



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı

Eğitimde Ölçme ve Değerlendirme Programı

DEĞİŞEN MADDE FONKSİYONUNUN MIMIC YÖNTEMLERİYLE FARKLI KOŞULLAR
ALTINDA İNCELENMESİ

Pelin TAMCI

Doktora Tezi

Ankara, 2024



Liderlik, araştırma, inovasyon, kaliteli eğitim ve değişim ile

Daha ileriye... En İyeye...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

DEĞİŞEN MADDE FONKSİYONUNUN MIMIC YÖNTEMLERİYLE FARKLI KOŞULLAR
ALTINDA İNCELENMESİ

EXAMINATION OF DIFFERENTIAL ITEM FUNCTIONING UNDER DIFFERENT
CONDITIONS USING MIMIC METHODS

Pelin TAMCI

Doktora Tezi

Ankara, 2024

Kabul ve Onay

Eğitim Bilimleri Enstitüsü Müdürlüğüne,

Ad SOYADI'nın hazırladığı "Tezin Başlığı Buraya Yazılacak" başlıklı bu çalışma jürimiz tarafından **Ana Bilim Dalı**, **Bilim Dalında Yüksek Lisans/Doktora Tezi** olarak kabul edilmiştir. (Transkriptte görünen Ana Bilim Dalı ve Bilim Dalı yazılmalıdır.)

Jüri Başkanı

Unvan Ad SOYADI

İmza

Jüri Üyesi (Danışman)

Unvan Ad SOYADI

İmza

Jüri Üyesi

Unvan Ad SOYADI

İmza

Jüri Üyesi

Unvan Ad SOYADI

İmza

Jüri Üyesi

Unvan Ad SOYADI

İmza

İkinci Tez Danışmanı

(İkinci tez danışmanı yoksa
bu bölümü tamamen
kaldırınız.)

Unvan Ad SOYADI

Enstitü Yönetim Kurulunun
....../....../..... Tarihli ve
sayılı kararı.

Bu tez Hacettepe Üniversitesi Lisansüstü Eğitim, Öğretim ve Sınav Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından / / tarihinde uygun görülmüş ve Enstitü Yönetim Kurulunca / / tarihi itibarıyla kabul edilmiştir.

Prof. Dr. İsmail Hakkı MİRİCİ

Eğitim Bilimleri Enstitüsü Müdürü

Öz

Bu çalışmanın amacı standart MIMIC (M-ST) ve arınık çapa maddeli MIMIC (M-AÇM) yöntemlerinin değişen madde fonksiyonu (DMF) belirleme performanslarının farklı koşullar altında incelenmesidir. Çalışma, farklı koşullara göre üretilmiş simülasyon veri setleri ve PISA 2012 matematik testinden elde edilen gerçek veri seti üzerinde yürütülmüştür. Simülasyon çalışması kapsamında ele alınan koşullar; MTK modeli (2PLM ve 3PLM), örneklem büyüklüğü (2000, 1000 ve 600), odak/referans grup oranı (1/1 ve 1/3), DMF büyüklüğü (0.3, 0.5 ve 0.7) ve DMF'li madde oranı (%20 ve %40)'dır. 72 koşul için 100 tekrarla üretilmiş her bir veri seti üzerinde analizler yapılmıştır. Veriler, 20 maddelik bir test üretilmiştir. Gerçek veri seti üzerinden yürütülen analizlerde örneklem büyüklüğü (900 ve 450) ve odak/referans grup oranı (350/550, 450/450, 225/225 ve 150/300) koşulları ele alınmıştır. DMF analizleri M-ST ve M-AÇM yöntemleriyle yapılmıştır. Simülasyon verisiyle gerçekleştirilen analizlerde yöntemlerin performanslarını incelemek amacıyla I. tip hata ve güç oranları hesaplanmıştır. I. tip hata ve güç oranlarının ele alınan koşullara göre nasıl değiştiğini incelemek için Faktöriyel ANOVA analizleri uygulanmıştır. I. tip hata ve güç oranlarının örneklem büyüklüğü, odak/referans grup oranı, DMF'li madde oranı ve DMF büyüklüğüne göre anlamlı farklılık gösterdikleri bulunmuştur. I. tip hata oranları açısından M-AÇM yönteminin M-ST'den daha üstün bir performans sergilediği görülmüştür. Gerçek veriyle gerçekleştirilen analizler sonucunda tüm koşullar için M-AÇM yöntemiyle belirlenen DMF'li madde sayısının M-ST yöntemiyle belirlenen DMF'li madde sayısından daha düşük olduğu görülmüştür.

Anahtar sözcükler: standart MIMIC, arınık çapa maddeli MIMIC, değişen madde fonksiyonu

Abstract

The aim of this study is to investigate the performance of the standard MIMIC method (M-ST) and the MIMIC method with a pure anchor (M-PA) methods in determining DIF under different conditions. The study was conducted on simulation datasets generated under different conditions and a real dataset obtained from the PISA 2012 mathematics test. The conditions considered in the simulation study were IRT model (2PLM and 3PLM), sample size (2000, 1000, and 600), focal/reference group ratio (1/1 and 1/3), DIF magnitude (0.3, 0.5, and 0.7), and the proportion of items with DIF (20% and 40%). Analyses were conducted on each of the 100 repetitions of 36 conditions generated for each dataset. The data were generated for a 20-item test. In analyses conducted on the real dataset, sample size (900 and 450) and focal/reference group ratio (350/550, 450/450, 225/225, and 150/300) conditions were considered. DIF analyses were performed using M-ST and M-PA methods. Type I error rates and power ratios were calculated to examine the performance of the methods in analyses conducted with simulation data. Factorial ANOVA analyses were applied to examine how Type I error rates and power ratios changed according to the conditions considered. It was found that Type I error rates and power ratios showed significant differences depending on sample size, focal/reference group ratio, the proportion of items with DIF, and DIF magnitude. M-PA method exhibited superior performance to M-ST method in terms of Type I error rates. In analyses conducted with real data, it was observed that the number of items identified with DIF using the M-PA method was lower than the number identified using the M-ST method for all conditions.

Keywords: standard MIMIC, MIMIC with a pure anchor, differential item functioning

Teşekkür

Lisansüstü eğitimimin başından bu yana bana rehberlik eden, bilgi ve tecrübeleriyle beni yönlendiren kıymetli tez danışmanım Prof. Dr. Hülya KELECİOĞLU'na sonsuz teşekkür ederim. Tez sürecim boyunca göstermiş olduğu disiplin, sabır ve anlayış için kendisine minnettarım.

Tez izleme komitelerimde yer alan, değerli fikir ve yönlendirmeleriyle tezimin son halini almasına katkı sağlayan Prof. Dr. Burcu ATAR ve Doç. Dr. Celal Deha DOĞAN'a en içten teşekkürlerimi sunarım.

Önerileriyle tezime katkı sağlayan jüri üyesi hocalarım Doç. Dr. Kübra ATALAY KABASAKAL ve Doç. Dr. Sümeyra SOYSAL'a çok teşekkür ederim.

Doktora eğitimim boyunca TÜBİTAK BİDEB 2211-A Genel Yurt İçi Doktora Burs Programı kapsamında burs desteği sağlayan TÜBİTAK'a teşekkür ederim.

Koşulsuz sevgi ve inançlarıyla her zaman yanımda olan, beni her konuda destekleyen annem Eda TAMCI'ya, babam Aytekin TAMCI'ya ve abim Ertan TAMCI'ya sonsuz teşekkür ederim.

Sadece varlığıyla bile beni çok mutlu eden, huzur ve neşe kaynağım kedim Mone'ye çok teşekkür ederim.

Bu süreçte sevgisi, sabrı ve anlayışıyla her zaman yanımda olan canım yol arkadaşım Emre TUNER'e sonsuz teşekkür ederim.

İçindekiler

Kabul ve Onay.....	ii
Öz.....	iii
Abstract.....	iv
Teşekkür.....	v
Tablolar Dizini.....	viii
Şekiller Dizini.....	x
Simgeler ve Kısaltmalar Dizini.....	xii
Bölüm 1 Giriş.....	1
Problem Durumu.....	1
Araştırmanın Amacı ve Önemi.....	3
Araştırma Problemi.....	5
Sınırlılıklar.....	6
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	7
Değişen Madde Fonksiyonu.....	7
Değişen Madde Fonksiyonu Belirleme Yöntemleri.....	10
İlgili Araştırmalar.....	19
Bölüm 3 Yöntem.....	23
Araştırmanın Türü.....	23
Araştırma Verileri.....	23
Çalışma Grubu.....	27
Verilerin Analizi.....	29
Bölüm 4 Bulgular, Yorumlar ve Tartışma.....	32
Alt Problem 1'e İlişkin Bulgular.....	39
Alt Problem 2'ye İlişkin Bulgular.....	45
Alt Problem 3'e İlişkin Bulgular.....	50
Alt Problem 4'e İlişkin Bulgular.....	56

Alt Problem 5'e İlişkin Bulgular.....	61
Tartışma.....	63
Bölüm 5 Sonuç ve Öneriler.....	67
Sonuçlar	67
Öneriler	70
Kaynaklar	73
EK-A: Simülasyon Verilerine Ait Örnek Paralel Analiz Sonuçları.....	82
EK-B: Gerçek Veriye Ait Paralel Analiz Sonucu	85
EK-C: M-ST Analizi için R Programı Kodu	86
EK-Ç: Alt Problem 1'e İlişkin Faktöriyel ANOVA Tablosu	88
EK-D: Alt Problem 2'ye İlişkin Faktöriyel ANOVA Tablosu	89
EK-E: Alt Problem 3'e İlişkin Faktöriyel ANOVA Tablosu	90
EK-F: Alt Problem 4'e İlişkin Faktöriyel ANOVA Tablosu	91
EK-G: Çalışma Grubu 1 için DMF Analiz Sonuçları.....	92
EK-H: Çalışma Grubu 2 için DMF Analiz Sonuçları.....	93
EK-I: Çalışma Grubu 3 için DMF Analiz Sonuçları	94
EK-İ: Çalışma Grubu 4 için DMF Analiz Sonuçları	95
EK-J: Araştırma Etik Komisyon İzin Muafiyeti Formu	96
EK-K: Etik Beyanı.....	97
EK-L: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu.....	98
EK-M: Thesis/Dissertation Originality Report	99
EK-N: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	100

Tablolar Dizini

Tablo 1	<i>i</i> Maddesi ve <i>j</i> Yetenek Düzeyi için MH Matrisi	10
Tablo 2	MH ile Elde Edilen ΔMH_i Değerleri İçin Kabul Edilen DMF Düzeyleri....	11
Tablo 3	LR ile Elde Edilen ΔR_2 Değerleri İçin Kabul Edilen DMF Düzeyleri	12
Tablo 4	SIBTEST ile Elde Edilen β Değerleri için Kabul Edilen DMF Düzeyleri..	14
Tablo 5	Simülasyon Koşulları	24
Tablo 6	Veri Üretiminde Kullanılan Madde Parametreleri	26
Tablo 7	Çalışma Grubuna Dâhil Edilen Ülkelerle İlgili Bilgiler	28
Tablo 8	Çalışma Gruplarının Frekans Dağılımı.....	28
Tablo 9	DMF Bağlamında I. Tip Hata ve Güç	31
Tablo 10	2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre I. Tip Hata Oranları.....	32
Tablo 11	3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre I. Tip Hata Oranları.....	33
Tablo 12	2PLM için M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre Güç Oranları	35
Tablo 13	3PLM için M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre Güç Oranları	37
Tablo 14	2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları	39
Tablo 15	2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları.....	41
Tablo 16	2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları	42
Tablo 17	2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları	43
Tablo 18	3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları	45
Tablo 19	3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları.....	46
Tablo 20	3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre I. Tip Hata Oranlarına İlişkin Sonuçlar	47

Tablo 21	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin Sonuçlar</i>	49
Tablo 22	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranlarına İlişkin ANOVA Sonuçları</i>	50
Tablo 23	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre Güç Oranlarına İlişkin ANOVA Sonuçları</i>	52
Tablo 24	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranlarına İlişkin ANOVA Sonuçları</i>	53
Tablo 25	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranlarına İlişkin ANOVA Sonuçları</i>	54
Tablo 26	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranlarına İlişkin Sonuçlar</i>	56
Tablo 27	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre Güç Oranlarına İlişkin Sonuçlar</i>	57
Tablo 28	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranlarına İlişkin ANOVA Sonuçları</i>	58
Tablo 29	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranlarına İlişkin Sonuçlar</i>	60
Tablo 30	<i>M-ST ve M-AÇM Yöntemleri için Farklı Koşullara Göre Elde Edilen DMF'li Madde Sayıları ve Numaraları</i>	61

Şekiller Dizini

Şekil 1	<i>Tek Biçimli DMF Gösteren Madde Grafiği</i>	8
Şekil 2	<i>Tek Biçimli Olmayan DMF Gösteren Madde Grafiği</i>	9
Şekil 3	<i>(a) M-ST Yöntemi ve (b) M-AÇM Yöntemi</i>	17
Şekil 4	<i>2 PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranları</i>	40
Şekil 5	<i>2 PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre I. Tip Hata Oranları</i>	41
Şekil 6	<i>2 PLM İçin M-ST ve M-AÇM Yöntemlerinin DMF'li Madde Oranlarına Göre I. Tip Hata Oranları</i>	43
Şekil 7	<i>2 PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranları</i>	44
Şekil 8	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranları</i>	45
Şekil 9	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı O/R Oranlarına Göre I. Tip Hata Oranları</i>	47
Şekil 10	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre I. Tip Hata Oranları</i>	48
Şekil 11	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranları</i>	49
Şekil 12	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranları</i>	51
Şekil 13	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı O/R Oranlarına Göre Güç Oranları</i>	52
Şekil 14	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranları</i>	54
Şekil 15	<i>2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranları</i>	55
Şekil 16	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranları</i>	57
Şekil 17	<i>3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı O/R Oranlarına Göre Güç Oranları</i>	58

Şekil 18 3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranları.....	59
Şekil 19 3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranları.....	60



Simgeler ve Kısaltmalar Dizini

AİÖÜ: Anadili İngilizce Olan Ülkeler

AİFOÜ: Anadili İngilizceden Farklı Olan Ülkeler

DMF: Değişen Madde Fonksiyonu

LR: Lojistik Regresyon Modeli

M-AÇM: Arınık Çapa Maddeli MIMIC Modeli

M-ST: Standart MIMIC Modeli

MH: Mantel-Haenszel Yöntemi

MIMIC: Çoklu Göstergeler Çoklu Nedenler

MTK: Madde Tepki Kuramı

O/R: Odak Grup/Referans Grup

PISA: Uluslararası Öğrenci Değerlendirme Programı

2PLM: 2 Parametrelili Lojistik Model

3PLM: 3 Parametrelili Lojistik Model

Bölüm 1

Giriş

Bu bölümde problem durumu, araştırmanın amacı ve önemi, araştırma problemi ve sınırlılıklar başlıklarına yer verilmiştir.

Problem Durumu

Testlerin okullarda, kliniklerde, endüstride ve devlet işlerinde kullanımı oldukça yaygınlaşmıştır. Modern toplumların bütünlük bir parçası olan standartlaştırılmış testlerden elde edilen sonuçlar eğitimde kabul, yerleştirme ve ilerleme; kariyer gelişimi ve özel hizmetlerin sağlanmasıyla ilişkili kararların alınmasında sıklıkla kullanılmaktadır. Bu kararların kişisel, sosyal ve politik açılardan önemli sonuçları vardır. Bu nedenle bu kararların alınmasında kullanılan test sonuçları büyük önem taşımaktadır (Clauser ve ark., 1991, Clauser & Mazor, 1998).

Test geliştirme sürecindeki temel kaygılardan biri testlerden elde edilen sonuçların ve bu sonuçlara dayalı yapılan yorumların geçerli olmasıdır (Bachman, 1990). Bu nedenle geçerlik açısından sorun yaratan madde yanlılığı, test geliştirme süreçlerinde dikkate alınıp incelenmesi gereken önemli bir husustur. Madde yanlılığı, testi alan bir grup öğrencinin maddeyi doğru cevaplama olasılığının, testin amacı dışındaki bir durumdan ya da madde özelliğinden dolayı testi alan diğer bir grup öğrenciye göre daha az olması durumunda ortaya çıkmaktadır (Zumbo, 1999). Aynı yetenek düzeyindeki kişilerin grup üyeliğine bakılmaksızın bir maddeyi doğru cevaplama olasılıkları aynıysa o maddede yanlılık söz konusu değildir (Woods, 2009). Bazı koşullar altında bir madde için sistematik olarak daha yüksek, bazı koşullar altında ise daha düşük puanlar gözlemlendiğinde ortaya çıkan puan farklılıkları ölçülen nitelikteki gerçek farklılıkları doğru şekilde yansıtmamaktadır. Bu durumda testi alanlardan belirli özelliklere sahip olanlar diğerlerine göre daha avantajlıdır. Bu nedenle test puanlarında ortaya çıkan farklılıklar, yetenekle ilgili farklılıkları doğru şekilde temsil etmemektedir (Bauer, 2017). Madde yanlılığı, hem ölçme sürecinde sistematik

hataya sebep olmasından hem de geçerlik açısından ciddi bir tehdit oluşturmamasından dolayı dikkate alınması gereken önemli bir kavramdır (Clauser & Mazor, 1998; Osterlind, 1983).

Madde yanlılığını incelemek için yapılan çalışmalar genellikle değişen madde fonksiyonu (DMF) belirleme olarak adlandırılan istatistiksel bir süreçle başlamaktadır (Holland & Thayer, 1998; Holland & Wainer, 1993). DMF, testi alan farklı gruplardaki öğrencilerin maddenin ölçmeyi amaçladığı temel özellik açısından eşleştirilmesinden sonra madde için farklı başarı olasılıkları göstermeleri durumunda ortaya çıkmaktadır (Zumbo, 1999). Bir başka deyişle DMF ifadesi, bir maddenin aynı yetenek düzeyindeki bireyler için farklı gruplarda farklı istatistiksel özellikler göstermesidir (Holland & Wainer, 1993). Maddenin DMF göstermesi maddenin yanlı olduğunu söyleyebilmek için gerekli bir koşuldur ancak tek başına yeterli değildir. Maddenin doğru cevaplanma olasılığının gruplar arasında farklılaşmasının yanlılıktan mı yoksa farklı gruplarda bulunan cevaplayıcıların maddenin ölçtüğü özellik bakımından gerçekten farklı olmalarından mı ortaya çıktığına dikkat edilmelidir (Clauser & Mazor, 1998; French & Miller, 1996).

Bir madde testi alanların bir grubunda diğer gruptan farklı işleyebilmektedir. DMF analizlerinde üzerinde çalışılan grup odak grup, odak grubun performansını karşılaştırmada standart olarak ele alınacak grup referans grup olarak adlandırılmaktadır (Holland & Thayer, 1998). Testteki maddeler bu gruplar arasında farklı işlerse, bu gruplar için elde edilen puanlar karşılaştırılmaz.

Test maddelerindeki DMF'nin tespit edilebilmesi amacıyla pek çok yöntem geliştirilmiştir. Bu yöntemler arasında Mantel-Haenszel yöntemi (MH) (Holland & Thayer, 1988), lojistik regresyon yöntemi (LR) (Swaminathan & Rogers, 1990), SIBTEST yöntemi (Shealy & Stout, 1993), Lord'un ki-kare testi yöntemi (Lord, 1980), Raju'nun alan ölçümleri yöntemi (Raju, 1988), olabilirlik oranı testi (Thissen ve ark., 1988; Wang & Yeh, 2003) ve çoklu göstergeler çoklu nedenler yöntemi (MIMIC) (Finch, 2005; Oort, 1998) gösterilebilir.

Testin adillik ve ölçme değişmezliği gibi özelliklerini değerlendirmek için MIMIC yönteminin kullanılması konusunda artan bir ilgi vardır. Bu da MIMIC yönteminin DMF'yi tespit etmek amacıyla geliştirilmiş çeşitli uygulamalarıyla sonuçlanmıştır (Lee ve ark., 2017).

Finch (2005), standart MIMIC (M-ST) yöntemini diğer DMF belirleme yöntemleriyle karşılaştırdığında çoğu koşulda diğer DMF belirleme yöntemlerinden daha iyi çalıştığını tespit etmiştir. DMF çalışmalarında odak ve referans grupları aynı metriğe yerleştirebilmek için eşleştirme kriteri kullanılmaktadır (Wainer ve ark., 1991). Fakat M-ST yönteminin de eşleştirme kriterinde yer alan DMF'li maddelerden etkilendiği görülmüştür. Eşleştirme kriterindeki DMF kirliliği arttıkça M-ST yönteminin performansı kötüleşmektedir (Finch, 2005; Navas-Ara & Gomez-Benito, 2002). Uygun bir DMF değerlendirmesinin yapılabilmesi için eşleştirme kriterinin saf/arınık olması gerekmektedir (Lord, 1980).

DMF kirliliğinin DMF analizleri üzerindeki olumsuz etkisini ortadan kaldırmanın uygun bir yolu, testteki DMF'li olmayan maddelerin oluşturduğu küçük bir madde setini belirleyip bu madde setini arınık bir eşleştirme kriteri (çapa) olarak kullanmak ve DMF analizlerini buna göre yapmaktır (Wang & Shih, 2010). Shih ve Wang (2009) bu amaçla, eşleştirme kriteri olarak arınık maddelerin kullanıldığı, arınık çapa maddeli MIMIC (M-AÇM) yöntemini geliştirmişlerdir.

DMF analizlerinin sağladığı bilgiler test geliştirme ve değerlendirme süreçlerinin kritik bir bileşenidir. O nedenle DMF analizinde kullanılacak yöntemlerin mevcut koşullara uygun seçilmesi elde edilecek sonuçların doğruluğu açısından büyük önem taşımaktadır. Bu bağlamda farklı koşullar altında M-ST ve M-AÇM yöntemlerinin performanslarının incelenip birbiriyle karşılaştırılması araştırmacılara yol gösterici bilgiler sağlayacaktır.

Araştırmanın Amacı ve Önemi

Sınıflarda öğretmenler tarafından yapılan testlerden, ulusal veya uluslararası düzeyde gerçekleştirilen büyük ölçekli sınavlara kadar her türlü testin, öğrenciler, veliler, eğitimciler ve politikacılar için büyük bir önemi vardır. Testler; bireylerin, ailelerin ve

toplumun kişisel ve sosyal alanları için olduğu gibi, politikacılar için de eğitim politikalarını belirleme noktasında kılavuz niteliğinde çıktılar sunarlar. Bu nedenle, testlerden elde edilen sonuçların güvenilir ve geçerli olması hem bireylerin hem de ülkelerin alacakları kararların doğruluğunu önemli ölçüde etkiler.

Test sonuçlarının karar alma süreçlerindeki kritik önemi göz önünde bulundurulduğunda, testlerin katılımcıların yeteneklerini adil bir şekilde değerlendirebilmesi gereklidir. Kabul, yerleştirme, mezuniyet, kariyer gelişimi ve işe alma gibi önemli kararların temellendirildiği sınavların adil olabilmesi için kullanılan test maddelerinin yanlılığı incelenmelidir.

Madde yanlılığını değerlendirme süreci maddelerin DMF açısından incelenmesiyle başlamaktadır. DMF'li maddelerin belirlenebilmesi için pek çok yöntem geliştirilmiştir. Bununla birlikte, testin özelliklerine ve koşullara uygun DMF yönteminin seçilmesi ve analizlerin bu yöntemle gerçekleştirilmesi oldukça önemlidir.

MIMIC yöntemleri birden fazla grup değişkeniyle eş zamanlı DMF analizlerinin yapılmasına ve grup üyeliği değişkeninin kategorik ya da sürekli bir değişken olmasına imkân sağlar. Bu nedenlerden MIMIC yöntemleri pek çok geleneksel DMF belirleme yöntemine göre daha esnektir (Wang ve ark., 2009). Bu esnekliklerden dolayı DMF analizleri için MIMIC yöntemlerinin kullanılmasıyla ilgili artan bir ilgi söz konusudur. Bununla birlikte MIMIC yöntemlerinden olan M-ST ve M-AÇM yöntemleri diğer DMF belirleme yöntemlerinden daha sonra geliştirildiği için bu iki yöntemin birlikte incelendiği çalışma sayısı yeterli değildir. Bu iki yöntemi ele alıp inceleyen bu çalışmanın literatüre önemli bir katkı sağlayacağı düşünülmektedir.

Çalışmanın amacı, M-ST ve M-AÇM yöntemleriyle gerçek veri ve yapay veri üzerinde farklı koşullarda DMF analizi yaparak bu iki yöntemin performanslarını inceleyip karşılaştırmaktır. Bu amaç doğrultusunda farklı koşullar altında bu iki yöntem için elde edilen I. tip hata ve güç değerleri belirlenip karşılaştırılmış ayrıca PISA 2012 matematik okuryazarlığı testindeki maddeler DMF açısından M-ST ve M-AÇM yöntemleriyle

incelenmiştir. Bu iki yöntemin performansının karşılaştırıldığı koşullar; MTK modeli, örneklem büyüklüğü, odak/referans grup oranı, testteki DMF'li madde oranı ve DMF büyüklüğüdür.

