1.2. Küçük örneklerde, örneklem büyüklüğünün, eşitlemenin yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi.....	59
1.1.3. Küçük örneklerde, örneklem büyüklüğünün, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	61
1.2.1. Küçük örneklerde, test formları arasındaki güçlük düzey farkının, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	63
1.2.2. Küçük örneklerde, test formları arasındaki ortalama güçlük düzeyleri farkının, eşitlemenin yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	64
1.2.3. Küçük örneklerde, test formları arasındaki ortalama güçlük düzeyleri farkının, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	66
1.3.1. Küçük örneklerde, şans başarısının, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi.....	68

1.3.2. Küçük örneklerde, şans başarısının, eşitlemenin yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	69
1.3.3. Küçük örneklerde, şans başarısının, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	71
1.4.1. Küçük örneklerde, test uzunluğunun, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi.....	73
1.4.2. Küçük örneklerde, test uzunluğunun, eşitleme yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	74
1.4.3. Küçük örneklerde, test uzunluğunun, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi .....	76

**2. Küçük örneklerde seçkisiz gruplar desenine dayalı kullanılan test eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına, yanlılığına ve hatasına; örneklem büyüklüğünün, test formları arasındaki ortalama güçlük düzeyleri farkının, şans başarısının ve test uzunluğunun ortak etkisi nasıldır? .....** 77

2.1.1. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının eşitlemenin standart hatasına (SEE) ortak etkisi .....	77
2.1.2. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme yanlılığına (BIAS) ortak etkisi .....	80
2.1.3. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme hatasına (RMSE) ortak etkisi .....	82

2.2.1. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının, eşitlemenin standart hatasına (SEE) ortak etkisi .....	84
2.2.2. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme yanlılığına (BIAS) ortak etkisi .....	86
2.2.3. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme hatasına (RMSE) ortak etkisi .....	88
<b>BÖLÜM IV: TARTIŞMA VE YORUM .....</b>	<b>90</b>
1. Örneklem büyüklüğünün, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine <i>temel etkisinin</i> değerlendirilmesi .....	90
2. Test formları arasındaki güçlük düzey farkının, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi	92
3. Şans başarısının, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi .....	93
4. Test uzunluğunun, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi .....	94
5. Çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE) ortak etkisinin değerlendirilmesi .....	95
6. Çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen eşitleme yanlılığına (BIAS) ortak etkisinin değerlendirilmesi .....	96
7. Çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) ortak etkisinin değerlendirilmesi .....	97
<b>ÖNERİLER.....</b>	<b>102</b>

<b>KAYNAKÇA.....</b>	<b>103</b>
<b>EKLER .....</b>	<b>110</b>
<b>EK-1:Simülatif Verilerden Oluşan X ve Y Formlarının Betimsel İstatistikleri .....</b>	<b>111</b>
<b>EK-2: Örneklem Büyüklüğünün, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri .....</b>	<b>113</b>
EK-2. 1-A: Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri .....	113
EK-2. 1-B: Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri .....	113
EK-2. 1-C: Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri.....	114
<b>EK-3: Test Formları Arasındaki Güçlük Düzey Farkının, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri .....</b>	<b>115</b>
EK-3-A: Çeşitli test formları arasındaki güçlük düzey farkına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri .....	115
EK-3-B: Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri .....	115
EK-3-C: Çeşitli Test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri.	116
<b>EK-4: Şans Başarısının, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri .....</b>	<b>117</b>



EK-4-A: Çeşitli şans başarısına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri .....	117
EK-4-B: Çeşitli şans başarısı parametrelerine göre eşitleme yöntemlerinden elde edilen ortalama yanlışlık (BIAS) değerleri .....	117
EK-4-C: Çeşitli şans başarı parametrelerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri.....	117
<b>EK-5: Test Uzunluğunun, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri .....</b>	<b>118</b>
EK-5-A: Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri .....	118
EK-5-B: Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama yanlışlık (BIAS) değerleri.....	118
EK-5-C: Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri .....	118
<b>ÖZGEÇMİŞ .....</b>	<b>119</b>

## **TABLULAR DİZİNİ**

Tablo 1: Tek Grup Deseni .....	6
Tablo 2: Seçkisiz Gruplar Deseni .....	7
Tablo 3: Karşıtdengelenmiş Desen .....	7
Tablo 4: NEAT deseni .....	9
Tablo 5: Araştırmada kullanılan faktörler ve düzeyleri .....	50
Tablo 6: Simülatif test formlarına ait faktör düzeyleri.....	52
Tablo 7: Araştırmada yer alan faktörlerin düzeylerinde kullanılması uygun olan eşitleme yöntemleri.....	101

## ŞEKİLLER DİZİNİ

Şekil 1: Dengelenmiş tek grup deseni.....	8
Şekil 2: Ankor Madde Güçlük Grafiği (Livingston, 2004; sf. 32).....	11

## GRAFİKLER DİZİNİ

Grafik 1: Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri.....	59
Grafik 2: Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri.....	60
Grafik 3: Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri .....	62
Grafik 4: Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri.....	64
Grafik 5: Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri.....	65
Grafik 6: Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri.....	67
Grafik 7: Çeşitli şans parametre değerlerine göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri.....	68
Grafik 8: Çeşitli şans parametre değerlerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri.....	70
Grafik 9: Çeşitli şans parametre değerlerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri .....	72
Grafik 10: Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri .....	73
Grafik 11: Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri .....	75

Grafik 12: Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri .....	76
Grafik 13: Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitlemenin standart hatasına (SEE) ortak etkisi (Test uzunluğu-25).....	79
<b>Grafik 14:</b> Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme yanlılığına (BIAS) ortak etkisi (Test uzunluğu-25)..	81
<b>Grafik 15:</b> Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme hatasına (RMSE) ortak etkisi (Test uzunluğu-25)....	83
<b>Grafik 16:</b> Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme ortalama standart hatasına (SEE) ortak etkisi (Test uzunluğu-50).....	85
<b>Grafik 17:</b> Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme yanlılığına (BIAS) ortak etkisi (Test uzunluğu-50)..	87
<b>Grafik 18:</b> Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme hatasına (RMSE) ortak etkisi (Test uzunluğu-50 madde) .....	89

## SİMGELER VE KISALTMALAR DİZİNİ

<b>NEAT</b>	: Denk olmayan gruplarda ortak madde deseni
<b>ID<sub>y</sub>(x)</b>	: Özdeş eşitleme fonksiyonu
<b>m<sub>y</sub>(x)</b>	: Ortalama eşitleme fonksiyonu
<b>Lin<sub>y</sub>(x)</b>	: Doğrusal eşitleme fonksiyonu
<b>x<sub>L</sub>(P*)</b>	: Eşit yüzdellikli eşitleme fonksiyonu
<b>x<sub>u</sub>(P*)</b>	: Eşit yüzdellikli eşitleme fonksiyonu
<b>LLC2</b>	: C=2 düzeyinde polinomial Loglinear Ön-düzgünleştirme
<b>LLC3</b>	: C=3 düzeyinde polinomial Loglinear Ön-düzgünleştirme
<b>e<sub>Lin-y</sub>(x)</b>	: Doğrusal bileşen fonksiyonu
<b>e<sub>Arc-y</sub>(x)</b>	: Eğrisel bileşen fonksiyonu
<b>e<sub>CIRC-y</sub></b>	: Dairesel yay fonksiyonu
<b>SEE</b>	: Eşitlemenin standart hatası
<b>BIAS</b>	: Eşitlemenin yanlılığı
<b>RMSE</b>	: Eşitleme Hatası (Hata kareleri ortalamasının karekökü)
<b>SMD</b>	: Ortalama standart sapma
<b>a</b>	: Madde ayırt edicilik parametresi
<b>b</b>	: Güçlük parametresi
<b>c</b>	: Şans parametresi
<b>GD</b>	Güçlük Düzeyi
<b>N</b>	Örneklem Büyüklüğü
<b>3PLM</b>	: Üç parametrelili lojistik model
<b><math>\bar{s}d_i</math></b>	: Eşitlemenin standart hata fonksiyonu
<b><math>\hat{d}_i</math></b>	: Eşitleme yanlılığı fonksiyonu
<b>RMSE<sub>i</sub></b>	: Eşitleme hata fonksiyonu

## BÖLÜM I: GİRİŞ

Eğitim ve psikolojide testler ve ölçekler, öğrencilerin öğrenme düzeylerinin izlenmesinde veya onların bir üst öğretim düzeyine yerleştirilmesinde, personel seçimlerinde, rehberlik ve klinik hizmetlerinin yürütülmesinde yaygın olarak kullanılmaktadır. Bireylerin testler ve ölçeklerden aldıkları puanlar dikkate alınarak, onlar hakkında önemli kararlar verilmektedir. Ancak bireyler hakkında doğru kararların verilebilmesi için, testler ve ölçeklerden elde edilen bu bilgilerin de doğru olması gerekmektedir.

Farklı durumlarda, farklı yöntemlerle veya araştırmacılarla yapılan testlerden elde edilen ölçümlerin karşılaştırılması tüm bilimlerin temel önkoşuludur (Dorans ve Holland, 2000). Bu koşul psikoloji ve eğitim alanlarında yapılan ölçmeler için de geçerlidir (Kolen ve Brennan, 2004). Güvenlik nedeniyle testler ve ölçekler, bazen farklı tarihlerde uygulanmakta veya aynı anda testin farklı paralel formları uygulanmaktadır. Bu durum bazı problemlere yol açabilmektedir. Testi geliştirenler, test formlarını aynı içerik ve istatistiksel özelliklere göre geliştirebilirler, test formları arasında güçlük açısından farklılıklar oluşacaktır. Bazı testler kolay maddelerden oluşurken, bazıları zor maddelerden oluşabilir. Bu sebeple, bireylerin aldıkları puanlarda farklılıklar oluşacaktır. Örneğin, sınavla öğrenci alan bir okula başvuran öğrenci, bir sonraki yıl tekrar sınava girdiğinde aynı test formunu alırsa, testten alacağı puanın daha yüksek olması, yüksek ihtimaldir. Bu durum, öğrencinin başarısını kesin olarak göstermez, çünkü aynı testi daha önceden aldığı için ezberlemiş olma ihtimali (hatırlama etkisi) söz konusu olabilir. Bu ihtimali ortadan kaldırmak için farklı

zamanlarda hazırlanan testler, aynı testin farklı formu şeklinde düzenlenmelidir. Bu durumda ise testin kolaylık veya güçlük sorunu ortaya çıkacaktır. Ülkemizde ÖSYM (Öğrenci Seçme ve Yerleştirme Merkezi) tarafından yapılan Akademik Lisansüstü Eğitim Sınavı (ALES) bu duruma örnek verilebilir. ALES yılda iki defa yapılmaktadır. Sınavın güvenliği ve gizliliğini sağlamak amacıyla aynı davranışı ölçen, farklı maddelerden oluşan iki farklı form şeklinde uygulanır. Farklı zamanlarda sınava giren iki bireyin puanı doğrudan karşılaştırılmaz. Çünkü aynı davranışı ölçen iki farklı test formu, farklı öğrenci gruplarına uygulandığında, test formlarında yer alan maddelerin güçlükleri eşit olmayabilir. Test eşitleme, bu güçlüklerin aşılmasında kullanılır ve test formlarından elde edilen puanların karşılıklı (interchangable) olarak yorumlanmasına imkan tanır (Kolen ve Brennan, 2004; Davier, Holland ve Thayer, 2004).

Kolen ve Brennan (2004) test eşitlemeyi, test formlarından elde edilen puanların yer değiştirilerek kullanılması için, test formlarına ait puanların ayarlanmasında kullanılan istatistiksel bir işlem olarak tanımlamıştır. Crocker ve Algina (1986) test eşitlemeyi, iki ölçme aracından eşdeğer puanlar oluşturma işlemi olarak tanımlamış ve eşit güvenilirliğe sahip, aynı özelliği ölçen, farklı testlerden elde edilen X ve Y puanlarına karşılık gelen yüzdeler eşit olduğunda, X ve Y puanlarının elde edildiği testlerin eşit olacağını belirtmiştir. Angoff (1984) test eşitlemeyi, bir test formunun birim sisteminin, diğer test formunun birim sistemine dönüştürme süreci olarak tanımlamış, böylelikle farklı formlardan elde edilen puanların dönüşümden sonra doğrudan eşitlenebileceğini belirtmiştir; ayrıca, test formlarının eşitlenmesini, inç'in santimetreye, poundun grama, Fahrenheit'in Santigrat'a dönüştürülmesine benzetmiştir. Angoff'un (1984), verdiği örneklerde, her iki ölçek de aynı yapıyı ölçmektedir, fakat ölçek birimleri farklı olduğundan farklı değerler vermektedir. Sonuç olarak test eşitleme, aynı



konu veya yapıyı ölçen, iki veya daha fazla testin, aynı birey için farklı puanlar üretmesinden ortaya çıkmıştır.

İki testin eşitlenebilmesi için bazı koşulların yerine getirilmesi gerekir. Literatürde, bu gerekli koşullara ilişkin birçok farklı görüş bulunmaktadır. Hambleton (1985), bu koşulları, *simetri, aynı değerler özelliği, eşitlik ve gruptan bağımsızlık* olarak sıralamıştır. Bu çalışmada Hambleton (1985) tarafından ortaya atılan koşullar temel alınmıştır.

## 1. TEST EŞİTLEMENİN ÖZELLİKLERİ

### 1.1. Simetri Özelliği

Lord (1980), eşitleme dönüşümlerinin simetrik olması gerektiğini belirtmiştir. Simetri özelliğinde, X formundan elde edilen puanlar, Y formu puanlarına dönüştürülürken, dönüştürme işlemi tersi için de geçerli olmalıdır (Kolen ve Brennan, 2004). Başka bir ifadeyle, X formundan Y formuna dönüştürülen puanlar, Y formundan X formuna dönüştürülen puanlara eşit olmalıdır.

### 1.2. Aynı Değerler Özelliği

Testlerin eşitlenebilmesi için aynı içerik ve istatistiksel değerlere sahip olması gerekir. Bu istatistiksel değerler göz ardı edilerek yapılan eşitlemelerden elde edilen puanlar karşılıklı kullanılmaz (Kolen ve Brennan, 2004). Angoff (1984), iki testin aynı değişkeni ölçmesi gerektiğini ve güvenilirliklerinin aynı olması gerektiğini belirtmiştir. Farklı güvenilirlikteki testler, farklı hatalara sahip olduklarında anlamlı bir şekilde eşitlenemez.

### 1.3.Eşitlik Özelliği

Lord (1980) tarafından ortaya atılan bu özelliğe, Bireylerin, X formunu veya Y formunu almalarının farklılık oluşturmaması gerektiği belirtilmiştir. Eşitlik özelliğine göre gerçek puanı verilen bir bireyin, gözlenen puan ortalamaları, standart sapmaları ve dönüştürülen X ve Y formunun dağılım şekilleri özdeşdir. Ancak Lord'un eşitlik özelliği, formların özdeş olduğu durumlarda mümkündür. Özdeş formlar oluşturmak pratikte oldukça zordur. Ayrıca, özdeş formlar oluşturulduğunda, eşitlemeye gerek kalmayacaktır. Sonuç olarak, Lord'un eşitlik özelliği ölçüt olarak kullanıldığında, eşitleme mümkün olmayacak veya eşitlemeye gerek kalmayacaktır (Crocker ve Algina, 1986; Kolen ve Brennan, 2004).

### 1.4. Gruptan Bağımsızlık Özelliği

Bu özellik sağlandığında, testlerin eşitlenmesi, gruptan bağımsız olacak ve eşitleme işleminden türetilen dönüştürme benzer tüm koşullara uygulanabilecektir. Örneğin; eşitleme işlemi gruptan bağımsız gerçekleştirildiğinde, kızlar ve erkekler için aynı eşitleme ilişkisi elde edilecektir (Kolen ve Brennan, 2004; Öztürk ve Anıl, 2012).

Peterson (2007), veri toplama işleminin test eşitleme çalışmalarında en önemli durum olduğunu belirtmiştir. Toplanan veriler ve verilerin toplanma şekli eşitleme sürecinde etkili olacaktır. Verileri toplamak için kullanılan birçok farklı desen bulunmaktadır. Aşağıdaki bölümde bu veri toplama desenlerinden kısaca bahsedilmektedir.

## 2. VERİ TOPLAMA DESENLERİ

Veri toplama desenleri, eşitleme için ihtiyaç duyulan verilerin toplanması için yapılan planlamadır. Test eşitlemenin yapılabilmesi için bazı bilgilere ihtiyaç duyulmaktadır. Bu bilgileri toplamak için, her iki testi alan ortak bireylere ya da her iki testte yer alan ortak maddelere ihtiyaç duyulur. Buna göre veri toplama desenleri ikiye ayrılır (Davies, Holland ve Thayer, 2004).

1. Ortak Bireyleri Kullanan Desenler (Common Examinees Designs)
2. Ortak Maddeleri Kullanan Desenler (Common Items Design)

### 2.1. Ortak Bireyleri (Common Examinees) Kullanan Desenler

Ortak bireyleri kullanan desenler,

- Tek Grup Deseni (Single-Group Design),
- Seçkisiz Gruplar Deseni ( Random Groups Design),
- Karşıt Dengelenmiş Tek Grup Deseni (Counterbalanced Single Group Design)

olmak üzere üçe ayrılır (Kolen ve Brennan, 2004).

#### 2.1.1. Tek Grup Deseni (Single Group Design)

Tablo 1’de görüldüğü gibi, tek evrenden, seçkisiz olarak seçilen bireylere, eşitlenecek her iki form (X ve Y formu) uygulanır. Bu desenin avantajı, her öğrenci her iki formu aldığından, X ve Y formlarından elde edilen puanlar arası korelasyon yüksek olması ve eşitleme fonksiyonu sonuçlarının standart hatalarının daha küçük olarak elde edilmesidir. Tek grup deseni iki koşula sahiptir:

- I. Testleri alan bireyler tek evrenden belirlenir.
- II. X ve Y formlarının uygulandığı seçkisiz tek örneklem vardır.

Tek grup deseninin basit ancak güçlü sayıltıları bulunmaktadır. Pratikte, ikinci sayıltı çok uygun olmayabilir.

**Tablo 1:** Tek Grup Deseni

Evren	Örneklem	X	Y
P	1	✓	✓

Bu desende, bireyler her iki formu sırasıyla aldıkları için, bireylerde yorgunluk etkisiyle ikinci form daha zor gelebilir. Tersi bir durumda ise, bireylerin test maddelerine aşına olmalarından dolayı, ikinci form daha kolay gelebilir. Bu sıralama etkilerinden dolayı, bu desen, pratikte sıklıkla tercih edilmemektedir (Kolen ve Brennan, 2004).

### 2.1.2. Seçkisiz Gruplar Deseni (Random Groups Design)

Hambleton ve Swaminathan (1991), iki test formunu eşitlemek için testi alan bireylerin aynı olması gerektiğini belirtmiştir. Ancak pratikte bu uygulama, test formlarını eş gruplara seçkisiz olarak verilerek uygulanmaktadır. Bu desende, Tablo 2’de görüldüğü gibi, bireylerin oluşturduğu P evreninden seçkisiz ve bağımsız olarak seçilmiş iki farklı örneklem ile birinci örnekleme yer alan bireylere uygulanacak olan X formu ve ikinci örnekleme yer alan bireylere uygulanacak olan Y formu bulunmaktadır (Davies, Holland ve Thayer, 2004).

**Tablo 2:** Seçkisiz Gruplar Deseni

Evren	Örneklem	X	Y
P	1	✓	
P	2		✓

Seçkisiz gruplar deseninde, bireyler uygulanacak forma seçkisiz olarak atanır. Bireyler, tek grup deseninin aksine tek form alır. Böylelikle, iki veya daha fazla form alan desenlere göre zaman tasarrufu sağlanır ve daha pratik şekilde uygulanabilir. Ayrıca bu desende, birden fazla form aynı anda uygulanabilmektedir (C, D formu gibi). Bu desende, farklı formları alan bireylerin grup performansları arasındaki fark, test formları arasındaki güçlük farkını ortaya koymaktadır. Her birey ayrı form aldığından dolayı, eşitleme yapabilmek için geniş örnekleme ihtiyaç duyulur (Kolen ve Brennan, 2004).

### 2.1.3. Karşıt dengelenmiş Desen (Counterbalanced Design)

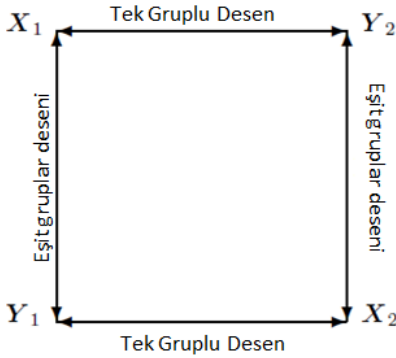
Tablo 3'te görüldüğü gibi, her iki test formunu alan, tek evrenden çekilmiş iki farklı örneklem bulunmaktadır. Tek grup desenindeki gibi, ilk örneklem grubu önce X formunu, ardından Y formunu alır. Diğer örneklem grubu ise, önce Y formunu, ardından X formunu alır. Testlerin sırasının bu şekilde yapılmasındaki amaç, testlerdeki sıra etkisini ortadan kaldırmaktır. Bu sebeple mümkünse, örneklem büyüklükleri eşit veya birbirine yakın olması istenir.

**Tablo 3:** Karşıtdengelenmiş Desen

Evren	Örneklem	X <sub>1</sub>	Y <sub>1</sub>	X <sub>2</sub>	Y <sub>2</sub>
P	1	✓			✓
P	2		✓	✓	

Şekil 1’de (Livingston, 2004) görüldüğü gibi, karşıt dengelenmiş desenler, seçkisiz grup desenini ve tek grup deseninin ikisini de içerir. Bu sebeple X ve Y testlerini eşitlemede kullanılacak birçok yol vardır. Karşıt dengelenmiş desenin iki koşul bulunmaktadır. Bunlar;

- I. Her iki desen için, testi alan bireylerin bulunduğu tek bir evren vardır.
- II. Birey evreninden seçkisiz ve bağımsız seçilmiş iki örneklem vardır.



Şekil 1: Dengelenmiş tek grup deseni

## 2.2. Ortak Maddeleri Kullanan Desenler ( Common Anchors Design)

### 2.2.1. Denk Gruplarda Ortak Madde Deseni

Eşitlenecek testler, seçkisiz olarak seçilen iki ayrı gruba uygulanır. Farklı test formlarını alan bireylerin puanlarını karşılaştırabilmek için her iki gruba da ortak bir A testi verilir. Bu test, grupların farklılıklarını kontrol etmek için kullanılır (Livingston, 2004).

### 2.2.2. Denk Olmayan Gruplarda Ortak Madde Deseni (Nonequivalent Anchor Tests-NEAT)

Bireylerin yeteneklerini kontrol altına almak için, hem testi alan bireylerin eşit düzeyde olmasına, hem de her iki testte ortak madde bulundurulmasına ihtiyaç duyulur. Önceki desenlerde ortak bireyler bulunurken, bu desende ortak maddeler bulunmaktadır (Davier, Holland ve Thayer, 2004).

Bu desende Tablo 4’te görüldüğü gibi, iki farklı evren (P ve Q) vardır ve her iki evrenden birer örneklem seçilir. P evreninden çekilen birey örneğine X testi, Q evreninden çekilen birey evrenine Y testi uygulanır. Ayrıca her iki gruba ortak maddeler yani ankor test (A) uygulanır (Davier, Holland ve Thayer, 2004).

**Tablo 4:** NEAT deseni

Evren	Örneklem	X	Y	A
P	1	✓		✓
Q	2		✓	✓

NEAT desende iki tek grup desen bulunmaktadır. Bu desenler, X-A ve Y-A olmak üzere iki test içerir. Bu testler genelde paralel değildirler. Ankor test (A), X ve Y testlerine göre daha kısa ve daha az güvenilirdir.

NEAT desende testler her iki örnekleme genellikle aynı sırada uygulanır. Böylelikle, eğer sıra etkisi varsa ankor testler ve diğer testler aynı yönde etkilenir. Ankor testler genellikle X ve Y testleri ile benzer içerik ve güçlükte hazırlanırlar. Ayrıca, eşitlenecek olan X ve Y testlerinden alınan puanlarla, ankor testten elde edilen puanlar arasındaki korelasyon ne kadar yüksek olursa, eşitlemede kullanılan ankor test o derece iyi olmaktadır.(Davier, Holland ve Thayer, 2004).

NEAT desenin koşulları aşağıda belirtilmiştir. Bunlar;

- I. Her bir testi ve ankor testi alan iki birey evreni vardır.
- II. İki farklı evrenden sırasıyla, seçkisiz ve birbirinden bağımsız çekilen iki örneklem vardır.

Ankor testlerin eşitlenecek testin içeriğini yansıtması beklenmektedir (Livingston, 2004). Ayrıca ankor testte yer alan maddelerin, eşitlenecek testteki maddelerin güçlüğü ile benzer olması istenir. Angoff (1984), Kolen ve Brennan (2004), ankor testlerin eşitlenecek testlerin minyatürü olması gerektiğini belirtmişlerdir. Kolen ve Brennan (2004), gruplar arasındaki farklılığı doğru yansıtması için ankor testlerin, eşitlenecek testlerle aynı içerik ve istatistiksel özelliklerle oluşturulması gerektiğini açıklamıştır. Holland ve Peterson (2006), ise minyatür ankorun (mini ankor) yanlılığı kaldırıp, test eşitleme fonksiyonunun kestirimindeki hassaslığı artırması gerektiğini ifade etmiştir. Ancak, Sinharay ve Holland (2007), Yi (2009), Cho, Wall, Lee, ve Harris (2010) gerçek ve simülatif veriler kullanarak yaptıkları çalışmada, madde güçlükleri eşit maddelerden oluşan ankor testle (midi ankor) yaptıkları eşitlemede, mini ankor testle yapılan eşitlemeyle benzer performans gösterdiğini, midi ankorların eşitlemenin kalitesini etkilemediğini ortaya koymuşlardır (Liu, Sinharay, Holland, Curley ve Feigenbaum, 2011). Angoff (1984), ankor testlerin uzunluğunun, eşitlenecek testin %20'sinin üzerinde olmasını tavsiye etmektedir. Ankor testin tek değişkeni ölçme zorunluluğu yoktur. Ankor testte farklı değişkenleri ölçen maddelerde (sayısal ve sözel yeteneği ölçen maddeler) yer alabilir (Kelecioğlu, 1994).

Ortak maddelere dayalı, eşit olmayan ankor testler (NEAT);

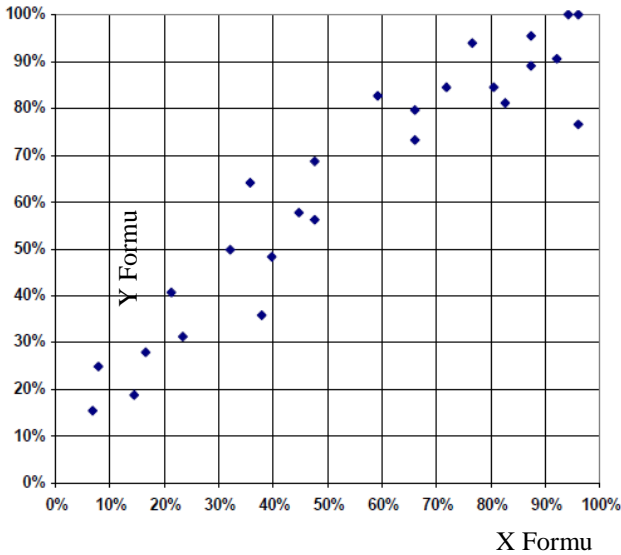
- İç Ankor (Internal Anchor)
- Dış Ankor (External Anchor)

testleri olarak ikiye ayrılmaktadır (Davier, Holland ve Thayer, 2004).



### 2.2.2.1. İç Ankor

İç ankor kullanılacak testlerde, ortak maddeler, eşitlenecek testlerin içerisine yerleştirilir. Eşitlenecek olan X ve Y test puanlarını hesaplamada, iç ankor madde puanları da hesaplamağa dahil edilir (Davier, Holland ve Thayer, 2004). İç ankor deseninde testin uygulaması kolaydır ancak test geliştirme süreci karmaşıktır. Ankor maddelerin her iki grup için de eşit güçlükte olması gerekmektedir. Ancak bu eşitlik tam olarak sağlanamayabilir. Bu sebeple, ankor maddelerin çoğunun güçlüğü sistematik olarak değişmez sayılısının eklenmesi gerekir.



Şekil 2: Ankor Madde Güçlük Grafiği (Livingston, 2004; sf. 32)

Şekil 2 incelendiğinde, her iki formda yer alan ankor maddelerin güçlük yüzdelerinin, her iki form için, hemen hemen eşdeğer güçlükte olduğu görülmektedir. Şekil 2 incelendiğinde, sadece sağ üst köşede gösterilen madde, X

formu için açık bir şekilde kolay, Y formu içinse biraz daha zordur. Bu durum yanlışlık oluşturacağı için bu madde ankor testten çıkartılmalıdır (Livingston, 2004).

İç ankor desenlerinin iki temel kısıtlaması vardır. Birincisi, ankor maddelerin güçlüğünün X ve Y formu için değişmesidir. İkincisi ise, yeterli ankor madde sayısıdır. Eğer testlerde, farklı güçlüklüklerde etkiye sahip maddeler varsa testten çıkartılmalıdır. Ancak, madde sayısının az olduğu testlerde bu durum problem oluşturmaktadır (Livingston, 2004). Örneğin, 6 maddelik bir test geliştirildiğinde, içerisine en az 2-3 ankor madde yerleştirilmesi gerekir. Ancak 2-3 tane ankor maddeye göre test eşitlemek bazı problemler oluşturmaktadır.