Araştırma Problemi

2PLM ve 3PLM'ye göre üretilen veriler için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata ve güç değerleri örneklem büyüklüğüne, odak/referans grubu oranlarına, testteki DMF'li madde oranına ve DMF büyüklüğüne göre nasıl değişmektedir? PISA 2012 matematik testinde, odak ve referans grup için farklı örneklem büyüklükleri ve farklı oranlar söz konusu olduğunda M-ST ve M-AÇM yöntemlerine göre belirlenen ortak DMF'li maddeler ve bunların sayısı nedir?

Alt Problemler

1. 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri,
 - a. Örneklem büyüklüğüne (2000, 1000 ve 600),
 - b. Odak/referans grup oranlarına (1/1 ve 1/3),
 - c. DMF'li madde oranına (%20 ve %40),
 - d. DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?
2. 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri,
 - a. Örneklem büyüklüğüne (2000, 1000 ve 600),
 - b. Odak/referans grup oranlarına (1/1 ve 1/3),
 - c. DMF'li madde oranına (%20 ve %40),
 - d. DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?

3. 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri,
 - a. Örneklem büyüklüğüne (2000, 1000 ve 600),
 - b. Odak/referans grup oranlarına (1/1 ve 1/3),
 - c. DMF'li madde oranına (%20 ve %40),
 - d. DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?
4. 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri,
 - a. Örneklem büyüklüğüne (2000, 1000 ve 600),
 - b. Odak/referans grup oranlarına (1/1 ve 1/3),
 - c. DMF'li madde oranına (%20 ve %40),
 - d. DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?
5. PISA 2012 matematik testinde M-ST ve M-AÇM yöntemlerine göre belirlenen ortak DMF'li maddeler ve bunların sayısı, örneklem büyüklüğüne (900 ve 450) ve odak/referans grup oranlarına (350/550, 450/450, 225/225 ve 150/300) göre nasıl değişmektedir?

Sınırlılıklar

1. Araştırma iki kategorili maddelerle sınırlıdır.
2. Araştırmada kullanılan gerçek veriler PISA 2012'de Birleşik Krallık, ABD, İzlanda, Çekya, Slovakya, Fransa ve Danimarka'da 4. kitapçığı çözen öğrencilerle sınırlıdır.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Bu bölümde değişen madde fonksiyonuna ilişkin kuramsal yapı aktarılmış, değişen madde fonksiyonu belirleme yöntemlerine, MIMIC modellerine ve literatürdeki ilgili araştırmalara yer verilmiştir.

Değişen Madde Fonksiyonu

DMF, ölçülen özellik açısından eşit yetenek düzeyine sahip olan ancak diğer özellikler açısından farklı alt gruplarda yer alan bireylerin maddeye verdikleri yanıtların farklılaşması durumunda ortaya çıkmaktadır. DMF, bir maddenin ölçülen özellikten bağımsız olarak testi alan farklı gruplarda farklı istatistiksel özelliklere sahip olması durumudur (Holland & Wainer, 1993; Woods, 2009; Zumbo, 1999). Gruplar arasında, ölçülen özellik açısından var olan yetenek farklılıkları kontrol altına alındıktan sonra maddenin bu gruplar için farklı istatistiksel özelliklere sahip olması o maddenin DMF'li olduğuna işaret etmektedir (Angoff, 1993).

Milsap ve Everson (1993), DMF'li olmayan bir maddeyi aşağıdaki denklemle tanımlamışlardır:

$$\{\text{Bir kişinin yeteneği ve ait olduğu grup dikkate alındığında bir maddeyi doğru cevaplama olasılığı}\} = \{\text{Bir kişinin yeteneği dikkate alındığında bir maddeyi doğru cevaplama olasılığı}\}$$

DMF'li maddelerin madde fonksiyonları, aynı yetenek seviyesine sahip alt gruplarda farklılaşmaktadır. Bu farklılaşmanın alt gruplar arasındaki gerçek farklılığı yansıtan madde etkisinden mi yoksa madde yanlılığından mı kaynaklandığını belirlemek önemlidir. Madde etkisi, farklı alt gruplarda yer alan cevaplayıcıların maddenin ölçtüğü temel yetenek açısından gerçekten farklı düzeyde olmalarından kaynaklı olarak maddeyi doğru cevaplama (veya onaylama) olasılıklarının farklı olması durumunda ortaya çıkmaktadır. Ortaya çıkan bu fark, alt gruplardan birinin sahip olduğu bilgi ve deneyimle ilişkilendirilebilir. Madde etkisi

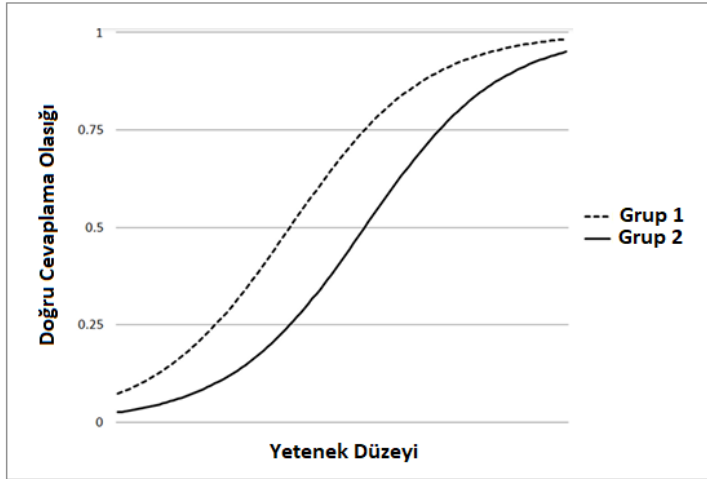
ve madde yanlılığı kavramları, grup farklılıklarının sebebinin testin ölçülmesi hedeflenen özelliklerinden mi yoksa ilgisiz durumlardan mı kaynaklandığına göre birbirinden farklılaşmaktadır. DMF, iki grup üyesinin başarı olasılıklarının farklı olup olmadığını belirlemeden önce bu iki grup üyesinin ölçülen yetenek açısından eşleştirilmesini gerektirir. (Gök ve ark., 2010; Zumbo, 1999).

DMF analizlerinde cevaplayıcılar iki gruba ayrılır. Bu gruplardan azınlık olan veya ilgili test açısından potansiyel olarak dezavantajlı olduğu düşünülen grup odak grup olarak adlandırılmaktadır. İlgili test açısından avantajlı olduğu düşünülen grup ise referans grup olarak adlandırılır (Karami & Nodoushan, 2011; Osterlind, 1983). Test geliştiriciler, çoğu DMF analizinde maddenin özellikle odak grup için olan zorluk seviyesinden endişe duymaktadır. DMF istatistikleri genellikle odak grubun madde üzerindeki performansını, kritere göre odak grupla eşleşen referans grubun performansı ile karşılaştırmaktadır. Bu nedenle DMF analizindeki kritere eşleştirme kriteri adı verilmektedir (Livingston, 2006).

Mellenbergh'e (1982) göre DMF, tek biçimli DMF ve tek biçimli olmayan DMF olmak üzere iki farklı şekilde gözlenmektedir. Tek biçimli DMF, tüm yetenek düzeyleri boyunca maddenin hep aynı gruba avantaj yarattığı, bir başka deyişle grup üyeliğiyle madde puanları arasında etkileşim olmadığı durumda ortaya çıkmaktadır (Ellis & Raju, 2003; Li & Stout, 1996). Tek biçimli DMF söz konusu olduğunda madde güçlük parametresi gruplar arasında farklılaşmaktadır (Zimbra, 2018). Bu durumda Şekil 1'de verildiği gibi iki gruba ait madde karakteristik eğrileri birbirine paralel olur ve herhangi bir noktada kesişmez.

Şekil 1

Tek Biçimli DMF Gösteren Madde Grafiği



Tek biçimli olmayan DMF, yetenek düzeyleri boyunca maddenin bazı düzeylerde bir gruba bazı düzeylerde ise diğer gruba avantaj yarattığı, bir başka deyişle grup üyeliğiyle madde puanları arasında etkileşimin olduğu durumda ortaya çıkmaktadır (Ellis & Raju, 2003; Li & Stout, 1996). Tek biçimli olmayan DMF söz konusu olduğunda madde ayırt edicilik parametreleri gruplar arasında farklılaşmaktadır (Zimbra, 2018). Bu durumda Şekil 2’de verildiği gibi iki gruba ait madde karakteristik eğrileri birbirine paralel değildir ve bazı yetenek düzeyi noktalarında kesişmektedir.

Şekil 2

Tek Biçimli Olmayan DMF Gösteren Madde Grafiği



Değişen Madde Fonksiyonu Belirleme Yöntemleri

DMF analizlerinde kullanılabilecek çok sayıda yöntem vardır. Mantel-Haenszel (Holland & Thayer, 1988), lojistik regresyon (Swaminathan & Rogers, 1990), SIBTEST (Shealy & Stout, 1993) yöntemleri klasik test kuramına dayalı yöntemler, Lord'un ki-kare testi (Lord, 1980), Raju'nun alan ölçümleri (Raju, 1988), olabilirlik oranı testi (Thissen ve ark., 1988; Wang & Yeh, 2003) yöntemleri madde tepki kuramına (MTK) dayalı yöntemler ve MIMIC yöntemi (Finch, 2005; Oort, 1998) yapısal eşitlik modeline dayalı yöntemler olarak örnek verilebilir.

Bu yöntemlerden sıklıkla kullanılan MH, LR, SIBTEST yöntemleri ve bu çalışmada kullanılan MIMIC yöntemlerinden olan M-ST ve M-AÇM yöntemleri aşağıda açıklanmıştır.

MH Yöntemi

Mantel ve Haenszel (1959) tarafından geliştirilmiş olan MH yöntemi, Holland ve Thayer (1988) tarafından DMF çalışmalarında kullanılmaya başlamıştır (Camilli & Shepard, 1994). MH yöntemi yalnızca tek biçimli DMF'yi belirlemek için kullanılmaktadır. MH yönteminde, aynı yetenek düzeyinde olduğu belirlenen bireylerin soruları doğru ya da yanlış cevaplama olasılıkları arasındaki farklar göz önünde bulundurularak incelemeler yapılmaktadır. Bu yöntem ölçülen özellik açısından var olan yetenek farklılıkları kontrol altına alındıktan sonra test performansı ve grup üyeliği arasında anlamlı bir ilişki yoktur hipotezini X^2 testiyle test etmektedir. Toplam test puanlarından yola çıkılarak aynı yetenek seviyesinde olduğu belirlenen odak ve referans grubundaki bireyler maddeyi doğru cevaplayabilme olasılıkları açısından karşılaştırılmaktadır. Analize başlarken her bir madde ve her bir yetenek düzeyi için odak ve referans grubundaki kişilerin maddeye verdikleri cevaplardan oluşan veri, 2x2'lik matris şeklinde düzenlenir (Gierl ve ark., 1999; Narayanan & Swaminathan, 1996).

Tablo 1

i Maddesi ve j Yetenek Düzeyi için MH Matrisi

Grup	Doğru	Yanlış	Toplam
Referans	A_j	B_j	n_{Rj}
Odak	C_j	D_j	n_{0j}
Toplam	m_{1j}	m_{0j}	T_j

Tablo 1'deki değerler kullanılarak maddenin olasılık oranı aşağıdaki formülle hesaplanmaktadır (Camilli & Shepard, 1994).

$$\alpha_{MH} = \frac{\sum_{j=1}^S A_j D_j / T_j}{\sum_{j=1}^S B_j C_j / T_j}$$

α_{MH} , $(0, \infty)$ aralığında değer alır. α_{MH} değerinin yorumlanmasını kolaylaştırmak amacıyla logaritmik bir dönüşüm uygulanır ve ΔMH_i elde edilir. ΔMH_i değeri ile de DMF'nin düzeyi belirlenir (Kelecioğlu ve ark., 2014). Bu logaritmik dönüşüm aşağıdaki formülle hesaplanmaktadır.

$$\Delta MH_i = -2.35 \ln(\alpha_{MH})$$

Zieky'e (1993) göre ΔMH_i değerlerinin işaret ettiği DMF düzeyleri Tablo 2'deki gibidir.

Tablo 2

MH ile Elde Edilen ΔMH_i Değerleri İçin Kabul Edilen DMF Düzeyleri

DMF Düzeyi	ΔMH_i
A Düzeyi (Göz ardı edilebilir)	$\Delta MH_i \leq 1$
B Düzeyi (Orta Düzey)	$1 < \Delta MH_i < 1.5$
C Düzeyi (Yüksek Düzey)	$\Delta MH_i \geq 1.5$

LR Yöntemi

LR yöntemi, toplam puanı, grup değişkenini ve grup değişkeniyle toplam puanın etkileşimini bağımsız değişkenler; madde yanıtlarını $(0,1)$ bağımlı değişken olarak kabul

eden bir yöntemdir. LR yöntemiyle hem tek biçimli DMF hem de tek biçimli olmayan DMF tespit edilebilmektedir. Bu yöntemle yapılan DMF incelemelerinde grup değişkeniyle toplam puanın arasındaki etkileşim tek biçimli olmayan DMF'ye, madde yanıtlarıyla toplam puan arasındaki etkileşim ise tek biçimli DMF'ye işaret etmektedir (Swaminathan & Rogers, 1990; Zumbo, 1999).

LR ile DMF'nin varlığı modele sırasıyla grup değişkeni, toplam puan ve grup değişkeni arasındaki etkileşimin eklenmesiyle test edilmektedir. Daha sonra sırasıyla modele etkilenen değişkenler test edilerek tek biçimli DMF ve tek biçimli olmayan DMF'nin varlığını incelemek için ki-kare testi kullanılmaktadır (Gierl ve ark., 2000).

$$\text{Model 1: } z = \beta_0 + \beta_1 X$$

$$\text{Model 2: } z = \beta_0 + \beta_1 X + \beta_2 G$$

$$\text{Model 3: } z = \beta_0 + \beta_1 X + \beta_2 G + \beta_3 XG$$

X'in toplam puanı temsil ettiği Model 1 temel modeldir. Temel modele, grup değişkeni (G) eklenerek elde edilen Model 2 ile tek biçimli DMF'nin varlığı incelenmektedir. Model 2'ye toplam puan ile grup değişkeni arasındaki etkileşim değişkeninin eklenmesiyle elde edilen Model 3 ile tek biçimli olmayan DMF incelenmektedir (Gierl ve ark., 2000). Her bir aşamada X^2 ve R^2 değerleri hesaplanmaktadır. Etki büyüklüğünü tespit etmek amacıyla modeller arasındaki R^2 farkları hesaplanarak ΔR^2 elde edilmekte ve ΔR^2 değerinin yorumlanmasıyla DMF'nin düzeyi belirlenmektedir (Zumbo, 1999).

Literatürde ΔR^2 değerinin yorumlanmasında kullanılan farklı ölçütler bulunmaktadır. Zumbo ve Thomas (1996), Jodoin ve Gierl'e (2001) göre ΔR^2 değerlerinin işaret ettiği DMF düzeyleri Tablo 3'teki gibidir.

Tablo 3

LR ile Elde Edilen ΔR^2 Değerleri İçin Kabul Edilen DMF Düzeyleri

DMF Düzeyi	Zumbo & Thomas (1996)	Jodoin & Gierl (2001)
------------	-----------------------	-----------------------

A Düzeyi (Göz ardı edilebilir)	$\Delta R^2 < .13$	$\Delta R^2 < .035$
B Düzeyi (Orta Düzey)	$.13 \leq \Delta R^2 \leq .26$	$.035 \leq \Delta R^2 \leq .07$
C Düzeyi (Yüksek Düzey)	$\Delta R^2 > .26$	$\Delta R^2 > .07$

SIBTEST Yöntemi

SIBTEST, bir maddedeki DMF miktarını kestirip ardından bu miktarın 0'dan farklı olup olmadığını istatistiksel olarak test eden bir yöntemdir. Diğer DMF belirleme yöntemlerinde olduğu gibi SIBTEST ile de odak ve referans grubu üyelerinin madde performanslarındaki farklılıkları yetenek düzeyi koşuluna bağlı olarak değerlendirilmektedir (Bolt, 2000; Shealy & Stout, 1993). Bu yöntemin I. tip hatayı kontrol edebilmek adına regresyon dayalı bir düzeltme kullanması en büyük avantajlarından biridir (Bolt, 2000; Clauser & Mazor, 1998).

SIBTEST analizlerinde test maddeleri iki alt teste ayrılmaktadır. Bunlardan ilki testin ölçmeyi hedeflediği yapıyı ölçtüğü varsayılan maddelerden oluşan geçerli bir alt test; diğeri ise DMF'li olup olmadığı test edilecek maddelerden oluşan şüpheli bir alt testtir. Geçerli alt testin belirlenebilmesi için öncelikle tüm maddelerin eşleştirme kriteri olarak kullanıldığı bir DMF analizi yapılır. DMF'li olduğu belirlenen maddelerin eşleştirme kriterinden çıkarılmasıyla devam eden süreç DMF göstermeyen maddelerin oluşturduğu geçerli bir alt test elde edilene ve böylelikle nihai eşleştirme kriteri belirlenene kadar devam eder (Bolt, 2000; Clauser & Mazor, 1998).

θ geçerli alt testin ölçmeyi hedeflediği yetenek olmak üzere, SIBTEST yönteminde aşağıdaki denklemler yoluyla β değeri bulunmakta ve DMF düzeyi belirlenmektedir (Bolt, 2000).

$$B(\theta) = P(\theta, R) - P(\theta, F)$$

$P(\theta, R)$: Maddenin referans grubundakiler tarafından doğru cevaplanma olasılığı

$P(\theta, F)$: Maddenin odak grubundakiler tarafından doğru cevaplanma olasılığı

$$\beta = \int B(\theta)f_F(\theta)d\theta$$

$f_F(\theta)$: θ için odak grubun yoğunluk fonksiyonu

β değerlerinin bulunduğu aralığa göre işaret ettiği DMF düzeyleri Tablo 4'te verilmiştir (Roussos & Stout, 1996).

Tablo 4

SIBTEST ile Elde Edilen β Değerleri için Kabul Edilen DMF Düzeyleri

DMF Düzeyi	β
A Düzeyi (Göz ardı edilebilir)	$ \beta < .059$
B Düzeyi (Orta Düzey)	$.059 \leq \beta < .088$
C Düzeyi (Yüksek Düzey)	$ \beta \geq .088$

MIMIC Yöntemleri

MIMIC modeli, Jöreskog ve Goldberger (1975) tarafından yapısal eşitlik modelinin özel bir uygulaması olarak geliştirilmiştir. MIMIC modeli, gizil bir değişkeni bir veya birden fazla gözlenen değişkene bağlayan yolların/ilişkilerin (path) kestirilmesine olanak sağlamaktadır. MIMIC modelleri, gizil grup ortalamalarını karşılaştırma yöntemi (Thompson & Green, 2006), gizil ve gözlenen değişkenler için regresyon analizi yapılmasına yönelik bir yöntem (Muthén, 1988) ve DMF belirleme yöntemi (Finch, 2005) olmak üzere farklı bağlamlarda çalışılmıştır. Farklı araştırma konularında kullanılabilme esnekliği nedeniyle son dönemlerde ilgi çeken bir yöntem olmuştur.

Camilli ve Shepard (1994), ikincil bir faktör açısından grup farklılıklarını incelemesi nedeniyle doğrulayıcı faktör analizinin DMF incelemelerinde kullanılabileceğini belirtmiştir. Doğrulayıcı faktör analizinin DMF incelemesinde kullanılmasına yönelik MIMIC modellerinin kullanımıyla ilgili yaklaşımlar geliştirilmiştir (Finch, 2005; Oort, 1996, 1998).

MIMIC modeli gizil bir değişkendeki çoklu grup farklılıklarını inceleyebildiğinden dolayı bu modelin geçerlik çalışmalarında kullanılması önemlidir. Doğrulayıcı faktör analizi

yoluyla bir veri kümesine teorik bir model uydurarak bir testin yapı geçerliği ölçmede, gizil faktör ortalamalarının gruplar arasında farklılık gösterip göstermediğini belirlemede ve gizil faktör ölçümlerini DMF açısından incelemeye kullanılabilmektedir (Muthén, 1988).

Standart doğrulayıcı faktör analizi modeli, gizil faktörleri etkileyen dışsal değişkenlerin eklenmesiyle MIMIC modeline genişletilir ve ilgilenilen tüm popülasyonların kombinasyonlarını içeren bir veri seti oluşturulur. Model, z değişkenlerinin (kovaryantların) sebep olduğu gözlemlenmeyen gizil faktörleri içerir ve y değişkenleriyle gösterilir. Modele ait denklemler aşağıda verilmiştir.

$$y = \lambda\theta + \varepsilon$$

$$\theta = \gamma'z + \zeta$$

Denklemden $y' = (y_1, y_2, \dots, y_k)$ gizil değişken θ 'nın göstergeleri, $z = (z_1, z_2, \dots, z_k)$ gizil değişken θ 'nın nedenleridir. Bu denklemlerden yola çıkılarak aşağıdaki denkleme ulaşılır.

$$y = \lambda\gamma'z + \lambda\zeta + \varepsilon$$

$$= \Pi z + w$$

Böylece $\Pi = \lambda\gamma'$, $w = \lambda\zeta + \varepsilon$ ve $\text{Cov}(w) = \lambda\lambda'\psi + \Theta_\varepsilon$ olur. Burada $\psi = \text{Var}(\zeta)$ 'dir ve Θ_ε , ε 'nin diagonal kovaryans matrisidir (Jöreskog & Sörbom, 2002).

MIMIC modeli DMF bağlamında ele alındığında aşağıdaki denklem tanımlanmıştır (Finch, 2005; Wang ve ark., 2009):

$$y_i^* = \lambda_i\theta + \beta_i z + \varepsilon_i$$

y_i^* , i maddesi için gizil cevap değişkenini; λ_i , i değişkeni için faktör yükünü; θ gizil özelliği; β_i cevapla grup değişkenini ilişkilendiren eğimi; z, DMF'yle ilişkili olduğu düşünülen grup üyeliği değişkenini; ε_i tesadüfi hatayı göstermektedir. β_i , z'nin y_i^* 'ye etkisini ifade etmektedir. $\beta_i=0$ olduğunda i maddesinin z grup üyeliği değişkeni boyunca homojen olduğu ve maddenin DMF'li olmadığı yorumu yapılmaktadır. $\beta_i \neq 0$ olması ise θ 'nın yanı sıra z'den de y_i^* 'ye direkt bir etkinin var olduğunu ve maddenin DMF'li olduğunu göstermektedir.

MIMIC modelleriyle DMF değerlendirilmesi sürecindeki temel yaklaşım, bir grup değişkenine ilişkin hem doğrudan (direct effect) hem de dolaylı etkilerin (indirect effect) tahmin edilmesine dayanmaktadır. Burada doğrudan etki, gruplar arasında madde cevap olasılıkları açısından bir fark olup olmadığını göstermektedir. Dolaylı etki ise gizil değişkenin ortalamasının gruplar arasında farklı olup olmadığını göstermektedir. Buna göre yukarıdaki formül gruplar için gizil özellik ortalamasındaki değişiklikler kontrol altına alındıktan sonra tek biçimli DMF'nin incelenmesinde kullanılmaktadır (Finch, 2005).

MIMIC yöntemi kullanılarak kategorik ya da sürekli değişkenlerin yer aldığı ikiden fazla grupta DMF incelemesi yapılabilen ve bunların gizil değişkenle olan ilişkisine dair kapsamlı bir inceleme elde edilebilmektedir (Glöckner-Rist & Hoijtink, 2003; Muthén, 1988). Bu sebeple MIMIC yöntemi, pek çok geleneksel DMF yöntemine göre daha esnektir.

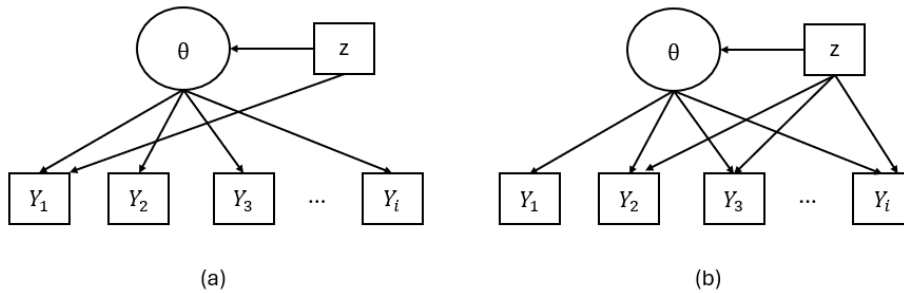
DMF incelemelerinde referans ve odak grupları aynı metriğe yerleştirebilmek adına eşleştirme kriteri kullanılmaktadır (Wainer ve ark., 1991). Eşleştirme kriteri olarak kullanılacak maddeler DMF açısından incelenen testte yer alan maddeler olduğunda eşleştirme kriteri arınık (DMF'li madde içermeyen) olmayabilir. İsabetli DMF analizlerinin yapılabilmesi içinde eşleştirme kriterinin arınık olması gerekmektedir (Holland & Thayer, 1988). Bu amaçla Shih ve Wang (2009), eşleştirme kriteri olarak kullanılmak üzere saf bir kısa çapa oluşturmayı önermiş ve M-AÇM yöntemini geliştirmişlerdir.

Standart MIMIC (M-ST) Yöntemi. Standart MIMIC (M-ST) yöntemiyle i maddesinin DMF incelemesi yapılırken, referans ve odak grubun bu maddedeki performansını karşılaştırmak için testteki diğer tüm maddeler gruplar arası ortak bir metrik oluşturabilmek amacıyla eşleştirme kriteri olarak kullanılmaktadır. Örneğin; 20 maddelik bir testte madde 1 için DMF incelemesi yapılırken β_1 'in kestiriminde madde 1'in gruplar arasında farklı parametrelere sahip olmasına izin verilir. Ancak madde 2'den madde 20'ye kadar diğer tüm maddeler gruplar arasında eşit parametrelere sahip olacak şekilde sınırlandırılır. Şekil 3(a)'da görüldüğü üzere grup üyeliği değişkeni z 'den Y_1 'e doğru bir ok vardır ve bu ok Y_1 'in β parametresinin olduğunu ve incelenen madde olduğunu

göstermektedir. z 'den diğer maddelere ok olmaması, bu maddelerin β parametresinin olmadığını dolayısıyla da DMF'siz ve de eşleştirme kriteri olarak görev alacakları şekilde sınırlandırıldıklarını göstermektedir. Bu inceleme aşamasında β_1 istatistiksel olarak 0'dan farklıysa madde 1'in DMF gösterdiği kabul edilmektedir. Benzer şekilde madde 2 için DMF incelemesi yapılırken madde 1 ve madde 3'ten madde 20'ye kadar olan maddeler gruplar arasında eşit parametrelere sahip olacak şekilde sınırlandırılırken β_2 'nin kestiriminde madde 2'nin gruplar arasında farklı parametrelere sahip olmasına izin verilir. Eğer β_2 istatistiksel olarak 0'dan farklıysa madde 2'nin DMF gösterdiği kabul edilmektedir. Testteki diğer maddeler için de analiz bu şekilde devam etmektedir (Wang ve ark., 2009, Wang & Shih, 2020). M-ST yönteminde bu şekilde uygulanan gruplar arasında ortak metrik oluşturma yöntemine all-other-item (AOI) denilmektedir (Wang, 2004; Wang & Yeh, 2003). Bu yöntemde çalışılan madde dışındaki tüm test maddeleri gruplar arasında eşit parametrelere sahip olacak şekilde sınırlandırılarak çapa görevi görür ve yalnızca DMF incelemesi yapılan madde gruplar arasında farklı parametrelere sahip olabilir. Çalışılan her maddenin kendine ait bir ortak metriği olduğundan testteki madde sayısı kadar ortak metrik olacaktır. Testte DMF gösteren bir madde olmadığı sürece eşleştirme değişkenlerinin hepsinin eş zamanlı olarak DMF'den arınık olması mümkün değildir. Testteki DMF'li madde sayısı ne kadar fazlaysa eşleştirme değişkenlerinin arınıklığı da o kadar kötü olmaktadır. Bu nedenle AOI yöntemi de aynı oranda kötü performans göstermektedir.

Şekil 3

(a) M-ST Yöntemi ve (b) M-AÇM Yöntemi



Arınık Çapa Maddeli MIMIC (M-AÇM) Yöntemi. Testteki tüm maddeler eşleştirme değişkeni olarak kullanılırsa eşleştirme değişkenindeki DMF'li olma durumu ölçek saflaştırma yöntemleriyle telafi edilemeyecek kadar büyük olabilir. Eşleştirme kriterindeki DMF'li maddelerin DMF incelemesindeki etkisini ortadan kaldırmanın uygun bir yolu testteki DMF göstermeyen küçük bir madde grubunu seçip arınık bir eşleştirme değişkeni (çapa) oluşturmak ve testteki diğer maddelerin DMF incelemesini bu arınık eşleştirme kriteriyle yapmaktır. Burada DMF analizi iki aşamada gerçekleşmektedir. Birinci adımda DMF göstermeme olasılığı en yüksek olan maddeler belirlenmektedir. İkinci adımda bu maddeler çapa görevi görmekte ve diğer maddeler için eş zamanlı olarak tek bir bilgisayar kalibrasyonu ile DMF analizi yapılmaktadır. Bu metoda Arınık Çapa Maddeli MIMIC Yöntemi (M-AÇM) denilmektedir (Shih & Wang, 2009; Wang & Shih, 2010). Gruplar arasında ortak bir metrik sağlamak amacıyla DMF göstermeyen çapa maddeleri seçmek için Shih ve Wang (2009) iteratif MIMIC (M-IT) yöntemini önermiştir. Yaptıkları simülasyon çalışmalarıyla bu yöntemin DMF'siz maddeleri belirlemede yüksek performans gösterdiğini, özellikle de 4 taneye kadar DMF'siz madde seti belirlemede mükemmele yakın performans gösterdiğini görmüşlerdir. Ayrıca çapa madde sayısının fazla olmaması gerektiğini (maddelerin hepsinin DMF'siz olması pek mümkün olmayacağı için) belirterek dört madde ya da test uzunluğunun %10 veya %20'si kadar madde içeren kısa çapalar kullanılmasını önermişlerdir.

M-IT yöntemiyle DMF göstermeme olasılığı en yüksek olan çapa maddelerin bulunduğu M-AÇM yönteminin ilk aşamasının adımları aşağıdaki gibidir (Shih & Wang, 2009):

1. Madde 1 çapa olarak kabul edilir ve diğer tüm maddeler için MIMIC yöntemiyle DMF incelemesi yapılır. Her bir madde için DMF katsayısı elde edilir.
2. Bir sonraki madde çapa olarak kabul edilir ve diğer tüm maddeler için MIMIC yöntemiyle DMF incelemesi yapılır. Her bir madde için DMF katsayısı elde edilir.
3. Testteki son madde de çapa kabul edilene kadar ikinci adım tekrar edilir.

4. Bütün tekrarların ardından her bir madde için DMF katsayılarının mutlak değerlerinin ortalaması hesaplanır. Ve en düşük ortalama değere sahip maddeler (çalışmada kullanılacak çapa madde sayısı kadar) çapa maddeler olarak belirlenir.

DMF göstermeme olasılığı yüksek olan çapa maddeler belirlendikten sonra ikinci aşamada M-AÇM yöntemi Şekil 3b'de gösterildiği gibi gerçekleştirilmektedir. Şekil 3b'de madde 1'in çapa madde olduğu ve testteki diğer tüm maddelerin DMF açısından incelendiği görülmektedir. Grup üyeliği değişkeni z 'den $Y_2 \dots Y_i$ 'ye doğru oklar varken Y_1 'e doğru ok yoktur. Bu da $Y_2 \dots Y_i$ maddelerinin kendi β parametrelerinin olduğunu ve de DMF açısından incelenen maddeler olduğunu, Y_1 'in ise β parametresinin olmadığını ve de çapa madde olduğunu göstermektedir. M-AÇM modelinde birden fazla çapa madde olabileceğinden z grup üyeliği değişkeninden çapa olan maddelere uzanan oklar kaldırılıp bu maddeler için β kestirimi yapılmayarak model birden fazla çapa madde olan durumlara uyarlanabilmektedir (Wang & Shih,2010).

İlgili Araştırmalar

Finch (2005), MIMIC yöntemini MH, SIBTEST ve MTK olabilirlik oranı yöntemleriyle karşılaştırmıştır. Çalışmada 2PLM ve 3PLM'ye göre üretilen simülasyon verisi kullanılmıştır. Çalışmada odak ve referans grubun yetenek düzeyi ortalamaları ((0,0) ve (0,1)), örneklem büyüklüğü (R500/O500 ve R500/O100), test uzunluğu (20 ve 50 madde), çapa maddelerdeki DMF'li madde oranı (%0 ve %15) ve DMF büyüklüğü (0 ve 0.6) koşulları incelenmiştir. Çalışmanın sonucunda MIMIC yönteminin testin uzunluğu 50 maddeden oluştuğunda veya veri 2PLM'e uygun olduğunda DMF'li maddeleri belirlemede etkili olduğu ancak 3PLM'ye göre üretilmiş 20 maddeli test koşulunda hatalı sonuçlar verdiği görülmüştür.

Shih ve Wang (2009) çalışmasında M-AÇM yöntemini tanıtmış ve farklı koşullar altında bu yöntemle elde edilen I. tip hata ve güç oranlarını incelemiştir. İncelemeler iki

kategorili simülasyon verisi üzerinden yürütülmüştür. Çalışmada MTK modeli (1PLM, 2PLM ve 3PLM), DMF deseni (tek yönlü, baskın ve dengeli), test uzunluğu (20, 30 ve 40 madde), DMF'li madde oranı (%0, %10, %20, %30 ve %40), örneklem büyüklüğü (R500/O500, R1000/O1000 ve R1500/O1500) ve çapa uzunluğu (1, 2, 4 ve 10 madde) olmak üzere altı bağımsız değişken manipüle edilerek yöntemler incelenmiştir. M-AÇM yöntemiyle çapa uzunluğundan bağımsız olarak testteki DMF'li madde oranı %40 olduğunda bile I. tip hata oranlarının kontrol altına alındığı görülmüştür. Çalışmada çapa uzunluğunun artmasının DMF'yi belirleme gücünün artmasına katkı sağladığı görülmekle birlikte 4 maddeli bir çapa uzunluğunun da yüksek bir DMF tespit gücü sağladığı gözlenmiştir. Ayrıca çalışmada arınık çapa olarak kullanılacak maddeleri belirlemek üzere tekrarlı bir yöntem önerilmiş ve diğer bir simülasyon çalışmasıyla bu yöntemin doğruluğu incelenmiştir. Bu yöntemin 4 maddeye kadar çapa madde setini belirlemede oldukça yüksek bir doğruluk oranı sağladığı bulunmuştur.

Wang ve ark. (2009), bir dizi simülasyon yoluyla iki kategorili veri için M-ST ve ölçek arındırmalı MIMIC yöntemini karşılaştırmıştır. Çalışmada DMF belirleme yöntemi (M-ST ve ölçek arındırmalı MIMIC yöntemi), MTK modeli (2PLM ve 3PLM), DMF deseni (tek yönlü ve dengeli), test uzunluğu (20 ve 50 madde), DMF'li madde oranı (%0, %10, %20, %30 ve %40), örneklem büyüklüğü (R500/O100, R500/O500 ve R1000/O500) olmak üzere altı bağımsız değişken manipüle edilerek yöntemler incelenmiştir. Ölçek arındırmalı MIMIC yönteminin I. tip hata oranlarını kontrol altına alma ve daha yüksek güç oranları elde etme açısından M-ST yöntemine göre daha iyi performans sergilediği görülmüştür. M-ST yönteminin yalnızca DMF deseninin dengeli olduğu veya DMF'li madde oranının düşük olduğu koşullarda ölçek arındırmalı MIMIC yöntemi kadar iyi performans gösterdiği bulunmuştur. Ayrıca her iki yöntem de 2PLM altında 3PLM'e göre daha yüksek güç değerleri sağlamıştır. Gerçek testlerde DMF deseninin düzgün şekilde dengeli olması ve DMF'li madde oranlarının düşük olması pek mümkün olmadığından M-ST'nin değil de ölçek arındırmalı MIMIC yönteminin kullanılmasının iyi olacağını söylemişlerdir.

Woods (2009) yaptığı simülasyon çalışmasında MIMIC ve MTK olabilirlik oranı yöntemlerinin performanslarını küçük odak grupları söz konusu olduğunda I. tip hata ve güç oranları açısından karşılaştırmıştır. Çalışma iki kategorili ve beş kategorili simülatif madde verisi üzerinde küçük odak grubu büyüklüğüne sahip örneklemeler üzerinde yürütülmüştür. Sonuçlar tek biçimli DMF belirlemede MIMIC yönteminin MTK olabilirlik oranı yöntemine göre daha iyi sonuçlar ortaya koyduğunu göstermiştir.

Woods ve ark. (2009) kişilik ölçeğinden elde ettikleri gerçek veri üzerinden cinsiyet ve etnik köken değişkenlerine göre MIMIC yöntemiyle DMF analizleri yapmışlardır. Çalışmanın sonucundaki ölçekteki pek çok maddenin DMF'li olduğu bulunmuştur ve bu maddelerin ölçekten çıkarılması ya da yeniden modellenmesi önerilmiştir.

Wang ve Shih (2010) çalışmasında, çok kategorili simülasyon verisi üzerinde M-ST, M-AÇM ve ölçek arındırmalı MIMIC yöntemini karşılaştırmıştır. Çalışmada DMF belirleme yöntemi (M-ST, ölçek arındırmalı MIMIC yöntemi ve M-AÇM (4 çapa maddeli)), MTK modeli (Kısmi Puan Modeli (Masters, 1982) ve Genelleştirilmiş Kısmi Puan Modeli (Muraki (1982))), test uzunluğu (10, 20 ve 30 madde), DMF'li madde oranı (%0, %10, %20, %30 ve %40) olmak üzere dört bağımsız değişken manipüle edilerek yöntemler incelenmiştir. Yapılan simülasyon çalışmaları sonucunda DMF'li madde oranı %0 olduğunda üç yöntemde I. tip hata oranları açısından iyi sonuçlar verdiği görülmüştür. Testteki DMF'li madde oranı %10 ve %20 olduğunda M-ST ve ölçek arındırmalı MIMIC yönteminin I. tip hata ve güç değerlerinin kötüleştiği tespit edilmiştir. Testteki DMF'li madde oranı %40 olduğunda bile M-AÇM modelinin I. tip hata ve güç değerlerinin kabul edilebilir seviyelerde olduğu belirlenmiştir.

Arıkan ve ark. (2016) çalışmasında, MIMIC, LR, MH ve LR yöntemlerinin benzerlik ve farklılıklarını belirlemeye yönelik incelemeler yapmıştır. Ayrıca elde edilen DMF sonuçlarıyla uzman görüşlerinden elde edilen sonuçların birbiriyle uyumu incelenmiştir. Araştırmada 2009 yılında gerçekleştirilmiş 8. sınıf Seviye Belirleme Sınavı fen ve teknoloji alt testinde yer alan 20 maddenin DMF'li olup olmadığı dört yöntemle cinsiyet değişkenine

göre incelenmiştir. Çalışma R150/O150, R500/O500, R1000/O1000, R500/F100, R1000/F200 olmak üzere beş alt örneklem üzerinde yapılmıştır. Sonuçlara bakıldığında dört yöntemin büyük örneklemelerde birbiriyle daha uyumlu sonuçlar verdiği görülmüştür. Ayrıca uzman görüşleriyle DMF analizi sonuçlarının da birbiriyle tutarlı olduğu belirlenmiştir.

Zimbra (2018), MIMIC yönteminin tek biçimli DMF ve tek biçimli olmayan DMF belirleme performansını MTK olabilirlik oranı yöntemi ve Wald testi yöntemleriyle karşılaştırmıştır. İki kategorili simülasyon verisiyle yapılan çalışmada manipüle edilen koşullar; örneklem büyüklüğü (R1000/O1000, R500/O500, R1500/O500, R750/F250), DMF türü (DMF'siz, tek biçimli DMF ve tek biçimli olmayan DMF), DMF'li madde oranı (%0, %20 ve %40) ve DMF büyüklüğü (0.3, 0.5 ve 0.7) dür. Çalışmanın sonucunda tek biçimli olmayan DMF için, örneklem büyüklüğünden bağımsız olarak yüksek DMF büyüklüğü ve %20 DMF oranında MIMIC yönteminin I. tip hata ve güç oranları açısından MTK olabilirlik oranı ve Wald testi yöntemlerinden daha iyi performans sergilediği görülmüştür. DMF oranı %40 olduğunda ise Wald testinin diğer tüm koşullardan bağımsız olarak iki yönteme göre daha iyi sonuçlar verdiği görülmüştür.

Yapılan araştırmalar incelendiğinde M-ST yönteminin geleneksel DMF belirleme yöntemleriyle karşılaştırıldığı, gerçek veriyle ya da simülasyon verisiyle uygulamaların yapıldığı çalışmalar olduğu görülmüştür. Bununla birlikte, M-AÇM yöntemiyle yapılan çalışmaların nadir olduğu belirlenmiştir. Ayrıca iki kategorili veri için M-AÇM yöntemiyle M-ST yönteminin birlikte ele alındığı bir çalışmaya rastlanamamıştır. Bu araştırmada iki kategorili veri için M-ST ve M-AÇM yöntemlerinin performanslarının incelenebilmesi için hem simülasyon verisiyle hem de gerçek veriyle çalışmalar yapılmıştır.

Bölüm 3

Yöntem

Bu bölümde araştırmanın türü, çalışma grubu, araştırma verileri ve verilerin analizi başlıklarına yer verilmiştir.

Araştırmanın Türü

Bu çalışmada farklı koşullara uygun olacak şekilde üretilmiş simülasyon verileri üzerinde MIMIC yöntemleriyle DMF incelemeleri yapılmış ve yöntemler birbiriyle karşılaştırılmıştır. Simülasyon verisi kullanılan ve farklı veri setlerinden elde edilen sonuçlar arasındaki ilişkilerin karşılaştırıldığı bu çalışma, simülatif bir ilişkisel araştırma çalışmasıdır. Çalışmanın ikinci bölümünde PISA 2012 matematik testi için MIMIC yöntemleriyle DMF incelemeleri yapılmıştır. Araştırma var olan durumu belirleyen bu yönüyle betimsel bir çalışma olarak nitelendirilebilir.

Araştırma Verileri

Simülasyon Koşulları

M-ST ve M-AÇM yöntemlerinin performanslarının karşılaştırmak amacıyla üretilen veri setlerinde farklı koşullar ele alınmıştır. Bu koşullar; MTK modeli, örneklem büyüklüğü, odak/referans grup oranı, DMF'li madde oranı ve DMF büyüklüğü şeklindedir.

MTK Modeli. MIMIC yöntemleriyle yapılan DMF çalışmalarında simülasyon verisi kullanıldığında, üretilen verinin uyum sağladığı MTK modelinin I. tip hata ve güç oranlarının üzerinde etkisinin olduğu görülmüştür. Finch (2005), Shih ve Wang (2009) çalışmalarında 2PLM ve 3PLM'ye göre üretilmiş veriler açısından farklı bulgular elde etmişlerdir. Bu nedenle bu çalışmanın simülasyon kısmında 2PLM ve 3PLM'ye uygun veriler üretilmiştir.

Örneklem Büyüklüğü. Örneklem büyüklüğü DMF yöntemlerinin performansını etkileyebilecek kritik bir faktördür (Meade & Bauer, 2007). Literatürdeki simülasyon çalışmaları incelendiğinde örneklem büyüklüğünün manipüle edilen bir değişken olduğu

görülmektedir (Navas-Ara & Gomez-Benito, 2002; Oort, 1998; Raju ve ark., 2002; Swaminathan & Rogers, 1990; Wang & Shih, 2010; Wang ve ark., 2009; Woods, 2009; Zimbra 2018). Bu çalışmada incelenen örneklem büyüklükleri 2000, 1000 ve 600 olarak belirlenmiştir. Örneklem büyüklükleri seçilirken Finch (2005), Shih ve Wang'ın (2009) çalışmaları dikkate alınmıştır.

Odak/Referans Grup Oranı. Odak ve referans grubu oranları, DMF analizi sonuçlarını etkileme potansiyeli olan önemli bir faktördür (Swaminathan & Rogers, 1990). Bu çalışmada incelenen odak ve referans grup oranları, oranların dengeli ve dengesiz olma durumunu yansıtmayı amacıyla 1/1 ve 1/3 olarak belirlenmiştir. Odak ve referans grup oranları seçilirken Lee ve ark. (2017), Ömür Sünbül ve Sünbül (2016), Zimbra'nın (2018) çalışmaları dikkate alınmıştır.

DMF'li Madde Oranı. Literatürdeki DMF'yle ilgili simülasyon çalışmaları incelendiğinde incelenen koşullardan birinin de testteki DMF'li madde oranı olduğu görülmektedir. Bu çalışmalarda testteki DMF'li madde oranının %0 ile %80 arasında değiştiği görülmektedir (Shih & Wang 2009; Swaminathan & Rogers, 1990; Wang ve ark., 2009; Wang & Shih, 2010; Woods, 2009; Zimbra, 2018). Bu çalışmada incelenen DMF'li madde oranları %20 (Madde 1 – Madde 4 DMF'li) ve %40 (Madde 1 – Madde 8 DMF'li) olarak belirlenmiştir. Bu oranlar seçilirken Wang (2009), Shih ve Wang (2009), Wang ve Shih (2010), Zimbra (2018)'in çalışmaları göz önünde bulundurulmuştur.

DMF Büyüklüğü. DMF yöntemlerinin performansını etkileyen unsurlardan biri DMF büyüklüğüdür. Bu çalışmada düşük düzeyde DMF büyüklüğünü temsilen 0.3, orta düzeyde DMF büyüklüğünü temsilen 0.5 ve yüksek düzeyde DMF büyüklüğünü temsilen 0.7 değerleri kullanılmıştır (Woods, 2013; Cao ve ark. 2017; Zimbra, 2018).

Çalışmada incelenen tüm koşullar Tablo 5'de verilmiştir.

Tablo 5

Simülasyon Koşulları

Model	Örneklem Büyükükleri	Odak/Referans Oranı	DMF'li Oranı	Madde	DMF Büyüküküğü
2PLM	600, 1000, 2000	1/1			0.3
				%20	0.5
					0.7
					0.3
				%40	0.5
					0.7
		1/3			0.3
				%20	0.5
					0.7
					0.3
				%40	0.5
					0.7
3PLM	600, 1000, 2000	1/1			0.3
				%20	0.5
					0.7
					0.3
				%40	0.5
					0.7
		1/3			0.3
				%20	0.5
					0.7
					0.3
				%40	0.5
					0.7

Araştırmada MTK modeli, örneklem büyüklüğü, odak/referans grup oranı, DMF'li madde oranı ve DMF büyüklüğü koşulları ele alınmıştır. Buna göre, çalışmadaki toplam koşul sayısı 72 (2x3x2x2x3)'dir.

Çalışmadaki test uzunluğu sabit koşul olarak ele alınıp analizler 20 madde üzerinden yürütülmüştür.

Verilerin Üretilmesi

Çalışmanın simülasyon verileri R programı (R Core Team, 2018) kullanılarak üretilmiştir. Çalışmada her bir koşul için literatürdeki pek çok DMF çalışmasında olduğu gibi (Lee ve ark., 2017; Shih, 2009; Wang & Shih 2010; Wang ve ark., 2009; Yıldıztekin, 2020) 100 replikasyon uygulanmıştır.

İki kategorili maddeler sıklıkla kullanılan madde türleri olduğundan bu çalışma, 20 maddelik bir test için 2PLM ve 3PLM'ye göre üretilen iki kategorili veriler üzerinden yürütülmüştür. İki kategorili madde cevapları üretilirken SAT sınavından (Donoghue & Allen, 1993; Finch, 2005; Lord, 1968) elde edilen ve Tablo 6'da verilen madde parametreleri kullanılmıştır. 2PLM'ye uygun veri üretilirken c parametresi 0 olarak sabitlenmiştir. Bireylerin yetenek parametreleri normal dağılım $N(0,1)$ dikkate alınarak üretilmiştir.

Tablo 6

Veri Üretiminde Kullanılan Madde Parametreleri

Madde No	a	b	c
1	1.10	-.70	.20
2	.70	-.60	.20
3	.90	-.40	.20
4	1.40	.10	.20
5	.90	.90	.16
6	1.20	.70	.12
7	.90	.30	.20

8	.40	.80	.20
9	1.60	1.10	.06
10	2.00	1.10	.05
11	.90	-1.50	.20
12	1.40	-.40	.20
13	1.60	-.10	.20
14	1.20	.50	.20
15	1.20	1.40	.11
16	1.80	1.40	.12
17	2.00	1.60	.16
18	1.00	1.60	.13
19	1.50	1.70	.09
20	1.20	1.60	.09

Çalışma, M-ST ve M-AÇM yöntemlerini tek biçimli DMF açısından incelemeyi amaçlamaktadır. Bu nedenle veri üretimi aşamasında tek biçimli DMF yaratmak adına odak gruba ait b parametreleri (ilk 4 ve ilk 8 madde için) koşullarda yer alan DMF büyüklükleri dikkate alınarak manipüle edilmiş, a parametreleri ise odak ve referans gruplar için aynı bırakılmıştır.

Üretilen verilerin tek boyutluluk koşuluna uygun olup olmadığını incelemek amacıyla, her bir koşul için üretilen 100 veri setinden rastgele 10 veri seti seçilmiş ve Factor.10.4.WIN64 programı kullanılarak faktör analizi yapılmıştır. Yapılan analizlerin sonucunda verilerin tek boyutluluk koşulunu sağladığı görülmüştür. Paralel analiz yöntemiyle elde edilen sonuçların bir kısmı Ek-A'da verilmiştir.