Livingston (2004), iç ankor testte yer alacak maddeleri seçerken izlenmesi gereken adımları ve dikkat edilmesi gereken durumları şöyle sıralamıştır:

- Eşitlenecek formlarda yeterli sayıda ankor madde yer alması gerekir. Eşitlenecek testte yer alan maddelerin %20'si kadar madde yeterli görülmektedir.
- Eşitlenecek testlerin içeriğini ve formatını yansıtacak maddeler seçilir.
- Testin güçlüklük ranjını yansıtan maddeler ankor teste dâhil edilir.
- Ankor testlerde yer alan madde değiştirilmez. Mutlaka değiştirilmesi isteniyorsa, testte yer alan maddelerden çıkarma yapılır.
- Ankor testte, madde kümeleri (okuma pasajı, grafik veya bir resim ile ilgili maddeler) yer alacaksa, kümede yer alan tüm maddelerin dâhil edilmesi gerekir.
- Bireylerin zamanı yetiştiremememe durumu olduğu için, ankor maddeler testin sonunda yer almamalıdır.

- Eşitlenecek olan X ve Y formlarında yer alan ankor maddeler, hemen hemen aynı şekilde konumlandırılmalıdır.
- Toplam puanla yüksek korelasyona sahip olacak ankor maddeler seçilmelidir.

#### **2.2.2.2. Dış Ankor**

Dış ankor test bireylere, eşitlenecek testlerden farklı olarak verilir. Dış ankor test puanları, eşitlenecek olan test puanlarını hesaplamada kullanılmaz (Davies, Holland ve Thayer, 2004). Dış ankor testten elde edilen puanların kullanılmamasının sebebi, dış ankor testini, testi alan tüm bireylerin alma zorunluluğunun olmamasıdır (Livingston, 2004). Dış ankor sıklıkla açık uçlu testlerin değerlendirmesinde tercih edilir. Örneğin; öğrencilere, açık uçlu maddelerden oluşan iki farklı test verilir ve bu testleri eşitlemek için aynı becerileri ölçen çoktan seçmeli maddelerden oluşan bir dış ankor test verilir. Eşitlenecek testler ile dış ankor test arasında yüksek korelasyon olması beklenir. Dış ankor deseni, diğer eşitleme desenlerinin uygulanamadığı durumlarda tercih edilir. Bu desenin temel dezavantajı, eşitlenecek testlerle aynı bilgi ve beceriyi ölçeceğinden, ankor testte yer alacak maddelerin hazırlanmasının güç olmasıdır (Livingston, 2004).

Ortak testle eşitlemenin bir başka türü de sözde ortak (quasi-common) testler kullanmaktır. Bu tür ortak testler, aynı maddelerden oluşturulmazlar. Bunlar aynı değişkeni ölçen bir testin farklı iki formu gibidir. U ve W gibi sözde ortak olan iki testin puanları aynı birimde tanımlanır. Yani U testi puanları, W ölçeğine dönüştürülür veya ikisi de ortak bir ölçeğe dönüştürülür. Bu işlemden sonra U ve W formları, ortak test olarak kullanılabilir (Kelecioğlu, 1994).

Kolen ve Brennan (2004) eşitleme desenini belirlemede, testin uygulama karmaşıklığı, test geliştirme zorluğu ve istatistiksel sayıtların yerine getirilebilirliğinin göz önünde bulundurulması gerektiğini belirtmiştir. Crocker ve Algina (1986), eşitleme desenini belirlemede, pratikliğin birinci ölçüt olduğunu belirtmiştir. Seçkisiz gruplar deseninde, test formlarının oluşturulması diğer desenlere göre daha az karmaşıktır. Çünkü, tüm testin içeriğini temsil edecek ortak madde oluşturmaya gerek yoktur. Formların bireylere seçkisiz olarak atanması ve sıra etkisinin olmamasından dolayı istatistiksel sayıtların yerine getirilmesinde en düşük probleme sahiptir ( Kolen ve Brennan, 2004). Kim (2006), pratikte sıklıkla NEAT deseninin kullanıldığını açıklamış olsa da, küçük örnekleme sahip öğretmen yapımı testlerde, kurs merkezlerinin yapmış olduğu sınavlarda vb. sıklıkla seçkisiz gruplar deseninin kullanılmaktadır. Ayrıca testlerin hazırlanmasının ve uygulanmasının kolay olması ve istatistiksel sayıtlarının yerine getirilmesinde en düşük probleme sahip olması nedeniyle bu çalışmada seçkisiz gruplar deseni tercih edilmiştir.

Eşitleme, testin bir formunun puanlarının diğer form ölçeğindeki puanlara dönüşümünde kullanılan istatistiksel bir işlemdir. Formların dönüşümlerini sağlayacak birçok yöntem bulunmaktadır. Dorans ve Holland (2000), eşitleme yöntemlerinin, doğrusal yöntemler-doğrusal olmayan yöntemler, Klasik Test Kuramına dayalı yöntemler-Madde Tepki Kuramına dayalı yöntemler ve gözlenen puanlara dayalı yöntemler-gerçek puanlara dayalı yöntemler gibi farklı sınıflandığını belirtmiştir.

Test eşitleme çalışmalarında kullanılacak birçok test eşitleme yöntemi bulunmaktadır. Doğru eşitleme yapılabilmesi için en uygun eşitleme yönteminin seçilmesi gerekmektedir. Crocker ve Algina (1986), eşitleme yöntemini seçmek için 3 ana ölçüt olduğunu belirtmiştir. Bu ölçütler; sayıtların sağlanabilirliği, kullanılabilirlik ve tutarlılıktır. Holland ve Dorans (2006) ise,

eşitleme yönteminin taksonomisini geliştirirken 3 faktörün belirleyici olduğunu belirtmiştir. Bu faktörler;

- 1- Veri toplama deseni ortak evrenden mi? Yoksa ortak kişilerden mi?
- 2- Gözlenen puanlar işlemi mi? Gerçek puanlar işlemi mi?
- 3-Doğrusal veya doğrusal olmayan yöntemler mi?

Aşağıdaki bölümde test eşitleme yöntemlerinden kısaca bahsedilmiştir.

### **3. TEST EŞİTLEME YÖNTEMLERİ**

Test eşitleme yöntemleri dayandığı kurama göre ikiye ayrılmaktadır:

- 1- Klasik Test Kuramına dayalı test eşitleme yöntemleri
- 2- Madde Tepki Kuramına dayalı test eşitleme yöntemleri

Madde Tepki Kuramına dayalı test eşitleme yöntemleri, bu çalışma kapsamı dışında tutulmuştur. Bu nedenle, aşağıda sadece Klasik Test Kuramına dayalı test eşitleme yöntemlerinden bahsedilmiştir.

#### **3.1. Özdeş Eşitleme (Identity Equating)**

Özdeş eşitleme, diğer yöntemlere göre en basit eşitleme yöntemidir. Formlar arasında puan dönüşümü yapılmadığından dolayı gerçek bir eşitleme yöntemi olarak düşünülmebilir. Özdeş eşitleme ile eşitlememe arasında hiçbir fark yoktur. Özdeş eşitleme yönteminin matematiksel fonksiyonu 1 no'lu eşitlikte ifade edilmektedir:

$$y = ID_y(x) = x \quad (1)$$

1 no'lu eşitlikte yer alan,  $x$ , X formundan elde edilen ham puanı;  $y$ , Y formu üzerinde  $x$  ham puanı için eşitlenmiş puanı belirtmektedir.

Skaggs (2005) ve Kim ve diğerleri (2006), özdeş eşitlemede, formlardan elde edilen puanlar birebir eşitlendiği için, seçkisiz hatanın sıfır olduğunu ancak formlar paralel olmadığı durumda sistematik hatanın (örneğin yanlılık) yüksek olacağını ifade etmişlerdir. Kolen ve Brennan (2004), küçük örneklemelerde, test formlarının paralel olduğuna inanıldığı zamanlarda, özdeş eşitleme yapılabileceğini belirtmiştir.

### 3.2. Ortalama Eşitleme (Mean Equating)

Eşitlenecek formlardan, X formunun Y formundan güçlük bakımından, ölçek puanı üzerinden sabit miktarda farklılaştığı düşünülür (Kolen ve Brennan, 2004). Ortalama eşitlemede, formu alan bireylerin yetenek düzeylerinde bir farklılık yoktur. Ortalama eşitleme yönteminde iki form 2 ve 3 no'lu eşitlikte belirtildiği şekilde eşitlenmektedir.

$$x - \mu_X = y - \mu_Y \quad (2)$$

$$m_Y(x) = y = x - \mu_X + \mu_Y \quad (3)$$

Eşitlik 3'te yer alan,  $x$ , X formundan elde edilen puanı,  $\mu_X$ , X formunun ortalamasını;  $y$ , Y formundan elde edilen puanı;  $\mu_Y$ , Y formunun ortalamasını ve  $m_Y(x)$ , X formunda yer alan ve ortalama eşitleme yöntemi kullanılarak  $x$  puanının Y formu ölçeğine dönüştürülmüş puanını belirtmektedir.

### 3.3. Doğrusal Eşitleme (Linear Equating)

Crocker ve Algina'ya (1986) göre doğrusal eşitleme yöntemi, X ve Y formlarının puan dağılımlarının aynı, ortalama ve standart sapmalarının farklı olduğu varsayımına dayanmaktadır. Bu formlardan türetilen standart puanların eşit olduğu düşünüldüğünde bu yöntem kullanılmaktadır. Donlon (1984), bir testin farklı formlarını alan gruplar eşit yetenek düzeyine sahip olduğu durumlarda doğrusal eşitleme yöntemlerinin kullanılabileceğini ifade etmiştir.

Angoff (1984, syf. 564) doğrusal eşitlemeyi, "İki test formundan elde edilen puanlar eşit standart puan sapmalarına karşılık geldiğinde, bu puanlar eşdeğerdir" şeklinde tanımlamıştır. Bu durumu 4 no'lu eşitlikte belirtilen şekilde ifade etmiştir;

$$\frac{Y-\mu_Y}{\sigma_Y} = \frac{X-\mu_X}{\sigma_X} \quad (4)$$

4 no'lu eşitlik yeniden düzenlendiğinde 5 no'lu eşitlik elde edilir.

$$Lin_Y(x) = \frac{\sigma_Y}{\sigma_X} X + \mu_Y - \frac{\sigma_Y}{\sigma_X} \mu_X \quad (5)$$

5 no'lu eşitlik,  $Y=AX+B$  şeklinde ifade ettiğimizde,  $A = \frac{\sigma_Y}{\sigma_X}$ , yani doğrunun eğimini,  $B = \mu_Y - \frac{\sigma_Y}{\sigma_X} \mu_X$  ise, doğrunun kesişim noktasını gösterir. Bu puan dönüşümü, regresyon eşitlemelerinin aksine, diğer eşitleme yöntemleri gibi simetrik değildir (Angoff, 1984).

Doğrusal eşitlemenin bazı sınırlılıkları bulunmaktadır. Bunlardan ilki, formlardan alınan en yüksek ve en düşük puanların, diğer formlarda maksimum ranjda karşılığının bulunmama durumunun olmasıdır (Livingston, 2004). Livingston (2004), bu durumu şu örnekle açıklamıştır. 100 maddelik X ve Y gibi iki farklı forma sahip bir test (her madde 1 ham puana sahiptir), doğrusal eşitleme

yöntemi ile eşitlendiğinde, eğer Y formu, X formundan daha zor ise, Y formundan alınan 99 ham puanı, X formundaki 103 ham puanına denk gelebilir. En yüksek 100 ham puanlık bir testten, 103 ham puan alınması bir sorunun olduğunu gösterir. Doğrusal eşitlemede diğer bir sınırlılık, doğrusal eşitleme sonuçlarının, testi alan gruplara çok bağlı olmasıdır (Livingston, 2004). Test formlarını alan grup değiştiğinde, örneğin X testini alan grubun çok başarılı ve Y testini alan grubun çok başarısız olduğu durumlarda, eşitleme sonuçlarında farklılıklar oluşabilir. Başarılı grup, kolay formu aldığı zaman puanların dağılımı küçük puan aralığında gerçekleşirken, zor formda puanlar geniş bir aralığa dağılır. Başarısız grupta ise tersi bir durum gerçekleşir. Kolay formda, puan dağılımları genişlerken, zor formda daralır (Livingston, 2004).

### **3.4. Eşit Yüzdellikli Eşitleme (Equipercentile Equating)**

Formların dağılımlarının (çarpıklık, basıklık), aynı olduğu durumlarda doğrusal eşitleme yöntemi kullanılır. Eğer formların dağılımı aynı değilse, formlar arasında doğrusal bir ilişkiden bahsedilemez. Bu durumda formlar arası eğrisel bir ilişki vardır (Livingston, 2004). Formların puan dağılımları aynı olmadığı durumlarda eşit yüzdellikli eşitleme yöntemi önerilmektedir. X formu, düşük ve yüksek puanlarda Y formundan çok daha zor olabilirken, orta puanlarda daha az zorlukta olabilir (Kolen ve Brennan, 2004). Eşit yüzdellikli eşitlemede, X ve Y formunu alan bireyler yüzdellik sıralarına göre eşitlenir (Kolen,1988). Y formuna dönüştürülen X formundaki puanların dağılımı, Y formundaki puan dağılımına eşit ise, iki form arasındaki eşitleme fonksiyonu eşit yüzdellikli eşitleme fonksiyonudur. Bu fonksiyon, Y formundaki puanlarla, aynı yüzdedeki sıraya sahip X formundaki puanları tanımlayarak geliştirilir (Kolen ve Brennan, 2004).



Eşit yüzdellikli eşitlemede, yapılması gereken ilk iş, her form için yüzdellik sıraları hesaplamaktır. Aynı yüzdellik sıraya gelen puanlar eşdeğerdir (Kolen, 1988; Livingston, 2004; Kolen ve Brennan, 2004). Ancak bu eşitlemeler, test puanlarının sürekli (continious) değişken olduğu varsayımına göre yapılmaktadır. Gerçekte test puanları, kesikli (discrete) değişkenlerdir. Test puanları kesikli olduğu zaman, eşit yüzdellikli eşitleme fonksiyonu tam olarak çalışmaz. Bu sınırlılığı aşmak için, kesikli değişkenler, yüzdellikli veya yüzdellikli sıralara dönüştürülerek sürekli hale getirilebilir (Kolen ve Brennan, 2004).

Eşit yüzdellikli yöntem ile eşitlemeler hem grafikte gösterim yoluyla hem de matematiksel formüllerle gerçekleştirilebilir. Grafikte gösterim, büyük örnekleme sahip çalışmalarda tercih edilmez. Bunun yerine matematiksel formüllerle hesaplama yapılır. Eşitleme işlemi için öncelikle yüzdellik sıralama fonksiyonunun hesaplanması gerekir. Yüzdellik sıralama fonksiyonunu hesaplamak için 6 no'lu eşitlik kullanılır.

$$P(x)=100\{F(x^*-1)+[x-(x^*-0,5)][F(x^*)-F(x^*-1)]\} \quad (6)$$

6 no'lu eşitlikte yer alan,

- $x \rightarrow X$  formunda yer alan puanı gösteren seçkisiz değişken
- $K_x \rightarrow X$  formundaki madde sayısı
- $F(x) \rightarrow$  evrendeki  $x$  puanı veya daha aşağısını alan bireylerin oranı
- $x^* \rightarrow x'$ 'e en yakın tamsayı (Örneğin  $x=2,1$  ise  $x^*=2$  veya  $x=3,6$  ise  $x^*=4$ )
- $P(x) \rightarrow X$  formu için Yüzdellik sıralama fonksiyonu

belirtmektedir.

$$-0,5 \leq x \leq K_x + 0,5$$

$$x < -0,5 \text{ ise } P(x) \text{ ve } x \geq K_x + 0,5 \text{ ise } P = 100 \quad (\text{Kolen ve Brennan, 2004}).$$

Yüzdeler sıralama fonksiyonu hesaplandıktan sonra formların dönüşümünü matematiksel olarak 7 ve 8 no'lu eşitliklerde belirtildiği şekilde hesaplanır.

$$x_L(P^*) = P^{-1}[P^*] = \frac{P^*/100 - F(x_L^*)}{F(x_L^* + 1) - F(x_L^*)} + (x_L^* + 0,5) \quad (7)$$

$$x_U(P^*) = P^{-1}[P^*] = \frac{P^*/100 - F(x_U^* - 1)}{F(x_U^*) - F(x_U^* - 1)} + (x_U^* - 0,5) \quad (8)$$

Bu eşitliklerde,

- $x_L^* \rightarrow P^*$ 'den büyük olan yüzdeler sıraya denk gelen en küçük madde ham puanı
- $x_U^* \rightarrow P^*$ 'den küçük olan yüzdeler sıraya denk gelen en büyük madde ham puanı
- $K_x \rightarrow X$  formundaki madde sayısı
- $F(x) \rightarrow$  evrendeki  $x$  puanı veya daha aşağısını alan bireylerin oranı
- $P^* \rightarrow$  Verilen yüzdeler sıra

ise;

Her iki eşitlikte de aynı değerler elde edilir (Kolen ve Brennan, 2004).

Eşit yüzdelerli eşitlemede, herhangi bir bireyin belirli bir puanı olmadığı durumlarda sorun ortaya çıkmaktadır. Bu durumda, aynı yüzdeler sıraya denk gelen puanların ranjı alınır ve bu puanların orta noktası eşdeğer olarak kabul edilir.

Eşit yüzdelerli eşitleme yönteminde, örneklem büyüklüğü önemli bir faktördür. Örneklem büyüklüğü yeterli büyüklükte olmadığı zamanlarda

dağılımın çarpık olma durumu vardır. Dağılımlar arasında aşırı farklılıklar bulunduğu durumda eşit yüzdelikli eşitleme yöntemini kullanmak anlamsız hale gelir (Kan, 2011).

### **3.4.1.Eşit Yüzdelikli Eşitleme Yönteminde Düzgünleştirme (Smoothing in Equipercntile Equating)**

Eşit yüzdelikli eşitleme yönteminde, test formlarını alan bireyler bir veya birden fazla evrenin örneklemleridir. Bu örneklemler çekilirken, örnekleme hatalarından dolayı ham puan dağılım grafiklerinde bazı düzensizlikler görülür (Kolen ve Brennan, 2004). Örnekleme hatalarını azaltmanın bir yolu örnekleme büyüklüğünü artırmaktır. Kolen ve Brennan (1995), eşit yüzdelikli eşitleme yöntemi için yeterli örnekleme büyüklüğünü 1500 olarak belirtmiştir. Ancak her zaman bu büyüklükte bir örnekleme çalışmak mümkün olmayabilir. Bu sebeple düzgünleştirme (smoothing) işlemine başvurulur (Dolan, 1984; Cui ve Kolen, 2009).

Düzgünleştirme, örnekleme hatalarını azaltmak için kullanılır. Dağılımın genişliği, şekli ve yeri değiştirilmeden düzensizlikleri gidererek yeni bir gözlenen puan dağılımı oluşturma işlemine düzgünleştirme (smoothing) adı verilir (Livingston, 2004).

Düzgünleştirme işlemi el ile veya analitik yöntemlerle yapılabilir. Kolen ve Brennan (2004), düzgünleştirme yöntemlerinin; evren dağılımını doğru kestirmesi gerektiğini, dağılımların çeşitliliği ile başa çıkabilecek kadar esnek olması gerektiğini, uyum çalışmaları için istatistiksel bir çerçevenin (statistical framework) olmasında fayda olacağını ve son olarak, kestirim gücünü artırması gerektiğini belirtmiş ve bu özelliklere analitik yöntemlerin sahip olduğunu belirtmiştir.

Düzgünleştirme işlemi düzgünleştirmenin yapılma zamanına göre ön-düzgünleştirme ve son-düzgünleştirme olarak ikiye ayrılmaktadır. Eşitleme işlemlerine başlamadan önce ham puanların frekans dağılımları üzerinde yapılan düzgünleştirme işlemine ön-düzgünleştirme (pre-smoothing) adı verilir (Livingston, 2004; Kolen ve Brennan, 2004). Ön-düzgünleştirme yöntemleri olarak, Ortalama Yuvarlama (Rolling Average), Log-Linear Yöntem ve Güçlü Gerçek Puan yöntemi (Strong True Score Method) kullanılmaktadır. Eğer puan dağılımları ile ilgili düzgünleştirme, eşitleme işlemlerinden sonra elde edilen dönüşümlere uygulanıyorsa son-düzgünleştirme (post-smoothing) olarak adlandırılır (Kolen ve Brennan, 2004; Öztürk, 2010). Son-düzgünleştirme yöntemi olarak kübik şeritler (cubics plines) yöntemi ve Polinomlar kullanılmaktadır.

Düzgünleştirme yeterli düzeyde olmazsa, puan dağılımındaki düzensizlikler kaldırılmayacaktır. Tersine durumda, düzgünleştirme çok güçlü düzeyde yapılırsa, puan dağılımının şekli değişecektir (Livingston, 2004). Bu sebeple düzgünleştirme işlemi yapılırken düzgünleştirmenin gücü doğru belirlenmelidir.

### **3.5.Eşit Yüzdellikli Eşitlemede Polinomal Loglinear Ön-düzgünleştirme Yöntemi (Polynomial Log-Linear Method in Equipercetile Equating)**

Eşit yüzdellikli eşitlemede büyük örneklem büyüklüğüne ihtiyaç duyulmaktadır. Kolen ve Brennan (1995), eşit yüzdellikli eşitlemenin, örneklem büyüklüğü 1500 ve üzeri olduğu durumlarda doğru sonuçlar verdiğini belirtmişlerdir. Hanson, Zeng ve Colton (1994), yapmış oldukları çalışmada düzgünleştirme işlemi ile küçük örneklerde eşit yüzdellikli eşitleme yönteminin doğruluğunu artırdığını ifade etmişlerdir. Benzer bir şekilde Cui ve

Kolen (2009), düzgünleştirmenin eşitleme kestirimini, toplam hatayı (seçkisiz hata ve yanlılık) azaltarak geliştirdiğini açıklamışlardır.

İlgili literatür incelendiğinde, birçok ön-düzgünleştirme çalışması yapıldığı görülmektedir. Çalışmalarda en sık kullanılan ön-düzgünleştirme yöntemi Darroch ve Ratcliff (1972), Haberman (1974a, b, 1978), Rosenbaum ve Thayer (1987) ve son olarak Holland ve Thayer (1987, 2000), tarafından tanımlanan Polinomial Loglinear ön-düzgünleştirme yöntemidir. Bu yöntemin çalışmalarda sıklıkla tercih edilme nedenlerinden birisi esnekliğidir. Ayrıca diğer yöntemlere nispeten uygulanması daha kolaydır (Holland ve Thayer, 2000; Hanson, 1991; Moses ve Holland, 2010). Bu sebeple bu çalışmada Polinomial Log-Linear ön-düzgünleştirme yöntemi tercih edilmiştir.

Kolen ve Brennan (2004) tarafından açıklanan Polinomial Log-Linear yöntem için 9 no'lu eşitlik kullanılmaktadır.

$$\log[N_x f(x)] = \omega_0 + \omega_1 x + \omega_2 x^2 + \dots + \omega_c x^c \quad (9)$$

Bu eşitlikte, N, örneklem büyüklüğünü,  $f(x)$ , bağıl frekans dağılımına uygulanan modeli, C, polinom derecesini ve  $\omega_0, \omega_1, \omega_2, \dots$  ve  $\omega_c$ , polinomial fonksiyon için kestirilen parametreleri temsil etmektedir. Bu yöntem, C düzeyinin alt sıra polinomu olarak açıklanmıştır. Örneğin, C=2 ise  $\log[N_x f(x)] = \omega_0 + \omega_1 x + \omega_2 x^2$  ve model 2. dereceden polinomialdir. Bu durumda düzgünleştirilen dağılım ile gözlenen puan dağılımının ortalaması ve standart sapması aynıdır, diğer momentler (çarpıklık, basıklık vb.) farklıdır. Modeldeki  $\omega$  parametresi maksimum olabilirlik yöntemi ile kestirilebilir (Kolen ve Brennan, 2004).

### 3.6. Dairesel Yay Yöntemi (Circle-Arc Method)

Livingston ve Kim (2009), küçük örneklerde test eşitlemeyi sağlayan yeni bir yöntem önermişlerdir ve küçük örneklerde verilerden kestirimde bulunabilmek için parametre sayısını azaltarak, yani verilerin bazı varsayımlarını yok sayarak eşitleme yapılabileceğini savunmuşlardır. Bu sebeple, Dairesel yay yöntemini geliştirmişlerdir. Bu yöntem, eşitleme ilişkisini doğrusal olarak varsaymayan güçlü bir modeldir. Bu model Divgi'nin (1987) fikrini temel almıştır. Divgi (1987), önceden belirlenmiş iki nokta ve görgül hesaplanmış bir orta nokta üzerinden geçen eşitleme eğrisi elde etmiştir (Livingston ve Kim, 2009). Livingston ve Kim (2010), Dairesel yay yöntemini “simetrik (symmetric) Dairesel yay” ve “basitleştirilmiş (simplified) Dairesel yay” olmak üzere ikiye ayırmıştır. Simetrik Dairesel yay yönteminde, iki bitiş noktası ve görgül olarak belirlenmiş bir orta nokta, dairesele yaya uydurulmaktadır. Basitleştirilmiş Dairesel yay yönteminde ise, eşitleme dönüşümü doğrusal ve eğrisel bileşenlerine ayrılmaktadır. Dairesel yay yönteminde alt son nokta, her formdaki anlamlı en düşük test puanıdır. Üst noktası ise, o testten alınabilecek maksimum puandır. Orta nokta ise puan dağılımların orta noktasıdır. Tek grup, dengelenmiş ve eşit gruplar deseninden elde edilen puanlarda, test formlarının puan ortalaması orta noktayı oluşturur. Eğer bu üç nokta aynı çizgi üzerindeyse, bu çizgi kestirilen eşitleme eğrisidir. Ancak bu üç nokta, doğrusal bir çizgi üzerinde değilse, bu noktalar dairesele bir yay oluştururlar (Livingston ve Kim, 2009).

Eşitleme fonksiyonun doğrusal ve eğrisel olmak üzere iki bileşeni vardır. Fonksiyonun doğrusal bileşenini hesaplamada alt ve üst noktalar kullanılır. X formundaki  $x_i$  puanı için 10 no'lu eşitlik ile y formundaki fonksiyonun doğrusal bileşeni hesaplanır.

$$E_{Lin-Y}(x) = y_1 + \frac{y_3 - y_1}{x_3 - x_1}(x_i - x_1) \quad (10)$$

Doğrusal bileşen hesaplandıktan sonra, üç nokta, doğrusal bileşen değerinden ( $E_{Lin-Y}(x)$ ) çıkarılarak puanların dönüşümleri hesaplanır. Dönüştürülen noktaların yüksekliği  $y^*$  şeklinde gösterilirse 11 no'lu eşitlik ile hesaplanır.

$$y^* = y - E_{Lin-Y}(x) \quad (11)$$

Eğrisel bileşen fonksiyonunu bulmak için; dairenin yarı (r) çapı ve ( $x_c$ ,  $y_c$ ) yarıçapın orta noktasının koordinatları ise 12 no'lu eşitlik kullanılır.

$$E_{Arc-Y}(x) = y_c \pm \sqrt{r^2 - (x_i - x_c)^2} \quad (12)$$

$y_2^*$  pozitif ise aradaki işaret "+", negatif ise aradaki işaret "-" olur. Ayrıca r ve r'nin koordinatlarını hesaplamak için 13, 14 ve 15 no'lu eşitlikler kullanılır.

$$x_c = \frac{(x_3^2 - x_1^2)}{2(x_3 - x_1)} \quad (13)$$

$$y_c = \frac{(x_1^2)(x_3 - x_2) - (x_2^2 + y_2^*)^2(x_3 - x_1) + (x_3^2)(x_2 - x_1)}{2[y_2^*(x_1 - x_2)]} \quad (14)$$

$$r = \sqrt{(x_c - x_1)^2 + (y_c)^2} \quad (15)$$

Sonuç olarak eşitlenmiş puanlar, 16 no'lu eşitlik ile hesaplanır (Livingston ve Kim, 2009).

$$E_{CIRC-Y} = E_{Lin-Y}(x) + E_{Arc-Y}(x) \quad (16)$$

### 3.7. Denk Olmayan Gruplarda Ortak Madde (NEAT) Desenlerinde Kullanılan Eşitleme Yöntemleri

NEAT desenlerde, bir testin X formu, aynı testin Y formuna, her iki testte bulunan ankor (ortak) maddelerle eşitlenir. Ankor maddeler, içerik açısından testin minyatürü, istatistiksel özellikler olaraksa aynı veya benzer

olması gerekmektedir. Testin farklı formları farklı evrenlere uygulanır. Ancak, eşitleme fonksiyonun belirtileceği tek bir evrene ihtiyaç vardır. Bu yüzden iki evren birleştirilerek hipotetik tek bir evren altında toplanmıştır. Braun ve Holland (1982), bu evreni yapay evren (synthetic population) olarak adlandırmıştır (Kolen, 1985). NEAT desen ile elde edilen veriler sentetik evren altında eşitleme yöntemlerine tabi tutulur.

### **3.7.1. Zincirlenmiş Doğrusal Eşitleme Yöntemi (Chained Linear Equation)**

P evreni üzerinde X formundan A formuna (ankor) ve Q evreni üzerinde A formundan Y formuna olmak üzere iki doğrusal ilişkilendirme fonksiyonunun birlikte zincirleme ile oluşan fonksiyona zincirlenmiş doğrusal eşitleme fonksiyonu adı verilir. Bu yöntemin sayılışı, X formundan A formuna ve A formundan Y formuna yapılan doğrusal ilişkilendirme fonksiyonu her iki evren için aynıdır (Davier, Holland ve Thayer, 2004).

### **3.7.2 Zincirlenmiş Eşit Yüzdellikli Eşitleme Yöntemi (Chained Equipercentile Equation)**

Zincirlenmiş doğrusal eşitleme yöntemi ile aynı işlem basamaklarına sahiptir. P evreninde yer alan X formunu, A formuna ve Q evreninde yer alan A formunu Y formuna olmak üzere iki eşit yüzdellikli eşitleme fonksiyonunun birlikte zincirlenmesiyle oluşan fonksiyona zincirlenmiş eşit yüzdellikli eşitleme fonksiyonu adı verilir. Zincirlenmiş doğrusal eşitleme yönteminden tek farkı, formlar arası kullanılan fonksiyonu eşit yüzdellikli eşitleme fonksiyonu olmasıdır (Davier, Holland ve Thayer, 2004).



### 3.7.3 Tucker Yöntemi

Gulliksen (1950), tarafından tanımlanan Tucker yöntemi, NEAT desen kullanılarak elde edilen verilerin eşitlenmesinde kullanılan doğrusal eşitleme yöntemidir. Tucker yöntemi, eşitlemenin parametrelerini kestiriminde iki tür sayılıya sahiptir. İlk sayılı, doğrusal regresyon sayılıdır. Ankor test (A) üzerindeki X ve ankor test (A) üzerindeki Y'nin regresyonları doğrusaldır ve her iki evren için aynıdır. Diğer sayılı ise, ankor test üzerinde verilen her hangi bir puan için X ve Y puanlarının koşullu varyansı her iki evren için aynıdır (Kolen ve Brennan, 2004; Livingston, 2004).

### 3.7.4 Levine Yöntemi

Tucker yönteminin sayılıları sadece gözlenen puanlar içindir. Gerçek puanlar için sayılısı bulunmamaktadır. Tucker yöntemi gibi doğrusal eşitleme yöntemi olan Levine yöntemi, gerçek puanlara dayalı sayılılara sahiptir. İlk sayılı korelasyonel sayılıdır. X, Y, ve A aynı yapıyı ölçmektedir ve X, Y ve A testlerinin gerçek puanları ( $T_X$ ,  $T_Y$ ,  $T_A$ ) arasında her iki evren için mükemmel bir korelasyon vardır. İkinci sayılı,  $T_X$  ile  $T_A$  ve  $T_Y$  ile  $T_A$  gerçek puanlarının regresyonları her iki evren için aynı doğrusal fonksiyona sahip olduğudur. Son sayılı ise X, Y ve A test puanlarının hata varyansları her iki evren için aynıdır (Kolen ve Brennan, 2004; von Davier, 2008).

### 3.7.5 Frekans Kestirim Yöntemi (Frequency Estimation Equipercentile Equating)

Angoff (1971), tarafından tanımlanan frekans kestirim yöntemi, NEAT desene dayalı eşit yüzdelikli eşitleme yöntemi altında kullanılmaktadır. Ankor testte verilen her puan için X ve Y testi toplam puanlarının koşullu dağılımları her

iki yapay evren için aynıdır sayılıısına sahiptir. Bu yöntem, yapay evrendeki X ve Y form puanlarının birikimli dağılımlarını kestirmede kullanılmaktadır. Yüzdeler sıralar birikimli dağılımlardan elde edilerek, formlar eşit yüzdelerli yönteme göre hesaplanır.

#### 4. EŞİTLEME HATASI

Eşitleme yapılırken dikkat edilmesi gereken önemli noktalardan biri istatistiksel hatadır. Kolen ve Brennan (2004), eşitleme hatasını seçkisiz ve sistematik eşitleme hatası olarak ikiye ayırmıştır. Seçkisiz eşitleme hatası, diğer bir ifade ile örneklem hatası, testi alan kişiler evreninden seçilen örneklem ortalama, standart sapma, yüzdeler sıralama gibi parametreleri kestirildiğinde oluşur (Kolen,1988). Seçkisiz hata, örneklemden ve evrenden kestirilen eşitleme ilişkisi arasındaki fark olarak da adlandırılabilir. Kolen (1988) örneklem büyüklüğü artırılarak ve uygun eşitleme deseni seçilerek seçkisiz örneklem hatasının azaltılabileceğini belirtmiştir. Örneklem yerine tüm evrenle çalışıldığında bu hata ortadan kalkacaktır (Kolen ve Brennan, 2004). Sistematik hata, kullanılan eşitleme yönteminin varsayımları veya koşulları karşılanmadığı durumlarda oluşur. Örneğin, tek grup desenlerde yorgunluk etkisi yüzünden başarısız olmak veya pratik etkisi yüzünden yüksek puan almak birer sistematik hatadır. Seçkisiz gruplar deseninde sarmalımsı işlem, grupların karşılaştırılmalarını başarmada etkisiz ise sistematik hata oluşur. Sistematik hatalar özellikle eşdeğer olmayan gruplar deseninde (NEAT) daha fazla sorun oluşturmaktadır. NEAT desende, Tucker ve Levine gibi eşitleme yöntemlerinin koşullarını karşılamak zor olduğu için sistematik hata yapma olasılığı yüksektir. Bu koşullar karşılanmadan yapılan eşitlemeler sonucunda sistematik hata oluşmaktadır. Örneğin ankor maddelerin tüm test formunun içeriğiyle aynı olması veya istatistiklerin benzer özellikte olması gerçekleştirmesi zor bir sayılıdır. Ayrıca NEAT desenlerde, ortak maddenin X formunda farklı Y formunda farklı

davranması da sistematik hataya yol açacaktır. Sonuç olarak, herhangi bir eşitleme deseninde aynı testin X ve Y formları arasında içerik, güçlük ve güvenilirlik farkı varsa sistematik hatanın olduğu söylenebilmektedir (Kolen,1988).

Örneklem büyüklüğü artırılarak seçkisiz hata azaltılır, ancak örneklem büyüklüğünü artırmak sistematik hataların azalmasına yol açmamaktadır. Düzgünleştirme (smoothing) işlemi doğru kullanılmadığında eşitleme hatasına yol açabilir. Düzgünleştirmenin doğru kullanımında seçkisiz hatalar azalırken, düzgünleştirme fazla yapıldığında sistematik hata oluşabilmektedir (Kolen ve Brennan, 2004). Ayrıca, eşitlenecek test formlarının sayısı arttıkça seçkisiz veya sistematik hatanın artma ihtimali artacaktır.

Livingston (2004), eşitlemenin standart hatasını (SEE), eşitlenen puanların örneklem dağılımlarının standart sapması şeklinde tanımlamıştır. Wang (2006) eşitlemenin standart hatasını, seçkisiz örneklem hatasının indeksi olduğunu belirtmiş ve tekrarlı örneklemelerden elde edilen eşitlenmiş puanların standart sapması olarak tanımlamıştır. Her tekrarlı yapılan eşitleme işleminde, örneklem evren veya evrenlerden seçkisiz olarak seçilir. Daha sonra, X formundan elde edilen ham puanlarının Y formunda eşitleme yöntemleri kullanılarak farklı düzeydeki eşdeğerlikleri bulunur. Her puan düzeyinde eşitlemenin standart hatası, tekrarlı eşitlemelerin standart sapmasına eşittir. Standart hata, dağılımın ortasında küçük olurken, dağılımın uçlarında daha büyük olmaktadır (Bozdağ, 2007). Kolen ve Brennan (2004), örneklem büyüklüğünün eşitlemenin standart hatasını doğrudan etkilediğini ve örneklem büyüklüğünün artırılarak, eşitlemenin standart hatasının sifıra yaklaşacağını ifade etmişlerdir.

Eşitlemenin standart hatası genellikle, test puanlarının raporlaştırılmasında eşitlemenin hatasını açıklamada, eşitlemenin doğruluk

düzeşini sonuçlandırma için gerekli olan örneklem büyüklüğünü hesaplamada ve eşitleme yöntemleri ve eşitleme desenlerini karşılaştırmada bir esas olarak kullanılır (Kolen, 1985).

#### 4.1 Önyükleme (Bootstrap) Yöntemi

Efron ve Tibshirani (1993) tarafından önerilen bu yöntem, “yeniden örnekleme” yöntemi olarak da adlandırılabilir (Kolen ve Brennan, 2004). Bu yöntemde öncelikle evrenden bir örneklem seçkisiz olarak seçilir. Seçilen örneklem verilerinin yerlerini deęiştirmeye, seçkisiz ve farklı birçok örneklem (bootstrap) çekilir ve her örneklem için eşitleme fonksiyonu tahmin edilir ve ilgili dięer istatistikler hesaplanır. Hesaplanan deęerlerin standart sapmaları eşitlemenin standart hatasını verir.

Bootstrap örneklemini oluşturmak için bilgisayarlardan faydalanılmaktadır. Efron ve Tibshirani (1993) replikasyon sayısının 25 ile 200 arasında yeterli olacağını belirtmişlerdir. (Kolen ve Brennan, 2004).

### 5. ÖLÇÜT EŞİTLEME

Eşitlemenin ne derece doğru yapıldığını belirleyebilmek için ölçüt eşitlemeye ihtiyaç duyulmaktadır. Harris ve Crouse (1993), yapılan eşitlemeyi deęerlendirmede kullanılabilecek ölçütleri *dairesel paradigmadaki eşitleme, simülasyon verileri, geniş örneklem ölçütü, standart hata, indisler ve tutarlık* olarak açıklamıştır. Kolen ve Brennan (2004) ise, *eşitlik, simetri ve evren deęişmezlięi* gibi eşitleme özelliklerinin ölçüt deęerlendirmeyi geliştirmede kullanılabileceğini ifade etmiştir.

Cope (1987), yapmış olduęu çalışmada dairesel eşitlemeyi ölçüt olarak kullanmıştır (Heh, 2007). Dairesel eşitlemede, X, Y ve Z formları için, X formu

Y formuna, Y formu Z formuna ve Z formu X formuna eşitlenir (Kolen & Brennan, 2004). Kolen ve Brennan (2004), dairesel eşitlemenin ölçüt olarak kullanıldığı durumlarda, özdeş eşitlemenin her zaman diğer eşitleme yöntemlerine göre daha iyi sonuçlar vereceğini ve daha az parametre kestiriminde bulunan eşitleme yönteminin daha çok parametre kestirimi yapan eşitleme yöntemine göre daha tercih edilebilir sonuç çıkacağını belirtmiştir. Wang ve diğerleri (2000), yapmış oldukları simülasyon çalışmasında, analitik yöntemleri kullanarak dairesel eşitleme ölçütünün eşitlemenin yeterliliğini değerlendirmede uygun olmayacağını ortaya koymuşlardır. Kolen ve Brennan (2004), eşitlemenin standart hatasını kullanarak kestirilen seçkisiz hatanın da ölçüt eşitleme geliştirmede kullanılabileceğini belirtmiştir. Harris ve Crouse (1993) bu değerlendirme ölçütünü standart hata ölçütü olarak adlandırmıştır. Bu yöntemin uygulaması ve yorumlaması kolaydır ancak sistematik hataları göz ardı ettiği için tek başına bir ölçüt olarak kullanılmamaktadır (Heh, 2007).

Bu çalışma simülasyon çalışması olduğu için tüm evren verilerine ulaşılabilmektedir. Bu sebeple bu çalışmada ölçüt eşitleme olarak geniş örneklem ölçütü kullanılmıştır. Geniş örneklem ölçütünde, evreni temsil eden geniş örneklem çekilir. Ardından aynı evrenden daha küçük örneklem çekilir ve sonuçları büyük örneklem sonuçları ile karşılaştırılır (Kolen ve Brennan, 2004). Küçük örneklemle ilgili yapılan birçok çalışmada geniş örneklem ölçütü kullanılmaktadır.

## **6.İLGİLİ ARAŞTIRMALAR**

### **6.1. Yurtdışında Yapılan Çalışmalar**

Livingston (1993) küçük örneklemelerde, NEAT deseni ile yapılan zincirlenmiş log-linear ön-düzgünleştirmenin, eşit yüzdelli eşitleme yönteminin

doğruluğuna etkisini incelemiştir. 93283 bireye her biri 58 maddeden ve 24 ortak maddeden oluşan iki farklı test formu uygulanmış, 25, 50, 100 ve 200 örneklem büyüklükleri üzerinde bir düzgünleştirme yapılmadan ve üç farklı derecede düzgünleştirme yapılmış olmak üzere dört farklı eşit yüzdellikli eşitleme yöntemi kullanılmıştır. 50 replikasyon yapılarak elde edilen RMSE istatistiğine bakılarak eşitlemenin doğruluğu değerlendirilmiştir. Sonuç olarak, düzgünleştirmenin, eşit yüzdellikli eşitleme yöntemi için gerekli örneklem büyüklüğünü yarı oranda azalttığını belirtmiştir.

Hanson, Zeng ve Colton (1994), seçkisiz gruplar deseni kullanılarak özdeş eşitleme, doğrusal eşitleme, düzgünleştirme yapılmamış, ön-düzgünleştirme yapılmış ve son-düzgünleştirme yapılmış eşit yüzdellikli eşitleme yöntemlerini, 100, 250, 500, 1000 ve 3000 örneklem büyüklükleri düzeyinde karşılaştırmıştır. Çalışma sonucunda, küçük örneklerde eşit yüzdellikli eşitlemede, ön düzgünleştirme ve son düzgünleştirme arasında açık bir fark olmadığı belirtilmiştir. Hanson ve diğerleri (1994), çalışmalarında, örneklem büyüklüğü 100 ve altı olduğu durumlarda özdeş eşitlemenin, doğrusal eşitleme ve eşit yüzdellikli eşitleme yöntemlerine göre daha az eşitleme hatası verdiğini belirtmişlerdir. Kolen ve Brennan (2004), Hanson ve diğerlerinin (1994) yapmış oldukları çalışmanın uzantısı olarak, özdeş eşitleme yöntemi ile diğer yöntemleri karşılaştırmıştır. Karşılaştırma sonucu, formlar arası ortalama güçlük farkı arttıkça (0.6 SMD), formları arası güçlüğü benzer olduğu (0.1 SMD) durumlara göre toplam hatanın arttığını açıklamışlardır (Skaggs, 2005). Bununla birlikte test güçlüklerinin benzer olduğu durumlarda (0.1 SMD) Hanson ve diğerlerinin (1994) yapmış olduğu çalışma sonuçlarına benzer şekilde, 100 örneklem büyüklüğünde özdeş eşitleme yönteminin diğer eşitleme yöntemlerine göre daha düşük hata verdiğini açıklamışlardır.

Parshall, Houghton ve Kromrey (1995) küçük örneklem için doğrusal eşitleme yöntemlerinde eşitleme hatası ve yanlılık üzerine çalışma yapmışlardır. Çalışmada veriler, birbirine paralel iki formdan oluşan NEAT desene göre hazırlanmış, formlar arası 0.0-0.4 SD arasında değişen güçlük farkına sahip beş testten elde edilmiştir. 15, 25, 50 ve 100 olmak üzere dört farklı örneklem üzerinde yapılan çalışmada, örneklem sayısı küçüldükçe yanlılıkta çok kayda değer değişiklik olmazken, eşitlemenin standart hatasında önemli ölçüde artma olduğu görülmüştür. Bununla birlikte, ankor test ile toplam test puanları arasında en yüksek korelasyona sahip eşitlemelerde en düşük eşitleme hatası hesaplanmıştır. Ayrıca, Efron'un (1982) çalışmasına benzer şekilde küçük örneklemde eşitlemenin standart hatasını kestirmede bootstrap yönteminin daha doğru sonuçlar verdiğini belirtmişlerdir (Kolen ve Brennan, 2004).

Skaggs (2005), küçük örneklemde seçkisiz gruplar deseni kullanılarak yapılan eşitlemelerin etkililiğini incelemiştir. Ayrıca geçme puanlarına göre eşitleme yöntemlerinin etkililiği ve doğruluğunu araştırmıştır. Araştırmada, 25, 50, 75, 100, 150 ve 200 örneklem büyüklükleri için ortalama eşitleme, doğrusal eşitleme, düzgünleştirme yapılmamış eşit yüzdelli eşitleme ve 2 ile 6 moment arasında değişen log-linear ön-düzgünleştirme yapılmış eşit yüzdelli test eşitleme yöntemlerini karşılaştırılmıştır. Gerçek veriler üzerinde yapılan çalışmada 0,9 ve daha üstü güvenilirliğe sahip, ortalama güçlük farkı 0,1 ortalama standart sapma olan çoktan seçmeli 50 maddeden oluşan iki form kullanılmıştır. Çalışmada örneklem, evrenden yeniden örnekleme yolu ile çekilmiştir. Ölçüt eşitleme olarak, büyük örneklemde sıklıkla tercih edilen eşit yüzdelli eşitleme yöntemi kullanılmıştır. Çalışma sonucunda, Parshall, Houghton ve Kromrey'in (1995) çalışmalarına paralel olarak, örneklem büyüklüğü arttıkça eşitlemenin standart hatasının azaldığı ve yanlılığın küçük miktarda değişim gösterdiğini belirtilmiştir. 25 ve daha küçük örneklem büyüklüklerinde, eşitleme

yapmanın, eşitleme yapmamaya göre daha kötü sonuçlar vereceğinden dolayı bu örneklem büyüklüğünde ve altında eşitleme yapılmamasını tavsiye etmiştir. Örneklem büyüklüğü 50 ve üzeri olduğu durumlarda geçme puanlarının bireylerin performansına göre olan durumu araştırılmıştır. Araştırma sonucunda, geçme puanı ortalamanın altında ise ortalama eşitleme yönteminin en doğru sonuçlar verdiğini, geçme puanı ortalamaya yakın ise tüm yöntemlerin en doğru eşitlemeyi veren doğrusal eşitlemeye benzer sonuçlar verdiğini, geçme puanı ortalamanın üzerinde ise 2 ve 3 momentli ön-düzgünleştirme yapılmış eşit yüzdelikli eşitleme yöntemlerinin en doğru sonucu verdiğini belirtmiştir.

Heh (2007) yapmış olduğu çalışmada, seçkisiz gruplar deseni için farklı düzeylerde ortalama güçlüğü sahip testlerde küçük örneklem test eşitleme yöntemlerinin doğruluğunu incelemiştir. Bununla birlikte, sonuçlar, performans standartlarına göre düzenlendiğinde, eşitleme hatasının sınıflama hatası üzerindeki etkisi de incelenmiştir. Çalışmada eşitlemelerin doğruluğu, RMSD (root means square deviation) ve bileşenleri olan yanlılık ve seçkisiz hata istatistiklerine göre belirlenmiştir. Çalışmada, seçkisiz gruplar deseni için küçük örneklemelerde kullanılan özdeş, ortalama, doğrusal, düzgünleştirme yapılmamış eşit yüzdelikli ve 2 ile 6 arası momentli çoklu log-linear ön-düzgünleştirme yapılmış eşit yüzdelikli eşitleme yöntemleri, örneklem büyüklükleri (25, 50, 75, 100, 150 ve 200) ve ortalama güçlük farkları (0, 0.15, 0.30, 0.45, 0.60 ve 0.75 SD) faktörlerine göre karşılaştırılmıştır. Eşitlemenin doğruluğunu kontrol etmek için 200000 bireyden oluşan evren üzerinde 6 moment log-linear ön-düzgünleştirme yapılmış eşit yüzdelikli eşitleme yöntemi kullanılarak ölçüt eşitleme istatistikleri elde edilmiştir. Sonuç olarak, 2 ve 3 momentli çoklu log-linear ön-düzgünleştirme yapılmış eşit yüzdelikli eşitleme yöntemleri birçok koşulda küçük örneklemelerde en doğru eşitleme sonucunu verdiği görülmüştür. Ayrıca, örneklem büyüklüğünün 25 ve üstü olduğu durumlarda ve test formları



arasındaki ortalama güçlük farkının en az 0,3 standart sapma olduğu durumlarda, test eşitlemenin faydalı olabileceğini belirtmiştir.

Kim, von Davier ve Haberman (2008), NEAT desene dayalı küçük örneklerde, örneklem hataları ve ilişkilendirme yanlılığını azaltmak için yeni bir yöntem önermişlerdir. Sentetik ilişkilendirme fonksiyonu ile özdeş ve zincirlenmiş doğrusal eşitleme yöntemlerini, farklı örneklem büyüklüklerinde (N=10, 25, 50, 100 ve 200) karşılaştırmışlardır. Ölçüt eşitleme yöntemi olarak zincirlenme eşitleme yöntemi kullanılmıştır. Yöntemleri karşılaştırmak için, ölçüt fonksiyon ile eşitleme yanlılığı (BIAS), eşitleme hatası (SEE) ve RMSE değerleri hesaplanmıştır. Çalışmada, sentetik fonksiyon yönteminin küçük örneklerde ve grup yetenek dağılımları farklılaştığında alternatif bir yöntem olarak kullanılabilmesini önermiştir. Sonuç olarak sentetik fonksiyon yönteminin örneklem büyüklüğü 100 ve üzeri olduğu durumlarda, özdeş eşitlemeye göre daha doğru ve daha az yanlı sonuçlar verdiğini belirtmişlerdir.

Livingston ve Kim (2009), küçük örneklerde test eşitleme yöntemleri için Dairesel yay yöntemini geliştirmişlerdir. Eşitleme ilişkisini, doğrusal olarak varsaymayan güçlü bir model olarak ele almışlardır. Eşitleme fonksiyonu, doğrusal ve eğrisel iki fonksiyonun toplamına denktir. Eşitleme için, NEAT desene dayalı zincirlenmiş eşit yüzdellikli eşitleme, Levine gözlenen puanlar, zincirlenmiş doğrusal eşitleme, Tucker, Ortalama ve Dairesel yay yöntemleri, grup yetenek düzeyleri eşit, 25 ve 75 kişilik gruplara, ortalama güçlük düzeyleri farklı 107 maddeden oluşan test formları kullanılmıştır. Ön-düzgünleştirme yapılmış eşit yüzdellikli eşitleme yöntemi ölçüt eşitleme yöntemi olarak kullanılmıştır. Yeniden örnekleme yoluyla kestirilen RMSE değerine göre eşitleme yöntemleri karşılaştırılmış ve sonuç olarak Dairesel yay yöntemi ile ortalama eşitleme yöntemlerinin daha doğru ve daha az yanlı sonuçlar verdiği görülmüştür.

Livingston ve Kim (2010), 2009'da yaptığı çalışmalara ek olarak NEAT desene dayalı eşitleme yöntemlerine, zincirlenmiş ortalama eşitleme yöntemi ve Dairesel yay yönteminin iki alt yöntemi olan simetrik dairesel yay ve basitleştirilmiş dairesel yay yöntemlerini çalışmalarına dahil etmişlerdir. Her biri en az 110 maddeden oluşan, ortalama güçlük düzeyleri 0.17-0.30 SD arasında değişen dört çift test 10,000 bireye uygulanmıştır. Ayrıca testleri alan bireylerin grup yetenekleri -0.30 ile 0.30 arasında değişimlenmiştir. Çalışmada, yeni form için 10, 25, 50 ve 100, eski form için ise yeni formdaki örneklem büyüklüğünün üç katı (30, 75, 150, 300) kullanılmıştır. Sonuç olarak çok küçük örneklerde (50 ve daha küçük) Dairesel yay yönteminin doğru eşitleme sonucu verdiği görülmüştür. Bununla birlikte sistematik ve basitleştirilmiş Dairesel yay yöntemleri arasında benzer sonuçlar elde edilmiştir.

Devdass (2011), yapmış olduğu simülasyon çalışmasında küçük örneklerde NEAT desen kullanarak, çeşitli yöntemlerinin doğruluğunu etkileyen koşulları incelemiştir. Çalışmada özdeş eşitleme, Dairesel yay, zincirlenmiş doğrusal eşitleme, düzgünleştirilmiş zincirlenmiş eşit yüzdelli eşitleme, düzgünleştirilmiş frekans kestirim, Tucker ve Levine-gözlenen puanlar yöntemleri kullanılmıştır. Devdass (2011), örneklem büyüklüğünü 50 veya altı olduğu durumlarda, gruplar arası yetenek farkının 0.1 standart sapma biriminden fazla olduğu durumlarda, formların ortalama madde güçlükleri arasındaki fark 0.25 standart sapma biriminden fazla farklılaştığında, formlarda yer alan madde sayılarının 30 veya daha az olduğu durumlarda eşitleme yapılmasının sakıncalı olabileceğini ifade etmiştir. Araştırma sonucunda eşitleme yöntemleri ortalama yanlılığa göre sıralandığında en düşük yanlılığı Levine gözlenen puanlar yöntemi, en yüksek yanlılığı özdeş eşitleme olduğunu açıklamıştır. Eşitlenen formlar benzer güçlükte olduğu durumlarda, özdeş eşitleme ve Dairesel yay eşitleme

yöntemleri en doğru ve kararlı sonuçlar verdiğini belirtmiştir. Ancak bu iki yöntem uzun testlerde daha yanlı sonuçlar ortaya çıkardığını bildirmiştir.

Babcock, Albano ve Raymond (2012), nominal ağırlıklı ortalama eşitleme yöntemini önermişlerdir. Tucker eşitleme yönteminin basitleştirilmiş versiyonu olan bu yöntem, küçük örneklem için geliştirilmiştir. NEAT desene dayalı yedi farklı eşitleme yöntemini (Nominal ağırlıklı ortalama eşitleme, Dairesel yay, özdeş, Tucker, log-linear ön-düzgünleştirme yapılmış (C=3) zincirleşmiş eşit yüzdelikli ve ortalama eşitleme) örneklem büyüklüğüne (N=20, 50, 80), testin ortalama güçlük düzeylerine (0.0 SD, 0.35 SD, 0.70 SD) ve grup yetenekleri faktörlerine göre kestirilen eşitleme hatasına ve yanlılığa göre karşılaştırılmıştır. Simülasyon verilerinin kullanıldığı bu çalışmada nominal ağırlık ortalama eşitleme yönteminin genel olarak en etkili sonuçlar verdiğini belirtmişlerdir. Dairesel yay eşitleme yöntemi de benzer bir şekilde etkili sonuçlar vermiştir. Özdeş eşitleme sadece test formlarının ortalama güçlüklerinin hemen hemen eşit olduğu durumlarda doğru eşitleme sonucu vermiştir.

## 6.2. Yurt İçinde Yapılan Çalışmalar

Ülkemizde test eşitleme ile ilgili çok fazla sayıda çalışma bulunmamaktadır. Kelecioğlu (1994), ÖSS puanlarının eşitlenmesi üzerine bir çalışma yapmıştır. Yaptığı çalışmada, Madde Tepki Kuramı'na dayalı test eşitleme yöntemleri ve eşitleme deseni olarak ankor maddeleri kullanmıştır. Çalışma sonucunda, ÖSS Türkçe alt testlerini eşitlemek için en iyi yöntemin doğrusal eşitleme, sosyal bilimler ve matematik testleri için Rasch model ve fen bilimleri testi için en iyi eşitleme yönteminin eşit yüzdelikli eşitleme yöntemi olduğunu belirtmiştir.

Bozdağ (2007), şans başarısının test eşitlemeye etkisi olup olmadığını incelemiştir. Araştırmada, eşitleme desenlerinden, tek grup deseni ve klasik eşitleme yöntemlerinden, doğrusal eşitleme ve eşit yüzdelliği eşitleme yöntemleri kullanmıştır. OKS 2003 ve 2005 Türkçe alt testlerine ait puanları bu yöntemlere göre eşitlemiştir. Araştırmanın sonucunda, şanstın arındırılmamış test puanlarının eşitlenmesinde doğrusal eşitlemenin, şanstın arındırılmış test puanlarının eşitlenmesinde ise eşit yüzdelliği eşitlemenin uygun yöntemler olduğu bulunmuştur.

Bir diğeri çalışma Öztürk (2010) tarafından yapılmış, 2008 ilkbahar ve sonbahar döneminde yapılan ALES puanları eşitlenmiştir. Yapılan çalışmada, veriler tek grup desene göre toplanmış ve ALES alt test puanları doğrusal ve eşit yüzdelliği eşitleme yöntemi kullanılarak eşitlenmiştir. Araştırma sonucunda, ALES alt testlerini eşitlemede eşit yüzdelliği eşitleme yönteminin en uygun olduğu bulunmuştur.

Kilmen (2010), yapmış olduğu araştırmada, Madde Tepki Kuramı'na dayalı "ortalama-ortalama", "ortalama-standart sapma", "Haebara" ve "Stocking-Lord" eşitleme yöntemlerinden kestirilen eşitleme hatalarının, yetenek dağılımı (benzer ve farklı yetenek dağılımı) ve örneklem büyüklüğü (500-1000 kişilik) değişkenlerine dayalı olarak karşılaştırmıştır. Araştırmada, 600 adet 1-0 şeklinde puanlanan 3 parametrelili modele uyumlu simülatif veriler üzerinde "ortak maddeli eşitlenmemiş gruplar eşitleme deseni" kullanılmıştır. Araştırma sonunda, 3 parametrelili modele uygunluk gösteren testler için; 500 ve 1000 kişilik benzer ve farklı yetenek dağılımına sahip grupların kullanıldığı durumlarda, Stocking-Lord yöntemiyle yapılan test eşitleme uygulamasının daha az hatalı eşitlemeler yaptığını saptamıştır.