Çalışma Grubu

PISA 2012 verisiyle yapılan çalışmada DMF incelemesi anadili İngilizce olma ve olmama değişkenine göre yapıldığından PISA 2012 uygulamasına katılan ülkeler bu açıdan ele alınmıştır. Çalışma grubu belirlenirken matematik başarı sıralaması ve uygulamaya

katılan öğrenci sayısı kriterleri bakımından yakın olan ülkeler belirlenmiştir. Buna göre Birleşik Krallık ve ABD anadili İngilizce olan ülkeleri (AİÖÜ) temsil etmek üzere; Danimarka, Çekya, Fransa, Slovakya ve İzlanda anadili İngilizceden farklı olan ülkeleri (AİFOÜ) temsil etmek üzere seçilmiştir.

Tablo 7

Çalışma Grubuna Dâhil Edilen Ülkelerle İlgili Bilgiler

Ülke	PISA 2012'ye Katılan Öğrenci Sayısı	PISA 2012 Matematik Okuryazarlığı Başarı Sırası	Anadil
ABD	4953	36	İngilizce
Birleşik Krallık	12239	26	İngilizce
Çekya	5327	24	İngilizce Değil
Danimarka	7481	22	İngilizce Değil
Fransa	4613	25	İngilizce Değil
İzlanda	3508	27	İngilizce Değil
Slovakya	4650	35	İngilizce Değil

PISA 2012'ye Tablo 7'de belirtilen ülkelerden katılan, 4 numaralı kitapçığı cevaplamış ve matematik testinde kayıp verisi olmayan katılımcıların oluşturduğu 566'sı AİÖÜ'den ve 513'ü AİFOÜ'den oluşan iki grup belirlenmiştir. Belirlenen bu iki grup üzerinden de Tablo 8'de verilen dört farklı çalışma grubu elde edilmiştir. Dört çalışma grubu belirlenirken örneklem büyüklükleri arasında yaklaşık $\frac{1}{2}$ 'lik oran olması ve odak/referans gruplarının eşit ve eşit büyüklükte olmaması koşulları dikkate alınmıştır. Çalışma grupları belirlenirken SPSS 15 paket programı kullanılmıştır.

Tablo 8

Çalışma Gruplarının Frekans Dağılımı

Çalışma Grubu	AİÖÜ Sayısı	AİFOÜ Sayısı	Toplam
Çalışma Grubu 1	550	350	900
Çalışma Grubu 2	450	450	900

Çalışma Grubu 3	225	225	450
Çalışma Grubu 4	300	150	450

Gerçek Verilerin Elde Edilmesi

Çalışmanın gerçek veriyle gerçekleştirilen uygulaması PISA 2012 matematik verisi üzerinden yürütülmüştür.

PISA, 15 yaşındaki öğrencilerin edinmesi beklenen bilgi ve becerilere ne ölçüde sahip olduklarının incelendiği, üç yıl arayla uygulanan uluslararası bir eğitim araştırmasıdır (MEB, 2013). PISA, zorunlu eğitim sürecini tamamlamak üzere olan öğrencilerin topluma tam entegrasyonu için kritik olan matematik, fen ve okuma alanlarında ne düzeyde bilgi ve beceriye sahip olduklarını değerlendirmektedir. Her bir PISA uygulamasında matematik, fen ya da okuma alanlarından birine ağırlık verilmektedir.

PISA 2012’de matematik alanına ağırlık verildiğinden tüm kitapçıklar matematik maddeleri göz önünde bulundurularak incelenmiştir. Buna göre, matematik alanını ölçen maddelerden uygulanmamış ve iki kategorili olmayan maddeler çıkarıldığında 4 numaralı kitapçığın en yüksek madde sayısına sahip olduğu görülmüştür. O nedenle çalışma, 4 numaralı kitapçıkta istenilen koşullara uygun olan 34 matematik maddesi için elde edilen veriler üzerinden yürütülmüştür. Verinin tek boyutluluk koşuluna uygun olup olmadığını incelemek amacıyla, Factor.10.4.WIN64 programı kullanılarak faktör analizi yapılmıştır. Yapılan analiz sonucunda verinin tek boyutluluk koşulunu sağladığı görülmüştür. Paralel analiz yöntemiyle elde edilen sonuç Ek-B’de verilmiştir.

Verilerin Analizi

Çalışmada DMF analizleri M-ST ve M-AÇM yöntemleri kullanılarak gerçekleştirilmiştir. Analizler R programında yer alan “MplusAutomation” (Hallquist & Wiley, 2018) paketi ve Mplus 6.12 (Muthén & Muthén, 1998, 2010) programı aracılığıyla yürütülmüştür. M-ST analizinde kullanılan örnek bir R programı kodu Ek-C’de verilmiştir.

100 tekrarla üretilmiş her bir yapay veri seti için ayrı ayrı analiz yapılmıştır. Elde edilen sonuçlar üzerinden I. tip hata ve güç oranları hesaplanmıştır.

MIMIC analizinde kestirim yöntemi olarak maksimum likelihood metodunun tercih edilmesi hatalı değerler bulunmasına neden olabilmektedir. Çünkü doğrulayıcı faktör analizinde gözlenen değişkenler kategoriktir ve maksimum likelihood yöntemi normallik sayılına dayanmaktadır. Bu nedenle MIMIC analizlerinde robust WLS yöntemlerinden olan WLSMV kestirim yöntemi kullanılmıştır (Arıkan ve ark., 2016; Shih, 2009; Wang ve ark., 2009).

M-AÇM yöntemiyle yapılan analizlerde çapa madde sayısı 4 olarak sabitlenmiştir. Shih ve Wang (2004), bu yöntem kullanıldığında test uzunluğunun %10'u veya %20'si kadar maddeden oluşan kısa çapalar kullanılmasını tavsiye etmişlerdir.

I. Tip Hata ve Güç Oranları

Her bir simülasyon koşulu için yöntemlerin performanslarını değerlendirebilmek amacıyla I. tip hata ve güç oranları hesaplanmıştır.

I. tip hata, hipotez testi sonucunda gerçekte doğru olan yokluk hipotezinin reddedilmesi nedeniyle ortaya çıkan hatadır. DMF bağlamında ele alındığında I. tip hata, gerçekte DMF'li olmayan maddenin DMF'li olarak belirlenmesidir (Vaughn & Wang, 2010). I. tip hatayı hesaplamak için DMF'li olmadığı halde DMF'li olarak belirlenmiş madde sayısı testteki DMF'li olmayan madde sayısına bölünmüştür. Bu çalışmada I. tip hata oranları yorumlanırken kriter olarak 0.05 değeri kabul edilmiş ve bu değer altındaki hata oranları için kabul edilebilir değerlendirilmesi yapılmıştır.

Testin gücü, hipotez testi sonucunda gerçekte yanlış olan yokluk hipotezinin reddedilmesiyle ortaya çıkan değerdir. DMF bağlamında ele alındığında güç, gerçekte DMF'li olan maddenin DMF'li olarak belirlenmesidir (Vaughn & Wang, 2010). Gücü hesaplamak için gerçekten DMF'li olan ve test sonucunda da DMF'li olarak belirlenmiş madde sayısı testteki DMF'li madde sayısına bölünmüştür. Bu çalışmada güç oranları

yorumlanırken kriter olarak 0.80 değeri kabul edilmiş ve bu değerin üzerindeki güç oranları için kabul edilebilir değerlendirilmesi yapılmıştır.

Tablo 9

DMF Bağlamında I. Tip Hata ve Güç

		Hipotez Testi Kararı	
		H_0 ret “DMF var”	H_0 kabul “DMF yok”
Gerçek	H_0 doğru “DMF yok”	I. tip hata	Doğru karar
	H_0 yanlış “DMF var”	Doğru karar (Güç)	II. tip hata

Bölüm 4

Bulgular, Yorumlar ve Tartışma

Bu bölümde, elde edilen bulgular ve bu bulgulara göre yapılan yorum ve tartışmalar yer almaktadır.

“2PLM ve 3PLM’ye göre üretilen veriler için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata ve güç değerleri örneklem büyüklüğüne, odak/referans grubu oranlarına, testteki DMF’li madde oranına ve DMF büyüklüğüne göre nasıl değişmektedir?” problem cümlesi için elde edilen bulgular Tablo 10 Tablo 11, Tablo 12 ve Tablo 13’te sunulmuştur.

Tablo 10

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre I. Tip Hata Oranları

Koşullar			Yöntem					
			M-ST			M-AÇM		
			DMF Büyüklüğü			DMF Büyüklüğü		
Örneklem Büyüklüğü	O/R Oranı	DMF’li Madde Oranı	0.3	0.5	0.7	0.3	0.5	0.7
2000	1/1	%20	.10	.19	.30	.04	.03	.04
		%40	.20	.46	.74	.21	.28	.37
	1/3	%20	.09	.15	.24	.04	.03	.04
		%40	.16	.36	.60	.16	.28	.38
1000	1/1	%20	.08	.10	.17	.03	.03	.03
		%40	.15	.28	.44	.11	.23	.27
	1/3	%20	.07	.10	.13	.03	.04	.03
		%40	.10	.22	.35	.08	.20	.30
600	1/1	%20	.06	.09	.12	.03	.04	.03
		%40	.09	.19	.31	.06	.17	.23
	1/3	%20	.06	.08	.11	.02	.03	.03
		%40	.09	.16	.23	.06	.12	.17

Tablo 10'da, 2PLM'e uygun veriyle ele alınan tüm koşullara göre M-ST ve M-AÇM yöntemleri için elde edilmiş I. tip hata oranları verilmiştir. Tablo 10 genel olarak incelendiğinde 18 koşul için elde edilen hata oranlarının, I. tip hata oranları için kritik değer olan 0.05'ten daha düşük çıktığı görülmektedir. Bulgulara bakıldığında testteki DMF'li madde oranı %20'den %40'a çıktığında genel olarak iki yöntemin de I. tip hata oranlarının yükselmekte olduğu ve kritik değeri aştığı görülmektedir. Her iki yöntemde de diğer koşullardan bağımsız olarak DMF büyüklüğü arttıkça I. tip hata oranlarının arttığı söylenebilir. Bu durum yalnızca M-AÇM yöntemi için DMF'li madde oranının %20 olduğu koşullarda ortaya çıkmamıştır. O/R oranı açısından bakıldığında oran 1/3 iken elde edilen I. tip hata oranlarının, oran 1/1 iken elde edilen I. tip hata oranlarından nispeten daha düşük olduğu görülmektedir.

M-ST yöntemi için elde edilen I. tip hata oranları incelendiğinde hiçbir değer kritik değer olan 0.05'ten düşük olmadığı ve 0.06-0.74 aralığında değerler aldığı gözlemlenmektedir. Testteki DMF'li madde oranı ve DMF büyüklüğü arttıkça M-ST yöntemi için elde edilen I. tip hata oranları yükselmektedir.

M-AÇM yöntemi için elde edilen I. tip hata oranları incelendiğinde 18 koşul için elde edilen değerlerin kritik değer olan 0.05'ten düşük olduğu ve tüm koşullar için elde edilen değerlerin 0.02-0.38 aralığında değerler aldığı görülmektedir. O/R oranı, örneklem büyüklüğü ve DMF büyüklüğü fark etmeksizin DMF'li madde oranı %20 olan tüm koşullarda I. tip hata oranının 0.05 altında olup kabul edilebilir değerler aldığı görülmektedir. DMF büyüklüğü 0.3 ve örneklem büyüklüğü 600 iken elde edilen I. tip hata oranlarının kritik değerin altında ve bu değere yakın sınırlarda olduğu söylenebilir.

Tablo 11

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre I. Tip Hata Oranları

Koşullar	Yöntem	
	M-ST	M-AÇM

Örneklem Büyükliği	O/R Oranı	DMF'li Madde Oranı	DMF Büyüklüğü			DMF Büyüklüğü		
			0.3	0.5	0.7	0.3	0.5	0.7
2000	1/1	%20	.08	.14	.14	.04	.04	.04
		%40	.15	.30	.50	.09	.20	.26
	1/3	%20	.06	.11	.16	.03	.03	.03
		%40	.12	.24	.41	.07	.02	.25
1000	1/1	%20	.07	.09	.13	.0	.04	.04
		%40	.09	.18	.30	.05	.15	.19
	1/3	%20	.06	.09	.11	.03	.04	.03
		%40	.09	.16	.24	.05	.09	.13
600	1/1	%20	.07	.09	.10	.03	.04	.03
		%40	.07	.14	.19	.04	.10	.10
	1/3	%20	.06	.07	.08	.03	.03	.03
		%40	.08	.12	.17	.04	.07	.13

Tablo 11’de, 3PLM’e uygun veriyle araştırmada ele alınan tüm koşullara göre M-ST ve M-AÇM yöntemleri için elde edilmiş I. tip hata oranları verilmiştir. Tablo 11 genel olarak incelendiğinde 21 koşul için elde edilen hata oranlarının, I. tip hata oranları için kritik değer olan 0.05’ten daha düşük çıktığı görülmektedir. Bulgular incelendiğinde testteki DMF’li madde oranı %20’den %40’a çıktığında genel olarak iki yöntemin de I. tip hata oranlarının yükselmekte olduğu görülmektedir. M-AÇM yöntemi için DMF’li madde oranı %20 olduğunda elde edilen değerler dışında, her iki yöntemde de diğer koşulların geneli için DMF büyüklüğü arttıkça I. tip hata oranlarının arttığı söylenebilir. Bulgular O/R oranı açısından ele alındığında oran 1/1 olduğunda elde edilen I. tip hata oranlarının, oran 1/3 olduğunda elde edilen I. tip hata oranlarına göre nispeten daha yüksek olduğu gözlenmektedir.

M-ST yöntemi için elde edilen I. tip hata oranları incelendiğinde tüm değerlerin kritik değer olan 0.05’ten daha yüksek olduğu ve 0.06 – 0.50 aralığında yer aldığı görülmektedir.

Testteki DMF'li madde oranı ve DMF büyüklüğü arttıkça M-ST yöntemi için elde edilen I. tip hata oranları yükselmektedir.

M-AÇM yöntemi için elde edilen I. tip hata oranları incelendiğinde değerlerin 0.00-0.26 aralığında yer aldığı ve 21 koşul için elde edilen değerlerin kritik değer olan 0.05'ten daha düşük çıktığı gözlemlenmektedir. DMF'li madde oranının %20 olduğu tüm koşullarda I. tip hata oranının 0.05 altında olup kabul edilebilir değerler aldığı görülmektedir. DMF büyüklüğünün 0.3 olduğu koşullar için elde edilen I. tip hata oranlarının DMF büyüklüğü 0.5 ve 0.7 olduğu koşullara göre nispeten daha düşük ve kabul edilebilir değerlerde ve de kritik değere yakın sınırlarda olduğu söylenebilir.

Tablo 12

2PLM için M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre Güç Oranları

Koşullar			Yöntem					
			M-ST			M-AÇM		
			DMF Büyüklüğü			DMF Büyüklüğü		
Örneklem Büyüklüğü	O/R Oranı	DMF'li Madde Oranı	0.3	0.5	0.7	0.3	0.5	0.7
2000	1/1	%20	.69	.94	.99	.65	.95	1
		%40	.42	.82	.93	.22	.69	.79
	1/3	%20	.55	.92	1	.44	.92	1
		%40	.37	.72	.89	.20	.53	.72
	1/1	%20	.45	.78	.95	.31	.76	.94
		%40	.29	.59	.80	.16	.39	.64
1000	1/1	%20	.35	.69	.90	.24	.64	.90
		%40	.20	.49	.71	.08	.27	.49
	1/3	%20	.27	.60	.84	.19	.51	.83
		%40	.21	.41	.65	.10	.22	.46
	1/3	%20	.27	.53	.76	.17	.40	.73
		%40	.16	.34	.54	.07	.17	.36
600	1/1	%20	.27	.53	.76	.17	.40	.73
		%40	.16	.34	.54	.07	.17	.36
	1/3	%20	.27	.53	.76	.17	.40	.73
		%40	.16	.34	.54	.07	.17	.36
	1/3	%20	.27	.53	.76	.17	.40	.73
		%40	.16	.34	.54	.07	.17	.36

Tablo 12’de, 2PLM’e uygun veriyle ele alınan tüm koşullara göre M-ST ve M-AÇM yöntemleri için elde edilmiş güç oranları verilmiştir. Tablo 12 genel olarak incelendiğinde 17 koşul için elde edilen güç oranlarının, güç oranları için kritik değer olan 0.80’den daha yüksek çıktığı görülmektedir. Bulgular incelendiğinde DMF büyüklüğü arttıkça iki yöntemin de güç oranlarının yükselmekte olduğu görülmektedir. Örneklem büyüklüğü 2000 ve DMF’li madde oranı %20 iken O/R oranı fark etmeksizin DMF büyüklüğü 0.5 ve 0.7 olan tüm koşullar için elde edilen güç değerlerinin her iki yöntemle göre de 0.80’den yüksek çıktığı gözlemlenmektedir. DMF büyüklüğü 0.3 iken diğer koşullardan bağımsız olarak her iki yöntem için elde edilen güç değerleri 0.80’den daha düşük değerler almıştır. Bulgulara bakıldığında testteki DMF’li madde oranı %20’den %40’a çıktığında iki yöntemin de güç oranlarının düşme eğiliminde olduğu görülmektedir. Sonuçlara O/R oranı açısından bakıldığında oran 1/1 olduğunda elde edilen güç oranlarının, 1/3 olduğunda elde edilen güç oranlarına göre nispeten daha yüksek olduğu söylenebilir. Örneklem büyüklüklerine bakıldığında örneklem büyüklüğü küçüldükçe her iki yöntem için de güç oranlarının düştüğü görülmektedir.

M-ST yöntemi için elde edilen güç oranları incelendiğinde değerlerin 0.16 – 1 aralığında yer aldığı ve 10 koşul için kritik değer olan 0,80’den daha yüksek olduğu görülmektedir. Örneklem büyüklüğü 2000, O/R oranı 1/1 iken DMF büyüklüğü 0.5 ve 0.7 olduğunda; O/R oranı 1/1, DMF’li madde oranı %20 ve DMF büyüklüğü 0.7 olduğunda tüm koşullar için elde edilen güç değerlerinin 0.80’den daha yüksek olduğu gözlemlenmektedir. Örneklem büyüklüğü 600 veya 1000 iken DMF’li madde oranının %40 olduğu tüm koşullarda elde edilen güç değerlerinin 0.80’den daha düşük olduğu görülmektedir. Örneklem büyüklüğü 600 iken yalnızca DMF’li madde oranı %20, O/R oranı 1/1 ve DMF büyüklüğü 0.7 olduğunda elde edilen güç değeri kabul edilebilir çıkmıştır.

M-AÇM yöntemi için elde edilen güç oranları incelendiğinde değerlerin 0.07 – 1 aralığında yer aldığı ve 7 koşul için kritik değer olan 0.80’den daha yüksek olduğu görülmektedir. Örneklem büyüklüğü 2000 ve DMF’li madde oranı %20 iken DMF büyüklüğü

0.5 ve 0.7 olduğunda tüm koşullar için elde edilen güç değerlerinin 0.80'den daha yüksek olduğu gözlemlenmektedir. DMF büyüklüğü 0.7 ve DMF'li madde oranı %20 iken, örneklem büyüklüğü 600; O/R oranı 1/3 koşulu dışındaki tüm koşullar için kabul edilebilir güç değerleri elde edilmiştir.

Tablo 13

3PLM için M-ST ve M-AÇM Yöntemlerinin Farklı Koşullara Göre Güç Oranları

Koşullar			Yöntem					
			M-ST			M-AÇM		
			DMF Büyüklüğü			DMF Büyüklüğü		
Örneklem Büyüklüğü	O/R Oranı	DMF'li Madde Oranı	0.3	0.5	0.7	0.3	0.5	0.7
2000	1/1	%20	.54	.87	.86	.43	.87	.86
		%40	.35	.67	.86	.21	.50	.76
	1/3	%20	.44	.78	.96	.31	.75	.95
		%40	.25	.57	.79	.15	.40	.65
1000	1/1	%20	.31	.66	.87	.21	.55	.84
		%40	.19	.42	.64	.09	.23	.47
	1/3	%20	.25	.52	.80	.15	.44	.76
		%40	.15	.34	.53	.08	.21	.39
600	1/1	%20	.20	.44	.72	.11	.32	.67
		%40	.13	.29	.45	.06	.17	.30
	1/3	%20	.18	.36	.54	.08	.24	.47
		%40	.11	.25	.36	.07	.13	.20

Tablo 13'te, 3PLM'e uygun veriyle ele alınan tüm koşullara göre M-ST ve M-AÇM yöntemleri için elde edilmiş güç oranları verilmiştir. Tablo 13 genel olarak incelendiğinde 9 koşul için elde edilen güç oranlarının, güç oranları için kritik değer olan 0.80'den daha yüksek çıktığı görülmektedir. Bulgular incelendiğinde örneklem büyüklüğü 2000 iken, DMF'li madde oranı %20 ve DMF büyüklüğü 0.7 olduğunda O/R oranı fark etmeksizin tüm koşullar

için elde edilen güç değerlerinin her iki yöntemle göre de 0.80'den yüksek çıktığı gözlemlenmektedir. DMF büyüklüğü 0.3 iken diğer koşullardan bağımsız olarak her iki yöntem için elde edilen güç değerleri 0.80'den daha düşük değerler almıştır. Bulgulara bakıldığında DMF büyüklüğü arttıkça güç değerlerinin de artma eğiliminde olduğu söylenebilir. Testteki DMF'li madde oranı %20'den %40'a çıktığında iki yöntemin de güç oranlarının düşmekte olduğu görülmektedir. Sonuçlar örneklem büyüklüğü açısından ele alındığında ise örneklem büyüklüğü küçüldükçe güç değerlerinin düştüğü söylenebilir. O/R oranlarına bakıldığında oran 1/1 olduğunda elde edilen güç oranlarının, 1/3 olduğunda elde edilen güç oranlarına göre nispeten daha yüksek olduğu gözlenmektedir.

M-ST yöntemi için elde edilen güç oranları incelendiğinde değerlerin 0.11 – 0.96 aralığında yer aldığı ve 5 koşul için kritik değer olan 0.80'den daha yüksek olduğu görülmektedir. Örneklem büyüklüğü 1000 ve 2000 iken DMF büyüklüğü 0.7 ve DMF'li madde oranı %20 olduğunda O/R oranı fark etmeksizin tüm koşullar için güç değerlerinin kabul edilebilir olduğu söylenebilir. Örneklem büyüklüğü 600 olduğunda tüm koşullar için elde edilen güç değerleri kritik değer olan 0.80'in altında kalmıştır.

M-AÇM yöntemi için elde edilen güç oranları incelendiğinde değerlerin 0.06 – 0.95 aralığında yer aldığı ve 4 koşul için kritik değer olan 0.80'den daha yüksek olduğu görülmektedir. DMF büyüklüğü 0.7 olduğunda, örneklem büyüklüğü 2000 ve testteki DMF'li madde oranı %20 iken O/R oranı fark etmeksizin tüm koşullar için güç oranı 0.80'in üstünde çıkmıştır. DMF'li madde oranı %40 olduğunda tüm koşullar için M-AÇM yöntemiyle elde edilen güç değerlerinin 0.80'den düşük çıktığı görülmektedir. Örneklem büyüklüğü 600 olduğunda tüm koşullar için elde edilen güç değerleri kritik değer olan 0.80'in altında kalmıştır.

Her iki yöntem için 2PLM'e göre elde edilen sonuçlara bakıldığında, M-AÇM yönteminin örneklem büyüklüğü 2000, O/R oranı 1/1, DMF'li madde oranı %20, DMF büyüklüğü 0.5 ve 0.7; örneklem büyüklüğü 2000, O/R oranı 1/3, DMF'li madde oranı %20, DMF büyüklüğü 0.5 ve 0.7; örneklem büyüklüğü 1000, O/R oranı 1/1 ve 1/3, DMF'li madde

oranı %20, DMF büyüklüğü 0.7; örneklem büyüklüğü 600, O/R oranı 1/1, DMF'li madde oranı %20 ve DMF büyüklüğü 0.7 olmak üzere toplam 7 koşulda hem I. tip hata oranlarının hem de güç oranlarının kabul edilebilir değerler aldığı görülmüştür.

Her iki yöntem için 3PLM'e göre elde edilen sonuçlara bakıldığında, M-AÇM yönteminin örneklem büyüklüğü 2000, O/R oranı 1/1, DMF'li madde oranı %20, DMF büyüklüğü 0.5 ve 0.7; örneklem büyüklüğü 2000, O/R oranı 1/3, DMF'li madde oranı %20, DMF büyüklüğü 0.7; örneklem büyüklüğü 1000, O/R oranı 1/1, DMF'li madde oranı %20 ve DMF büyüklüğü 0.7 olmak üzere toplam 4 koşulda hem I. tip hata oranlarının hem de güç oranlarının kabul edilebilir değerler aldığı görülmüştür.

Hem 2PLM hem de 3PLM'e göre elde edilen sonuçlara bakıldığında, M-ST yöntemi için elde I. tip hata değerlerinin kabul edilebilir aralıkta olduğu bir koşul olmadığından M-ST yöntemi için hem I. tip hata hem de güç oranlarının kabul edilebilir değerler aldığı bir koşul gözlemlenmemiştir.

Alt Problem 1'e İlişkin Bulgular

Alt problem 1'e yanıt bulmak amacıyla 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle elde edilen I. tip hata oranları faktöriyel ANOVA ile incelenmiştir. Elde edilen sonuçlar alt problem 1'in altındaki problemlere göre ayrı ayrı sunulmuştur. Faktöriyel ANOVA sonuçlarının tablosu Ek- Ç'de verilmiştir.

a) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri örneklem büyüklüğüne (2000, 1000 ve 600) göre nasıl değişmektedir?