Kan (2011), 2003 ve 2005 yıllarında OKS Türkçe alt testlerinin istatistiksel eşitliğini test etmiştir. 2003 ve 2005 yılında, 25 adet çoktan seçmeli maddeden oluşan ve Türkçe'yi kullanma gücünü ölçen OKS Türkçe alt testlerini, tek grup desene dayalı doğrusal eşitleme yöntemi kullanarak eşitlemiştir. Her bir işlem için eşitlemenin standart hatası kestirilerek, testlerin eşitlikleri değerlendirilmiştir. 2003 ve 2005 yıllarında yapılan Türkçe alt testlerinin yanlı olduğu ve testlerin birbirine paralel olmadığı sonucuna ulaşılmıştır.

Gök (2012), yapmış olduğu araştırmada farklı koşullara (örneklem büyüklüğü, yetenek dağılımı, test uzunluğu ve model türü) göre üretilen test formlarını Madde Tepki Kuramı'na dayalı olan kestirim yöntemlerini (ortalama-ortalama, ortalama-standart sapma ve Stocking-Lord) kullanarak eşitlemiş ve bu yöntemlerden elde edilen sonuçları karşılaştırmıştır. Bununla birlikte, araştırmada hangi koşullarda hangi yöntemin daha iyi sonuç verdiği de araştırılmıştır. Araştırmada simülatif veriler üretilerek NEAT desen kullanılmıştır. Araştırma sonucunda, en iyi eşitlemelerin 3000 kişilik örneklem, 80 maddelik testler, benzer yetenek dağılımına sahip gruplar, 2PLM ve SL yöntemi kullanılarak elde edildiği sonucuna ulaşılmıştır. Ayrıca, araştırmada ele alınan koşullar doğrultusunda, büyük örneklem ile daha uzun testler kullanıldığında ve benzer yetenek dağılımına sahip gruplarda yöntemlerin daha az hatalı ve yanlı olduğu sonucuna ulaşılmıştır.

Kelecioğlu ve Gübeş (2013), seçkisiz gruplar deseni kullanılarak yapılan doğrusal eşitleme ve eşit yüzdelli yöntemleri karşılaştırmıştır. 2009 yılında yapılan Öğrenci Başarılarının Belirlenmesi Sınavının sosyal bilimler iki alt testinden elde edilen ham puanları, doğrusal eşitleme ve üç farklı eşit yüzdelli eşitleme (düzenleştirilmemiş, ön-düzenleştirilmiş ve son-düzenleştirilmiş eşit yüzdelli eşitleme) yöntemleri kullanılarak eşitlenmiştir. Yöntemler, RMSD değeri ile bootstrap standart hatalarının ortalaması hesaplanarak karşılaştırılmıştır.

Araştırma sonucunda, en küçük BSHO ve RMSD katsayısının doğrusal eşitleme yöntemiyle elde edildiği eşdeğer puanların sahip olduğu, ancak, ham puanlar ile eşitlenmiş puanlar arasında eğrisel bir ilişki olmasına dayalı olarak puanların eşitlenmesi için göreceli olarak daha az seçkisiz hataya sahip olan son-düzgünleştirilmiş eşit yüzdelikli eşitlemenin en uygun eşitleme yöntemi olduğu kararına varmışlardır.

## 7. ARAŞTIRMANIN AMACI VE ÖNEMİ

Bu araştırmada, küçük örneklerde seçkisiz gruplar desenine dayalı yapılan test eşitleme yöntemlerinin, örneklem büyüklüğüne, formlar arasındaki ortalama güçlük düzeyleri farkına, şans başarısına ve test uzunluğuna göre karşılaştırılması amaçlanmıştır. Aynı davranışı ölçen maddelerden oluşan bir testin farklı formlarının kullanılması gerektiği durumlarda, elde edilen puanların birbirlerinin yerine kullanılması için öncelikle eşitlenmesi gerekmektedir. Test eşitlemede sıklıkla geniş örneklem için uygulamalar ve araştırmalar yapılmış ve yöntemler geniş örneklem için önerilmiştir. Örneklem büyüklüğü ne kadar büyük olursa, o derece evreni temsil etme gücü artar ve eşitleme o derece doğru gerçekleştirilir. Ancak her zaman büyük örneklemle çalışılmayabilir. Küçük örneklemle çalışmak kaçınılmaz olduğu durumlar da bulunmaktadır. Parshall (1995), testi alan birey sayısının küçük olduğu durumlarda da, test eşitleme ihtiyacının söz konusu olacağını belirtmiştir. Örneğin; öğretmen yapımı sınavlarda, üniversitelerde öğrencilere uygulanan sınavlarda veya bazı kurs merkezlerinin yaptıkları sınavlarda küçük örneklemle çalışılmak zorunda kalılabilmektedir. Örneklem sayısı küçük olduğu durumlarda, örneğin evreni temsil etmemesi ve kullanılacak yöntemlerin sayıltılarını karşılamama problemleri ortaya çıkmaktadır. Bu problemleri aşmak için Livingston (2009), eşitleme ile ilgili bazı sayıltıları ortadan kaldıran güçlü modellerin kullanılabileceğini ifade etmiştir. Her ne kadar Harris (1993), Kolen ve Brennan

(2004), bazı arařtırmacıların küçük örneklemlere uygulanan testlerde, seçkisiz hataların yüksek olma durumu nedeniyle eşitleme yapmamayı tercih ettiklerini belirtse de, bu durum küçük örnekleme sahip testlerde eşitleme yapılamayacağı anlamına gelmemektedir. Bu tür durumlarda, eşitleme yapılıp yapılmayacağına, yapılacaksa hangi şartlarda ve hangi eşitleme yönteminin kullanılması gerektiğine karar vermek gerekmektedir.

Kim ve diğerleri (2006), küçük örneklemlerde önemli olan özdeş eşitleme fonksiyonun kullanılıp kullanılmayacağına karar vermek olduğunu ifade etmiştir. Bu karar, örneklem büyüklüğü, kullanılacak eşitleme deseni, testin uzunluğu, şans başarısı, formlar arası güçlük düzeyi farkı gibi birçok değişkene bağlıdır. Test formları hemen hemen paralel olduğu zamanlarda, test formları arası güçlük düzey farkı çok düşük olacaktır. Ancak pratikte, paralel test formları hazırlamak çok güçtür. Özellikle, daha küçük sayıda bireyden madde parametreleri kestirildiği durumlarda, paralel formlar hazırlamak daha da güçleşecektir. Örneklem büyüklüğü küçük ve güçlük düzey farkı yüksek olduğu durumlarda nasıl bir işlem yapılacağı hakkında net bir bilgi yoktur. Ayrıca, küçük örneklemlerde şans başarının test eşitlemeye etkisini inceleyen herhangi bir çalışmaya rastlanamamıştır. Bu arařtırmada, farklı test uzunluklarında, farklı şans parametre değerlerinde, farklı formlar arası güçlük düzey farklarında ve çeşitli örneklem büyüklüklerinde hangi eşitleme yönteminin en küçük hatayı vereceğini belirlemek amaçlanmaktadır. Böylelikle, pratikte küçük örneklemlerde test eşitleme ile ilgili çalışacak kişilere rehberlik etmesi amaçlanmıştır.

İlgili literatür incelendiğinde, ülkemizde küçük örneklemlerle ilgili yapılan bir çalışma yapılmadığı görülmüştür. Bu çalışmaya benzer çalışmaların yurtdışında yapıldığı görülmüştür. Ancak, bu çalışma kapsamında değişimlenen faktörler ve bu faktörlerin düzeyleri çeşitlendirilerek alana özgül değer katacağı öngörülmektedir. Ayrıca, bu çalışmanın, küçük örneklemlerde test eşitleme

yöntemleriyle ilgili ülkemizde yapılan ilk araştırma olması nedeniyle alana katkı getireceği düşünülmektedir.

## 8. PROBLEM CÜMLESİ

Küçük örneklerde, seçkisiz gruplar desenine dayalı kullanılan test eşitleme yöntemlerinin (özdeş, ortalama, doğrusal, Dairesel yay,  $C=2$  ve  $C=3$  düzeyinde Log-Linear ön-düzgünleştirme yapılmış eşit yüzdelliği eşitleme yöntemleri), örneklem büyüklüğüne, test formları arasındaki ortalama güçlük düzeyleri farkına, şans başarısına ve test uzunluğuna göre eşitlemenin standart hatası (SEE), yanlılığı (BIAS) ve eşitleme hata (RMSE) kestirimleri nasıl değişmektedir?

### 8. 1. Alt Problemler

1. Çeşitli faktörlerin, küçük örneklerde seçkisiz gruplar desenine dayalı kullanılan test eşitleme yöntemlerinden (özdeş, ortalama, doğrusal, Dairesel yay, LLC2 ve LLC3 eşit yüzdelliği eşitleme) elde edilen eşitlemenin standart hatasına, yanlılığına ve eşitleme hatasına temel etkisi nasıldır?
  - a. Örneklem büyüklüğünün, eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?
  - b. Test formları arası güçlük düzey farkının, eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?
  - c. Şans başarısının, eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?



- d. Test uzunluğunun, eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?
2. Çeşitli faktörlerin, küçük örneklerde seçkisiz gruplar desenine dayalı kullanılan test eşitleme yöntemlerinden (özdeş, ortalama, doğrusal, Dairesel yay, LLC2 ve LLC3 eşit yüzdelikli eşitleme) elde edilen eşitlemenin standart hatasına, yanlılığına ve eşitleme hatasına ortak etkisi nasıldır?
- a. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının, eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?
- b. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme yanlılığına (BIAS) ortak etkisi nasıldır?
- c. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme hatasına (RMSE) ortak etkisi nasıldır?
- d. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?
- e. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme yanlılığına (BIAS) ortak etkisi nasıldır?
- f. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme hatasına (RMSE) ortak etkisi nasıldır?

## 9. SINIRLILIKLAR

1. Arařtırmada kullanılan eřitleme deseni sekisiz gruplar deseni ile sınırlıdır.
2. Verilerin Madde Tepki Kuramına dayalı üretilmesine rağmen, bu çalışma sadece Klasik Test Kuramı'na dayalı eřitleme yöntemleri ile sınırlıdır.
3. Arařtırmada kullanılan yöntemler özdeş, ortalama, doğrusal, LLC2 ve LLC3 eşit yüzdelikli ve Dairesel yay eřitleme yöntemleri ile sınırlıdır..
4. Arařtırma, simülasyon düzeneğinde yer alan faktörlerle (örneklem büyüklüğü, test formları arasındaki ortalama güçlük düzeyleri farkı, şans başarısı ve test uzunluęu) ve bu faktörlerin çeřitli düzeyleriyle sınırlıdır.

## BÖLÜM II: YÖNTEM

### 1. Araştırmanın Türü

Bu araştırmada, küçük örneklerde seçkisiz gruplar desenine dayalı yapılan test eşitleme yöntemlerinin, örneklem büyüklüğüne, formlar arasındaki ortalama güçlük düzeyleri farkına, şans başarısına ve test uzunluğuna göre karşılaştırılması amaçlanmıştır. Böylece, çeşitli faktör düzeylerinde, (örneklem büyüklüğü, formlar arasındaki ortalama güçlük düzeyleri farkı, şans başarısı ve test uzunluğu), küçük örneklerde seçkisiz gruplar desenine dayalı kullanılan test eşitleme yöntemlerinden elde edilen en düşük standart hata (SEE), yanlılık (BIAS) ve eşitleme hata (RMSE) değerleri saptanarak, bu konuda yapılan kuramsal çalışmalara katkıda bulunulması amaçlanmaktadır. Araştırma bu yönüyle temel araştırma olarak değerlendirilebilir.

### 2. Araştırma Kapsamında Değişimlenen Faktörler

Bu çalışma, küçük örneklerde, test eşitleme yöntemlerinden elde edilen standart hata, yanlılık ve eşitleme hata değerlerini çeşitli faktörlere göre inceleyen simülatif bir çalışmadır. Araştırma kapsamında yer alan faktörler belirlenirken, ilgili literatürde yer alan benzer çalışmalarda kullanılan faktörler ve bu çalışmalardan elde edilen bulgular dikkate alınmıştır. Çalışmada yer alan faktörler aşağıda belirtilmiştir;

**Örneklem büyüklüğü:** İlgili literatür araştırıldığında, örneklem büyüklüğünün eşitleme hatasını etkileyen önemli bir faktör olduğu görülmektedir.

Parshall ve diğerleri (1995), Skaggs (2005), Kim ve diğerleri (2008), Livingston ve Kim (2009, 2011), yapmış oldukları arařtırmalarda 25 ile 400 arasında deęiřen örneklem büyüklüklerinde çalışmışlardır. Bununla birlikte, Hanson ve diğerleri (1994), yaptıkları çalışmada, 1000 ve 3000 örneklem büyüklüklerini, Livingston ve Kim (2010) ise 10 örneklem büyüklüğünü arařtırmalarına dahil etmiştir.

Bu çalışma kapsamında, eşitlemenin standart hatasına, eşitleme yanlışlığına ve eşitleme hatasına örneklem büyüklüklerinin etkisini belirlemek amacıyla, örneklem büyüklükleri 10, 25, 50, 75, 100, 150 ve 200 olmak üzere yedi düzey olarak deęişimlenmiştir.

***Test formları arasındaki ortalama güçlük düzeyleri farkı:*** Bu faktör, eşitleme hatasını kestirmede önemli bir etkiye sahiptir (Heh, 2007). Kim, von Davier ve Heberman (2006), formlar ve puan dağılımları arasındaki farklılığın önemli olduđu durumlarda, özdeş eşitlemenin uygun olmayacağını belirtmişlerdir. Bununla birlikte Skaggs (2005), formlar arasındaki ortalama güçlük düzeylerinin derecesinin, seçilecek eşitleme yöntemini etkilediğini belirtmiştir. İlgili literatür incelendiğinde, test formları arasındaki ortalama güçlük düzeyleri arasındaki farklılıklar ortalama standart sapma (SMD) birimi ile ifade edilmiştir. Hanson ve diğerleri (1994), 0.08, 0.09, 0.11, 0.59 ve 0.64 SMD ve Parshall ve diğerleri (1995), ortalama güçlük farkı 0.02, 0.17, 0.26, 0.33, 0.44 SMD olan beş çift form kullanmışlardır. Kim ve diğerleri (2006), 0.08 ve 0.39 SMD olan iki çift, Livingston (1993), 0.77 SMD; Skaggs (2005), 0.10 SMD, Livingston ve Kim (2009), 0.36 SMD güçlük farkına sahip bir çift test kullanmışlardır. Bunlarla birlikte, Livingston ve Kim (2010), güçlük düzeyleri 0.17-0.30 SMD arası deęişen dört çift test, Devdass (2011) 0.0, 0.10, 0.25, .0.50 ve 0.75 SMD, Babcock (2012) 0.0, 0.35 ve 0.75 SMD ortalama güçlük düzeyine sahip test formlarını kullanmışlardır.

Yapılan çalışmalarda, eşitlenecek formlar arası ortalama güçlük düzeyinin 0.0 ile 0.77 SMD arasında değiştiği görülmektedir. Bu çalışmada, test formları arasındaki ortalama güçlük düzeyleri farkı 0,1 SMD, 0,4 SMD ve 0.7 SMD olmak üzere üç düzey olarak değişimlenmiştir.

**Test uzunluğu:** İlgili çalışmalar incelendiğinde, testin uzunluğu arttıkça, testin güvenilirliğinin arttığı görülmektedir. Düşük güvenilirliğe sahip testlerde seçkisiz hata yüksek, yüksek güvenilirliğe sahip testlerde ise, seçkisiz hata küçük olacaktır. Bu durum da hataların büyüklükleri, bu puanları eşitlemede kullanılacak eşitleme yöntemlerinin doğruluğunu doğrudan etkilemektedir. Yapılan çalışmalar incelendiğinde, eşitlenecek testin uzunluğu 50 ile 110 madde sayısı arasında değişmektedir. Skaggs (2005), 50 maddelik; Heh (2007), 60 maddelik; Livingston ve Kim (2009 ve 2010), 110 maddelik; Babcock (2012), 100 maddelik; Kurtz (2013), 75 maddelik test formlarını kullanmışlardır. Ayrıca, Devdass (2012), 30 ve 60 maddeden, Gök (2012), 30, 60 ve 80 maddeden oluşan test formlarını kullanmış ve test uzunluğunun, test eşitlemenin doğruluğuna etkisini incelemişlerdir.

Bu çalışma kapsamında, düşük test uzunluğuna sahip testlerin de eşitleme hata kestirimindeki etkisini incelemek amacıyla, test uzunluğu 25 ve 50 maddeden oluşan testler olmak üzere iki düzey olarak değişimlenmiştir.

**Şans Başarısı:** Şans başarısının eşitleme hata kestiriminde etkili olduğu düşünülmektedir. Yurtdışında yapılan çalışmalar incelendiğinde, şans başarısının test eşitlemeye etkisini inceleyen herhangi bir çalışmaya rastlanmamıştır. Yurtiçinde yapılan çalışmalar incelendiğinde ise, Bozdağ (2007), şans puanından arındırılmış ve arındırılmamış test puanlarının eşitleme yöntemlerine etkisini araştırdığı görülmüştür.

Bu çalışmada, şans başarısının eşitleme hata kestirimindeki etkisini incelemek amacıyla, c parametre değerleri 0, 0.10, 0.20 ve 0.25 olmak üzere dört düzey olarak değişimlenmiştir.

**Eşitleme Yöntemleri;** Kullanılacak eşitleme yöntemleri, kullanılan kurama, veri toplama desenlerine ve örneklem büyüklüklerine göre değişmektedir. İlgili literatür incelendiğinde, Livingston (1993), düzgünleştirme yapılmamış ve farklı düzeylerde düzgünleştirme yapılmış NEAT desenler için zincirlenmiş eşit yüzdelikli yöntemleri, Hanson ve diğerleri (1994), ön-düzgünleştirme ve son düzgünleştirme yapılmış eşit yüzdelikli yöntemleri, Parshall ve diğerleri (1995), Levine-Angoff doğrusal yöntemini, Sakggs (2005) ve Heh (2007) düzgünleştirme yapılmamış ve farklı düzeylerde düzgünleştirme yapılmış zincirlenmiş eşit yüzdelikli, özdeş, ortalama ve doğrusal eşitleme yöntemlerini, Kim ve diğerleri (2008), zincirlenmiş doğrusal, özdeş ve sentetik fonksiyon yöntemlerini, Livingston ve Kim (2009), zincirlenmiş eşit yüzdelikli ve doğrusal, ortalama, Tucker, Levine ve Dairesel yay yöntemlerini kullanmışlardır. Benzer şekilde, Livingston ve Kim (2010), düzgünleştirme yapılmış eşit yüzdelikli, doğrusal, ortalama simetrik ve basitleştirilmiş Dairesel yay yöntemlerini, bir başka çalışmada ise zincirlenmiş eşit yüzdelikli, doğrusal, ortalama, simetrik ve basitleştirilmiş Dairesel yay ve özdeş eşitleme yöntemlerini kullanmıştır.

Bu çalışmada, seçkisiz gruplar veri toplama desenine uygun yeni yöntem (Dairesel yay) diğer yöntemler (özdeş eşitleme, ortalama eşitleme, doğrusal eşitleme, LLC2 ve LLC3 eşit yüzdelikli eşitleme yöntemleri) birlikte kullanılmıştır.

Tablo 5'te araştırma kapsamında kullanılan tüm faktörler ve bu faktörlerin düzeyleri gösterilmektedir.

**Tablo 5:** Araştırmada kullanılan faktörler ve düzeyleri

Faktör	Düzyey Sayısı	Düzyey Deęerleri
Örneklem Büyüklüęü	7	10
		25
		50
		75
		100
		150
		200
Test formları arasındaki ortalama güçlük düzeyleri farkı	3	0.1 SMD
		0.4 SMD
		0.7 SMD
Test uzunluęu	2	25
		50
Şans Başarısı (c parametre düzeyi)	4	0
		0.10
		0.20
		0.25
Eşitleme Yöntemleri	6	-Özdeş Eşitleme
		-Ortalama Eşitleme
		-Doęrusal Eşitleme
		-LLC2 eşit yüzdellikli eşitleme
		- LLC3 eşit yüzdellikli eşitleme
		- Dairesel yay

### 3. Veri Üretimi

Bu çalışma kapsamında, değişimlenen faktörler ve bu faktörlerin düzeylerine göre, Madde Tepki Kuramı modellerinden üç parametrelili lojistik modele (3PLM) uygun olarak veriler üretilmiştir. Veri üretiminde, R 3.1 programı kullanılmıştır. Veri üretim aşamasında ilk olarak, seçkisiz gruplar desenine göre, her bir form için, 5000'er kişilik, 1-0 şeklinde iki kategorili olarak evren verisi üretilmiştir. Formu alan bireylerin yetenek dağılımları, ortalaması 0 ve standart sapması 1 ve o formda yer alan maddelerin ayırt edicilik parametreleri ( $a$ ), ortalaması 1.00 ve standart sapması 0.05 olan normal dağılımdan elde edilmiştir.

X formuna uygun verilerin üretiminde,  $b$  (güçlük farkı) parametre değerleri, ortalaması 0, standart sapması 1 olan uniform dağılımdan elde edilmiştir. Ayrıca,  $c$  parametre değeri dört düzey (0, 0.1, 0.2 ve 0.25), test uzunluğu iki düzey (25 ve 50) olarak değişimlenmiş ve toplam  $4 \times 2 = 8$  farklı X formu üretilmiştir.

Y formuna uygun verilerin üretiminde ise,  $b$  parametre değerleri, X formundaki  $b$  parametre değerlerine, 0.1, 0.4 ve 0.7 eklenerek elde edilmiştir. Ayrıca,  $c$  parametre değeri dört düzey (0, 0.1, 0.2, 0.25), test uzunluğu iki düzey (25 ve 50) olarak değişimlenmiştir. Sonuç olarak, toplam  $3 \times 2 \times 4 = 24$  farklı Y formu üretilmiştir. Simülatif olarak üretilen formlara ait faktör düzeyleri Tablo 6'da verilmiştir.



**Tablo 6:** Simülatif test formlarına ait faktör düzeyleri

<b>Form</b>	<b>Madde Sayısı</b>	<b>Güçlük düzey farkı</b>	<b>Şans Başarısı</b>	<b>Form</b>	<b>Madde Sayısı</b>	<b>Güçlük düzey farkı</b>	<b>Şans Başarısı</b>
<b>X</b>	25	0.0	0.00	<b>Y</b>	25	0.7	0.00
<b>X</b>	25	0.0	0.10	<b>Y</b>	25	0.7	0.10
<b>X</b>	25	0.0	0.20	<b>Y</b>	25	0.7	0.20
<b>X</b>	25	0.0	0.25	<b>Y</b>	25	0.7	0.25
<b>X</b>	50	0.0	0.00	<b>Y</b>	50	0.1	0.00
<b>X</b>	50	0.0	0.10	<b>Y</b>	50	0.1	0.10
<b>X</b>	50	0.0	0.20	<b>Y</b>	50	0.1	0.20
<b>X</b>	50	0.0	0.25	<b>Y</b>	50	0.1	0.25
<b>Y</b>	25	0.1	0.00	<b>Y</b>	50	0.4	0.00
<b>Y</b>	25	0.1	0.10	<b>Y</b>	50	0.4	0.10
<b>Y</b>	25	0.1	0.20	<b>Y</b>	50	0.4	0.20
<b>Y</b>	25	0.1	0.25	<b>Y</b>	50	0.4	0.25
<b>Y</b>	25	0.4	0.00	<b>Y</b>	50	0.7	0.00
<b>Y</b>	25	0.4	0.10	<b>Y</b>	50	0.7	0.10
<b>Y</b>	25	0.4	0.20	<b>Y</b>	50	0.7	0.20
<b>Y</b>	25	0.4	0.25	<b>Y</b>	50	0.7	0.25

#### 4. İşlem ve Verilerin Analizi

Her faktör düzeyine uygun olarak X ve Y formlarına ait toplam 32 adet evren verisi üretilmiştir. Tüm formlar için, 1-0 şeklinde iki kategorili üretilen bu verilerin satırları toplanmış ve bireylerin testten aldıkları toplam ham puanlar hesaplanmıştır. Toplam puan hesaplama işleminden sonra, formların betimsel istatistikleri hesaplanmış ve bu istatistikler Ek-1 de verilmiştir.

Çalışma kapsamında seçkisiz gruplar deseni kullanılmış, X formu eski, Y formu yeni form olarak ele alınmış ve X formu, Y formuna dönüştürülmüştür.

Verilerin eşitlenmesi, R 3.1 programında “equate” paketi kullanılarak gerçekleştirilmiştir. Eşitlemenin doğruluğunu kontrol etmek için, 10.000 bireyden oluşan evren üzerinde eşit yüzdeliği eşitleme yöntemi kullanılarak ölçüt eşitleme istatistikleri elde edilmiştir. Tüm örneklem büyüklükleri için, 100 replikasyon yapılarak örneklem çekilmiştir. Her replikasyon için X formu, Y formu ölçeğine eşitlenerek, ham puan düzeyinde eşitlenmiş puanlar elde edilmiştir. X formu, Y formu ölçeğine eşitlenirken, aynı madde uzunluğu ve aynı şans parametre değerindeki formlar birbirine eşitlenmiştir. Örneğin 25 test uzunluğuna, 0.1 şans parametre değerine sahip X formu, 25 test uzunluğuna, 0.1 şans parametre değerine ve 0.1, 0.4. ve 0.7 güçlük düzey farkına sahip üç farklı Y formu ile ayrı ayrı eşitlenmiştir. Eşitleme işleminden sonra, çalışmada kullanılan tüm örneklem büyüklükleri düzeyinde, eşitlenmiş formların, araştırmada kullanılan her bir yöntem için, eşitlemenin standart hataları (SEE), eşitleme yanlışlıkları (BIAS) ve eşitleme hataları (RMSE) hesaplanmıştır.

Çalışmada yer alan faktörlerin (örneklem büyüklüğü, formlar arasındaki güçlük düzey farkı, şans başarısı, test uzunluğu) eşitlemenin standart hatasına, yanlışlığına ve eşitleme hatasına temel etkisini incelemek için, tüm faktörlerin her yöntem için ortalama standart hataları, ortalama yanlışlıkları ve ortalama eşitleme

hataları hesaplanmış ve temel etki grafikleri oluşturulmuştur. Faktörlerin, eşitleme hatalarına temel etki tabloları Ek 2 'de verilmiştir. Bu işlemten sonra, çalışmada yer alan faktörlerin (örneklem büyüklüğü, formlar arası güçlük düzey farkı, şans başarısı, test uzunluğu) eşitlemenin standart hatasına, yanlılığına ve eşitleme hatasına ortak etkisi incelenerek ortak etki grafikleri oluşturulmuştur.

Değerlendirme ölçütü olarak kullanılan eşitlemenin standart hatası, eşitleme yanlılığı ve RMSE değerleri hesaplanmıştır. Bu değerlerin hesaplanmasına ilişkin formüller aşağıda belirtilmiştir. Yöntemleri karşılaştırmada, çalışmada yer alan yöntemlerin sahip oldukları hata değerleri ile özdeş eşitleme yönteminin sahip olduğu hata değerleri karşılaştırılmıştır.

### ***Eşitlemenin Standart Hatası (Standard Error of Equating)***

Eşitlemenin Standart Hatası, 100 replikasyon sonucu, eşitleme yöntemi ile ölçüt eşitleme arasındaki standart sapma farkını gösterir. Eşitlemenin standart hatası her ham x puanı için cebirsel olarak; 17 no'lu eşitlikte belirtildiği şekilde hesaplanır (Kolen ve Breannan, 2004).

$$\hat{e}_Y(x_i) = \frac{\sum_r \hat{e}_{Yr}(x_i)}{R} \text{ ise,}$$

$$\bar{s}d_i = SEE_i = \sqrt{\frac{\sum_r [\hat{e}_{Yr}(x_i) - \hat{e}_Y(x_i)]^2}{R-1}} \quad (17)$$

Eşitlikte,  $\bar{s}d_i$ , eşitlemenin standart hatasını,  $i$ , puanı,  $R$ , replikasyon sayısını,  $\hat{e}_{Yr}(x_i)$  r örneklemindeki hücrelerin birinden hesaplanmış ham puan eşdeğerini göstermektedir.

### ***Eşitleme Yanlılığı (Equating Bias)***

Eşitleme yanlılığı her puan için, 100 replikasyonla elde edilen eşitleme yöntemi ile ölçüt eşitleme arasındaki ortalama fark olarak tanımlanmıştır.

Kim ve diğerleri (2006) cebirsel olarak yanlılığı, 18 no'lu eşitlikte belirtildiği şekilde hesaplandığını ifade etmiştir.

$$\hat{d}_i = \frac{\sum_{r=1}^R [\hat{e}_{yr}(x_i) - e_y(x_i)]}{R} \quad (18)$$

Eşitlikte,  $i$ , puan sayısını,  $R$ , replikasyon sayısını,  $\hat{e}_{yr}(x_i)$   $r$  örneklemindeki hücrelerin birinden hesaplanmış ham puan eşdeğerini,  $e_y(x_i)$  ölçüt eşitleme fonksiyonundan elde edilen ham puan eşdeğerini göstermektedir.

### ***Eşitleme Hatası (Root Mean Square Error)***

100 replikasyon sonucu oluşan toplam eşitleme hata varyanslarının kareköküdür. RMSE, eşitleme yanlılığı ve eşitlemenin standart hatasının kareleri toplamının kareköküne eşittir (Skaggs, 2005, Kim ve diğerleri, 2006). Eşitleme hatası, 19 no'lu eşitlik ile hesaplanmıştır.

$$RMSE_i = \sqrt{\hat{d}_i^2 + \bar{s}d_i^2} \quad (19)$$

## BÖLÜM III: BULGULAR

Bu bölümde, çeşitli faktörlerin (örneklem büyüklüğü, test formları arasındaki ortalama güçlük düzeyleri farkı, şans başarısı ve test uzunluğu), test eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve hata kestirimlerine (RMSE), temel etkilerine ve ortak etkilerine ait bulgulara yer verilmiştir. Temel etkilerin gösterilmesinde tek boyutlu grafiklerden, ortak etkilerin gösterilmesinde ise çok boyutlu grafiklerden yararlanılmıştır.

**1. Küçük örneklerde, çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen, eşitlemenin standart hatasına (SEE), yanlılığına (BIAS) ve hata kestirimlerine (RMSE) temel etkileri nasıl değişmektedir?**

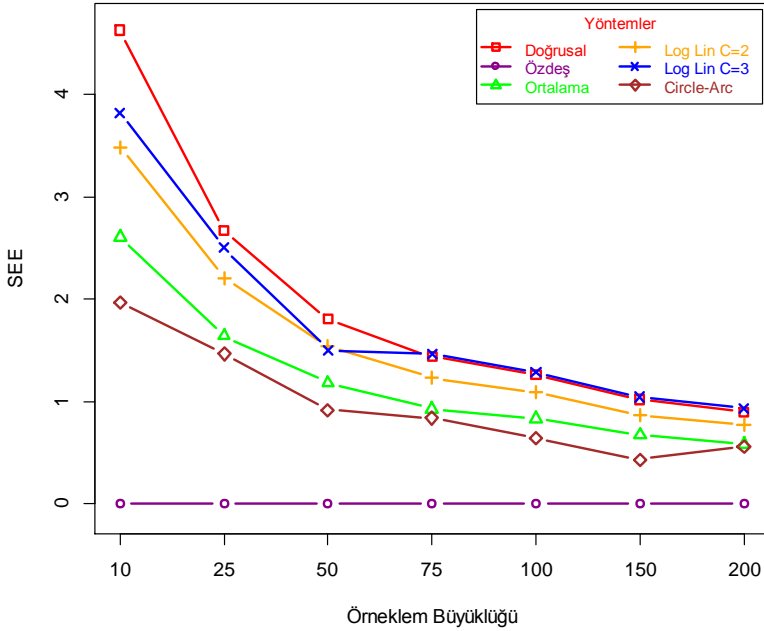
**1.1.1. Küçük örneklerde, örneklem büyüklüğünün, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama standart hataya (SEE) ilişkin değerler Ek 2 -A' da verilmiştir. Bu değerler grafiksel olarak Grafik 1'de gösterilmektedir.

Grafik 1 incelendiğinde, özdeş eşitleme yönteminden elde edilen ortalama standart hata değerlerinde herhangi bir değişme olmadığı ve tüm örneklem büyüklüklerinde ortalama standart hatanın sıfır olduğu görülmektedir. Örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama standart hatalar karşılaştırıldığında, en düşük ortalama standart hatanın Dairesel

yay eşitleme yöntemiyle elde edildiği görülmektedir. En yüksek ortalama standart hata değerinin, 10-75 örneklem büyüklükleri arasında doğrusal eşitleme yöntemiyle, 75-200 örneklem büyüklükleri arasında ise LLC3 eşit yüzdelliği eşitleme yöntemiyle elde edildiği görülmektedir. Özdeş eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinin tümünden elde edilen ortalama standart hata değerlerinin, örneklem büyüklükleri arttıkça azaldığı görülmektedir. 10-50 örneklem büyüklükleri arasında ortalama standart hata değerlerinde hızlı bir azalma görülürken, özellikle 150-200 örneklem büyüklükleri arasındaki değişimin çok az olduğu görülmektedir.

### Örneklem Büyüklüğünün Eşitlemenin Standart Hatasına(SEE) Etkisi

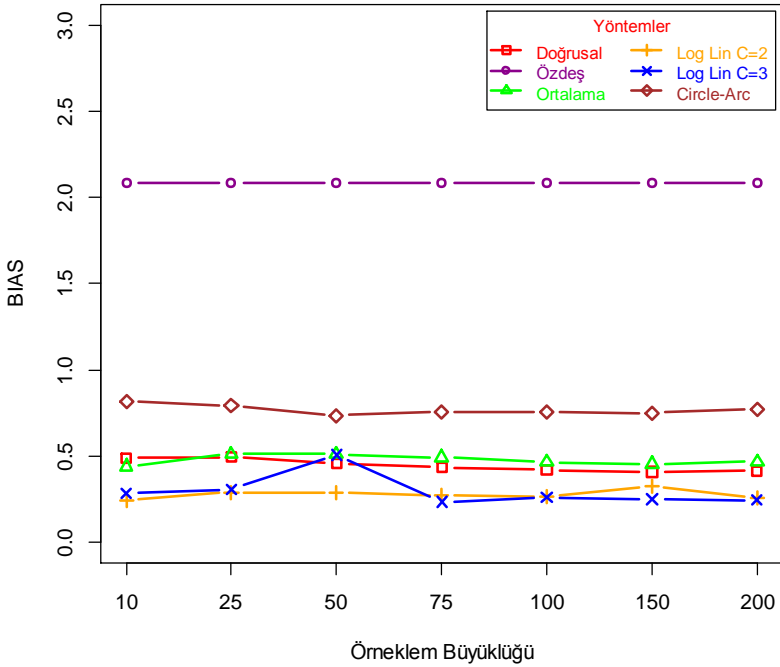


**Grafik 1:** Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

#### 1.1.2. Küçük örneklerde, örneklem büyüklüğünün, eşitlemenin yanlışlığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi

Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama eşitleme yanlışlığına (BIAS) ilişkin değerler Ek 2-B de verilmiştir. Bu değerler grafiksel olarak Grafik 2'de gösterilmektedir.

### Örneklem Büyüklüğünün Eşitlemenin Yanlılığına (BIAS) Etkisi



**Grafik 2:** Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

Grafik 2 genel olarak incelendiğinde, özdeş eşitleme yönteminden elde edilen ortalama yanlılık değerlerinde herhangi bir değişme olmadığı görülmektedir. Tüm örneklem büyüklüğü düzeylerinde en yüksek ortalama yanlılık değerlerinin özdeş eşitleme yöntemiyle elde edildiği görülmektedir. En düşük ortalama yanlılık değerlerinin, 10-75 örneklem büyüklükleri arasında LLC2 eşit yüzdelikli eşitleme yöntemiyle, 75-200 örneklem büyüklüğü arasında LLC3 eşit yüzdelikli eşitleme yöntemiyle elde edildiği görülmektedir. Özdeş

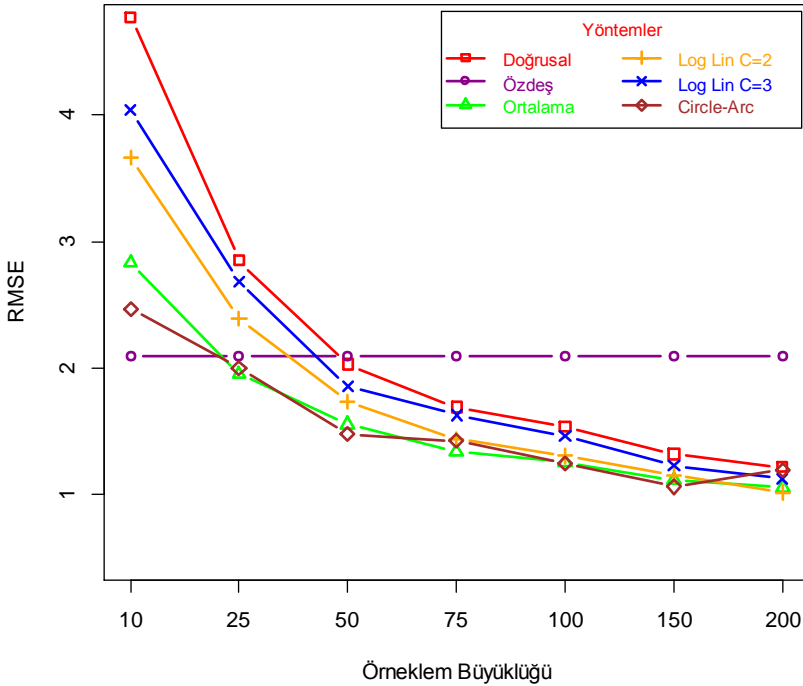


eşitlemeden sonra en yüksek ortalama yanlılık değerlerinin Dairesel yay yöntemiyle elde edildiği ve bu değerlerin örneklem büyüklüğü artıkça çok küçük miktarda azaldığı görülmektedir. LLC2 eşit yüzdeliği eşitleme yöntemiyle 50 örneklem büyüklüğünde elde edilen yanlılık değerinin dışında, genel olarak ortalama yanlılık değerlerinin örneklem büyüklüklerine göre önemli ölçüde değişmediği görülmektedir.

### **1.1.3. Küçük örneklerde, örneklem büyüklüğünün, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama eşitleme hatasına (RMSE) ilişkin değerler Ek 2-C de verilmiştir. Bu değerler grafiksel olarak Grafik 3'te gösterilmektedir.

### Örneklem Büyüklüğünün Eşitleme Hatasına(RMSE) Etkisi



**Grafik 3:** Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

Grafik 3 incelendiğinde, örneklem büyüklüğü arttıkça, özdeş eşitleme dışındaki diğer eşitleme yöntemleriyle elde edilen ortalama eşitleme hatasının (RMSE) azaldığı görülmektedir. 25 örneklem büyüklüğünde, Dairesel yay ve ortalama eşitleme yöntemlerinin eşitleme hataları (RMSE), özdeş eşitleme yönteminin eşitleme hatasına göre daha küçük hata verdiği görülmektedir. Örneklem büyüklüğünün 50 olduğu durumda, ortalama eşitleme, Dairesel yay ve LLC2 eşit yüzdelli eşitleme yöntemleri, özdeş eşitleme yöntemine göre daha az

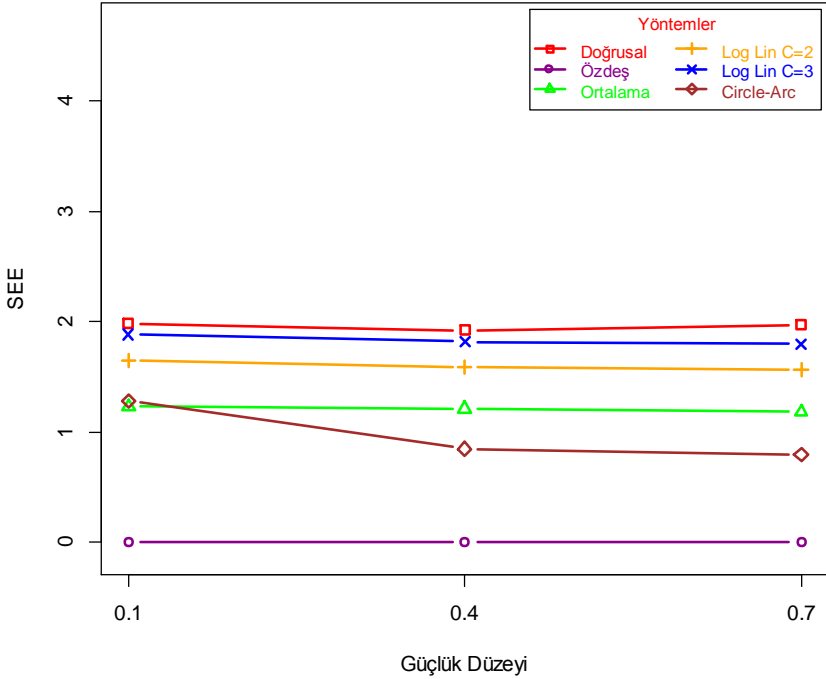
RMSE değerine sahiptir. 75 ve daha büyük örneklem büyüklüklerinde, doğrusal eşitleme ve LLC3 eşit yüzdelli eşitleme yöntemi özdeş eşitleme yöntemine göre daha az RMSE değerine sahip olduğu görülmektedir. 100 ve 200 örneklem büyüklüklerinde en düşük RMSE değeri ortalama eşitleme yöntemiyle elde edilirken, diğer örneklem büyüklüklerinde, en düşük RMSE değerinin Dairesel yay yöntemiyle elde edildiği görülmektedir.

### **1.2.1. Küçük örneklerde, test formları arasındaki güçlük düzey farkının, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli formlar arası güçlük düzey farkına göre eşitleme yöntemlerinden elde edilen ortalama standart hataya (SEE) ilişkin değerler Ek 3-A'da verilmiştir. Bu değerler grafiksel olarak Grafik 4'te gösterilmektedir.

Grafik 4 incelendiğinde, formlar arasındaki güçlük düzeyleri farklılaştığında, Dairesel yay yöntemi dışındaki diğer yöntemlerin, ortalama standart hatalarının önemli ölçüde değişmediği görülmektedir. Dairesel yay yönteminde ise, formlar arasındaki güçlük düzeyi arttıkça, eşitlemenin ortalama standart hatasının azaldığı görülmektedir. Özdeş eşitleme yönteminde eşitlemenin ortalama standart hatası sıfırdır. Diğer yöntemler formlar arasındaki güçlük düzeyi farkına göre karşılaştırıldığında, tüm koşul düzeylerinde, en düşük ortalama standart hatanın Dairesel yay yöntemiyle elde edildiği görülmektedir. Bu yöntemi sırasıyla, ortalama, LLC2, LLC3 eşit yüzdelli ve doğrusal eşitleme yöntemleri izlemektedir.

### Güçlük Düzeyinin Eşitlemenin Standart Hatasına (SEE) Etkisi

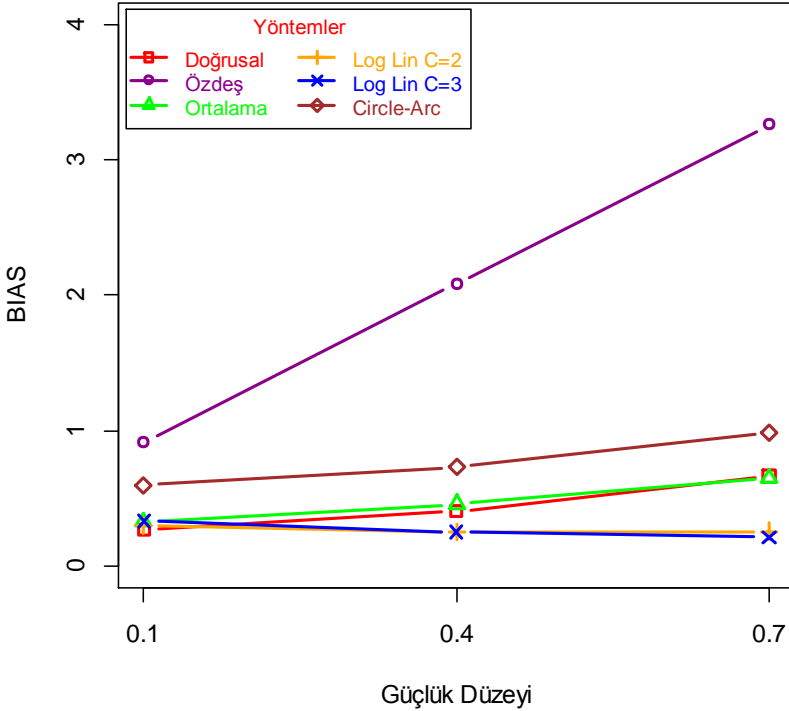


**Grafik 4:** Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

#### 1.2.2. Küçük örneklerde, test formları arasındaki ortalama güçlük düzeyleri farkının, eşitlemenin yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi

Çeşitli formlar arası güçlük düzey farkına göre eşitleme yöntemlerinden elde edilen ortalama eşitleme yanlılığına (BIAS) ilişkin değerler Ek 3-B de verilmiştir. Bu değerler grafiksel olarak Grafik 5'de gösterilmektedir.

### Güçlük Düzeyinin Eşitlemenin Yanlılığına (BIAS) Etkisi



**Grafik 5:** Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

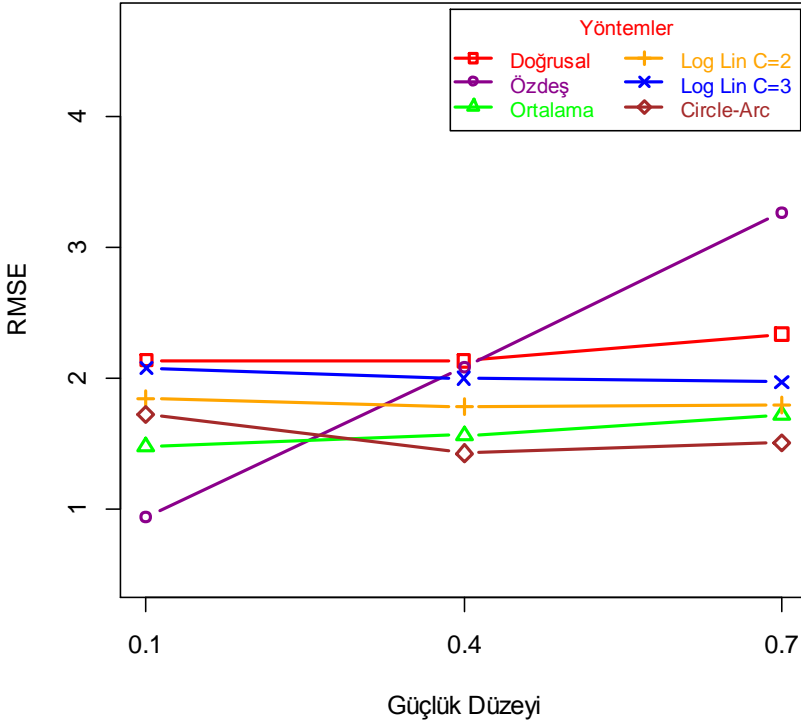
Grafik 5 test formları arasındaki güçlük düzeyleri arasındaki fark açısından incelendiğinde, yöntemlerin eşitleme yanlılığına etkilerinin farklı olduğu görülmektedir. Tüm güçlük düzeylerinde, en yüksek ortalama yanlılık değerlerinin özdeş eşitleme daha sonra ise Dairesel yay yöntemiyle elde edildiği görülmektedir. Özdeş eşitlemede, güçlük düzeyleri arasındaki fark arttığında

ortalama yanlılık değerlerinin de arttığı görülmektedir. Dairesel yay, ortalama ve doğrusal eşitleme yöntemlerinde, güçlük düzeyleri arasındaki fark arttıkça yanlılık değerlerindeki değişimin çok az olduğu görülmektedir. LLC2 ve LLC3 eşit yüzdeli eşitleme yöntemlerinde, güçlük düzeyleri arasındaki fark arttıkça ortalama yanlılığın azaldığı görülmektedir.

### **1.2.3. Küçük örneklerde, test formları arasındaki ortalama güçlük düzeyleri farkının, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli formlar arası güçlük düzey farkına göre eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) ilişkin değerler Ek 3-C’de verilmiştir. Bu değerler grafiksel olarak Grafik 6’da gösterilmektedir.

## Güçlük Düzeyinin Eşitleme Hatasına(RMSE) Etkisi

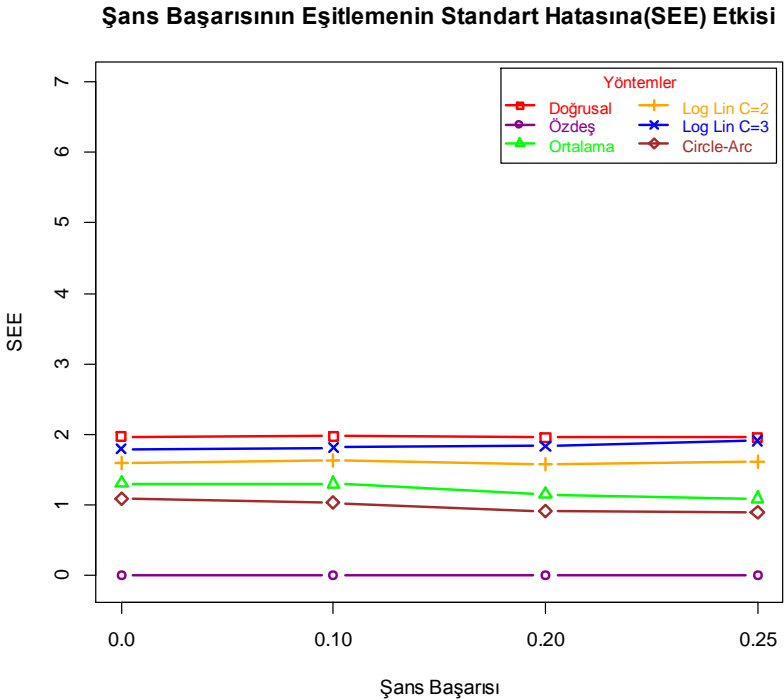


**Grafik 6:** Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

Grafik 6 incelendiğinde, güçlük düzeyleri arasındaki farkın 0.1 olduğu durumda, en düşük ortalama eşitleme hatasının özdeş eşitleme yöntemiyle; 0.4 ve 0.7 olduğu durumda ise, Dairesel yay yöntemiyle elde edildiği görülmektedir. Formlar arasındaki güçlük düzeyleri arasındaki fark arttıkça, Dairesel yay yöntemi hariç diğer yöntemlerin ortalama eşitleme hatalarında çok küçük değişimler olduğu görülmektedir.

### 1.3.1. Küçük örneklerde, şans başarısının, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi

Çeşitli şans başarısı düzeylerine göre eşitleme yöntemlerinden elde edilen ortalama standart hataya (SEE) ilişkin değerler Ek 4-A da verilmiştir. Bu değerler grafiksel olarak Grafik 7'de gösterilmektedir.



**Grafik 7:** Çeşitli şans parametre değerlerine göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri



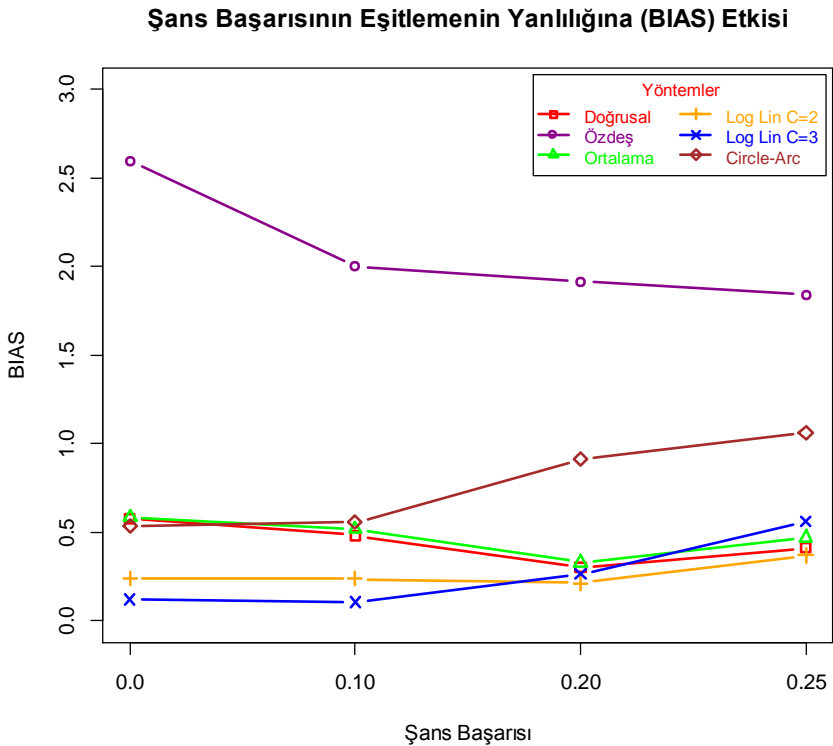
Grafik 7 incelendiğinde, şans parametresinin eşitlemenin ortalama standart hatasına doğrudan bir etkisinin olmadığı görülmektedir. Tüm şans parametre düzeylerinde, Dairesel yay ve ortalama eşitleme yöntemlerinin en düşük ortalama standart hataya sahip oldukları görülmektedir. Bu yöntemleri sırasıyla, LLC2 ve LLC3 eşit yüzdelliği eşitleme ve doğrusal eşitleme yöntemleri izlemektedir. Özdeş eşitleme yönteminin ortalama standart hatasının tüm şans parametre düzeylerinde sıfır olduğu görülmektedir.

### **1.3.2. Küçük örneklerde, şans başarısının, eşitlemenin yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli şans başarısı düzeylerine göre eşitleme yöntemlerinden elde edilen ortalama eşitleme yanlılığına (BIAS) ilişkin değerler Ek 4-B de verilmiştir. Bu değerler grafiksel olarak Grafik 8'de gösterilmektedir.

Grafik 8 incelendiğinde, şans parametre değeri arttıkça, eşitleme yanlılığındaki değişimler yöntemlere göre farklılık göstermektedir. Tüm şans parametre değerlerinde en yüksek yanlılığın özdeş eşitleme yöntemiyle elde edildiği görülmektedir. Özdeş eşitleme yönteminde, şans parametre değeri arttıkça yanlılık değerlerinin azaldığı görülmektedir. Şans parametre değeri arttıkça Dairesel yay yönteminin yanlılık değerleri diğer yöntemlere göre daha fazla değişim göstermektedir. Dairesel yay yönteminde, şans parametre değeri arttıkça eşitleme yanlılığı da artmaktadır. Doğrusal ve ortalama eşitleme yöntemlerinin ortalama yanlılık değerleri birbirine çok yakındır ve şans parametre değeri 0.0-0.20 aralığındayken, bu yöntemlerinin yanlılığının azaldığı, 0.2-0.25 aralığında, ortalama eşitleme yönteminin yanlılığının küçük miktarda arttığı görülmektedir. Sonuç olarak, şans başarısı parametresi 0.0 ve 0.10 olduğu durumlarda LLC3 eşit yüzdelliği eşitleme ve doğrusal eşitleme yönteminin en düşük ortalama yanlılığa, 0.20 ve 0.25 olduğu durumlarda LLC3 eşit yüzdelliği

eşitleme ve doğrusal eşitleme yöntemlerinin en düşük ortalama yanlılığa sahip olduğu görülmektedir.



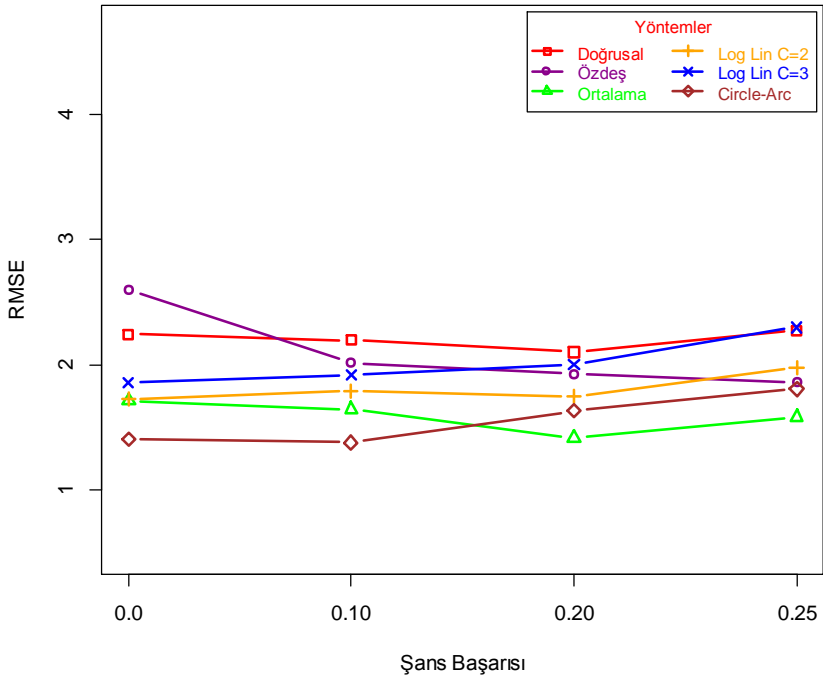
**Grafik 8:** Çeşitli şans parametre değerlerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

### **1.3.3. Küçük örneklerde, şans başarısının, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli şans başarısı düzeylerine göre eşitleme yöntemlerinden elde edilen ortalama eşitleme hatasına (RMSE) ilişkin değerler Ek 4-C de verilmiştir. Bu değerler grafiksel olarak Grafik 9'da gösterilmektedir.

Grafik 9 incelendiğinde, farklı şans parametre değerlerinde RMSE değerleri yöntemlere göre farklılık göstermektedir. Şans parametresi 0-0.1 aralığında özdeş eşitleme dışındaki diğer yöntemlerin RMSE değerlerinin önemli ölçüde değişmediği, özdeş eşitleme yönteminin RMSE değerinin küçük miktarda arttığı görülmektedir. Şans parametre değeri arttıkça, Dairesel yay yönteminin RMSE değeri artış göstermektedir. 0.20-0.25 parametreleri aralığında özdeş eşitleme yöntemi hariç diğer yöntemlerin RMSE değerlerinde artış olduğu görülmektedir. Şans parametre değerleri 0 ve 0.10 olduğu durumlarda Dairesel yay yönteminin, 0.20 ve 0.25 olduğu durumlarda ise ortalama eşitleme yönteminin en küçük RMSE değerine sahip olduğu görülmektedir. Doğrusal eşitleme yönteminin tüm şans parametre değerlerinde en yüksek RMSE değerine sahip olduğu görülmektedir.