Tablo 14

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları

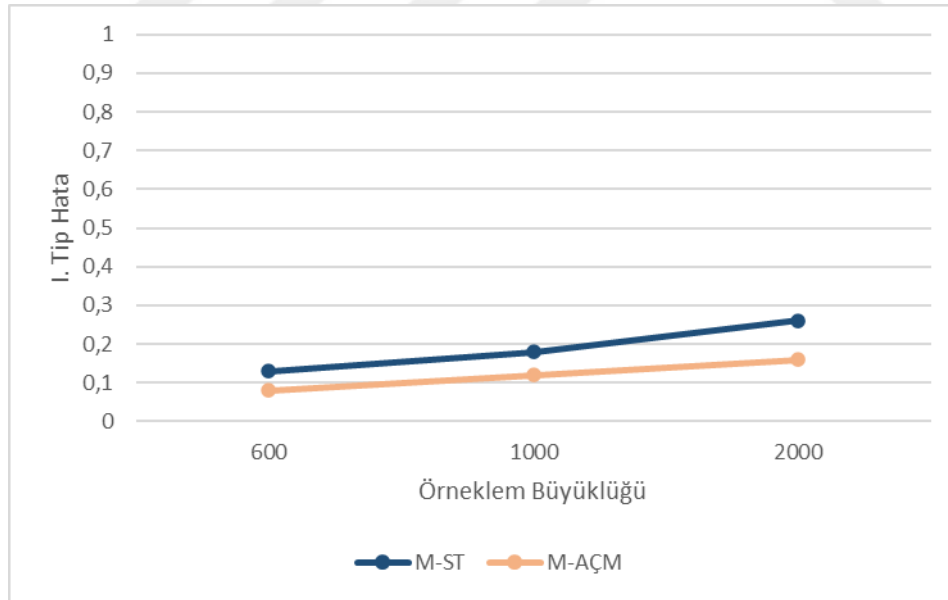
Örneklem Büyüklüğü			Yöntem	Örneklem Büyüklüğü	Etkileşim
600	1000	2000			

Yöntem	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.13	.08	.18	.12	.26	.19	$F_{1;45}=22.07$	$F_{2;45}=14.33$	$F_{2;45}= .80$
M-AÇM	.08	.07	.12	.11	.16	.14	$p=.00$	$p=.00$	$p=.46$
							$\omega^2=.03$	$\omega^2=.02$	

Tablo 14 incelendiğinde I. tip hata oranlarının örneklem büyüklüklerine göre anlamlı farklılık gösterdiği görülmektedir. Buna göre örneklem büyüklüğünün I. tip hatadaki değişimi %2 oranında etkilediği söylenebilir. Örneklem büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında da I. tip hata oranları açısından anlamlı bir farklılık görülmemiştir. Örneklem büyüklüğü açısından ortaya çıkan farklılığın hangi örneklem büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın 600 – 2000 ve 1000 - 2000 örneklem büyüklükleri arasında olduğu tespit edilmiştir.

Şekil 4

2 PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranları



Şekil 4 incelendiğinde örneklem büyüklüğü arttıkça her iki yöntemin de I. tip hata oranlarının arttığı görülmektedir. Tüm örneklem büyüklükleri boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

b) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri odak/referans grup oranlarına (1/1 ve 1/3) göre nasıl değişmektedir?

Tablo 15

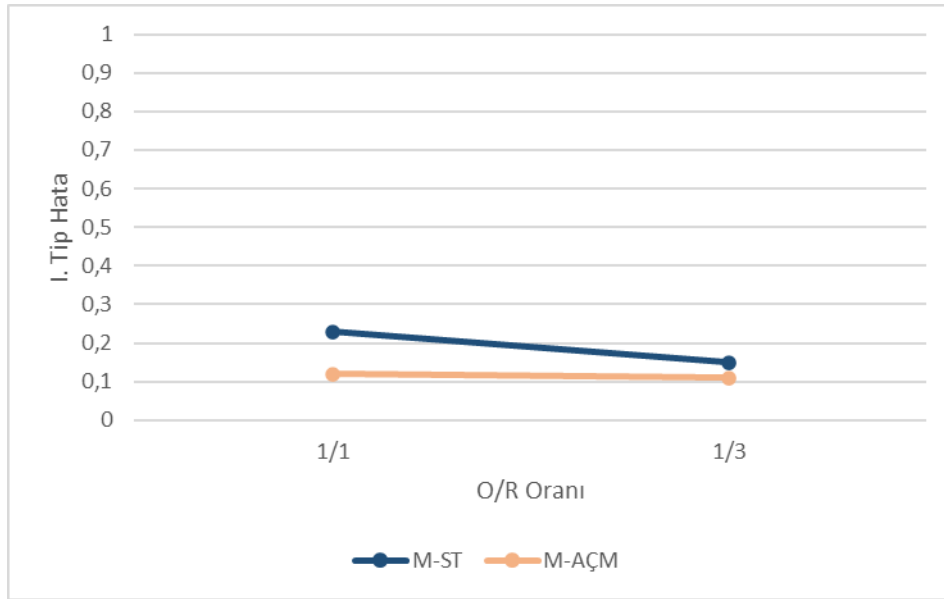
2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları

Yöntem	O/R Oranı				Yöntem	O/R Oranı	Etkileşim
	1/1		1/3				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.23	.18	.15	.09	$F_{1;45}=22.07$	$F_{1;45}=7.42$	$F_{1;45}=4.12$
					p=.00	p=.01	p=.05
M-AÇM	.12	.11	.11	.11	$\omega^2=.03$	$\omega^2=.01$	

Tablo 15 incelendiğinde I. tip hata oranlarının O/R oranına göre anlamlı farklılık gösterdiği görülmektedir. Buna göre O/R oranının I. tip hatadaki değişimi %1 oranında etkilediği söylenebilir. O/R oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 5

2 PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre I. Tip Hata Oranları



Şekil 5 incelendiğinde O/R oranı 1/1 olduğunda her iki yöntem için elde edilen I. tip hata oranlarının O/R oranı 1/3 olduğunda elde edilen I. tip hata oranlarından daha yüksek olduğu görülmektedir. Tüm O/R oranları boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

c) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri DMF'li madde oranlarına (%20 ve %40) göre nasıl değişmektedir?

Tablo 16

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları

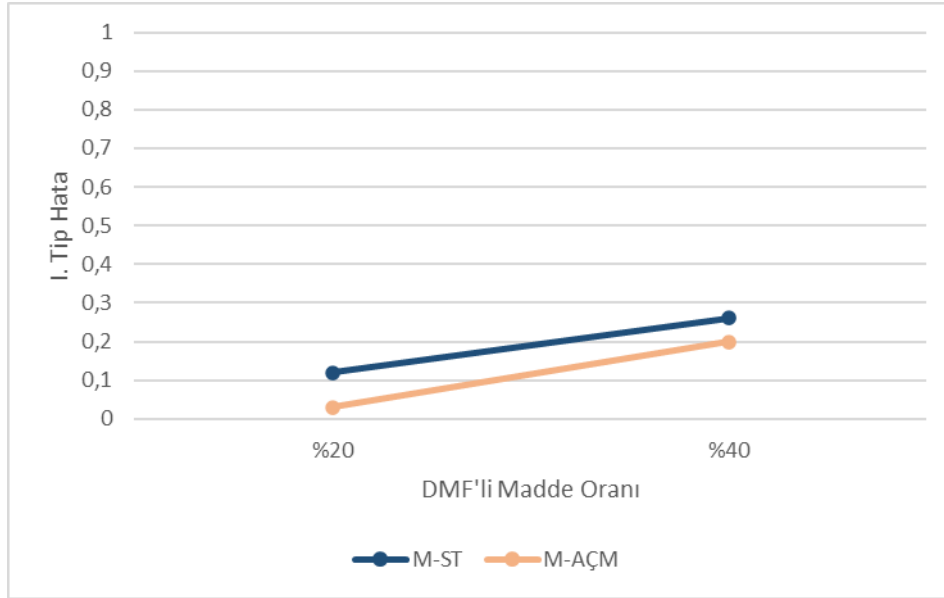
Yöntem	DMF'li Madde Oranı				Yöntem	DMF'li Madde Oranı	Etkileşim
	%20		%40				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.12	.07	.26	.17	$F_{1;45}=22.07$	$F_{1;45}=99.28$	$F_{1;45}=1.78$
M-AÇM	.03	.01	.20	.10	p=.00	p=.00	p=.19
					$\omega^2=.03$	$\omega^2=.14$	

Tablo 16 incelendiğinde I. tip hata oranlarının DMF'li madde oranına göre anlamlı farklılık gösterdiği görülmektedir. Buna göre DMF'li madde oranının I. tip hatadaki değişimi

%14 oranında etkilediği söylenebilir. DMF'li madde oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 6

2 PLM İçin M-ST ve M-AÇM Yöntemlerinin DMF'li Madde Oranlarına Göre I. Tip Hata Oranları



Şekil 6 incelendiğinde DMF'li madde oranı %40 olduğunda her iki yöntem için elde edilen I. tip hata oranlarının DMF'li madde oranı %20 olduğunda elde edilen I. tip hata oranlarından daha yüksek olduğu görülmektedir. Tüm DMF'li madde oranları boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

d) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?

Tablo 17

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları

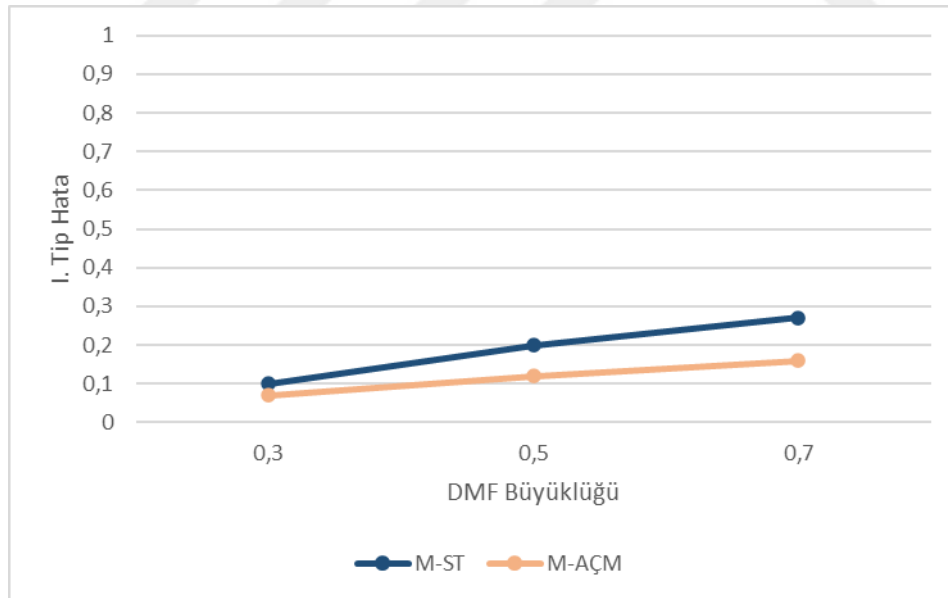
DMF Büyüklüğü	Yöntem	DMF Büyüklüğü	Etkileşim
---------------	--------	---------------	-----------

	0.3		0.5		0.7				
Yöntem	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.10	.04	.20	.12	.27	.19	$F_{1,45}=22.07$	$F_{2,45}=22.89$	$F_{1,45}=2.09$
M-AÇM	.07	.06	.12	.10	.16	.14	p=.00	p=.00	p=.14
							$\omega^2=.03$	$\omega^2=.06$	

Tablo 17 incelendiğinde I. tip hata oranlarının DMF büyüklüğüne göre anlamlı farklılık gösterdiği görülmektedir. Buna göre, DMF büyüklüğünün I. tip hatadaki değişimi %6 oranında etkilediği söylenebilir. DMF büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmemiştir. DMF büyüklüğü açısından ortaya çıkan farklılığın hangi DMF büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm gruplar arasında olduğu tespit edilmiştir.

Şekil 7

2 PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranları



Şekil 7 incelendiğinde DMF büyüklüğü arttıkça her iki yöntemin de I. tip hata oranlarının arttığı görülmektedir. Tüm DMF büyüklükleri boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

Alt Problem 2'ye İlişkin Bulgular

Alt problem 2'ye yanıt bulmak amacıyla 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle elde edilen I. tip hata oranları faktöriyel ANOVA ile incelenmiştir. Elde edilen sonuçlar alt problem 2'nin altındaki problemlere göre ayrı ayrı sunulmuştur. Faktöriyel ANOVA sonuçlarının tablosu Ek-D 'de verilmiştir.

a) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri örneklem büyüklüğüne (2000, 1000 ve 600) göre nasıl değişmektedir?

Tablo 18

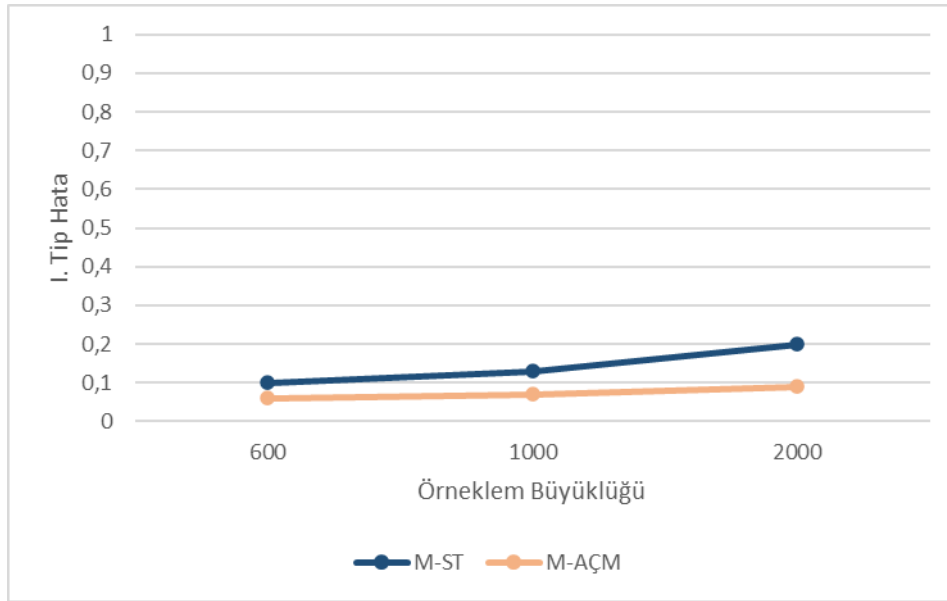
3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları

Yöntem	Örneklem Büyüklüğü						Yöntem	Örneklem Büyüklüğü	Etkileşim
	600		1000		2000				
	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.10	.04	.13	.07	.20	.14	$F_{1;45}=122.04$	$F_{2;45}=34.54$	$F_{2;45}=7.64$
M-AÇM	.06	.04	.07	.06	.09	.09	p=.00	p=.00	p=.00
							$\omega^2=.07$	$\omega^2=.04$	$\omega^2=.01$

Tablo 18 incelendiğinde I. tip hata oranlarının örneklem büyüklüklerine göre anlamlı farklılık gösterdiği görülmektedir. Buna göre, örneklem büyüklüğünün I. tip hatadaki değişimi %4 oranında etkilediği söylenebilir. Örneklem büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmektedir. İkili etkileşim I. tip hatadaki değişimi %1 oranında etkilemektedir. Örneklem büyüklüğü açısından ortaya çıkan farklılığın hangi örneklem büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm gruplar arasında olduğu tespit edilmiştir.

Şekil 8

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre I. Tip Hata Oranları



Şekil 8 incelendiğinde örneklem büyüklüğü arttıkça her iki yöntemin de I. tip hata oranlarının arttığı görülmektedir. Tüm örneklem büyüklükleri boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

b) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri odak/referans grup oranlarına (1/1 ve 1/3) göre nasıl değişmektedir?

Tablo 19

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre I. Tip Hata Oranlarına İlişkin ANOVA Sonuçları

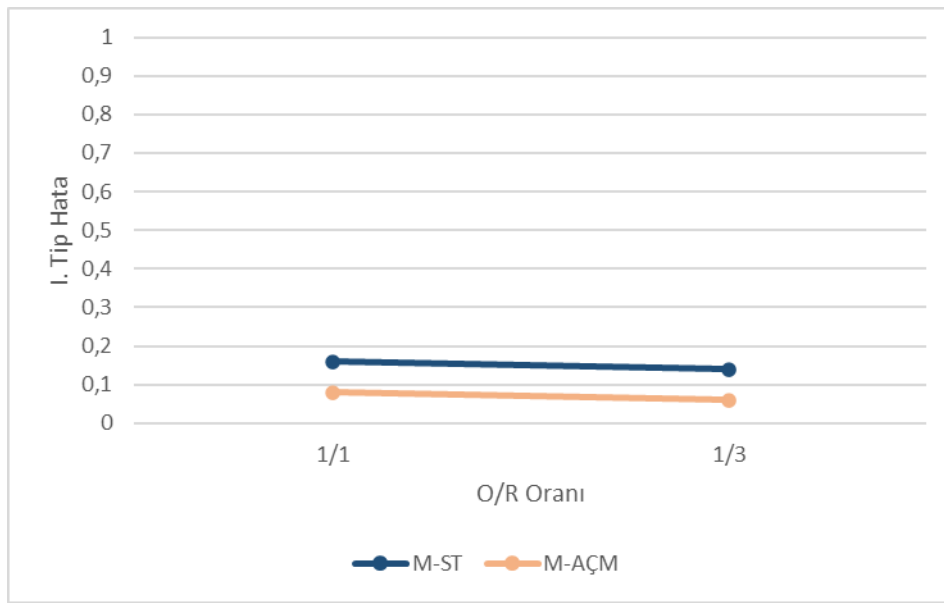
Yöntem	O/R Oranı				Yöntem	O/R Oranı	Etkileşim
	1/1		1/3				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.16	.11	.14	.09	$F_{1;45}=122.04$	$F_{1;45}=9.78$	$F_{2;45}=.04$
M-AÇM	.08	.07	.06	.06	$p=.00$	$p=.00$	$p=.84$
					$\omega^2=.07$		

$$\omega^2=.01$$

Tablo 19 incelendiğinde I. tip hata oranlarının O/R oranına göre ise anlamlı farklılık gösterdiği görülmektedir. Buna göre O/R oranının I. tip hatadaki değişimi %1 oranında etkilediği söylenebilir. O/R oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 9

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı O/R Oranlarına Göre I. Tip Hata Oranları



Şekil 9 incelendiğinde O/R oranı 1/1 olduğunda her iki yöntem için elde edilen I. tip hata oranlarının O/R oranı 1/3 olduğunda elde edilen I. tip hata oranlarından daha yüksek olduğu görülmektedir. Tüm O/R oranları boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

c) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri DMF'li madde oranlarına (%20 ve %40) göre nasıl değişmektedir?

Tablo 20

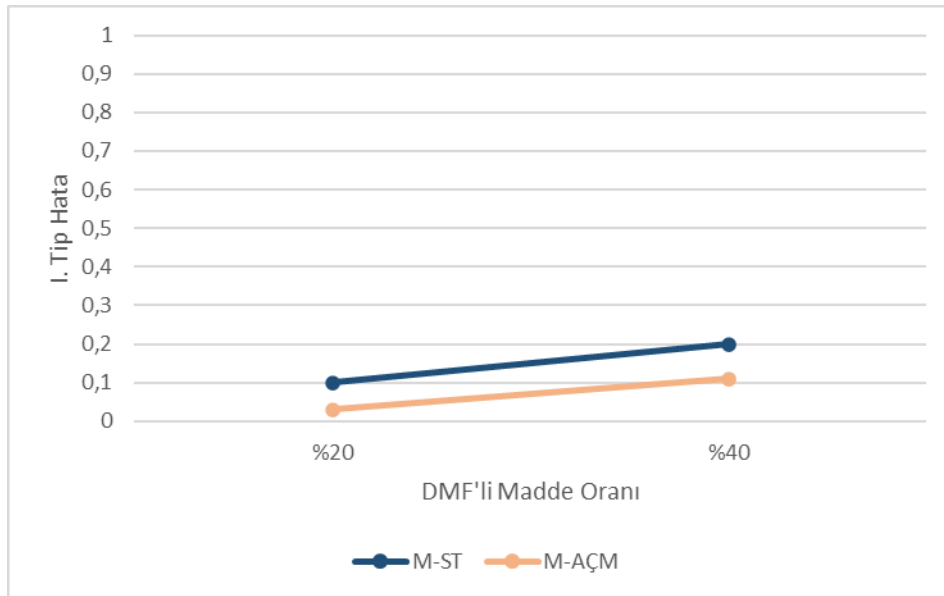
3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre I. Tip Hata Oranlarına İlişkin Sonuçlar

Yöntem	DMF'li Madde Oranı				Yöntem	DMF'li Madde Oranı	Etkileşim
	%20		%40				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.10	.03	.20	.12	$F_{2;45}=122.05$	$F_{1;45}=188.11$	$F_{1;45}=2.64$
M-AÇM	.03	.01	.11	.07	p=.00	p=.00	p=.11
					$\omega^2=.07$	$\omega^2=.10$	

Tablo 20 incelendiğinde I. tip hata oranlarının DMF'li madde oranına göre anlamlı farklılık gösterdiği görülmektedir. Buna göre, DMF'li madde oranının I. tip hatadaki değişimi %10 oranında etkilediği söylenebilir. DMF'li madde oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 10

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre I. Tip Hata Oranları



Şekil 10 incelendiğinde DMF'li madde oranı %40 olduğunda her iki yöntem için elde edilen I. tip hata oranlarının DMF'li madde oranı %20 olduğunda elde edilen I. tip hata oranlarından daha yüksek olduğu görülmektedir. Tüm DMF'li madde oranları boyunca M-

ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

d) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata değerleri DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?

Tablo 21

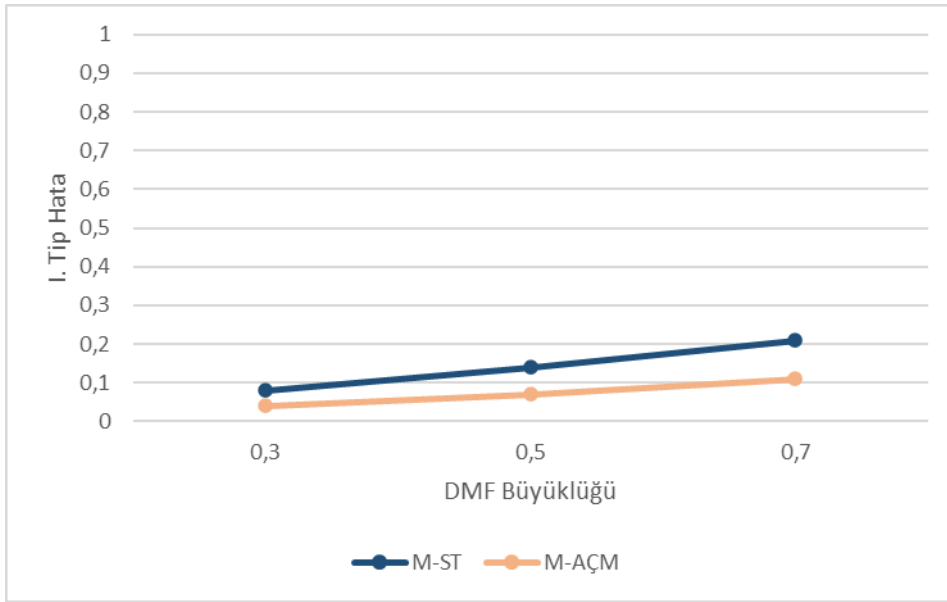
3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranlarına İlişkin Sonuçlar

Yöntem	DMF Büyüklüğü						Yöntem	DMF Büyüklüğü	Etkileşim
	0.3		0.5		0.7				
	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.08	.03	.14	.07	.21	.13	$F_{1;45}=122.05$	$F_{2;45}=68.42$	$F_{2;45}=7.73$
M-AÇM	.04	.02	.07	.06	.11	.09	p=.00 $\omega^2=.07$	p=.00 $\omega^2=.07$	p=.00 $\omega^2=.01$

Tablo 21 incelendiğinde I. tip hata oranlarının DMF büyüklüğüne göre anlamlı farklılık gösterdiği görülmektedir. Buna göre, DMF büyüklüğünün I. tip hatadaki değişimi %7 oranında etkilediği söylenebilir. DMF büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise I. tip hata oranları açısından anlamlı bir farklılık görülmüştür. İkili etkileşim I. tip hatadaki değişimi %1 oranında etkilemektedir. DMF büyüklüğü açısından ortaya çıkan farklılığın hangi DMF büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm DMF büyüklükleri arasında olduğu tespit edilmiştir.

Şekil 11

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre I. Tip Hata Oranları



Şekil 11 incelendiğinde DMF büyüklüğü arttıkça her iki yöntemin de I. tip hata oranlarının arttığı görülmektedir. Tüm DMF büyüklükleri boyunca M-ST yöntemine ait I. tip hata oranlarının M-AÇM yöntemine ait I. tip hata oranlarından daha yüksek olduğu gözlemlenmektedir.

Alt Problem 3'e İlişkin Bulgular

Alt problem 3'e yanıt bulmak amacıyla 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle elde edilen güç oranları faktöriyel ANOVA ile incelenmiştir. Elde edilen sonuçlar alt problem 3'ün altındaki problemlere göre ayrı ayrı sunulmuştur. Faktöriyel ANOVA sonuçlarının tablosu Ek-E'de verilmiştir.

a) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri örneklem büyüklüğüne (2000, 1000 ve 600) göre nasıl değişmektedir?

Tablo 22

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranlarına İlişkin ANOVA Sonuçları

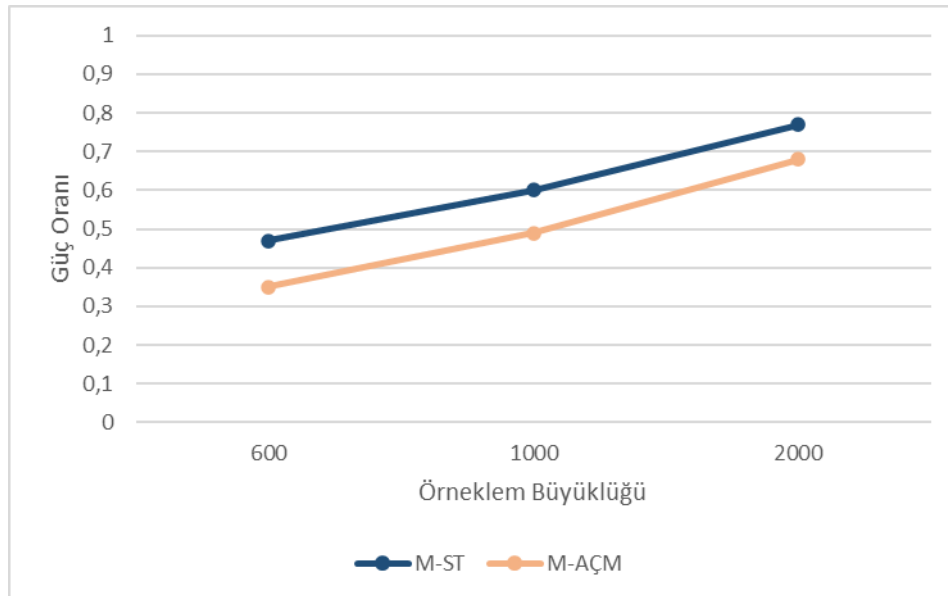
Örneklem Büyüklüğü			Yöntem	Örneklem Büyüklüğü	Etkileşim
600	1000	2000			

Yöntem	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.47	.22	.60	.25	.77	.22	$F_{1;45}=81.06$	$F_{2;45}=232.45$	$F_{1;45}=.32$
M-AÇM	.35	.25	.49	.29	.68	.28	p=.00	p=.00	p=.73
							$\omega^2=.01$	$\omega^2=.04$	

Tablo 22 incelendiğinde güç oranlarının örneklem büyüklüklerine göre anlamlı farklılık gösterdiği görülmektedir. Buna göre örneklem büyüklüğünün güç oranlarındaki değişimi %4 oranında etkilediği söylenebilir. Örneklem büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise güç oranları açısından anlamlı bir farklılık görülmemiştir. Örneklem büyüklüğü açısından ortaya çıkan farklılığın hangi örneklem büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm örneklem büyüklükleri arasında olduğu tespit edilmiştir.

Şekil 12

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranları



Şekil 12 incelendiğinde örneklem büyüklüğü arttıkça her iki yöntemin de güç oranlarının arttığı görülmektedir. Tüm örneklem büyüklükleri boyunca M-ST yöntemine ait

güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

b) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri odak/referans grup oranlarına (1/1 ve 1/3) göre nasıl değişmektedir?

Tablo 23

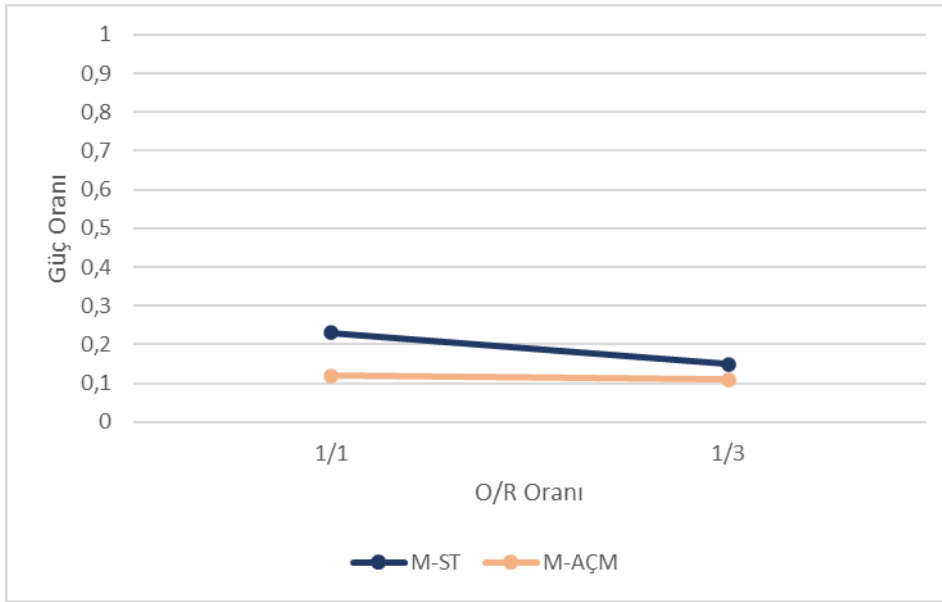
2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre Güç Oranlarına İlişkin ANOVA Sonuçları

Yöntem	O/R Oranı				Yöntem	O/R Oranı	Etkileşim
	1/1		1/3				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.65	.25	.58	.26	$F_{1,45}=81.06$	$F_{1,45}=39.84$	$F_{1,45}=.31$
M-AÇM	.55	.30	.46	.30	$p=.00$ $\omega^2=.01$	$p=.00$ $\omega^2=.00$	$p=.58$

Tablo 23 incelendiğinde güç oranlarının O/R oranına göre anlamlı farklılık gösterdiği görülmektedir. Ancak O/R oranının güç oranındaki değişime etkisi %0 çıkmıştır. O/R oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında da güç oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 13

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı O/R Oranlarına Göre Güç Oranları



Şekil 13 incelendiğinde O/R oranı 1/1 olduğunda her iki yöntem için elde edilen güç oranlarının O/R oranı 1/3 olduğunda elde edilen güç oranlarından daha yüksek olduğu görülmektedir. Tüm O/R oranları boyunca M-ST yöntemine ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

c) 2PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri DMF'li madde oranlarına (%20 ve %40) göre nasıl değişmektedir?

Tablo 24

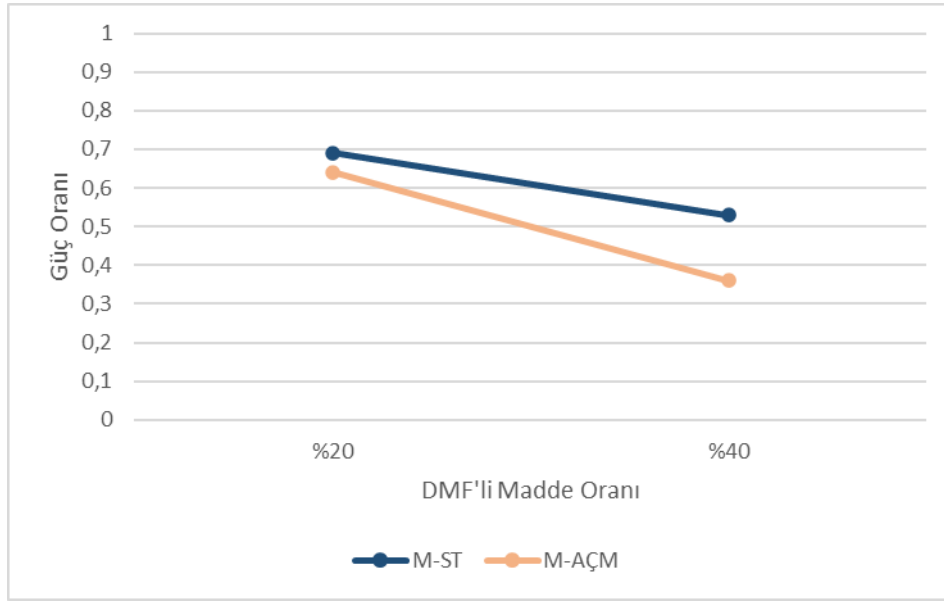
2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranlarına İlişkin ANOVA Sonuçları

Yöntem	DMF'li Madde Oranı				Yöntem	DMF'li Madde Oranı	Etkileşim
	%20		%40				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.69	.25	.53	.25	$F_{1;45}=81.64$	$F_{1;45}=341.18$	$F_{1;45}=23.30$
M-AÇM	.64	.29	.36	.23	p=.00	p=.00	p=.00
					$\omega^2=.01$	$\omega^2=.03$	$\omega^2=.00$

Tablo 24 incelendiğinde güç oranlarının DMF'li madde oranına göre anlamlı farklılık gösterdiği görülmektedir. Buna göre DMF'li madde oranının güç oranındaki değişimi %3 oranında etkilediği söylenebilir. DMF'li madde oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise güç oranları açısından anlamlı bir farklılık görülmüştür. Ancak ikili etkileşim I. tip hatadaki değişimi %0 oranında etkilemektedir.

Şekil 14

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranları



Şekil 14 incelendiğinde DMF'li madde oranı %20 olduğunda her iki yöntem için elde edilen güç oranlarının DMF'li madde oranı %40 olduğunda elde edilen güç oranlarından daha yüksek olduğu görülmektedir. Tüm DMF'li madde oranları boyunca M-ST yöntemine ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

d) 2PLM'ye göre üretilen veri için, M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?

Tablo 25

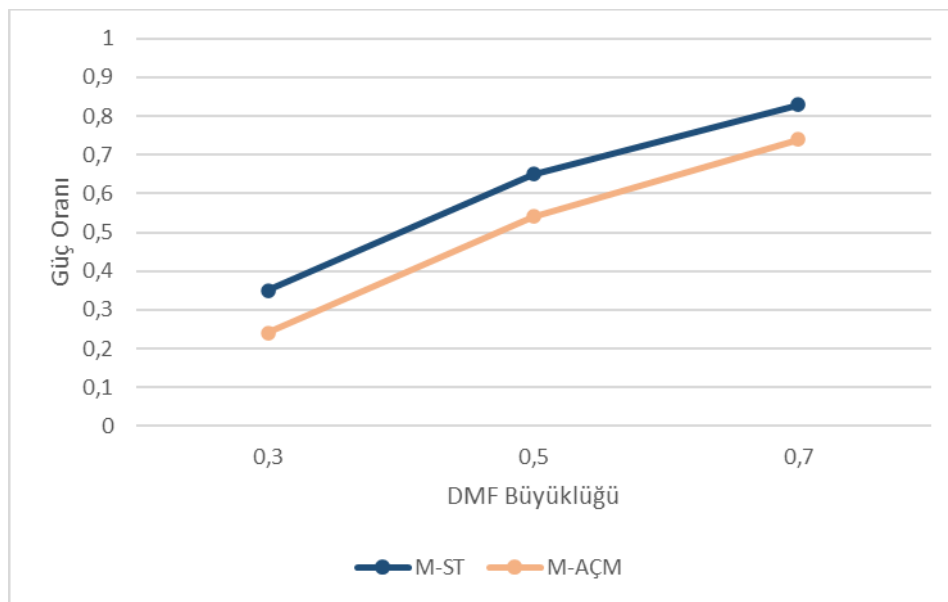
2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranlarına İlişkin ANOVA Sonuçları

Yöntem	DMF Büyüklüğü						Yöntem	DMF Büyüklüğü	Etkileşim
	0.3		0.5		0.7				
	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.35	.15	.65	.19	.83	.14	$F_{1;45}=81.06$	$F_{2;45}=568.19$	$F_{2;45}=.45$
M-AÇM	.24	.17	.54	.26	.74	.21	p=.00 $\omega^2=.01$	p=.00 $\omega^2=.10$	p=.64

Tablo 25 incelendiğinde güç oranlarının DMF büyüklüğüne göre anlamlı farklılık gösterdiği görülmektedir. Buna göre; DMF büyüklüğünün güç oranlarındaki değişimi %10 oranında etkilediği söylenebilir. DMF büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise güç oranları açısından anlamlı bir farklılık görülmemiştir. DMF büyüklüğü açısından ortaya çıkan farklılığın hangi DMF büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm DMF büyüklükleri arasında olduğu tespit edilmiştir.

Şekil 15

2PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranları



Şekil 15 incelendiğinde DMF büyüklüğü arttıkça her iki yöntemin de güç oranlarının arttığı görülmektedir. Tüm DMF büyüklükleri boyunca M-ST yöntemine ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

Alt Problem 4'e İlişkin Bulgular

Alt problem 4'e yanıt bulmak amacıyla 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle elde edilen güç oranları faktöriyel ANOVA ile incelenmiştir. Elde edilen sonuçlar alt problem 3'ün altındaki problemlere göre ayrı ayrı sunulmuştur. Faktöriyel ANOVA sonuçlarının tablosu Ek-F'de verilmiştir.

a) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri örneklem büyüklüğüne (2000, 1000 ve 600) göre nasıl değişmektedir?

Tablo 26

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranlarına İlişkin Sonuçlar

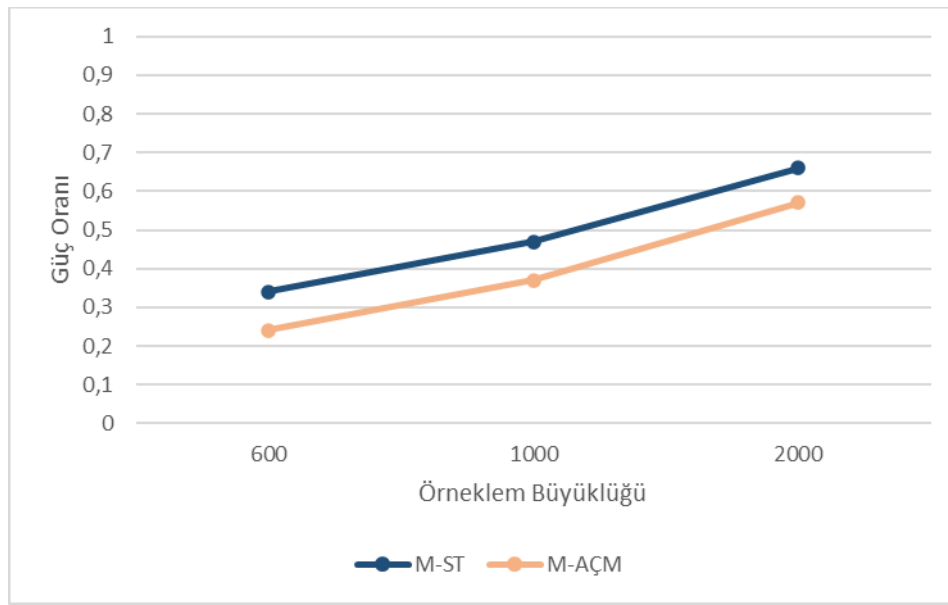
Yöntem	Örneklem Büyüklüğü						Yöntem	Örneklem Büyüklüğü	Etkileşim
	600		1000		2000				
	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.34	.18	.47	.24	.66	.23	$F_{1,45}=55.24$	$F_{2,45}=206.62$	$F_{1,45}=.09$
M-AÇM	.24	.18	.37	.25	.57	.27	p=.00 $\omega^2=.01$	p=.00 $\omega^2=.07$	p=.92

Tablo 26 incelendiğinde güç oranlarının örneklem büyüklüklerine göre anlamlı farklılık gösterdiği görülmektedir. Buna göre örneklem büyüklüğünün güç oranlarındaki değişimi %7 oranında etkilediği söylenebilir. Örneklem büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise güç oranları açısından anlamlı bir farklılık görülmemiştir. Örneklem büyüklüğü açısından ortaya çıkan farklılığın hangi örneklem büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar

yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm örneklem büyüklükleri arasında olduğu tespit edilmiştir.

Şekil 16

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Örneklem Büyüklüklerine Göre Güç Oranları



Şekil 16 incelendiğinde örneklem büyüklüğü arttıkça her iki yöntemin de güç oranlarının arttığı görülmektedir. Tüm örneklem büyüklükleri boyunca M-ST yöntemine ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

b) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri odak/referans grup oranlarına (1/1 ve 1/3) göre nasıl değişmektedir?

Tablo 27

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı Odak/Referans Grup Oranlarına Göre Güç Oranlarına İlişkin Sonuçlar

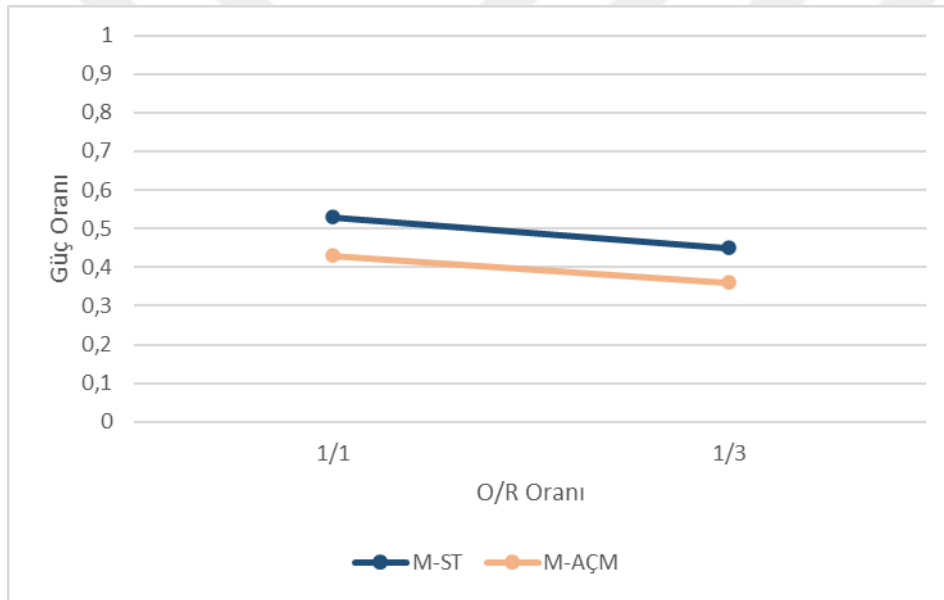
O/R Oranı						
1/1		1/3		Yöntem	O/R Oranı	Etkileşim
Yöntem	$\bar{X}$	SS	$\bar{X}$	SS		

M-ST	.53	.25	.45	.25	$F_{1;45}=55.24$	$F_{1;45}=27.30$	$F_{1;45}=.02$
M-AÇM	.43	.28	.36	.27	$p=.00$	$p=.00$	$p=.89$
					$\omega^2=.01$	$\omega^2=.01$	

Tablo 27 incelendiğinde güç oranlarının O/R oranına göre anlamlı farklılık gösterdiği görülmektedir. Buna göre O/R oranının güç oranlarındaki değişimi %1 oranında etkilediği söylenebilir. O/R oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında güç oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 17

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı O/R Oranlarına Göre Güç Oranları



Şekil 17 incelendiğinde O/R oranı 1/1 olduğunda her iki yöntem için elde edilen güç oranlarının O/R oranı 1/3 olduğunda elde edilen güç oranlarından daha yüksek olduğu görülmektedir. Tüm O/R oranları boyunca M-ST yöntemine ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

c) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri DMF'li madde oranlarına (%20 ve %40) göre nasıl değişmektedir?

Tablo 28

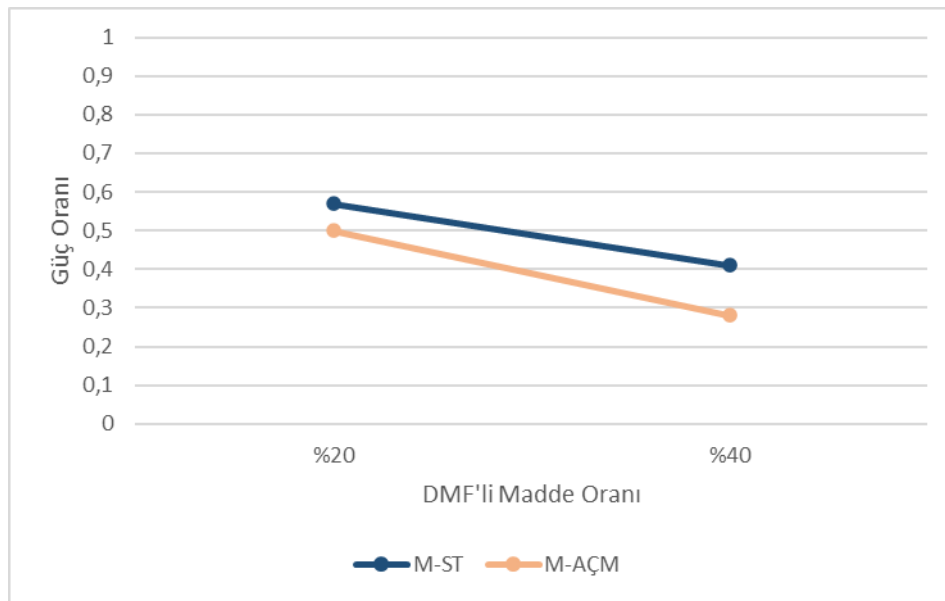
3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranlarına İlişkin ANOVA Sonuçları

Yöntem	DMF'li Madde Oranı				Yöntem	DMF'li Madde Oranı	Etkileşim
	%20		%40				
	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.57	.25	.41	.23	$F_{1,45}=55.24$	$F_{1,45}=205.74$	$F_{1,45}=4.25$
M-AÇM	.50	.01	.28	.21	p=.00	p=.00	p=.05
					$\omega^2=.01$	$\omega^2=.03$	

Tablo 28 incelendiğinde güç oranlarının DMF'li madde oranına göre anlamlı farklılık gösterdiği görülmektedir. Buna göre DMF'li madde oranının güç oranındaki değişimi %3 oranında etkilediği söylenebilir. DMF'li madde oranı ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise güç oranları açısından anlamlı bir farklılık görülmemiştir.

Şekil 18

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF'li Madde Oranlarına Göre Güç Oranları



Şekil 18 incelendiğinde DMF'li madde oranı %20 olduğunda her iki yöntem için elde edilen güç oranlarının DMF'li madde oranı %40 olduğunda elde edilen güç oranlarından daha yüksek olduğu görülmektedir. Tüm DMF'li madde oranları boyunca M-ST yöntemine

ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

d) 3PLM'ye göre üretilen veri için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen güç değerleri DMF büyüklüğüne (0.3, 0.5 ve 0.7) göre nasıl değişmektedir?

Tablo 29

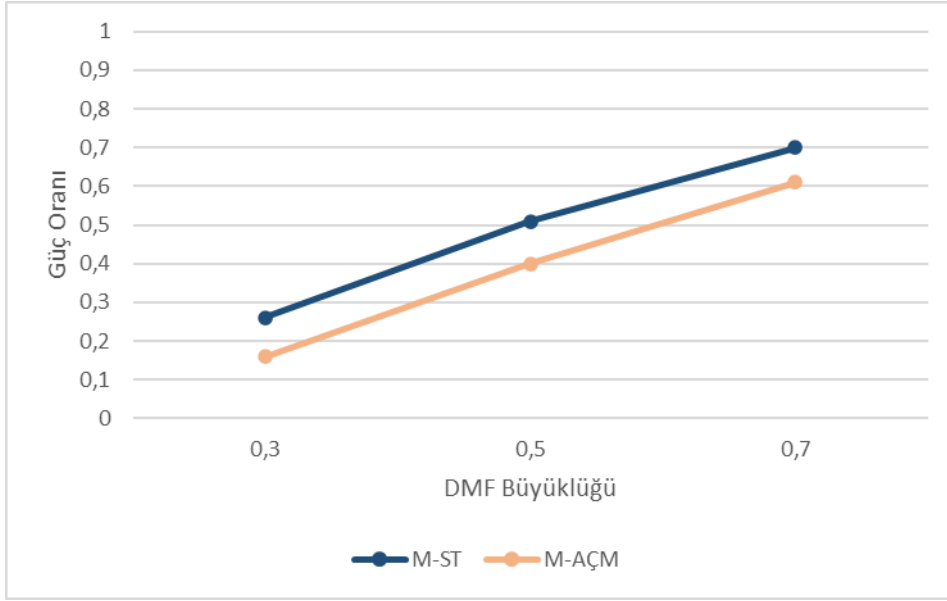
3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranlarına İlişkin Sonuçlar

Yöntem	DMF Büyüklüğü						Yöntem	DMF Büyüklüğü	Etkileşim
	0.3		0.5		0.7				
	$\bar{X}$	SS	$\bar{X}$	SS	$\bar{X}$	SS			
M-ST	.26	.13	.51	.20	.70	.19	$F_{1;45}=55.24$	$F_{2;45}=370.26$	$F_{1;45}=.31$
M-AÇM	.16	.11	.40	.23	.61	.24	p=.00 $\omega^2=.01$	p=.00 $\omega^2=.13$	p=.74

Tablo 29 incelendiğinde güç oranlarının DMF büyüklüğüne göre anlamlı farklılık gösterdiği görülmektedir. Buna göre DMF büyüklüğünün güç oranlarındaki değişimi %13 oranında etkilediği söylenebilir. DMF büyüklüğü ve DMF belirleme yöntemlerinin ortak etkileşimine bakıldığında ise güç oranları açısından anlamlı bir farklılık görülmemiştir. DMF büyüklüğü açısından ortaya çıkan farklılığın hangi DMF büyüklükleri arasında olduğunu belirlemek için Post-Hoc analiziyle ikili karşılaştırmalar yapılmıştır. Bu karşılaştırmalar sonucunda anlamlı farklılığın tüm DMF büyüklükleri arasında olduğu tespit edilmiştir.