### Şans Başarısının Eşitleme Hatasına(RMSE) Etkisi

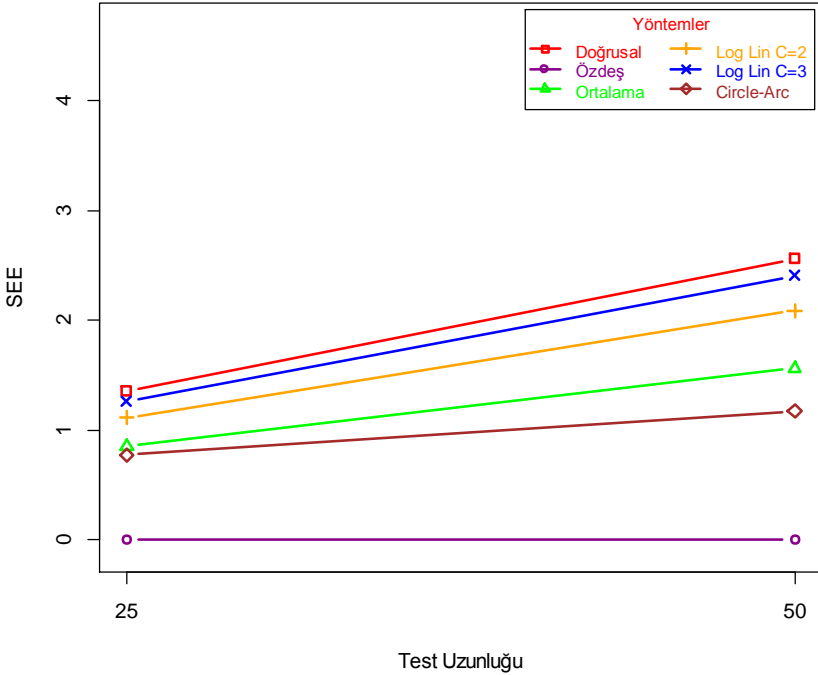


**Grafik 9:** Çeşitli şans parametre değerlerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

### 1.4.1. Küçük örneklerde, test uzunluğunun, eşitlemenin standart hatasına (SEE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi

Çeşitli test uzunluğu düzeylerine göre eşitleme yöntemlerinden elde edilen ortalama standart hataya (SEE) ilişkin değerler Ek 5-A da verilmiştir. Bu değerler grafiksel olarak Grafik 10'da gösterilmektedir.

**Test Uzunluğunun Eşitlemenin Standart Hatasına (SEE) Etkisi**



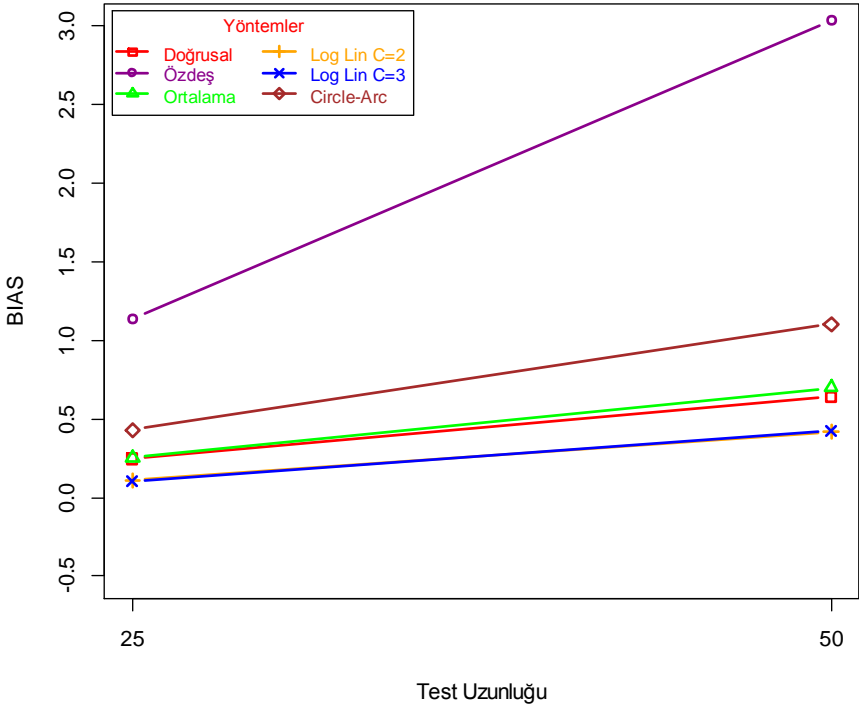
**Grafik 10:** Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

Grafik 10 incelendiğinde, eşitlemenin ortalama standart hatasının, test uzunluğu arttıkça tüm yöntemlerde artış gösterdiği gözlenmektedir. Özdeş eşitleme yönteminin ortalama standart hatası sıfırdır. Tüm test uzunluğu düzeylerinde, en düşük ortalama standart hata değerlerinin Dairesel yay yöntemiyle, en yüksek ortalama standart hata değerlerinin ise, doğrusal eşitleme yöntemiyle elde edildiği görülmektedir.

#### **1.4.2. Küçük örneklerde, test uzunluğunun, eşitleme yanlılığına (BIAS) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi**

Çeşitli test uzunluğu düzeylerine göre eşitleme yöntemlerinden elde edilen ortalama eşitleme yanlılığına (BIAS) ilişkin değerler Ek 5-B de verilmiştir. Bu değerler grafiksel olarak Grafik 11’de gösterilmektedir.

### Test Uzunluęunun Eşitlemenin Yanlılıęına (BIAS) Etkisi



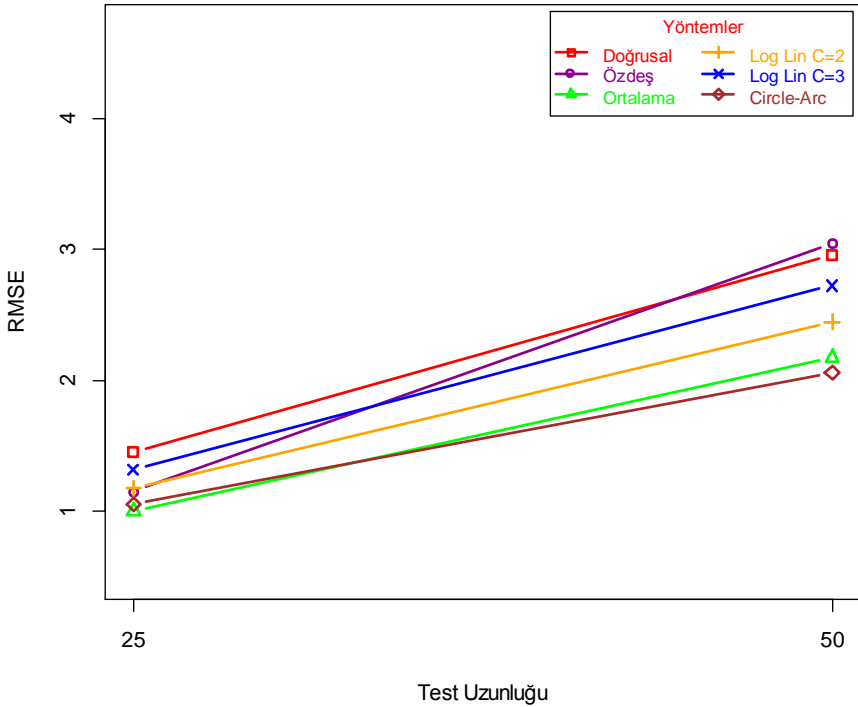
**Grafik 11:** Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

Grafik 11 incelendiğinde, test uzunluęu arttıkça, tüm eşitleme yöntemlerinden elde edilen yanlılık değerlerinin arttığı görülmektedir. En yüksek ortalama yanlılık değerlerinin özdeş eşitleme yöntemiyle, en düşük ortalama yanlılık değerinin ise, LLC3 eşit yüzdeliđli eşitleme yöntemiyle elde edildięi görülmektedir. Bu yöntemleri sırasıyla, LLC2, doğrusal eşitleme, ortalama eşitleme, Dairesel yay ve özdeş eşitleme yöntemleri takip ettięi görülmektedir.

### 1.4.3. Küçük örneklerde, test uzunluğunun, eşitleme hatasına (RMSE) temel etkisinin çeşitli eşitleme yöntemlerine göre incelenmesi

Çeşitli test uzunluğu düzeylerine göre eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) ilişkin değerler Ek 5-C de verilmiştir. Bu değerler grafiksel olarak Grafik 12'de gösterilmektedir.

**Test Uzunluğunun Eşitleme Hatasına (RMSE) Etkisi**



**Grafik 12:** Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri



Grafik 12 incelendiğinde, test uzunluğu arttıkça eşitleme hatasının da arttığı görülmektedir. Test uzunluğunun 25 olduğu durumda ortalama eşitleme yöntemiyle, 50 olduğu durumda Dairesel yay eşitleme yöntemiyle en düşük eşitleme hatasının elde edildiği görülmektedir. 25 test uzunluk düzeyinde en yüksek eşitleme hata değerinin doğrusal eşitleme yöntemiyle 50 test uzunluk düzeyinde ise özdeş eşitleme yöntemiyle elde edildiği görülmektedir.

**2. Küçük örneklerde seçkisiz gruplar desenine dayalı kullanılan test eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına, yanlılığa ve hatasına; örneklem büyüklüğünün, test formları arasındaki ortalama güçlük düzeyleri farkının, şans başarısının ve test uzunluğunun ortak etkisi nasıldır?**

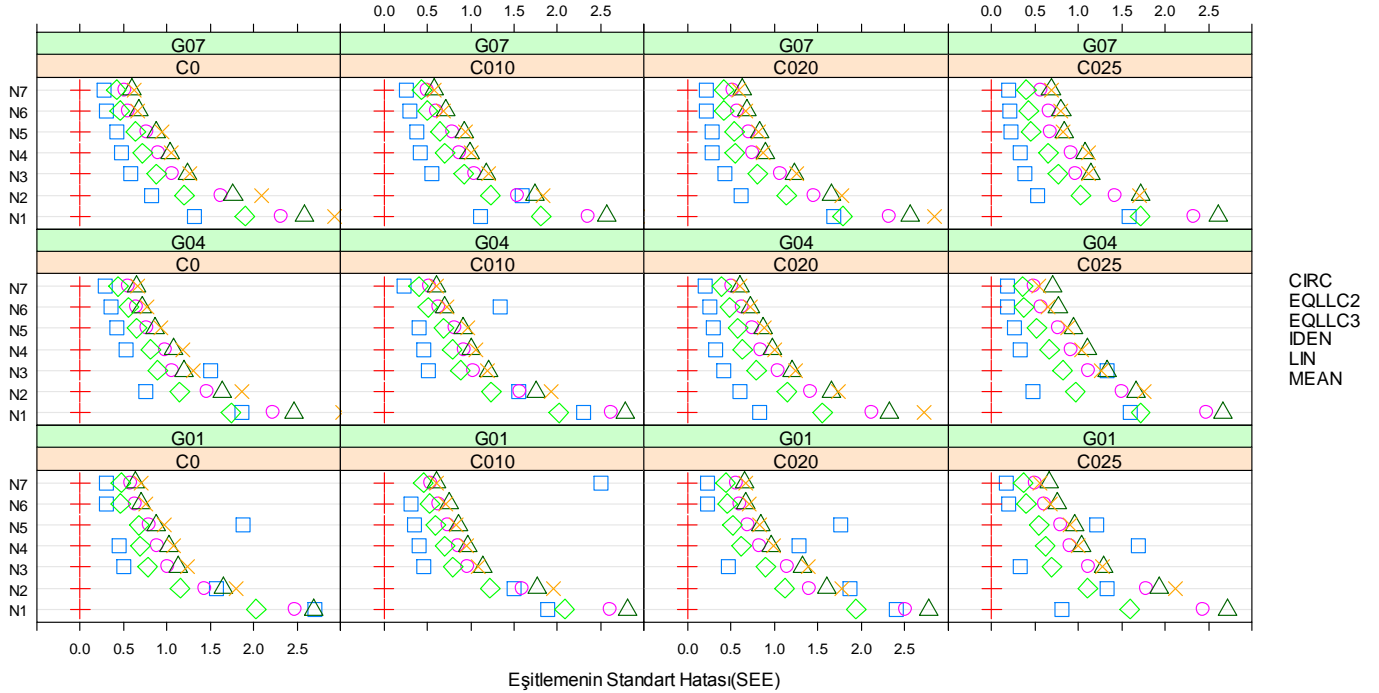
**2.1.1. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının eşitlemenin standart hatasına (SEE) ortak etkisi**

Grafik 13 incelendiğinde, özdeş eşitleme dışındaki, diğer tüm yöntemlerin ortalama standart hatalarının, formlar arasındaki güçlük düzeyleri farkının ve şans başarısının tüm düzeylerinde, örneklem büyüklüğünün değişiminden etkilendiği görülmektedir. Örneklem büyüklüğü arttıkça, özdeş eşitleme yönteminin standart hatası sıfır değerinde sabit kalırken, diğer yöntemlerin ortalama standart hatalarının azaldığı görülmektedir.

Formlar arasındaki güçlük düzeyindeki değişimlerin, Dairesel yay yöntemi dışındaki yöntemlerin ortalama standart hatalarını önemli ölçüde etkilemediği görülürken, Dairesel yay yönteminde, güçlük düzey farkının 0.1 olduğu durumlarda ortalama standart hataların düzensiz değişimler gösterdiği görülmektedir. Şans parametre değerinin değişimi, eşitleme yöntemlerinin

standart hatalarını önemli ölçüde deęiřtirmedięi görölmektedir. Ancak, güçlük düzey farkı 0.1 olduęu durumda, řans parametreleri arttıkça, Dairesel yay yönteminin standart hatasında düzensiz deęiřimler görölmektedir.

### Ortak Etki M.S=25



**Grafik 13:** Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitlemenin standart hatasına (SEE) ortak etkisi (Test uzunluğu-25)

Diğer güçlük düzeylerinde, şans parametresinin değişimi, ortalama standart hatayı önemli ölçüde değiştirmedeği görülmektedir.

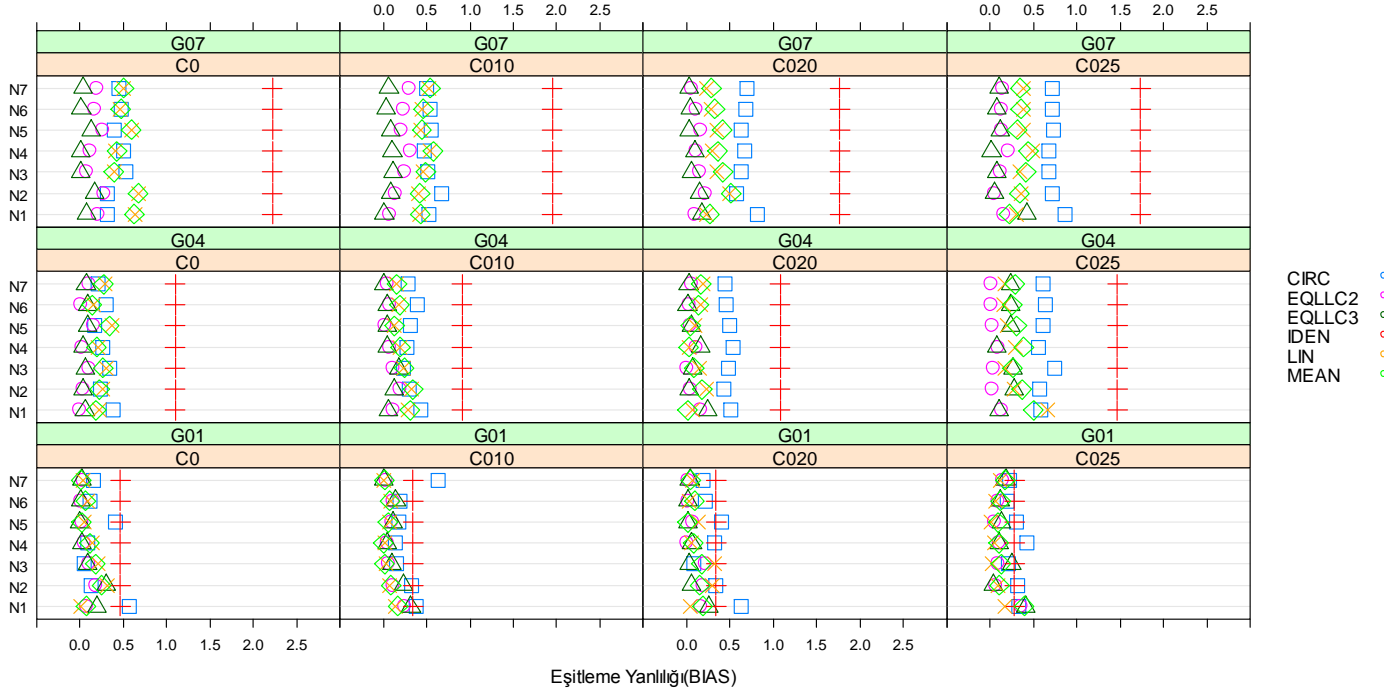
### **2.1.2. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme yanlılığına (BIAS) ortak etkisi**

Grafik 14 incelendiğinde, örneklem büyüklüğünün, test formları arasındaki güçlük düzey farkının ve şans başarısının, yöntemlerin ortalama yanlılık değerlerini etkilemediği görülmektedir.

Formlar arasındaki güçlük düzey farkının artmasıyla, yöntemlerin eşitleme yanlılığında küçük miktarda artış olduğu görülmektedir. Güçlük düzeyleri farkının artmasından en çok etkilenen yöntem özdeş eşitleme yöntemidir. Güçlük düzey farkının arttığı durumlarda, özdeş eşitlemenin ortalama yanlılığında büyük miktarda artış görülmektedir. Güçlük düzey farkı arttıkça ortalama eşitleme, doğrusal eşitleme, özdeş eşitleme ve Dairesel yay eşitleme yöntemlerinin ortalama yanlılık değerlerinin arttığı, LLC2 ve LLC3 eşit yüzdelikli eşitleme yöntemlerinin ortalama yanlılık değerlerinde önemli değişim olmadığı görülmektedir. Şans parametre değerlerinin arttığı durumlarda, güçlük düzey farkı arttıkça eşitleme yanlılıklarının daha çok etkilendiği görülmektedir.

Dairesel yay eşitleme yöntemi dışındaki eşitleme yöntemlerinin ortalama yanlılıkları şans parametre değerinin değişiminden çok etkilenmediği görülmektedir. Dairesel yay yönteminde şans parametresi arttıkça yanlılık değerlerinde küçük bir artış olduğu görülmektedir.

Ortak Etki M.S=25



**Grafik 14:** Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme yanlılığına (BIAS) ortak etkisi (Test uzunluğu-25)

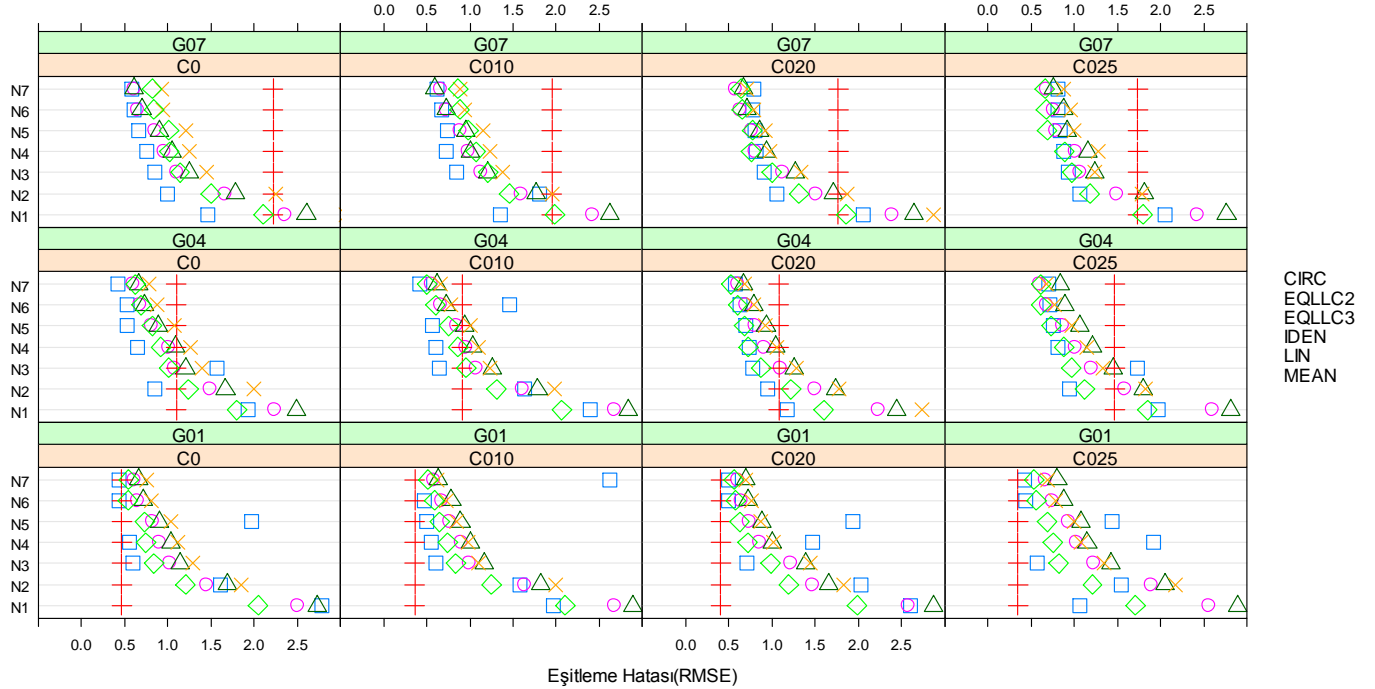
### **2.1.3. Test uzunluğu 25 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme hatasına (RMSE) ortak etkisi**

Grafik 15 incelendiğinde, örneklem büyüklükleri arttıkça güçlük düzey farkının ve şans başarısının tüm düzeylerinde, özdeş eşitleme dışındaki diğer eşitleme yöntemlerinin RMSE değerlerinin azaldığı, özdeş eşitleme yönteminin RMSE değerlerinin ise sabit kaldığı görülmektedir

Formların güçlük düzeyleri farkının arttığı durumlarda, özdeş eşitleme yönteminin RMSE değerinde büyük bir artış olduğu görülürken, diğer yöntemlerde çok küçük düzeyde artış olduğu görülmektedir. Formların güçlük düzey farkı küçük olduğu durumlarda (0.1 SMD) özdeş eşitleme yönteminin en düşük hataya sahip olduğu görülmektedir. Formlar arası güçlük düzey farkının 0.4 olduğu ve şans parametre değerleri 0.00 ve 0.10 olduğu durumlarda 100 örneklem büyüklüğü ve üstünde, şans parametre değerleri 0.20 ve 0.25 olduğu durumlarda 75 örneklem büyüklüğü ve üstünde özdeş eşitleme yöntemiyle, diğer yöntemlere göre daha yüksek hata elde edildiği görülmektedir. Güçlük düzey farkının çok olduğu durumlarda (0.7 SMD), 50 ve üstü örneklem büyüklüklerinde özdeş eşitleme yöntemiyle diğer yöntemlere göre daha büyük hata elde edildiği görülmektedir.

Formların güçlük düzey farkı 0.1 SMD olduğu durumda, şans başarısı arttıkça Dairesel yay yönteminin RMSE değerlerinde düzensiz değişimler görülmektedir. Şans parametresi değiştikçe diğer yöntemlerin RMSE değerlerinde önemli değişimlerin olmadığı görülmektedir.

### Ortak Etki M.S=25



**Grafik 15:** Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme hatasına (RMSE) ortak etkisi (Test uzunluğu-25)

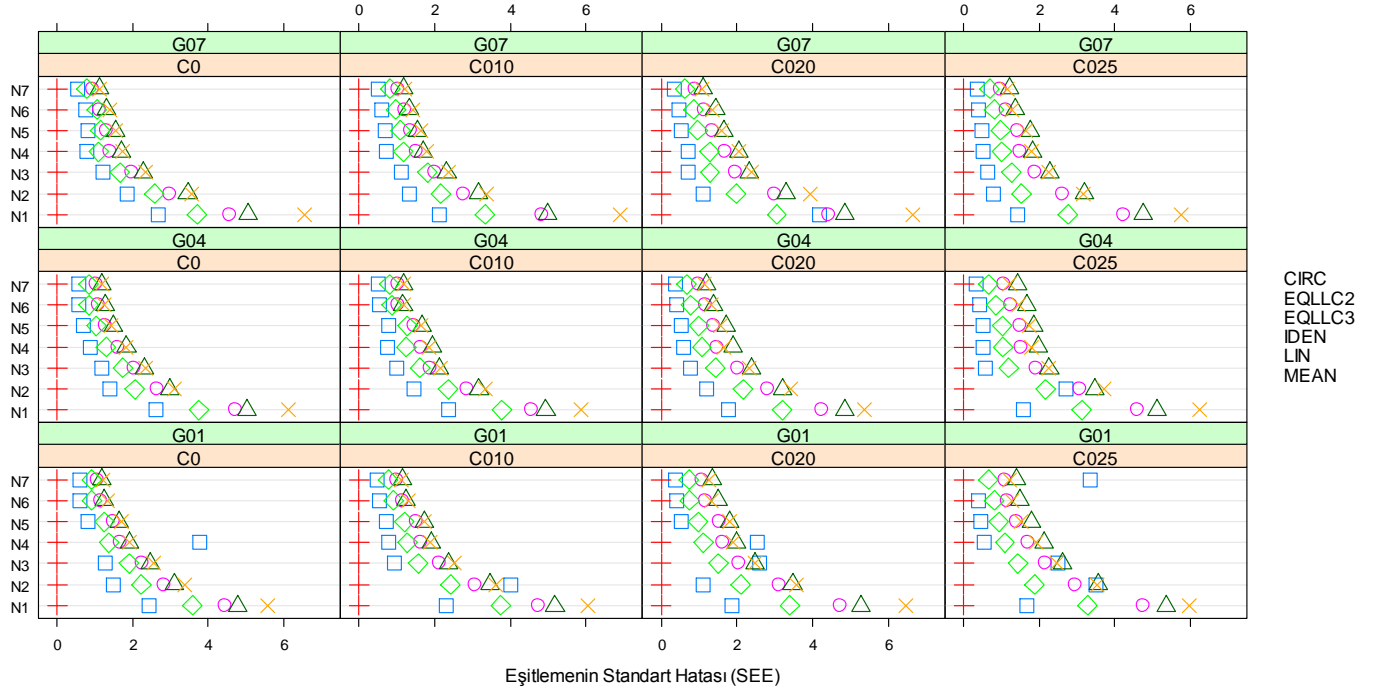
### **2.2.1. Test uzunluđu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının, eşitlemenin standart hatasına (SEE) ortak etkisi**

Grafik 16 incelendiğinde, özdeş eşitleme dışındaki tüm yöntemlerin ortalama standart hataları, formlar arasındaki güçlük düzey farkı ve şans başarısının tüm düzeylerinde örneklem büyüklüğünün deđişiminden etkilendiđi görölmektedir. Örneklem büyüklüğü arttıkça, özdeş eşitleme yönteminin standart hatası sıfır deđerinde sabit kalırken, diđer yöntemlerin ortalama standart hatalarının azaldığı görölmektedir.

Formlar arası güçlük düzeyindeki deđişimlerin, Dairesel yay yöntemi dışındaki yöntemlerin ortalama standart hatalarını önemli ölçüde etkilemediđi görölrken, Dairesel yay yönteminde, güçlük düzey farkı 0.1 olduđu durumlarda ortalama standart hatalar düzensiz deđişimler göstermektedir. Şans parametre deđerinin deđişimi eşitleme yöntemlerinin ortalama standart hatalarını önemli ölçüde deđiştirmediđi görölmektedir. Ancak, güçlük düzey farkı 0.1 olduđu durumda, Dairesel yay yöntemi, şans parametreleri arttıkça düzensiz deđişimler göstermektedir. Diđer güçlük düzeylerinde, şans parametresinin deđişimi, ortalama standart hatayı önemli ölçüde deđiştirmediđi görölmektedir.