Şekil 19

3PLM İçin M-ST ve M-AÇM Yöntemlerinin Farklı DMF Büyüklüklerine Göre Güç Oranları



Şekil 19 incelendiğinde DMF büyüklüğü arttıkça her iki yöntemin de güç oranlarının arttığı görülmektedir. Tüm DMF büyüklükleri boyunca M-ST yöntemine ait güç oranlarının M-AÇM yöntemine ait güç oranlarından daha yüksek olduğu gözlemlenmektedir.

Alt Problem 5'e İlişkin Bulgular

PISA 2012 matematik testinde, M-ST ve M-AÇM yöntemlerine göre belirlenen ortak DMF'li maddeler ve bunların sayısı örneklem büyüklüğüne (900 ve 405) ve odak/referans grup oranlarına (350/550, 450/450, 225/225 ve 150/300) nasıl değişmektedir?

Bu alt probleme yanıt bulmak amacıyla çalışma grubu 1, çalışma grubu 2, çalışma grubu 3 ve çalışma grubu 4 için M-ST ve M-AÇM yöntemleriyle DMF analizi yapılmıştır. Analizler sonucunda elde edilen bulgular Tablo 30'da verilmiştir. Madde bazında DMF analizi sonuçları EK-G, EK-H, EK-I ve EK-İ'de sunulmuştur.

Tablo 30

M-ST ve M-AÇM Yöntemleri için Farklı Koşullara Göre Elde Edilen DMF'li Madde Sayıları ve Numaraları

Örneklem Büyüklüğü	O/R Oranı	M-ST		M-AÇM	
		DMF'li Madde Sayısı	DMF'li Madde Numaraları	DMF'li Madde Sayısı	DMF'li Madde Numaraları

900	350/550	11	3, 8, 9, 14, 15, 16, 22, 24, 26, 28, 31	7	3, 8, 9, 14, 22, 24, 26
	450/450	10	3, 6, 8, 14, 15, 22, 24, 26, 31, 34	7	6, 8, 14, 15, 22, 24, 26
450	225/225	9	8, 9, 14, 15, 20, 22, 26, 31, 34	6	9, 20, 22, 26, 31, 34
	150/300	6	3, 8, 14, 24, 26, 28	4	3, 11, 24, 26

Tablo 30 incelendiğinde örneklem büyüklüğü 900 olduğunda M-ST yöntemiyle belirlenen DMF'li madde sayılarının 10 ve 11, M-AÇM yöntemiyle belirlenen DMF'li madde sayılarının 7 olduğu görülmektedir. Bu örneklem büyüklüğünde tüm koşullar için 8, 14, 22, 24 ve 26 nolu maddeler hem M-ST ve hem de M-AÇM yöntemiyle DMF'li olarak tespit edilmiştir. Örneklem büyüklüğü 450 olduğunda M-ST yöntemiyle belirlenen DMF'li madde sayısının 6 ve 9, M-AÇM yöntemiyle belirlenen DMF'li madde sayısının 4 ve 6 olduğu görülmektedir. Bu örneklem büyüklüğünde tüm koşullar için 26 nolu madde hem M-ST ve hem de M-AÇM yöntemiyle DMF'li olarak tespit edilmiştir.

Tablo 30 odak/referans grubu oranları açısından incelendiğinde O/R oranı 350/500 olduğunda M-ST yöntemiyle belirlenen DMF'li madde sayısının 11, M-AÇM yöntemiyle belirlenen DMF'li madde sayısının 7 olduğu görülmektedir. O/R oranı 450/450 olduğunda M-ST yöntemiyle belirlenen DMF'li madde sayısının 10, M-AÇM yöntemiyle belirlenen DMF'li madde sayısının 7 olduğu gözlenmektedir. O/R oranı 225/225 olduğunda M-ST yöntemiyle belirlenen DMF'li madde sayısı 9, M-AÇM yöntemiyle belirlenen DMF'li madde sayısı 6 olarak bulunmuştur. O/R oranı 150/300 olduğunda M-ST yöntemiyle belirlenen DMF'li madde sayısının 6, M-AÇM yöntemiyle belirlenen DMF'li madde sayısının 4 olduğu gözlenmektedir. 26 nolu madde tüm O/R oranlarında DMF'li olarak belirlenmiştir.

Tartışma

Bu bölümde simülasyon verisi üzerinde M-ST ve M-AÇM yöntemleriyle uygulanan DMF analizleri sonucunda elde edilen I. tip hata oranları ve güç değerlerinin 2PLM ve 3PLM'e göre örneklem büyüklüğü, O/R oranı, DMF'li madde oranı ve DMF büyüklüğü açısından nasıl sonuçlar verdiği ve gerçek veri üzerinde M-ST ve M-AÇM yöntemleriyle uygulanan DMF analizleri sonucunda belirlenen DMF'li madde sayılarının örneklem büyüklüğü ve O/R oranına göre nasıl sonuçlar verdiği tartışılmıştır.

Simülasyon verisiyle yürütülen çalışmadan elde edilen sonuçlara bakıldığında, M-ST yöntemine ait I. tip hata oranlarının hem 2PLM'ye hem de 3PLM'ye göre üretilen verilerde tüm koşullar için 0.05'ten yüksek çıktığı görülmüştür. Bu durumun, Finch (2005)'in de belirttiği gibi testin 20 maddeden oluşan kısa bir test olmasından kaynaklandığı söylenebilir. Finch (2005) çalışmasında veri modeli 3PLM ve test uzunluğu 20 madde olduğu durumda MIMIC yöntemi için şişirilmiş I. tip hata oranlarıyla karşılaştığını ve bu durumun eşleştirme kriterindeki DMF'li maddelerin varlığı nedeniyle ortaya çıktığını raporlamıştır. Zimbra (2018), 3PLM'ye göre ürettiği veriyle yaptığı çalışmada M-ST yöntemiyle elde ettiği I. tip hata değerlerinin tüm koşullar için 0.05'ten büyük olduğunu belirtmiştir. Wang ve ark. (2009), test uzunluğu 20 iken hem 2PLM hem de 3PLM'ye uygun verilerde tüm koşullar için elde ettikleri I. tip hata değerlerinin 0.05'ten yüksek olduğu raporlamıştır. Bu çalışmada hem 2PLM'ye hem de 3PLM'ye göre üretilen veriler için M-ST yöntemiyle elde edilen yüksek I. tip hata oranlarının sebebinin, testin 20 maddelik kısa bir test olmasından ve eşleştirme kriterindeki DMF'li maddelerden kaynakladığı yorumu yapılabilir.

M-AÇM yöntemi için hem 2PLM hem de 3PLM için O/R 1/1 ve DMF'li madde oranı %20 olduğunda diğer koşullardan bağımsız olarak tüm durumlarda I. tip hata değerlerinin 0.05'in altında çıktığı görülmüştür. Testteki DMF'li madde oranı %20'den %40'a çıktığında ve DMF büyüklüğü arttığında M-AÇM yöntemi için elde edilen I. tip hata oranlarında artış meydana geldiği tespit edilmiştir. Shih ve Wang (2009) çalışmasındaki tüm koşullar için M-

AÇM yöntemiyle elde edilen I. tip hata oranlarının kritik değerin altında veya kritik değere yakın değerler alarak kontrol altına alındığını belirtmişlerdir. Shih ve Wang (2009) çalışmasında DMF büyüklüğünü 0.4 olarak belirleyip sabit koşul olarak ele almıştır. Bu çalışmada testteki DMF'li madde oranının %40 olduğu durumda M-AÇM yöntemiyle elde edilen I. tip hata oranlarının DMF büyüklüğü yükseldikçe arttığı belirlenmiştir.

M-ST ve M-AÇM yöntemleri için elde edilen güç oranları incelendiğinde, güç oranlarının modelden, örneklem büyüklüğünden ve DMF büyüklüğünden etkilendiği gözlenmektedir. Buna göre, modelden bağımsız olarak her iki yöntem için de örneklem büyüklüğünün ya da DMF büyüklüğünün artmasıyla güç oranlarının da arttığı görülmektedir. Sonuçlar bu anlamda Zimbra'nın (2018) çalışmasıyla paralellik göstermektedir. Her iki yöntem için de 2PLM'ye göre elde edilen güç değerlerinin 3PLM'ye göre elde edilen güç değerlerinden daha yüksek olduğu görülmüştür. Finch (2005) M-ST yöntemine ait en yüksek güç değerlerini 2PLM'ye uygun veriye ait koşullarda, Shih ve Wang (2009) M-AÇM yöntemine ait en yüksek güç değerlerini 2PLM'ye uygun veriye ait koşullarda bulmuştur. Bu çalışmada 2PLM için 3PLM'ye göre daha yüksek güç oranları elde edildiğinden sonuçların Finch (2005) ve Shih ve Wang'ın (2009) çalışmalarıyla tutarlı olduğu yorumu yapılabilir. MIMIC modellerinin 2PLM ile ilişkili olması sebebiyle bu sonucun beklenen bir durum olduğu söylenebilir.

Yapılan incelemeler sonucunda örneklem büyüklüğünün hem 2PLM hem de 3PLM için I. tip hata oranları açısından düşük düzeyde anlamlı etki yarattığı tespit edilmiştir. Örneklem büyüklüğü arttıkça her iki yönteme ait I. tip hata oranlarının arttığı gözlenmiştir. Bu sonuç, örneklem büyüklüğü arttıkça DMF belirleme yöntemlerinden elde edilen I. tip hata oranlarının da arttığı diğer çalışmalarla paralellik göstermektedir (Finch & French, 2007; Li ve ark., 2012; Wood, 2011). İncelemeler sonucunda örneklem büyüklüğünün güç oranı açısından 2PLM için düşük düzeyde, 3PLM için orta düzeyde anlamlı etki yarattığı belirlenmiştir. Örneklem büyüklüğü arttıkça güç oranlarının da arttığı gözlenmiştir. Bu durum

Shih ve Wang (2009), Wang ve ark., (2013) ve Woods ve Grimm (2011)'in çalışmalarında elde edilen bulgularla tutarlılık göstermektedir.

Çalışmanın bulgularında O/R oranının hem 2PLM hem de 3PLM için, I. tip hata ve güç oranları açısından düşük düzeyde anlamlı etki yarattığı tespit edilmiştir. Ancak 2PLM için O/R oranının güç oranları üzerindeki etki büyüklüğü %0 çıktığından bu etki yok sayılabilir. O/R oranının 1/1 olduğu durumda elde edilen I. tip hata ve güç değerlerinin O/R oranının 1'den farklılaştığı (1/3) durumda elde edilen değerlerden daha yüksek olduğu görülmüştür. Bu sonuç, odak ve referans grup büyüklükleri eşit olduğunda elde edilen I. tip hata ve güç değerlerinin odak ve referans grup büyüklüklerinin eşit olmadığı durumda elde edilen değerlerden daha yüksek olduğunu belirten çalışmalarla paralellik göstermektedir (Atalay Kabasakal ve ark., 2014; Ayva Yörü, 2019; Kristjansson ve ark., 2005). Zimbra (2018), eşit olmayan odak ve referans grubu örneklemeleri için elde edilen güç oranlarının eşit olan odak ve referans grubu örneklemeleri için elde edilen güç oranlarından daha düşük olduğunu ancak bu farkın küçük bir fark olduğunu belirtmiştir.

Bulgular DMF'li madde oranı açısından incelendiğinde DMF'li madde oranının 2PLM için I. tip hata üzerinde yüksek, 3PLM için I. tip hata üzerinde orta düzeyde anlamlı etkiye sahip olduğu görülmektedir. Güç oranı açısından bakıldığında DMF'li madde oranının hem 2PLM hem de 3PLM için düşük düzeyde anlamlı etkiye sahip olduğu belirlenmiştir. Shih ve Wang (2009) çalışmasında DMF'li madde oranının M-AÇM yöntemi için elde edilen güç oranları üzerinde düşük düzeyde etkisinin olduğunu raporlamışlardır. Çalışmalarında DMF'li madde oranı arttığında I. tip hata ve güç oranlarının kabul edilebilir değerler almaya devam ettiğini bulmuşlardır. Bu çalışmanın sonuçları bu anlamda Shih ve Wang (2009)'ın çalışmasıyla tutarlılık göstermemektedir. Bu çalışmada DMF'li madde oranının güç oranı üzerinde orta ve yüksek düzeyde etkisinin olduğu ve DMF'li madde oranı arttıkça I. tip hata değerlerinin arttığı, güç değerlerinin ise düştüğü gözlemlenmiştir. Zimbra (2018), Wang ve Shih (2010) DMF'li madde oranı arttıkça M-ST yöntemi için elde edilen I. tip hata oranlarının arttığını ve gücün düştüğünü belirtmişlerdir. Bu çalışmanın M-ST yöntemi için elde ettiği

bulgular açısından Zimbra (2018), Wang ve Shih (2010)'un çalışmalarıyla tutarlılık gösterdiği söylenebilir. M-ST yöntemi için kaynaklanan bu durumun test uzunluğunun kısa olmasından ve eşleştirme değişkeninde DMF'li maddelerin olabilmesi ihtimalinden kaynaklandığı söylenebilir.

DMF büyüklüğü bağlamında yapılan incelemeler doğrultusunda, 2PLM ve 3PLM için DMF büyüklüğünün hem I. tip hata oranı hem de güç oranı üzerinde orta düzeyde anlamlı etkiye sahip olduğu görülmektedir. Çalışma sonuçlarına göre DMF büyüklüğü arttıkça I. tip hata oranları ve güç oranları da artmaktadır. Hou ve ark. (2014), Lee ve ark. (2017), Zimbra (2018) çalışmalarında DMF büyüklüğü arttıkça diğer koşullardan bağımsız olarak güç oranlarının da arttığını belirtmiştir. Bu durum DMF büyüklüğü arttıkça DMF'li maddelerin belirlenmesinin kolaylaştığı fikrini desteklemektedir. Ayva Yörü (2019), Su ve Wang, (2005), Wang ve Su, (2004) çalışmalarında DMF büyüklüğü arttıkça I. tip hatanın arttığını belirtmişlerdir.

Gerçek veri üzerinde yapılan analizler incelendiğinde dört çalışma gurubu için de M-AÇM yöntemiyle daha az sayıda DMF'li madde tespit edilmiştir. M-AÇM yöntemiyle yapılan analizlerde tüm koşullarda 26 nolu madde, M-ST yöntemiyle yapılan analizlerde tüm koşullarda 8, 14 ve 26 nolu maddelerin DMF'li olarak belirlendiği gözlemlenmektedir. Buna göre tüm koşullar için her iki yöntemle yapılan analizlere bakıldığında 26 nolu maddenin tüm durumlarda DMF'li olarak belirlendiği görülmektedir. En yüksek DMF'li madde sayısının M-ST yöntemiyle 900 kişilik örneklem ve 350/550 O/R oranı koşulu için yapılan analizde elde edildiği belirlenmiştir. Simülasyon verisiyle yapılan çalışma sonucunda M-ST yöntemiyle elde edilen I. tip hata oranlarının, ele alınan koşulların tümünde kabul edilebilir değerlerin üstünde çıktığı görülmüştür. Buna göre, gerçek veriyle yapılan çalışmada M-ST yöntemiyle elde edilen DMF'li madde sayısının M-AÇM yöntemiyle elde edilen DMF'li madde sayısından yüksek olması durumunun M-ST yönteminin I. tip hata oranlarının daha yüksek olmasından kaynaklandığı yorumu yapılabilir.

Bölüm 5

Sonuç ve Öneriler

Araştırmanın bu bölümünde simülasyon verisine ait sonuçlar, gerçek veriden elde edilen sonuçlar ve bu sonuçlara göre ortaya koyulan öneriler yer almaktadır.

Sonuçlar

Simülasyon Verilerine İlişkin Sonuçlar

Araştırmada 2PLM ve 3PLM'ye göre üretilen veriler için M-ST ve M-AÇM yöntemleriyle belirlenen DMF sonuçlarından elde edilen I. tip hata ve güç değerlerinin örneklem büyüklüğüne, odak/referans grubu oranlarına, testteki DMF'li madde oranına ve DMF büyüklüğüne göre nasıl değiştiğini incelemek ve yöntemlerin hangi koşullarda daha iyi performans sergilediğini belirlemek amaçlanmıştır. Bu amaç doğrultusunda elde edilen bulgulara dayalı sonuçlar aşağıda sunulmuştur.

1) Hem 2PLM hem de 3PLM için M-ST yöntemiyle elde edilen I. tip hata oranları tüm koşullar için 0.05'ten yüksek çıkmıştır.

2) M-AÇM yönteminin 2PLM'de örneklem büyüklüğü 2000, O/R oranı 1/1, DMF'li madde oranı %20, DMF büyüklüğü 0.5 ve 0.7; örneklem büyüklüğü 2000, O/R oranı 1/3, DMF'li madde oranı %20, DMF büyüklüğü 0.5 ve 0.7; örneklem büyüklüğü 1000, O/R oranı 1/1 ve 1/3, DMF'li madde oranı %20, DMF büyüklüğü 0.7; örneklem büyüklüğü 600, O/R oranı 1/1, DMF'li madde oranı %20 ve DMF büyüklüğü 0.7 olmak üzere toplam 7 koşulda; 3PLM'de örneklem büyüklüğü 2000, O/R oranı 1/1, DMF'li madde oranı %20, DMF büyüklüğü 0.5 ve 0.7; örneklem büyüklüğü 2000, O/R oranı 1/3, DMF'li madde oranı %20, DMF büyüklüğü 0.7; örneklem büyüklüğü 1000, O/R oranı 1/1, DMF'li madde oranı %20 ve DMF büyüklüğü 0.7 olmak üzere bu 4 koşulda hem I. tip hata oranlarının hem de güç oranlarının kabul edilebilir değerler aldığı görülmüştür. Buna göre M-AÇM yöntemi bu 11 koşul için M-ST'ye göre daha iyi performans göstermiştir. M-ST yöntemi için hem I. tip hata oranının hem de güç oranının kabul edilebilir değerler aldığı bir koşul gözlemlenmemiştir.

3) Hem 2PLM hem de 3PLM için I. tip hata oranlarının örneklem büyüklüklerine göre anlamlı farklılık gösterdiği belirlenmiştir. Örneklem büyüklüğünün I. tip hata üzerinde düşük düzeyde etkiye sahip olduğu görülmüştür. Örneklem büyüklükleri arttıkça M-ST ve M-AÇM yöntemlerine ait I. tip hata oranlarının arttığı gözlemlenmiştir.

4) Hem 2PLM hem de 3PLM için I. tip hata oranlarının O/R oranlarına göre anlamlı farklılık gösterdiği belirlenmiştir. O/R oranının I. tip hata üzerinde düşük düzeyde etkiye sahip olduğu görülmüştür. O/R oranının 1/1 olduğu durumda elde edilen I. tip hata değerlerinin O/R oranının 1'den farklılaştığı (1/3) durumda elde edilen değerlerden daha yüksek olduğu görülmüştür.

5) Hem 2PLM hem de 3PLM için I. tip hata oranlarının DMF'li madde oranına göre anlamlı farklılık gösterdiği belirlenmiştir. DMF'li madde oranının I. tip hata üzerinde 2PLM için yüksek, 3PLM için orta düzeyde etkiye sahip olduğu görülmüştür. DMF'li madde oranı arttıkça M-ST ve M-AÇM yöntemlerine ait I. tip hata oranlarının arttığı gözlemlenmiştir.

6) Hem 2PLM hem de 3PLM için I. tip hata oranlarının DMF büyüklüğüne göre anlamlı farklılık gösterdiği belirlenmiştir. DMF büyüklüğünün I. tip hata üzerinde orta düzeyde etkiye sahip olduğu görülmüştür. DMF büyüklüğü arttıkça M-ST ve M-AÇM yöntemlerine ait I. tip hata oranlarının arttığı gözlemlenmiştir.

7) Hem 2PLM hem de 3PLM için güç oranlarının örneklem büyüklüklerine göre anlamlı farklılık gösterdiği belirlenmiştir. Örneklem büyüklüğünün güç üzerinde 2PLM için düşük, 3PLM için orta düzeyde etkiye sahip olduğu görülmüştür. Örneklem büyüklükleri arttıkça M-ST ve M-AÇM yöntemlerine ait güç oranlarının arttığı gözlemlenmiştir.

8) Hem 2PLM hem de 3PLM için güç oranlarının O/R oranlarına göre anlamlı farklılık gösterdiği belirlenmiştir. O/R oranının güç üzerinde düşük düzeyde etkiye sahip olduğu görülmüştür. O/R oranının 1/1 olduğu durumda elde edilen güç değerlerinin O/R oranının 1'den farklılaştığı (1/3) durumda elde edilen değerlerden daha yüksek olduğu görülmüştür.

9) Hem 2PLM hem de 3PLM için güç oranlarının DMF'li madde oranına göre anlamlı farklılık gösterdiği belirlenmiştir. DMF'li madde oranının güç üzerinde düşük düzeyde etkiye sahip olduğu görülmüştür. DMF'li madde oranı arttıkça M-ST ve M-AÇM yöntemlerine ait güç oranlarının azaldığı gözlemlenmiştir.

10) Hem 2PLM hem de 3PLM için güç oranlarının DMF büyüklüğüne göre anlamlı farklılık gösterdiği belirlenmiştir. DMF büyüklüğünün güç üzerinde orta düzeyde etkiye sahip olduğu görülmüştür. DMF büyüklüğü arttıkça M-ST ve M-AÇM yöntemlerine ait güç oranlarının arttığı gözlemlenmiştir.

11) Bağımsız değişkenlerin I. tip hata üzerindeki etkilerine bakıldığında 2PLM'de DMF'li madde oranının, 3PLM'de DMF'li madde oranı ve DMF büyüklüğünün diğer değişkenlere göre daha yüksek etkiye sahip olduğu görülmektedir.

12) Tüm koşullar için bağımsız değişkenlerin güç üzerindeki etkilerine bakıldığında 2PLM'de DMF büyüklüğünün, 3PLM'de örneklem büyüklüğü ve DMF büyüklüğünün diğer değişkenlere göre daha yüksek etkiye sahip olduğu görülmektedir.

12) Yöntemin; örneklem büyüklüğü, O/R oranı, DMF'li madde oranı ve DMF büyüklüğüyle ikili ortak etkileşimlerine bakıldığında 2PLM'de etkileşimlerin hiçbirinin I. tip hata oranları üzerinde anlamlı bir etkisinin olmadığı, 3PLM'de örneklem büyüklüğü ve DMF büyüklüğüyle olan ikili etkileşimlerin I. tip hata oranları üzerinde anlamlı bir etkisi olduğu bulunmuştur. Güç oranları açısından bakıldığında 2PLM'de yöntemin DMF'li madde oranıyla ikili etkileşiminin güç oranları üzerinde anlamlı bir etkisi olduğu, 3PLM'de etkileşimlerin hiçbirinin güç oranları üzerinde anlamlı bir etkisinin olmadığı bulunmuştur.

Gerçek Verilere İlişkin Sonuçlar

Araştırmada, PISA 2012 uygulamasında 4 numaralı kitapçıkta yer alan 34 adet çoktan seçmeli matematik sorusu için M-ST ve M-AÇM yöntemleriyle DMF analizi uygulanmış ve elde edilen sonuçların örneklem büyüklüğü ve O/R grubu oranlarına göre

nasıl farklılaştığını incelemek amaçlanmıştır. Bu amaç doğrultusunda elde edilen bulgulara dayalı sonuçlar aşağıda sunulmuştur.

1) M-ST yöntemiyle yapılan analizlerde, örneklem büyüklüğü 450 iken O/R oranı 1 olduğu durumda belirlenen DMF'li madde sayısının O/R oranı 1/3 olduğu durumda belirlenen DMF'li madde sayısından daha fazla olduğu bulunmuştur. Örneklem büyüklüğü 900 iken, O/R oranı 1 olduğu durumda belirlenen DMF'li madde sayısının daha fazla olduğu görülmektedir.

2) M-AÇM yöntemiyle yapılan analizlerde, örneklem büyüklüğü 450 iken O/R oranı 1 olduğu durumda belirlenen DMF'li madde sayısının O/R oranı 1/3 olduğu durumda belirlenen DMF'li madde sayısından daha fazla olduğu bulunmuştur. Örneklem büyüklüğü 900 iken, her iki O/R oranı için belirlenen DMF'li madde sayısının eşit olduğu görülmektedir.