### Ortak Etki M.S=50



**Grafik 16:** Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme ortalama standart hatasına (SEE) ortak etkisi (Test uzunluğu-50)

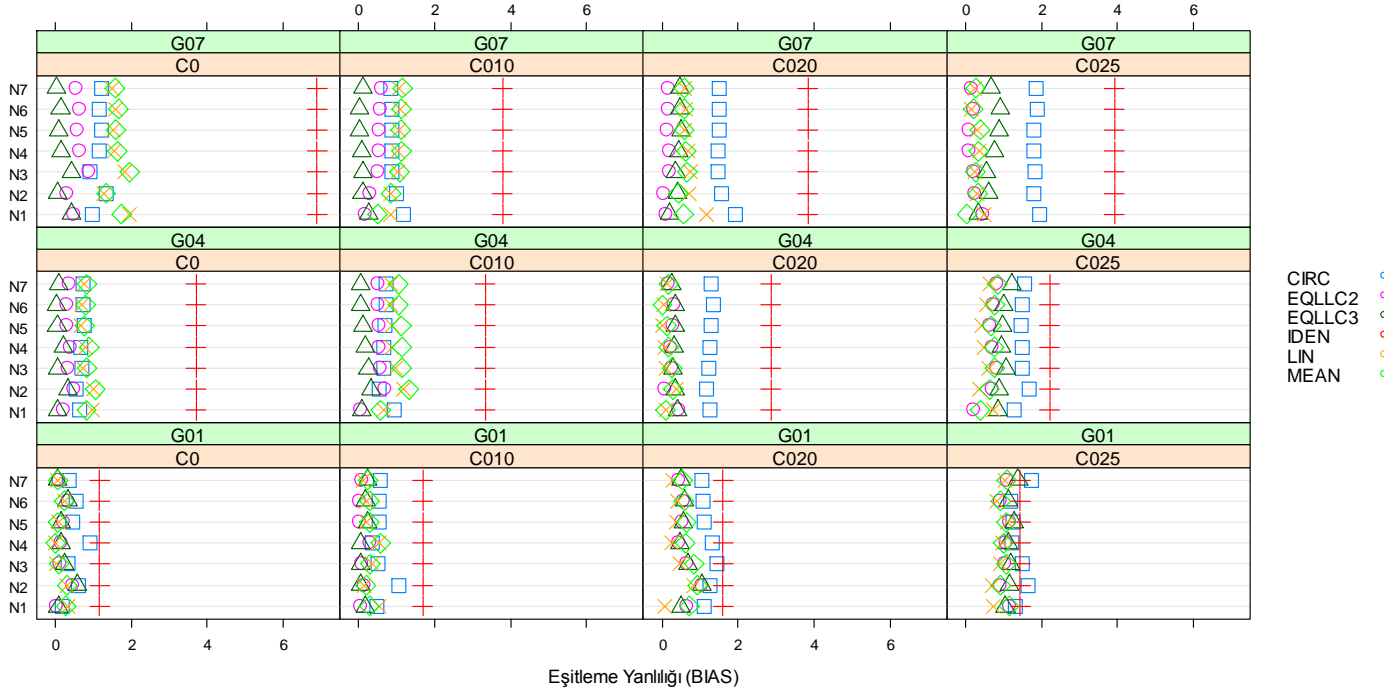
### **2.2.2. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme yanlılığına (BIAS) ortak etkisi**

Grafik 17 incelendiğinde, örneklem büyüklüğünün, test formları arasındaki güçlük düzey farkının ve şans başarısının, eşitleme yöntemlerinin ortalama yanlılık değerlerini etkilemediği görülmektedir.

Formlar arasındaki güçlük düzey farkı arttıkça, yöntemlerin eşitleme yanlılığında küçük miktarda artış olduğu görülmektedir. Güçlük düzeyleri farkının artmasından en çok etkilenen yöntem özdeş eşitleme yöntemidir. Güçlük düzeyinin arttığı durumlarda, özdeş eşitlemenin ortalama yanlılığının büyük miktarda arttığı görülmektedir. Güçlük düzey farkı arttıkça ortalama, doğrusal, özdeş ve Dairesel yay eşitleme yöntemlerinin ortalama yanlılık değerlerinin arttığı, LLC2 ve LLC3 eşitleme yöntemlerinin ortalama yanlılık değerlerinin ise önemli ölçüde değişmediği görülmektedir.

Şans parametre değeri arttıkça, 0.1 ve 0.4 güçlük düzey farklarında özdeş eşitleme yöntemi dışındaki diğer yöntemlerin ortalama yanlılıklarında çok küçük artış olduğu görülmektedir. 0.7 güçlük düzeyinde, şans parametresinin değişiminin ortalama yanlılık değerlerini etkilemediği görülmektedir.

Ortak Etki M.S=50



**Grafik 17:** Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme yanlılığına (BIAS) ortak etkisi (Test uzunluğu-50)

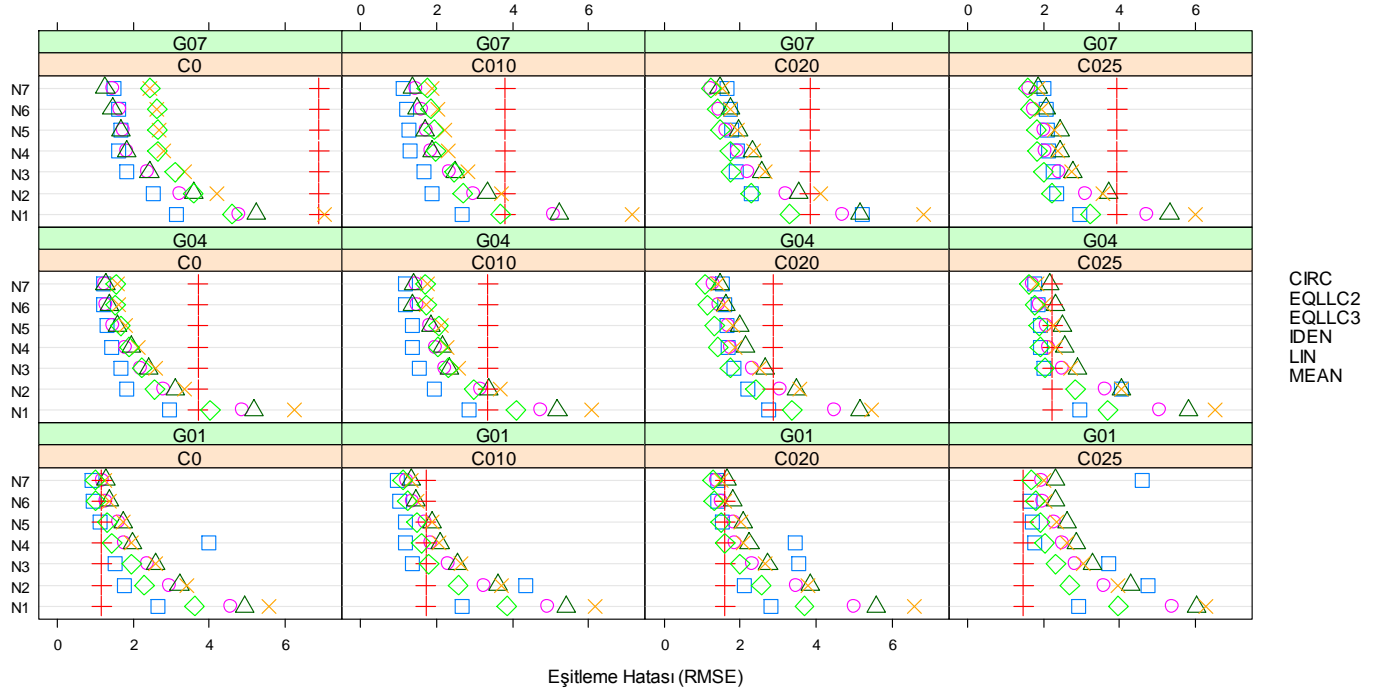
### **2.2.3. Test uzunluğu 50 olan testlerde, örneklem büyüklüğünün, şans başarısının ve test formları arasındaki güçlük düzey farkının ortalama eşitleme hatasına (RMSE) ortak etkisi**

Grafik 18 incelendiğinde, örneklem büyüklükleri arttıkça formlar arası güçlük düzey farkının ve şans başarısının tüm düzeylerinde, özdeş eşitleme dışındaki diğer eşitleme yöntemlerinin RMSE değerlerinin azaldığı, özdeş eşitlemenin ise sabit kaldığı görülmektedir

Formların güçlük düzeyleri farkının arttığı durumlarda, özdeş eşitleme yönteminin RMSE değerinde büyük bir artış olduğu görülürken, diğer yöntemlerin RMSE değerlerinde önemli düzeyde artış olmadığı görülmektedir. Formların güçlük düzey farkının küçük olduğu durumlarda (0.1 SMD), şans parametresi 0 ve 0.10 olduğu durumlarda 10 ve 25 örneklem büyüklüklerinde; şans parametresi 0.20 olduğu durumlarda 10-50 örneklem büyüklükleri arasında ve şans parametresi 0.25 olduğu durumda tüm örneklem büyüklüklerinde özdeş eşitleme yönteminin en düşük hataya sahip olduğu görülmektedir. Formlar arası güçlük düzey farkının 0.4 ve 0.7 olduğu ve şans parametre değerlerinin 0, 0.10 ve 0.20 olduğu durumlarda 50 örneklem büyüklüğü ve üstünde özdeş eşitleme yöntemiyle diğer yöntemlerden daha yüksek hata elde edildiği görülmektedir..

Şans parametre değeri arttıkça, 0.1 ve 0.4 güçlük düzey farklarında özdeş eşitleme yöntemi dışındaki diğer yöntemlerin eşitleme hatalarında artış olduğu görülmektedir. Güçlük düzey farkı 0.7 olduğu durumlarda, şans parametre değerlerinin artışı yöntemlerin eşitleme hatalarını etkilemediği görülmektedir.

### Ortak Etki M.S=50



**Grafik 18:** Örneklem büyüklüğü, şans başarısı ve test formları arası güçlük düzey farkı faktörlerinin eşitleme hatasına (RMSE) ortak etkisi (Test uzunluğu-50 madde)

## BÖLÜM IV: TARTIŞMA ve YORUM

Bu bölümde elde edilen bulgular araştırma soruları çerçevesinde ele alınıp ilgili araştırmalarla tartışılıp yorumlanmıştır.

### 1. Örneklem büyüklüğünün, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi

Çalışma sonucunda, örneklem büyüklüğü arttıkça özdeş eşitleme haricindeki diğer eşitleme yöntemlerinin standart hatalarının (SEE) azaldığı görülmektedir. Elde edilen bu bulguların, literatürdeki diğer çalışmaların sonuçlarıyla örtüştüğü görülmektedir (Livingston, 1993; Kolen ve Brennan, 2004; Skaggs, 2005; Kim ve diğerleri, 2006; Heh, 2007; Devdass, 2011). Örneklem büyüklüğü değiştiğinde, özdeş eşitleme yönteminde, formlardan elde edilen puanların standart hatalarının sıfır olduğu ve elde edilen bu sonucun Skaggs (2005), Kim ve diğerleri (2006), Heh (2007) tarafından yapılan çalışmalarla paralel sonuçlar gösterdiği görülmektedir. Çalışma sonucunda, 10-75 örneklem büyüklükleri arasında en yüksek ortalama standart hatanın doğrusal eşitleme yöntemiyle, , 75-200 örneklem büyüklükleri arasında ise, LLC3 eşit yüzdeliği eşitleme yöntemiyle elde edildiği görülmektedir. Elde edilen bulgular, Parshall ve diğerleri (1995), Skaggs (2005), Heh (2007), Devdass (2011) tarafından yapılan araştırmaların bulguları ile benzerlik göstermektedir. Sonuç olarak, örneklem büyüklüğünün arttığı durumlarda, özdeş eşitleme yöntemi dışındaki

tüm yöntemlerin standart hatalarının azaldığı, özdeş eşitleme yönteminin standart hatasının ise sıfır olduğu görülmüştür.

Örneklem büyüklüğünün arttığı durumlarda, eşitleme yöntemlerinin yanlılıklarında (BIAS) önemli ölçüde değişme olmadığı görülmektedir. Özdeş eşitleme yöntemi, örneklem büyüklüğünün değişiminden hiç etkilenmezken, diğer yöntemlerde çok küçük değişimler olduğu görülmektedir. Livingston (1993), örneklem büyüklüğünün eşitleme yanlılığına etkisinin olmadığını belirtirken, Skaggs (2005), Heh (2007) ve Devdass (2011) ise örneklem büyüklüğünün arttığı durumlarda eşitleme yanlılığının çok küçük düzeyde artış gösterdiği sonucuna ulaşmışlardır. Bu çalışmadan elde edilen bulgular, Skaggs (2005), Heh (2007) ve Devdass (2011) tarafından yapılan çalışmaların sonuçlarıyla örtüşmektedir.

Örneklem büyüklüğü arttıkça, özdeş eşitleme yönteminin eşitleme hatası (RMSE) sabit kalırken, diğer yöntemlerin eşitleme hata değerlerinin azaldığı görülmektedir. Özdeş eşitleme yöntemiyle, 10 ve 25 örneklem büyüklüklerinde en düşük RMSE değerinin elde edildiği görülmektedir. 25 örneklem büyüklüğünde, ortalama ve Dairesel yay eşitleme yöntemleriyle, özdeş eşitleme yöntemlerinden elde edilen hata değerlerinin birbirlerine yakın olduğu görülmektedir. 50 örneklem büyüklüğünde, ortalama, Dairesel yay, LLC2 ve LLC3 eşit yüzdeli eşitleme yöntemlerinin özdeş eşitlemeye göre daha küçük eşitleme hatasına sahip olduğu, 75 örneklem büyüklüğünde ise özdeş eşitleme yönteminin tüm yöntemlerden daha yüksek hataya sahip olduğu görülmektedir. Elde edilen sonuçlar, Devdass (2011) tarafından yapılan çalışmaların bulgularıyla örtüşmektedir. Hanson (1994) çalışmasında, eşitleme yapılabilmesi için gerekli en düşük örneklem büyüklüğünü 100 olarak belirlemiştir. Skaggs (2005), 25 örneklem büyüklüğünde dahi eşitlemenin yapılabileceği sonucuna ulaşmıştır. Kim, von Davier ve Haberman (2008), eşitleme yapılabilmesi için en küçük

örneklem büyüklüğünün 50 olması gerektiğini ifade etmişlerdir. Çalışma sonucunda, 50 örneklem büyüklüğünde ortalama, Dairesel yay, LLC2 ve LLC3 eşit yüzdelikli eşitleme yöntemleri ve 75 örneklem büyüklüğü ve üstünde çalışmada yer alan tüm yöntemler, özdeş eşitleme yöntemi yerine kullanılabilceği bulgusuna ulaşılmış ve elde edilen bu bulgular Hanson (1994), Skaggs (2005) ve Kim, von Davier ve Haberman (2008) tarafından yapılan çalışmaların bulgularıyla benzerlik göstermektedir. 10-50 örneklem büyüklükleri arasında RMSE değerlerinde hızlı bir düşüş olduğu görülmektedir. Bu sonuç, Skaggs (2005), Heh (2007) ve Devdass (2011) tarafından yapılan çalışmaların bulgularıyla paralellik göstermektedir. Ayrıca, Livingston (1993), Parshall (1995), Skaggs (2005) ve Livingston ve Kim (2009, 2010) yaptıkları çalışmalarda, farklı momentlerde Log-Linear ön düzgünleştirme yapmış ve ön düzgünleştirme yapmanın eşitleme hatasını (RMSE) önemli ölçüde azalttığını belirtmişlerdir. Çalışma kapsamında kullanılan ön düzgünleştirme işleminden elde edilen bulguların, Livingston (1993), Parshall (1995), Skaggs (2005) ve Livingston ve Kim (2009, 2010) tarafından yapılan çalışmaların bulgularıyla örtüştüğü görülmektedir.

## **2. Test formları arasındaki güçlük düzey farkının, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi**

Çalışmada, formlar arasındaki güçlük düzey farkının artmasının, Dairesel yay eşitleme yöntemi dışındaki diğer eşitleme yöntemlerin standart hatalarını (SEE) önemli ölçüde etkilemediği sonucuna ulaşılmıştır.

Formlar arasındaki güçlük düzey farkının artması, özdeş eşitleme yanlılığını (BIAS) önemli miktarda artırdığı, Dairesel yay, ortalama ve doğrusal eşitleme yöntemlerinin yanlılığını önemli miktarda değiştirmedeği görülmektedir.



Ayrıca, güçlük düzey farkının artması, LLC2 ve LLC3 eşit yüzdelikli eşitleme yöntemlerinin yanlılıklarında düzensiz değişimler gösterdiği sonucuna ulaşılmıştır. Elde edilen bu bulgular, Devdass (2011) tarafından elde edilen bulgularla benzerlik göstermektedir.

Formlar arasındaki güçlük düzey farkının artması, özdeş eşitlemenin RMSE değerlerini artırdığı, diğer eşitleme yöntemlerin RMSE değerlerinde önemli değişimlere yol açmadığı sonucuna ulaşılmıştır. Formların güçlük düzeyleri birbirine yakinken, özdeş eşitleme yöntemi kullanılmasının uygun olduğu yorumu yapılabilir. Harris ve Crouse (1993) ve Kolen ve Brennan (2004) test formlarının güçlükleri birbirine yakın olduğu durumlarda eşitleme yapmamanın daha az hataya sahip olacağını belirtmiştir. Harris ve Crouse (1993) ve Kolen ve Brennan (2004) tarafından yapılan çalışmanın sonuçlarıyla, bu çalışmanın sonuçları tutarlılık göstermektedir. Formlar arasındaki güçlük düzey farkı arttıkça özdeş eşitleme yöntemi yerine diğer yöntemlerin kullanılmasının uygun olacağı yorumu yapılabilir. Elde edilen bu sonuç, Kim, von Davier, ve Heberman (2006) tarafından yapılan çalışmanın sonuçlarıyla paralellik göstermektedir. Formlar arasındaki güçlük farkının yüksek olduğu durumlarda kullanılacak en uygun yöntemin, Dairesel yay yöntemi ve ortalama eşitleme yöntemi olduğu söylenebilir. Babcock (2012), eşit yetenek düzeyinde, formlar arasındaki güçlük düzey farkı yüksek olduğu durumlarda ortalama eşitleme yönteminin kullanılmasını önermiştir. Babcock (2012) tarafından yapılan çalışmanın bulgularıyla, bu çalışmanın bulguları benzerlik göstermektedir.

### **3. Şans başarısının, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi**

Çalışma sonucunda, şans parametre değeri arttıkça eşitleme yöntemlerinin standart hatalarında (SEE) önemli miktarda değişim olmadığı

görülmektedir. Böylelikle, şans parametresinin değişimlenmesi, eşitlemenin standart hatasını etkilemediği sonucuna ulaşılmıştır.

Şans parametresinin değişimlenmesinin eşitleme yanlılığına (BIAS) etkisi yöntemden yonteme farklılık göstermektedir. Ancak, yanlılık açısından değerlendirildiğinde, şans parametresinden en çok Dairesel yay eşitleme yönteminin etkilendiği yorumu yapılabilir.

Şans parametre değeri arttıkça, eşitleme hatasının (RMSE) değişimi yöntemlere göre farklılık göstermektedir. Şans başarısından en çok etkilenen yöntem Dairesel yay yöntemidir. Kim ve Livingston (2009), Dairesel yay yönteminin, başlangıç ve bitiş noktasına ve görgül olarak hesaplanmış bir orta noktaya sahip eşitleme eğrisi içerdiğini ve bu eğrinin üst noktasının, testten alınabilecek maksimum puan, alt noktasının ise şans başarısı ile testten alınabilecek minimum puan olduğunu ifade etmişlerdir. Dairesel yay yönteminin, şans başarısından etkilenmesinin, şans başarısının Dairesel yay eşitleme fonksiyonunu doğrudan etkilediğinden kaynaklandığı düşünülmektedir. Sonuç olarak, testlerin şans başarısından arındırılmadan eşitlendiği durumlarda ve şans parametre değeri 0.10 olduğu durumlarda Dairesel yay yöntemi, diğer durumlarda ise ortalama eşitleme yönteminin kullanılabilceği yorumu yapılabilir.

#### **4. Test uzunluğunun, test eşitleme yöntemlerinden elde edilen SEE, BIAS, RMSE değerlerine temel etkisinin değerlendirilmesi**

Test uzunluğu arttıkça, eşitlemenin standart hatası (SEE) artmaktadır. bu nedenle, test uzunluğunun eşitlemenin standart hatasına doğrusal etkisi olduğu yorumu yapılabilir.

Test uzunluğu arttıkça, yöntemlerin küçük miktarda yanlılıkları (BIAS) artmaktadır. Elde edilen bu bulgu, Devdass (2011) tarafından yapılan çalışmanın

bulgusu ile paralellik göstermektedir. Test uzunluğu arttıkça, LLC2 ve LLC3 eşit yüzdelikli eşitleme yöntemlerinin en düşük, özdeş eşitleme yönteminin ise en yüksek yanlılığa sahip olduğu sonucuna ulaşılmıştır.

Test uzunluğu arttıkça, eşitleme hatası (RMSE) artmaktadır. 25 maddeden oluşan testlerde ortalama eşitleme yöntemi, 50 maddeden oluşan testlerde ise Dairesel yay eşitleme yöntemi en düşük hataya sahiptir. 25 maddeden oluşan test formlarında, ortalama ve Dairesel yay yöntemleri, 50 maddeden oluşan test formlarında, LLC2 eşit yüzdelikli eşitleme yöntemi, ortalama ve Dairesel yay eşitleme yöntemleri özdeş eşitlemeden daha düşük hataya sahip oldukları için bu test uzunluklarında kullanılması önerilmektedir.

##### **5. Çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE) ortak etkisinin değerlendirilmesi**

Yapılan çalışmada, örneklem büyüklüğü arttıkça, tüm faktörlerin tüm düzeylerinde özdeş eşitleme haricindeki diğer yöntemlerin standart hatalarının azaldığı görülmektedir. Elde edilen bu bulgu, Livingston, (1993), Parshall (1995), Skaggs (2005), Heh (2007), Devdass (2011), Babcock (2012) çalışmalarından elde edilen bulgularla paralellik göstermektedir.

Formların günlük düzeyleri farkı birbirine yakın olduğu durumlarda ve şans parametresi değiştiğinde, Dairesel yay yönteminin standart hatasının düzensiz değişimler gösterdiği görülmektedir. Sonuç olarak, örneklem büyüklüğü ve test uzunluğunun arttığı durumlarda, eşitlemenin standart hatasının azaldığı, günlük düzey farkları ve şans başarısı düzeyleri değiştiğinde, eşitlemenin standart hatasında önemli ölçüde değişimlerin olmadığı görülmektedir. Eşitleme yöntemleri, standart hata değerlerine göre incelendiğinde, en küçük standart hata değerine özdeş eşitleme yöntemi sahiptir. Bu yöntemi sırasıyla, Dairesel yay,

ortalama eşitleme, LLC2 eşit yüzdellikli eşitleme, LLC3 eşit yüzdellikli eşitleme ve doğrusal eşitleme yöntemleri izlemektedir.

## **6. Çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen eşitleme yanlılığına (BIAS) ortak etkisinin değerlendirilmesi**

Örneklem büyüklüğünün arttığı durumlarda, tüm faktörlerin tüm düzeylerinde, özdeş eşitlemenin yanlılık (BIAS) değerleri sabit kalırken, diğer yöntemlerin yanlılık değerlerinde çok küçük değişimlerin olduğu görülmektedir. Elde edilen bu bulgular, Skaggs (2005), Heh (2007) ve Devdass (2011) tarafından yapılan çalışmalardan elde edilen bulgularla tutarlılık göstermektedir.

Test formlarının güçlük düzeyleri farkının arttığı durumlarda özdeş eşitleme ve Dairesel yay yöntemlerinin yanlılıklarının artmakta olduğu görülmektedir. Güçlük düzey farkı arttıkça ortalama ve doğrusal eşitleme yöntemlerinin yanlılık değerlerinde küçük bir artış görülürken, LLC2 ve LLC3 eşit yüzdellikli eşitleme yöntemlerinin yanlılıkları düzensiz küçük değişimler göstermektedir. Testler arası güçlük düzey farkı arttığı durumlarda, LLC2 ve LLC3 eşit yüzdellikli eşitleme yöntemleri, doğrusal, ortalama ve Dairesel yay yöntemlerine göre daha düşük yanlılık değerine sahiptir. Ancak formlar arası güçlük düzeyleri farkı az olduğu durumlarda tüm yöntemler yakın yanlılık değerlerine sahiptir. Elde edilen bu bulgular, Devdass (2011) tarafından elde edilen bulgularla örtüşmektedir.

Şans parametre değeri 0 olduğu durumlarda, eşitleme yöntemlerinin yanlılıklarında önemli değişimler görülmezken, 0.10-0.25 aralığında Dairesel yay yönteminin yanlılığının arttığı, diğer yöntemlerde ise, çok küçük değişimler olduğu görülmektedir.

Sonuç olarak, örneklem büyüklüğü, eşitleme yöntemlerinin yanlılıklarını önemli ölçüde etkilemediği, formlar arası güçlük düzey farkının arttığı durumlarda, özdeş ve dairesel yay yöntemlerinin yanlılıklarının arttığı, diğer yöntemlerin yanlılıklarında çok küçük değişimler gösterdiği yorumu yapılabilir. Şans başarı düzeyi arttıkça dairesel yay yönteminin yanlılığının arttığı diğer yöntemlerin ise önemli değişimler göstermediği sonucuna ulaşılmıştır. Test uzunluğu arttıkça, yöntemlerin eşitleme yanlılığı küçük miktarda artış olduğu sonucuna ulaşılmıştır. Eşitleme yöntemleri, standart hata değerlerine göre incelendiğinde, en küçük standart hataya özdeş eşitleme yöntemi sahiptir. Bu yöntemi sırasıyla, Dairesel yay, ortalama eşitleme, LLC2 eşit yüzdelliği eşitleme, LLC2 eşit yüzdelliği eşitleme ve doğrusal eşitleme yöntemleri izlemektedir.

#### **7. Çeşitli faktörlerin, test eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) ortak etkisinin değerlendirilmesi**

Örneklem büyüklüğü arttığı durumlarda, tüm faktörlerin tüm düzeylerinde, özdeş eşitleme yönteminin RMSE değerleri sabit kalırken, diğer yöntemlerin RMSE değerleri azalmıştır. Böylelikle, örneklem büyüklüğü arttıkça eşitleme hatasının (RMSE) azaldığı yorumu yapılabilir.

Güçlük düzeyinin değişimlenmesinden en çok etkilenen yöntem özdeş eşitleme yöntemidir. Güçlük düzey farkının arttığı durumlarda, özdeş eşitlemenin RMSE değerleri artarken, diğer yöntemlerin RMSE değerlerinde önemli değişimler olmadığı görülmektedir. Formların güçlük düzeyleri yakın olduğu durumlarda, dairesel yay yönteminin RMSE değerleri düzensiz değişimler gösterdiğinden, bu düzeyde kullanılmaması önerilmektedir.

Şans parametresinin değişimlenmesi RMSE değerlerini önemli ölçüde etkilememektedir. Şans parametresinin değişimlenmesinden en çok etkilenen yöntem Dairesel yay eşitleme yöntemidir. Özellikle, güçlük düzeyi yakın olan formlarda şans parametre değeri arttıkça, dairesel yay yönteminin RMSE değerlerinde düzensiz değişimler görülmektedir.

Formlar arasındaki güçlük düzey farkı 0.1 olduğu durumlarda, özdeş eşitleme yöntemi en düşük hataya sahiptir. Ancak güçlük düzey farkı 0.1, şans parametresi 0 olduğu durumlarda, 150 ve 200 örneklem büyüklüklerinde dairesel yay yöntemi daha düşük RMSE değerine sahiptir. Bu sebeple, test formlarının eşitlenmesinde, güçlük düzey farkı 0.1, şans parametresi 0 olduğunda 150 ve 200 örneklem büyüklüklerinde Dairesel yay, diğer koşullarda özdeş eşitlemenin kullanılmasının daha doğru sonuçlar vereceği yorumu yapılabilir.

Güçlük düzeyi 0.4, şans parametresi 0 ve 0.20 olduğu durumlarda, eşitleme yöntemlerinin kullanılması için gerekli minimum örneklem büyüklükleri, Dairesel yay yöntemi için 25, Ortalama eşitleme yöntemi için 50, LLC2 eşit yüzdelli eşitleme yöntemi için 75, LLC3 eşit yüzdelli eşitleme yöntemi ve doğrusal eşitleme yöntemi için 100 olduğu sonucuna ulaşılmıştır. 150 örneklem büyüklüğünde tüm yöntemler özdeş eşitlemeden daha az RMSE değerlerine sahip oldukları için 100 ve üstü örneklem büyüklüklerinde özdeş eşitleme dışındaki tüm yöntemlerin test formlarının eşitlenmesinde kullanılabileceği sonucuna ulaşılmıştır. Güçlük düzeyi 0.4, şans parametresi 0.10 olduğu durumlarda, eşitleme yöntemlerinin kullanılması için gerekli minimum örneklem büyüklükleri, Dairesel yay yöntemi için 50, Ortalama eşitleme yöntemi için 75, LLC2 eşit yüzdelli eşitleme yöntemi için 100, LLC3 eşit yüzdelli eşitleme yöntemi ve doğrusal eşitleme yöntemi için 150 olduğu sonucuna ulaşılmıştır. 150 örneklem büyüklüğünde tüm yöntemler özdeş eşitlemeden daha az RMSE değerlerine sahip oldukları için 150 ve üstü örneklem büyüklüklerinde

özdeş eşitleme dışındaki tüm yöntemlerin test formlarının eşitlenmesinde kullanılabileceği sonucuna ulaşılmıştır. Güçlük düzeyi 0.4, şans parametresi 0.25 olduğu durumlarda, eşitleme yöntemlerinin kullanılması için gerekli minimum örneklem büyüklükleri, Dairesel yay ve Ortalama eşitleme yöntemleri için 25, doğrusal eşitleme ve LLC2 eşit yüzdelliği eşitleme yöntemleri için 50 ve LLC3 eşit yüzdelliği eşitleme yöntemi için 75 olduğu sonucuna ulaşılmıştır. 75 örneklem büyüklüğünde tüm yöntemler özdeş eşitlemeden daha az RMSE değerlerine sahip oldukları için 75 ve üstü örneklem büyüklüklerinde özdeş eşitleme dışındaki tüm yöntemlerin test formlarının eşitlenmesinde kullanılabileceği sonucuna ulaşılmıştır.

Formların güçlük düzey farkı 0.7 ve şans parametresi 0 olduğu durumlarda, Dairesel yay ve ortalama eşitleme yöntemlerinin tüm örneklem büyüklüklerinde özdeş eşitleme yöntemine göre daha düşük RMSE değerine sahip olduğu görülmektedir. 25 örneklem büyüklüğünde, doğrusal eşitleme haricindeki tüm yöntemlerin kullanılmasının uygun olacağı sonucuna ulaşılmıştır. 50 örneklem büyüklüğünde özdeş eşitleme dışındaki tüm yöntemler test formlarının eşitlenmesinde kullanılmasının uygun olacağı sonucuna ulaşılmıştır. Formların güçlük düzey farkı 0.7, şans parametresi 0.10 ve 0.20 olduğu durumlarda, Dairesel yay, ortalama, LLC2 ve LLC3 eşit yüzdelliği eşitleme yöntemlerinin 25 ve üstü, doğrusal eşitleme yönteminin ise 50 ve üstü örneklem büyüklüklerinde test formlarının eşitlenmesinde kullanılmasının uygun olacağı sonucuna ulaşılmıştır. Formların güçlük düzey farkı 0.7, şans parametresi 0.25 olduğu durumlarda, Dairesel yay, ortalama ve LLC2 eşit yüzdelliği eşitleme yöntemlerinin 25 ve üstü, doğrusal eşitleme ve LLC3 eşit yüzdelliği eşitleme yöntemlerinin ise 50 ve üstü örneklem büyüklüklerinde test formlarının eşitlenmesinde kullanılmasının uygun olacağı sonucuna ulaşılmıştır.

Test uzunluęu arttıķa, eřitleme yntemlerinin RMSE deęerlerinde doęrusal artıř olduęu grlmektedir. Test uzunluęunun arttırılması, Dairesel yay eřitleme ynteminin RMSE hatalarını dięer yntemlere gre daha az etkilemektedir.

Çalıřmada yer alan faktrlerin eřitli dzeylerinde kullanılması uygun olan eřitleme yntemleri Tablo 7’de gsterilmektedir. Tablo 7 incelendięinde, formlar arası glk dzey farkı 0.1 olduęu durumlarda, zdeř eřitleme (IDEN) ynteminin uygun olduęu grlmektedir. Glk dzeyi 0.1 olduęu durumlarda, dięer faktrlerin eřitli dzeylerinde Circle-Arc (CIRC) ve ortalama eřitleme (MEAN) yntemlerinin de (C-M) kullanılabileceęi grlmektedir. Formlar arası glk dzey farkı 0.4 ve 0.7 olduęu durumlarda rnekleme byklę arttıķa, Tm yntemlerin, zdeř eřitleme yntemine gre daha dřk hata vermesinden dolayı belirtilen kořullarda tm yntemlerin kullanılabileceęi grlmektedir.



**Tablo 7:** Araştırmada yer alan faktörlerin düzeylerinde kullanılması uygun olan eşitleme yöntemleri

		G.D.		0.1				0.4				0.7			
T.U.	N	0.0	0.10	0.20	0.25	0.0	0.10	0.20	0.25	0.0	0.10	0.20	0.25		
25	200	CIRC	IDEN	IDEN	IDEN	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM		
	150	CIRC	IDEN	IDEN	IDEN	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM		
	100	IDEN	IDEN	IDEN	IDEN	TÜM	C-M-LLC2	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM		
	75	IDEN	IDEN	IDEN	IDEN	C-M-LLC2	C-M	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM		
	50	IDEN	IDEN	IDEN	IDEN	MEAN	CIRC	C-M	TÜM	TÜM	TÜM	TÜM	TÜM		
	25	IDEN	IDEN	IDEN	IDEN	CIRC	IDEN	CIRC	C-M	C-M-LLC2-3	C-M-LLC2-3	C-M-LLC2-3	C-M-LLC2		
	10	IDEN	IDEN	IDEN	IDEN	IDEN	IDEN	IDEN	IDEN	C-M	CIRC	IDEN	IDEN		
50	200	C-M	TÜM	C-M-D-LLC2	IDEN	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM	TÜM		
	150	C-M	TÜM	C-M-LLC2	IDEN	TÜM	TÜM	TÜM	C-M-D-LLC2	TÜM	TÜM	TÜM	TÜM		
	100	IDEN	C-M	C-M	IDEN	TÜM	TÜM	TÜM	C-M-LLC2	TÜM	TÜM	TÜM	TÜM		
	75	IDEN	C-M	IDEN	IDEN	TÜM	TÜM	TÜM	C-M-LLC2	TÜM	TÜM	TÜM	TÜM		
	50	IDEN	CIRC	IDEN	IDEN	TÜM	TÜM	TÜM	C-M	TÜM	TÜM	TÜM	TÜM		
	25	IDEN	IDEN	IDEN	IDEN	TÜM	C-M-LLC2	C-M	IDEN	TÜM	TÜM	C-M-LLC2-3	TÜM		
	10	IDEN	IDEN	IDEN	IDEN	CIRC	CIRC	CIRC	IDEN	C-M-LLC2-3	C-M	MEAN	C-M		

## ÖNERİLER

1. Bu çalışmada, eşitleme yöntemlerinin karşılaştırılmasında simülatif veriler kullanılmıştır. Aynı eşitleme yöntemleri gerçek veriler üzerinde de karşılaştırılabilir.
2. Bu çalışma, Klasik Test Kuramına dayalı özdeş, ortalama, doğrusal, LLC2 ve LLC3 eşit yüzdelikli ve Dairesel yay eşitleme yöntemleriyle sınırlıdır. Aynı çalışma, Klasik Test Kuramı'na dayalı farklı eşitleme yöntemleri de dahil edilerek veya Madde Tepki Kuramı'na dayalı eşitleme yöntemleri kullanılarak yapılabilir.
3. Çalışmada seçkisiz gruplar deseni ve bu desene uygun eşitleme yöntemleri kullanılmıştır. Aynı çalışma, NEAT desen ve bu desene uygun eşitleme yöntemleri kullanılarak da yapılabilir.
4. Bu çalışmada, örneklem büyüklüğü, formlar arası güçlük düzey farkı, şans başarısı ve test uzunluğu faktörlerinin farklı düzeylerinin eşitlemenin standart hatasına, yanlılığına ve eşitleme hatasına temel etkisi ve ortak etkisi incelenmiştir. Aynı çalışma, bu faktörlerin düzeyleri farklılaştırılarak da yapılabilir. Bununla birlikte, çalışmada yer alan faktörlerin yanı sıra, madde ayırt ediciliği, grup yetenek dağılımı, geçme/kalma puanı gibi farklı faktörler araştırmaya dahil edilerek, bu faktörlerin eşitleme hatalarına etkileri incelenebilir.

## KAYNAKÇA

- Angoff, W. H. (1984). Scales, norms, and equivalent scores. Princeton, NJ: Educational Testing Service.
- Babcock, B., Albano, A., & Raymond, M. (2012). Nominal weights mean equating: A method for very small samples. *Educational And Psychological Measurement*, 72(4), s. 608-628.
- Baykul, Yaşar (2010). Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması. Ankara: Pegem Yayınları
- Bozdağ, S. (2010). Şans başarısının test eşitlemeye etkisi. Yayınlanmamış Yüksek Lisans Tezi, Mersin Üniversitesi, Mersin.
- Bozdağ, S. ve Kan, A. (2010). Şans başarısının test eşitlemeye etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. 39, ss.91-108.
- Crocker, L. ve Aigina, J. (1986). Introduction to classical and modern test theory. Ohio: Cengage Learning.
- Cui, Z. ve Kolen, M. J. (2009). Evaluation of two new smoothing methods in equating: the cubic b-spline presmoothing method and the direct presmoothing method. *Journal of Educational Measurement*, 46(2), s. 25–44.

- Cui, Z. ve Kolen, M. J. (2009). Evaluation of two new smoothing methods in equating: the cubic b-spline presmoothing method and the direct presmoothing method. *Journal of Educational Measurement*, 46 (2), s. 135-158.
- Devdass, S. (2011). Conditions Affecting the Accuracy of Classical Equating Methods for Small Samples Under the NEAT Design: A Simulation Study. Doctoral dissertation, The University of North Carolina.
- Dorans, N. J., Moses, T. P. ve Eignor, D. R. (2010). Principles and practices of test score equating (ETS Research Rept. RR-10-29). Princeton, NJ: ETS.
- Gök, B. (2012). Denk Olmayan Gruplarda Ortak Madde Deseni Kullanılarak Madde Tepki Kuramına Dayalı Eşitleme Yöntemlerinin Karşılaştırılması. Doktora Tezi, Hacettepe Üniversitesi.
- Heh, V.K., (2007). Equating Accuracy Using Small Samples in the Random Groups Design. Doctoral dissertation, the College of Education of Ohio University.
- Hanson, B. A., Zeng, L., & Colton, D. (1994). A Comparison of Presmoothing and Postsmoothing Methods in Equipercentile Equating (ACT Research Report 94-4). Iowa City, IA: American College Testing.
- Kan, A. (2010). Test eşitleme: Aynı davranışları ölçen, farklı madde formlarına sahip testlerin istatistiksel eşitliğinin sınanması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), ss. 16-21.
- Kan, A. (2011). Test eşitleme: OKS testlerinin istatistiksel eşitliğinin sınanması. *Eğitim ve Bilim*, 36 160.

- Keleciođlu, H. (1994). Öğrenci seçme sınavı puanlarının eşitlenmesi üzerine bir çalışma. Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Kilmen, S. (2010). Madde tepki kuramına dayalı test eşitleme yöntemlerinden kestirilen eşitleme hatalarının örneklem büyüklüğü ve yetenek dağılımına göre karşılaştırılması. Yayınlanmamış Doktora Tezi, Ankara Üniversitesi, Ankara.
- Kim, S. ve Walker, M. E. (2012). Investigating repeater effects on chained equipercetile equating with common anchor items. *Applied Measurement in Education*, 25(1), s. 41-57.
- Kim, S., vonDavier, A. A., ve Haberman, S. (2008). Small-sample equating using a synthetic linking function. *Journal Of Educational Measurement*,45(4), s. 325-342.
- Kim, S.,vonDavier A. A. ve Haberman, S. (2011). Practical application of a synthetic linking function on small-sample equating. *Applied Measurement in Education*, 24(2), s. 95-114.
- Kolen, M. J. (1985). Standard errors of Tucker equating. *Applied Psychological Measurement*, 9(2), s. 209-223.
- Kolen, M. J. (1988). Traditional equating methodology. *Educational Measurement: Issues and Practice*, 7(4).
- Kolen, M. J.,& Brennan R. L. (2004). Test equating, scaling, and linking: Method and practice (2nd ed.). New York, NY: Springer-Verlag.

- Lee, E. (2013). Equating multidimensional tests under a random groups design: a comparison of various equating procedures. Doctoral dissertation, University of Iowa.
- Liu, C. (2011). A comparison of statistics for selecting smoothing parameters for Loglinear presmoothing and cubic spline postsmoothing under a random groups design. Doctoral dissertation, University of Iowa.
- Liu, J., Sinharay, S., Holland, P. W., Curley, E. ve Feigenbaum, M. (2011). Test score equating using a mini-version anchor and a midi anchor: A case study using sat data. *Journal of Educational Measurement*, 48(4), s. 361-379.
- Livingston, S. A. (1993). Small-sample equating with log-linearsmoothing. *Journal of Educational Measurement*, 30, s. 23-39.
- Livingston, S. A. (2004). *Equating test scores (without IRT)*. Princeton, NJ: ETS.
- Livingston, S. A. & Kim, S. (2008). Small-sample equating by the circle-arc method (ETS Research Rept. RR-08-39). Princeton, NJ: ETS.
- Livingston, S. A. & Kim, S. (2009). Comparisons among Small Sample Equating Methods in a Common-Item Design. *Journal of Educational Measurement*, 47(3), s. 286–298.
- Livingston, S. A., & Kim, S. (2009). The circle-arc method for equating in small samples. *Journal of Educational Measurement*, 46(3), s. 330–343.
- Livingston, S. A., & Kim, S. (2010). Random-Groups equating with samples of 50 to 400 test takers. *Journal of Educational Measurement*, 47, s. 175–185.

- Livingston, S. A., ve Kim, S. ve Lewis, C. (2011). Collateral Information for Equating in Small Samples: A Preliminary Investigation, *Applied Measurement in Education*, 24(4), s. 302-323.
- Moses, T. ve Liu, J. (2011). Smoothing and equating methods applied to different types of test score distributions and evaluated with respect to multiple equating criteria (ETS Research Rept. RR-11-20). Princeton, NJ: ETS.
- Öztürk, N. (2010). Akademik personel ve lisansüstü eğitimi giriş sınavı puanlarının eşitlenmesi üzerine bir çalışma. Yayınlanmamış Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.
- Öztürk, N. ve Anıl, D. (2012). Akademik personel ve lisansüstü eğitimi giriş sınavı puanlarının eşitlenmesi üzerine bir çalışma. *Eğitim ve Bilim*, 37 165.
- Parshall, C. G., Houghton, P. D. B., & Kromrey, J. D. (1995). Equating error and statistical bias in small sample linear equating. *Journal of Educational Measurement*, 32(1), s. 37-54.
- Puhan, G., Moses, T. P., Grant, M.C., McHale, F. (2009). Small-Sample Equating Using a Single-Group Nearly Equivalent Test (SiGNET) Design. *Journal of Educational Measurement*, 47(3), s. 344–362.
- Puhan, G. (2010). Chained versus post-stratification equating in a linear context: An evaluation using empirical data (ETS Research Rept. RR-10-06). Princeton, NJ: ETS.
- Puhan, G. (2011). Futility of Log-Linear Smoothing when Equating with Unrepresentative Small Samples. *Journal of Educational Measurement*, 48(3), s. 274-292.

- Puhan, G. (2011). Can Smoothing Help When Equating With Unrepresentative Small Samples? (ETS Research Rept. RR-11-09). Princeton, NJ: ETS.
- Sinharay, S., Holland, P. W. (2007). Is it necessary to make anchor tests mini-versions of the tests being equated or can some restrictions be relaxed? *Journal of Educational Measurement*, 44(3), s.249-275.
- Sinharay, S. ve Holland, P. W. (2010). A new approach to comparing several equating methods in the context of the neat design. *Journal of Educational Measurement*, 47(3), s. 261-285.
- Skaggs, G. & Lissitz, R. W. (1986) The effect of examinee ability on test equating invariance. American Educational Research Association
- Skaggs, G. (2005). Accuracy of random groups equating with very small samples. *Journal of Educational Measurement*, 42(4), 309-330.
- van der Linden, W. J. (2006). Evaluating equating error in observed-score equating. Law School Admission Council Computerized Testing Report, Newtown: Law School Admission Council.
- van der Linden, W. J. (2010). On bias in linear observed-score equating. *Measurement: Interdisciplinary Research and Perspectives*, 8(1), s. 21–26.
- vonDavier, A. A. ve Kong, N. (2005). A unified approach to linear equating for the nonequivalent groups design. *Journal of Educational and Behavioral Statistics*, 30(3), s. 313-342.
- vonDavier, A. A. (2008). New results on the linear equating methods for the non-equivalent-groups design. *Journal of Educational and Behavioral Statistics*, 33(2), s. 186-203.



- vonDavier, A. A., Holland, P. W., ve Thayer, D. T. (2004). The kernel method of test equating. New York, NY: Springer-Verlag.
- vonDavier, M. (2010). Statistical models for test equating, scaling, and linking. New York, NY: Springer-Verlag.
- Wang, C. (2011). An investigation of bootstrap methods for estimating the standard error of equating under the common-item nonequivalent groups design. Doctoral dissertation, University of Iowa.
- Wang, T. (2006). Standard errors of equating for equipercentile equating with log-linear pre-smoothing using the delta method. Iowa City: Center for Advanced Studies in Measurement and Assessment (CASMA).
- Yu, C. H. ve Osborn Popp, S. E. (2005). Test equating by common items and common subjects: concepts and applications. *Practical Assessment, Research & Evaluation*, 10(4).
- Zu, J. ve Liu, J. (2010). Observed score equating using discrete and passage-based anchor items. *Journal of Educational Measurement*, 47(4), s. 395-412.

**EKLER**

**EK-1:Simülatif Verilerden Oluşan X ve Y Formlarının Betimsel İstatistikleri**

<b>Form</b>	<b>Madde Sayısı</b>	<b>Şans</b>	<b>Güçlük Düzeyi</b>	<b>N</b>	<b>Ortalama</b>	<b>Standart Sapma</b>	<b>Medyan</b>	<b>Min</b>	<b>Max</b>	<b>Ranj</b>	<b>Çarpıklık</b>	<b>Basıklık</b>	<b>Standart Hata</b>
<b>X</b>	25	0	-	5000	12,90	4,44	13	0	25	25	-0,009	-0,454	0,063
<b>X</b>	25	0,1	-	5000	12,96	4,33	13	1	25	24	0,0510	-0,469	0,061
<b>X</b>	25	0,2	-	5000	14,32	3,89	14	2	25	23	-0,015	-0,474	0,055
<b>X</b>	25	0,25	-	5000	15,93	3,77	16	3	25	22	-0,180	-0,384	0,053
<b>X</b>	50	0	-	5000	23,26	8,60	23	1	46	45	0,049	-0,548	0,122
<b>X</b>	50	0,1	-	5000	28,52	7,98	29	5	49	44	-0,115	-0,523	0,113
<b>X</b>	50	0,2	-	5000	30,80	7,01	31	8	50	42	-0,080	-0,510	0,099
<b>X</b>	50	0,25	-	5000	31,26	6,72	31	11	49	38	-0,084	-0,459	0,095
<b>Y</b>	25	0	0,1	5000	12,38	4,68	12	0	25	25	0,021	-0,595	0,066
<b>Y</b>	25	0	0,4	5000	11,54	4,64	11	0	24	24	0,285	-0,455	0,065
<b>Y</b>	25	0	0,7	5000	10,18	4,37	10	0	25	25	0,129	-0,525	0,066
<b>Y</b>	25	0,1	0,1	5000	12,66	4,17	13	0	25	25	0,062	-0,435	0,059
<b>Y</b>	25	0,1	0,4	5000	11,86	4,30	12	0	25	25	0,136	-0,462	0,061
<b>Y</b>	25	0,1	0,7	5000	10,53	4,12	10	0	25	25	0,329	-0,228	0,058
<b>Y</b>	25	0,2	0,1	5000	13,99	4,12	14	2	25	23	-0,003	-0,505	0,058
<b>Y</b>	25	0,2	0,4	5000	13,12	3,96	13	0	25	25	0,077	-0,353	0,056

<b>Y</b>	25	0,2	0,7	5000	12,20	3,76	12	1	24	23	0,131	-0,324	0,053
<b>Y</b>	25	0,25	0,1	5000	15,75	3,81	16	4	25	21	-0,148	-0,496	0,054
<b>Y</b>	25	0,25	0,4	5000	14,16	3,62	14	2	25	23	0,050	-0,321	0,051
<b>Y</b>	25	0,25	0,7	5000	13,84	3,81	14	2	25	23	0,040	-0,404	0,054
<b>Y</b>	50	0	0,1	5000	25,25	8,46	25	1	47	46	-0,067	-0,569	0,120
<b>Y</b>	50	0	0,4	5000	21,88	8,01	22	1	49	48	0,118	-0,440	0,113
<b>Y</b>	50	0	0,7	5000	17,74	8,05	17	0	45	45	0,406	-0,214	0,114
<b>Y</b>	50	0,1	0,1	5000	26,48	7,72	26	5	47	42	0,002	-0,564	0,109
<b>Y</b>	50	0,1	0,4	5000	24,15	7,46	24	5	47	42	0,160	-0,455	0,106
<b>Y</b>	50	0,1	0,7	5000	23,58	7,70	23	4	48	44	0,169	-0,499	0,109
<b>Y</b>	50	0,2	0,1	5000	29,86	7,27	30	6	49	43	-0,079	-0,498	0,103
<b>Y</b>	50	0,2	0,4	5000	27,87	6,84	28	7	48	41	-0,024	-0,390	0,097
<b>Y</b>	50	0,2	0,7	5000	26,42	7,03	26	5	47	42	0,170	-0,387	0,099
<b>Y</b>	50	0,25	0,1	5000	30,85	6,76	31	9	49	40	-0,048	-0,465	0,096
<b>Y</b>	50	0,25	0,4	5000	29,79	6,96	30	7	49	42	-0,005	-0,506	0,098
<b>Y</b>	50	0,25	0,7	5000	27,09	6,64	27	7	48	41	0,148	-0,424	0,094

**EK-2: Örneklem Büyüklüğünün, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri**

**EK-2. 1-A:** Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

<b>Örneklem Büyüklüğü</b>	<b>Özdeş</b>	<b>Ortalama</b>	<b>Doğrusal</b>	<b>LLC2</b>	<b>LLC3</b>	<b>Dairesel yay</b>
<b>10</b>	0,000	2,595	4,585	3,482	3,819	2,070
<b>25</b>	0,000	1,653	2,684	2,231	2,522	1,468
<b>50</b>	0,000	1,170	1,803	1,540	1,503	0,903
<b>75</b>	0,000	0,941	1,456	1,254	1,483	0,723
<b>100</b>	0,000	0,829	1,272	1,099	1,295	0,722
<b>150</b>	0,000	0,659	1,016	0,863	1,047	0,422
<b>200</b>	0,000	0,588	0,895	0,778	0,936	0,558

**EK-2. 1-B:** Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

<b>Örneklem Büyüklüğü</b>	<b>Özdeş</b>	<b>Ortalama</b>	<b>Doğrusal</b>	<b>LLC2</b>	<b>LLC3</b>	<b>Dairesel yay</b>
<b>10</b>	1,872	0,470	0,489	0,310	0,373	0,802
<b>25</b>	1,872	0,508	0,507	0,314	0,319	0,714
<b>50</b>	1,872	0,477	0,441	0,288	0,489	0,687
<b>75</b>	1,872	0,473	0,430	0,281	0,251	0,677
<b>100</b>	1,872	0,475	0,439	0,293	0,289	0,681
<b>150</b>	1,872	0,444	0,413	0,345	0,271	0,674
<b>200</b>	1,872	0,460	0,421	0,281	0,280	0,709

**EK-2. 1-C:** Çeşitli örneklem büyüklüklerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

<b>Örneklem Büyüküğü</b>	<b>Özdeş</b>	<b>Ortalama</b>	<b>Doğrusal</b>	<b>LLC2</b>	<b>LLC3</b>	<b>Dairesel yay</b>
<b>10</b>	1,914	2,827	4,733	3,677	4,053	2,563
<b>25</b>	1,914	1,963	2,870	2,429	2,719	1,986
<b>50</b>	1,914	1,528	2,027	1,749	1,852	1,458
<b>75</b>	1,914	1,337	1,712	1,476	1,664	1,294
<b>100</b>	1,914	1,249	1,558	1,340	1,489	1,291
<b>150</b>	1,914	1,105	1,325	1,175	1,246	1,031
<b>200</b>	1,914	1,054	1,224	1,046	1,146	1,174

**EK-3: Test Formları Arasındaki Güçlük Düzey Farkının, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri**

**EK-3-A:** Çeşitli test formları arasındaki güçlük düzey farkına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

<b>Güçlük düzey farkı</b>	<b>Özdeş</b>	<b>Ortalama</b>	<b>Doğrusal</b>	<b>LLC2</b>	<b>LLC3</b>	<b>Dairesel yay</b>
<b>0.1</b>	0,000	1,222	1,993	1,648	1,883	1,226
<b>0.4</b>	0,000	1,209	1,916	1,594	1,824	0,931
<b>0.7</b>	0,000	1,184	1,967	1,586	1,818	0,785

**EK-3-B:** Çeşitli test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

<b>Güçlük düzey farkı</b>	<b>Özdeş</b>	<b>Ortalama</b>	<b>Doğrusal</b>	<b>LLC2</b>	<b>LLC3</b>	<b>Dairesel yay</b>
<b>0.1</b>	0,981	0,361	0,296	0,297	0,326	0,585
<b>0.4</b>	1,733	0,422	0,387	0,273	0,286	0,658
<b>0.7</b>	2,903	0,634	0,663	0,308	0,268	0,876

**EK-3-C:** Çeşitli Test formları arasındaki ortalama güçlük düzeyleri farkına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

<b>Güçlük düzey farkı</b>	<b>Özdeş</b>	<b>Ortalama</b>	<b>Doğrusal</b>	<b>LLC2</b>	<b>LLC3</b>	<b>Dairesel yay</b>
<b>0.1</b>	1,008	1,488	2,157	1,852	2,085	1,674
<b>0.4</b>	1,787	1,551	2,132	1,81	2,025	1,489
<b>0.7</b>	2,947	1,702	2,332	1,847	2,012	1,463



**EK-4: Şans Başarısının, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri**

**EK-4-A:** Çeşitli şans başarısına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

Şans parametresi	Özdeş	Ortalama	Doğrusal	LLC2	LLC3	Dairesel yay
<b>0.00</b>	0,000	1,298	1,958	1,635	1,821	1,109
<b>0.10</b>	0,000	1,293	1,970	1,621	1,808	1,023
<b>0.20</b>	0,000	1,146	1,954	1,577	1,831	0,903
<b>0.25</b>	0,000	1,083	1,953	1,605	1,907	0,889

**EK-4-B:** Çeşitli şans başarısı parametrelerine göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

Şans Parametresi	Özdeş	Ortalama	Doğrusal	LLC2	LLC3	Dairesel yay
<b>0.00</b>	1,73	0,57	0,598	0,352	0,244	0,295
<b>0.10</b>	2,001	0,519	0,483	0,237	0,107	0,557
<b>0.20</b>	1,917	0,33	0,301	0,213	0,263	0,912
<b>0.25</b>	1,841	0,471	0,412	0,37	0,56	1,063

**EK-4-C:** Çeşitli şans başarı parametrelerine göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

Şans parametresi	Özdeş	Ortalama	Doğrusal	LLC2	LLC3	Dairesel yay
<b>0.00</b>	1,851	1,68	2,256	1,826	1,94	1,353
<b>0.10</b>	2,015	1,644	2,196	1,794	1,92	1,38
<b>0.20</b>	1,928	1,416	2,102	1,749	2,005	1,631
<b>0.25</b>	1,863	1,582	2,274	1,975	2,299	1,806

**EK-5: Test Uzunluğunun, Eşitlemenin Standart Hatası (SEE), Yanlılığı (BIAS) ve Eşitleme Hatasına (RMSE) Temel Etkileri**

**EK-5-A:** Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama standart hata (SEE) değerleri

Madde Sayısı	Özdeş	Ortalama	Doğrusal	LLC2	LLC3	Dairesel yay
25	0,000	0,849	1,356	1,109	1,259	0,773
50	0,000	1,561	2,561	2,110	2,424	1,189

**EK-5-B:** Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama yanlılık (BIAS) değerleri

Madde Sayısı	Özdeş	Ortalama	Doğrusal	LLC2	LLC3	Dairesel yay
25	1,139	0,253	0,246	0,11	0,103	0,429
50	2,606	0,692	0,651	0,476	0,484	0,984

**EK-5-C:** Çeşitli test uzunluklarına göre eşitleme yöntemlerinden elde edilen ortalama hata (RMSE) değerleri

Madde Sayısı	Özdeş	Ortalama	Doğrusal	LLC2	LLC3	Dairesel yay
25	1,151	1,003	1,453	1,173	1,316	1,053
50	2,677	2,158	2,961	2,499	2,766	2,031

## ÖZGEÇMİŞ

### 1. KİŞİSEL BİLGİLER

**1.1. Adı Soyadı:** Semih AŞİRET

**1.2. Doğum Yeri ve Tarihi:** Mersin – 15/09/1983

**1.3. İletişim Bilgileri:**

**Adres:** Tozkoparan Mah. Atike Akel Ortaokulu  
Toroslar/MERSİN

**E-posta:** semihasiret@gmail.com

**1.4. Yabancı Dil:** İngilizce

**1.5. Eğitimi**

Derece	Alan	Üniversite	Yıl
Lisans	Bilgisayar ve Öğretim Teknolojileri Eğitimi	Orta Doğu Teknik Üniversitesi	2001-2006
Yükseklisans	Eğitimde Ölçme ve Değerlendirme	Mersin Üniversitesi	2011-2014

## 1.6. Çalıştığı Kurumlar

Unvanı	Görevi	Görev Yeri Yılı
Bilişim Teknolojileri Öğretmeni	Milli Eğitim Bakanlığı Gazi İlköğretim Okulu Mut/MERSİN	2006-2010
Bilişim Teknolojileri Öğretmeni	Milli Eğitim Bakanlığı Atatürk İlköğretim Toroslar/MERSİN	2010-2012
Bilişim Teknolojileri Öğretmeni	Milli Eğitim Bakanlığı Nihal Erdem Ticaret Meslek Lisesi Toroslar/MERSİN	2012-2013
Bilişim Teknolojileri Öğretmeni	Milli Eğitim Bakanlığı Atike Akel Ortaokulu Toroslar/MERSİN	2013-Devam ediyor

## 1.7. Tezleri

**Yüksek Lisans:** “Küçük Örneklerde Test Eşitleme Yöntemlerinin Çeşitli Faktörlere Göre İncelenmesi” (Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü,2014). (Danışman: Yrd. Doç. Dr. Seçil ÖMÜR SÜNBÜL)

## 2. AKADEMİK İLGİ ALANLARI

Klasik Test Teorisi

Test Geliştirme ve Madde Analizi

Test Eşitleme

Ölçek Geliştirme

### **3. BİLGİSAYAR BECERİLERİ**

**3.1. İşletim Sistemi:** Windows, Linux, Unix

**3.2. Kullanılan Programlar:** Microsoft Office Programları, SPSS, ITEMAN, R, C++, PHP