3) Dört çalışma gurubu için elde edilen sonuçlar, M-AÇM yöntemiyle elde edilen DMF'li madde sayısının M-ST yöntemiyle elde edilen DMF'li madde sayılarından daha düşük olduğunu göstermektedir.

4) M-ST yöntemi için dört farklı koşulda ortaya DMF'li maddelere bakıldığında 8, 14 ve 26 nolu maddelerin, M-AÇM yönteminde ise 26 nolu maddenin ortak şekilde DMF gösterdiği görülmüştür. Buna göre 26 nolu maddenin tüm koşullarda her iki DMF belirleme yöntemiyle de DMF gösterdiği sonucu ortaya çıkmıştır.

5) En yüksek DMF'li madde sayısının M-ST yöntemiyle örneklem büyüklüğü 900, O/R oranı 350/550 koşulunda gerçekleştirilen analiz sonucunda elde edildiği görülmüştür.

Öneriler

Uygulayıcılara Yönelik Öneriler

1) M-ST ve M-AÇM yöntemi kullanılacağında verinin uyum sağladığı MTK modeli göz önünde bulundurulmalıdır.

2) DMF analizi yapılacak veri 2PLM'ye uyum sağladığından M-ST ve M-AÇM yöntemlerinin kullanılması önerilmektedir.

3) DMF analizi uygulanan test kısa bir test olduğunda DMF'li madde oranı %40 ve üzerindeyse M-ST ve M-AÇM yöntemlerinin kullanılmaması önerilmektedir.

4) Testteki DMF'li madde oranının yaklaşık %20 oranlarında ve DMF büyüklüğünün 0.5-0.7 gibi orta ve yüksek olmasının beklendiği durumlarda M-ST yerine M-AÇM yönteminin kullanılması önerilmektedir.

Araştırmacılara Yönelik Öneriler

1) Araştırma iki kategorili veri üzerinden yürütülmüştür. Yöntemlerin çok kategorili maddeler için performanslarını değerlendirebilmek adına iki kategorili ve çok kategorili verilerin bir arada yer aldığı bir çalışma yapılabilir.

2) Araştırmadaki madde sayısı 20 olarak sabitlenmiştir. Yöntemlerin uzun testlerdeki performanslarını görmek amacıyla madde sayısının daha fazla olduğu çalışmalar yapılabilir. Test uzunluğu incelenen bir koşul olarak araştırmaya eklenebilir.

3) M-AÇM yönteminde çapa madde sayısı 4 olarak sabitlenmiştir. Farklı çapa madde sayılarının olduğu çalışmalar yapılarak çapa madde sayısının M-AÇM yönteminin performansına etkisi incelenebilir.

4) Çalışma tek biçimli DMF açısından ele alınmıştır. DMF türü, çalışmaya bir koşul olarak eklenip tek biçimli olmayan DMF için elde edilen sonuçlar incelenebilir.

5) M-ST ve M-AÇM yöntemiyle elde edilen sonuçların diğer DMF belirleme yöntemleriyle karşılaştırıldığı çalışmalar yapılabilir.

6) Çalışmaya ölçek arındırma yaklaşımına uygun başka bir DMF yöntemi eklenerek M-AÇM yöntemiyle karşılaştırılabilir.

7) M-AÇM yöntemiyle yapılan analizler arınık çapa maddeleri belirleme aşamaları nedeniyle uzun vakit alabilmektedir. M-AÇM yöntemiyle yapılan analizleri daha pratik hale getirecek bir R paketi üretilebilir.

8) Gerçek veriyle yapılan analizlerde grup değişkeni olarak anadili İngilizce olup olmama durumu ele alınmıştır. Grup değişkeninin cinsiyet, sosyoekonomik durum, okul türü vb. olarak ele alındığı çalışmalar yapılabilir.



Kaynaklar

- Angoff, W. (1993). Perspectives on Differential Item Functioning Methodology. In P. Holland, & H. Wainer, *Differential Item Functioning* (pp. 3-23). Hillsdale, NJ: Erlbaum.
- Arikan, Ç. A., Uğurlu, S., & Atar, B. (2016) Mimic, Sibtest, Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle gerçekleştirilen DMF ve yanlılık çalışması. *Hacettepe Eğitim Fakültesi Dergisi*, 31: 34- 52. <https://doi.org/10.16986/huje.2015014226>
- Atalay Kabasakal, K., Arsan, N., Gök, B., & Kelecioğlu, H. (2014). Comparing performances (Type I error and power) of IRT likelihood ratio SIBTEST and Mantel Haenszel methods in the determination of differential item functioning. *Educational Science: Theory&Practice* ,14(6), 2186-2193. <https://doi.org/10.12738/estp.2014.6.2165>
- Ayva Yörü, F. G., (2019). *Değişen Madde Fonksiyonu Belirleme Yöntemlerinin Madde ve Yetenek Parametre Kestirimine Etkisi* (Doktora Tezi). Gazi Üniversitesi, Ankara.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, 22(3), 507-526. <https://doi.org/10.1037/met0000077>
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In: Lord, F.M. & Novick, M.R. (Eds.), *Statistical theories of mental test scores* (pp. 397-479). Reading, Mass.: Addison-Wesley.
- Bolt, D. M. (2000). A SIBTEST approach to testing DIF hypotheses using experimentally designed test items. *Journal of Educational Measurement*, 37(4), 307-327. <https://doi.org/10.1111/j.1745-3984.2000.tb01089.x>

- Bolt, D. M. (2002). A Monte Carlo comparison of parametric and nonparametric polytomous DIF detection methods. *Applied Measurement in Education*, 15(2), 113-141.
<https://eric.ed.gov/?id=EJ654822>
- Cao, M., Tay, L., & Liu, Y. (2017). A Monte Carlo Study of an Iterative Wald Test Procedure for DIF Analysis. *Educational and Psychological Measurement*, 77(1), 104–118.
- Clauser, B. E., Mazor, K. & Hambleton, R. K. (1991). Influence of the criterion variable on the identification of differential functioning test items using the Mantel-Haenszel statistic. *Applied Psychological Measurement*, 15, 353-359.
- Clauser, B. E., & Mazor, K. M. (1998). Using Statistical Procedures To Identify Differentially Functioning Test Items. *An NCME Instructional Module. Educational Measurement: issues and practice*, 17(1), 31-44.
- Donoghue, J. R., & Allen, N. L. (1993). Thin versus thick matching in the Mantel-Haenszel procedure for detecting DIF. *Journal of Educational Statistics*, 18, 131-154.
- Ellis, B. B., & Raju, N. S. (2003). Test and Item Bias: What They Are, What They Aren't, and How to Detect Them. *Educational Resources information center (ERIC)*.
<https://files.eric.ed.gov/fulltext/ED480042.pdf>
- Finch, H. (2005). The MIMIC model as a method for detecting DIF: Comparison with Mantel-Haenszel, SIBTEST, and the IRT likelihood ratio. *Applied Psychological Measurement*, 29, 278-295.
- Finch, W. H., & French, B. F. (2007). Detection of crossing differential item functioning a comparison four methods. *Educational Psychological Measurement*, 67(4), 565-582. <http://doi.org/10.1177/0013164406296975>
- French, A. W., & Miller, T. R. (1996). Logistic regression and its use in detecting differential item functioning in polytomous items. *Journal of Educational Measurement*, 33, 315-332.

- Gierl, M. J., Jodoin, M. G., & Ackerman, T. A. (2000). *Performance of Mantel-Haenszel, Simultaneous Item Bias Test, and Logistic Regression When the Proportion of DIF Items is Large*. Paper Presented at the Annual Meeting of the American Educational Research Association (AERA) New Orleans, Louisiana, USA April 24–27.
- Gierl, M. J., Khaliq, S. N., & Boughton, K. (1999). *Gender differential item functioning in mathematics and science: prevalence and policy implications*. Paper presented at the Annual Meeting of the Canadian Society for the Study of Education. Canada.
- Glöckner-Rist, A., & Hoijtink, H. (2003). The best of both worlds: Factor analysis of dichotomous data using item response theory and structural equation modeling. *Structural Equation Modeling*, 10(4), 544–565.
- Gök, B., Kelecioğlu, H., & Doğan, N. (2010). Değişen Madde Fonksiyonunu Belirlemede Mantel-Haenszel ve Lojistik Regresyon Tekniklerinin Karşılaştırılması. *Eğitim ve Bilim*, 35(156), 3-16.
- Hallquist, M. N. & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in mplus. *Structural Equation Modelling*, 25, 621-638. <http://doi.org/10.1093/geronj/49.6.p251>
- Holland, P. W., & Thayer, D. T. (1988). Differential item functioning and the Mantel Haenszel procedure. H. Wainer, and H. I. Braun (Eds.), *Test validity* (pp. 129 145). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.
- Hou, L., la Torre, J. D., & Nandakumar, R. (2014). Differential item functioning assessment in cognitive diagnostic modeling: Application of the Wald test to investigate DIF in the DINA model. *Journal of Educational Measurement*, 51(1), 98–125.
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I Error and Power Rates Using an Effect Size Measure With the Logistic Regression Procedure for DIF

Detection. *Applied Measurement in Education*, 14(4), 329–349.
https://doi.org/10.1207/S15324818AME1404_2

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70(351a), 631–639.

Karami, H., & Salmani Nodoushan, M. A. (2011). Differential Item Functioning (DIF): Current problems and future directions. *International Journal of Language Studies*, 5(3), 133-142.

Kelecioğlu, H., Karabay, B. & Karabay, E. (2014). Investigation of placement test in terms of item biasness. *Elementary Education Online*, 13(3), 934-953.

Kristjansson, E., Aylesworth, R., Mcdowell, I., & Zumbo, B. D. (2005). A comparison of four methods for detecting differential item functioning in ordered response items. *Educational and Psychological Measurement*, 65(6), 935-953.
<https://doi.org/10.1177/0013164405275668>

Lee, S., Bulut, O., & Suh, Y. (2017). Multidimensional Extension of Multiple Indicators Multiple Causes Models to Detect DIF. *Educational and Psychological Measurement*, 77(4), 545-569. <https://doi.org/10.1177/0013164416651116>

Li, HH. & Stout, W. A (1996). A new procedure for detection of crossing DIF. *Psychometrika* 61, 647–677. <https://doi.org/10.1007/BF02294041>

Livingston, S. A. (2006). Item analysis. In S. M. Downing & T. M. Haladyna(Eds.), *Handbook of test development* (pp. 421-441). Mahwah, New Jersey:Lawrence Erlbaum Associates Inc.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

- Lord, F. M. (1968). An Analysis of the Verbal Scholastic Aptitude Test Using Birnbaum's Three-Parameter Logistic Model. *Educational and Psychological Measurement*, 28(4), 989-1020. <https://doi.org/10.1177/001316446802800401>
- MacIntosh, R., & Hashim, S. (2003). Variance estimation for converting MIMIC model parameters to IRT parameters in DIF analysis. *Applied Psychological Measurement*, 27, 372-379.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Meade, A. W., & Bauer, D. J. (2007). Power and precision in confirmatory factor analytic tests of measurement invariance. *Structural Equation Modeling*, 14(4), 611-635.
- Mellenbergh, G. J. (1982). Contingency Table Models for Assessing Item Bias. *Journal of Educational Statistics*, 7(2), 105-118.
- Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias. *Applied Psychological Measurement*, 17, 297-334.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159-176.
- Muthén, B. O. (1985). A method for studying the homogeneity of test items with respect to other relevant variables. *Journal of Educational Statistics*, 10, 121-132.
- Muthén, B. O. (1988). Some uses of structural equation modeling in validity studies: Extending IRT to external variables. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 213-238). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54(4), 557-585.

- Muthén, B. O., Kao, C.-F., & Burstein, L. (1991). Instructionally sensitive psychometrics: Application of a new IRT-based detection technique to mathematics achievement test items. *Journal of Educational Measurement*, 28, 1-22.
- Muthén, L. K. & Muthén, B. O. (1998, 2010). *Mplus user's guide* (6th ed.) Los Angeles, CA: Muthén & Muthén.
- Narayanan, P., & Swaminathan, H. (1996). Identification of Items that Show Nonuniform DIF. *Applied Psychological Measurement*, 20(3), 257-274.
<https://doi.org/10.1177/014662169602000306>
- Navas-Ara, M. J., & Gomez-Benito, J. (2002). Effects of ability scale purification on identification of DIF. *European Journal of Psychological Assessment*, 18, 9-15.
- Oort, F. J. (1998). Simulation study of item bias detection with restricted factor analysis. *Structural Equation Modeling*, 5, 107-124.
- Osterlind, J. S. (1983). *Test item bias*. London: Sage Publications.
- Ömür Sünbül, S., & Sünbül, Ö. (2016). Değişen Madde Fonksiyonunun Belirlenmesinde Kullanılan Yöntemlerde I. Tip Hata ve Güç Çalışması. *İlköğretim Online*, 15(3).
<https://doi.org/10.17051/io.2016.10640>
- Raju, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, 53(4), 495-502. <https://doi.org/10.1007/BF02294403>
- Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: a comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, 87(3), 517.
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Shealy, R. T., & Stout, W. F. (1993). A model-biased standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as

well as item bias/DIF. *Psychometrika*, 58(2), 159-194.
<https://doi.org/10.1007/BF02294572>

- Shih, C. L., & Wang, W. C. (2009). Differential item functioning detection using the multiple indicators, multiple causes method with a pure short anchor. *Applied Psychological Measurement*, 33(3), 184–199. <https://doi.org/10.1177/0146621608321758>
- Su, Y. H., & Wang, W. C. (2005). Efficiency of the mantel, generalized mantel-haenszel, and logistic discriminant function analysis methods in detecting differential item functioning for polytomous items. *Applied Measurement in Education*, 18(4), 313–350.
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group differences in trace lines. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 147-169). Hillsdale, NJ: Lawrence Erlbaum.
- Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent variable means. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 119–169). Charlotte, NC: Information Age.
- Vaughn, B. K., & Wang, Q. (2010). DIF trees: Using classification trees to detect differential item functioning. *Educational and Psychological Measurement*, 70, 941-952.
- Wainer, H., Sireci, S., & Thissen, D. (1991). Differential testlet functioning: Definitions and detection. *Journal of Educational Measurement*, 28, 197-219.
- Wang, W. C., & Yeh, Y. L. (2003). Effects of anchor item methods on differential item functioning detection with the likelihood ratio test. *Applied Psychological Measurement*, 27, 479-498.
- Wang, W. C., & Su, Y. H. (2004). Factors influencing the mantel and generalized Mantel-Haenszel methods for the assessment of differential item functioning in polytomous

items. *Applied Psychological Measurement*, 28, 450-481
<https://doi.org/10.1177/0146621604269792>

Wang, W., Shih, C., & Yang, C. (2009). The MIMIC method with scale purification for detecting differential item functioning. *Educational and Psychological Measurement*, 69, 713-731. <https://doi.org/10.1177/0013164409332228>

Wang, W. C., & Shih, C. (2010). MIMIC methods for assessing differential item functioning in polytomous items. *Applied Psychological Measurement*, 34(3), 166-180.
<https://doi.org/10.1177/0146621609355279>

Woods, C. M. (2009). Evaluation of MIMIC-model methods for DIF testing with comparison to two group analysis. *Multivariate Behavioral Research*, 44(1), 1–27.
<http://dx.doi.org/10.1080/00273170802620121>

Woods, C. M., Oltmanns, T. F., & Turkheimer, E. (2009). Illustration of MIMIC-model DIF testing with the schedule for nonadaptive and adaptive personality. *J Psychopathol Behav Assess*, 31, 320-330.

Wood, W. S. (2011). Differential item functioning procedures for polytomous items when examinee sample sizes are small (Doctoral dissertation). Graduate College of The University of Iowa, Iowa.

Woods, C. M., & Grimm, K. J. (2011). Testing for nonuniform differential item functioning with multiple indicator multiple cause models. *Applied Psychological Measurement*, 35, 339-361.

Woods, C. M., Cai, L., & Wang, M. (2013). The Langer-Improved Wald Test for DIF Testing With Multiple Groups Evaluation and Comparison to Two-Group IRT. *Educational and Psychological Measurement*, 73(3), 532–547.

Yıldıztekin, B. (2020). *Test boyutluluğunun değişen madde fonksiyonuna etkisinin farklı koşullar altında incelenmesi* (Doktora tezi). Hacettepe Üniversitesi, Ankara.

- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Lawrence Erlbaum Associates, Inc.
- Zimbra, D. J. (2018). *An examination of the MIMIC method for detecting DIF and comparison to the IRT likelihood ratio and Wald tests* (Doctoral dissertation). The University of Hawaii Manoa, USA
<https://scholarspace.manoa.hawaii.edu/items/640e35f9-2f0b-4fe0-9e80-957b7e98bf8c>
- Zumbo, B. D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zumbo, B. D., & Thomas, D. R. (1996). *A measure of DIF effect size using logistic regression procedures*. Paper presented at the National Board of Medical Examiners, Philadelphia, PA

EK-A: Simülasyon Verilerine Ait Örnek Paralel Analiz Sonuçları

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	43.3*	10.0	11.1	1	27.8*	9.9	11.1
2	5.4	9.4	10.5	2	6.3	9.3	10.4
3	5.2	8.8	9.7	3	6.0	8.8	9.7
4	4.9	8.3	9.1	4	5.9	8.3	9.1
5	4.7	7.8	8.4	5	5.5	7.7	8.4
6	4.6	7.3	7.9	6	5.3	7.2	7.8
7	4.2	6.7	7.3	7	5.0	6.7	7.3
8	4.1	6.2	6.8	8	4.6	6.2	6.8
9	3.7	5.7	6.3	9	4.6	5.8	6.2
10	3.5	5.2	5.8	10	4.5	5.3	5.8
11	3.3	4.7	5.3	11	4.2	4.8	5.3
12	3.1	4.2	4.9	12	4.0	4.3	4.9
13	2.7	3.8	4.4	13	3.4	3.8	4.4
14	2.5	3.2	4.0	14	3.4	3.3	3.9
15	2.1	2.8	3.5	15	3.2	2.8	3.5
16	1.2	2.2	2.9	16	2.4	2.3	3.0
17	0.8	1.7	2.4	17	2.1	1.7	2.4
18	0.5	1.2	1.9	18	1.2	1.2	1.8
19	0.2	0.7	1.2	19	0.4	0.7	1.3
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	28.7*	9.9	11.2	1	30.3*	9.9	10.9
2	6.7	9.3	10.4	2	6.8	9.2	10.0
3	5.9	8.8	9.6	3	6.6	8.7	9.4
4	5.5	8.3	9.0	4	5.8	8.2	8.8
5	5.3	7.7	8.4	5	5.4	7.7	8.3
6	5.1	7.2	7.9	6	5.2	7.2	7.7
7	4.8	6.7	7.3	7	5.0	6.7	7.1
8	4.8	6.2	6.8	8	4.5	6.2	6.6
9	4.6	5.7	6.3	9	4.3	5.7	6.2
10	4.4	5.2	5.8	10	4.1	5.3	5.7
11	4.2	4.7	5.4	11	3.7	4.8	5.2
12	4.0	4.3	4.8	12	3.5	4.3	4.8
13	3.9	3.8	4.3	13	3.2	3.8	4.3
14	3.5	3.3	3.9	14	2.8	3.3	3.8
15	2.7	2.8	3.4	15	2.7	2.8	3.4
16	2.0	2.3	2.9	16	2.2	2.3	3.0
17	1.9	1.8	2.4	17	2.0	1.8	2.5
18	1.5	1.2	1.9	18	1.4	1.3	1.9
19	0.4	0.7	1.2	19	0.6	0.8	1.3
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	31.7*	9.9	11.0	1	32.0*	10.0	11.0
2	7.1	9.2	10.1	2	7.2	9.3	10.1
3	6.8	8.7	9.5	3	6.7	8.7	9.5
4	6.5	8.2	8.9	4	6.1	8.2	8.8
5	5.8	7.7	8.3	5	5.9	7.7	8.2
6	5.5	7.2	7.7	6	5.8	7.2	7.6
7	5.4	6.7	7.2	7	5.3	6.7	7.2
8	4.5	6.2	6.7	8	5.3	6.2	6.6
9	4.4	5.7	6.2	9	4.6	5.7	6.2
10	4.1	5.3	5.7	10	4.2	5.2	5.7
11	3.6	4.8	5.2	11	3.8	4.7	5.2
12	3.4	4.3	4.8	12	3.4	4.3	4.8
13	2.9	3.8	4.3	13	3.1	3.8	4.3
14	2.4	3.3	3.9	14	2.0	3.3	3.8
15	1.9	2.8	3.4	15	1.8	2.8	3.4
16	1.6	2.3	2.9	16	1.2	2.3	2.9
17	1.2	1.8	2.4	17	1.1	1.8	2.4
18	0.9	1.3	1.9	18	0.4	1.2	1.8
19	0.3	0.7	1.3	19	0.3	0.7	1.2
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	28.5*	10.1	11.0	1	29.7*	10.1	11.0
2	7.4	9.3	10.0	2	6.9	9.3	10.1
3	6.4	8.7	9.3	3	6.2	8.7	9.3
4	5.8	8.1	8.8	4	5.8	8.1	8.7
5	5.7	7.6	8.2	5	5.5	7.6	8.2
6	5.4	7.2	7.6	6	5.2	7.1	7.6
7	5.2	6.7	7.1	7	4.9	6.6	7.1
8	4.8	6.2	6.6	8	4.8	6.2	6.6
9	4.6	5.7	6.1	9	4.4	5.7	6.1
10	4.4	5.2	5.7	10	3.9	5.2	5.7
11	4.0	4.8	5.2	11	3.8	4.8	5.2
12	3.4	4.3	4.8	12	3.5	4.3	4.8
13	3.3	3.8	4.3	13	3.4	3.8	4.3
14	2.8	3.3	3.9	14	3.2	3.3	3.8
15	2.5	2.8	3.4	15	2.7	2.9	3.4
16	2.1	2.3	2.9	16	2.1	2.4	2.9
17	1.8	1.8	2.4	17	1.9	1.8	2.4
18	1.2	1.3	1.9	18	1.2	1.3	1.9
19	0.6	0.7	1.3	19	0.9	0.7	1.3
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	27.8*	10.1	11.2	1	34.6*	9.9	11.2
2	7.9	9.3	10.1	2	7.0	9.3	10.4
3	7.7	8.7	9.4	3	6.7	8.8	9.7
4	7.1	8.1	8.8	4	6.6	8.2	9.0
5	6.3	7.6	8.1	5	6.2	7.7	8.4
6	5.7	7.1	7.6	6	5.3	7.2	7.8
7	5.5	6.6	7.1	7	4.5	6.7	7.2
8	5.1	6.2	6.6	8	4.0	6.2	6.7
9	4.5	5.7	6.1	9	3.7	5.8	6.3
10	4.1	5.2	5.6	10	3.6	5.3	5.8
11	3.9	4.8	5.2	11	3.4	4.8	5.4
12	3.1	4.3	4.8	12	2.9	4.3	4.9
13	2.9	3.8	4.3	13	2.7	3.8	4.4
14	2.6	3.3	3.8	14	2.4	3.3	3.9
15	2.2	2.8	3.3	15	1.9	2.8	3.5
16	1.6	2.3	2.9	16	1.8	2.3	3.0
17	1.0	1.8	2.4	17	1.2	1.8	2.4
18	0.9	1.3	1.8	18	0.9	1.2	1.9
19	0.3	0.7	1.2	19	0.6	0.7	1.2
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	31.0*	9.9	11.2	1	29.6*	9.9	11.1
2	7.3	9.3	10.4	2	6.4	9.3	10.6
3	6.9	8.8	9.6	3	6.3	8.8	9.9
4	6.2	8.3	9.1	4	5.8	8.3	9.1
5	6.0	7.7	8.4	5	5.6	7.7	8.4
6	5.3	7.2	7.9	6	5.2	7.2	7.8
7	5.2	6.7	7.3	7	5.0	6.7	7.3
8	4.8	6.2	6.8	8	5.0	6.2	6.8
9	4.4	5.8	6.3	9	4.4	5.7	6.2
10	4.3	5.3	5.8	10	4.1	5.2	5.7
11	3.9	4.8	5.3	11	3.7	4.7	5.3
12	3.6	4.3	4.8	12	3.6	4.3	4.9
13	2.7	3.8	4.5	13	3.1	3.8	4.4
14	2.3	3.3	4.0	14	3.0	3.3	3.9
15	1.9	2.8	3.5	15	2.7	2.8	3.5
16	1.6	2.3	3.0	16	2.3	2.3	3.0
17	1.3	1.7	2.5	17	1.5	1.8	2.4
18	0.9	1.2	1.9	18	1.4	1.2	1.8
19	0.3	0.6	1.2	19	1.2	0.7	1.2
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	32.8*	10.0	11.0	1	43.9*	10.1	11.0
2	7.4	9.3	10.2	2	6.4	9.3	10.1
3	7.0	8.7	9.5	3	5.8	8.7	9.5
4	6.2	8.2	8.9	4	5.2	8.2	8.8
5	5.8	7.7	8.3	5	4.8	7.6	8.2
6	5.2	7.2	7.7	6	4.2	7.2	7.7
7	4.8	6.7	7.2	7	4.1	6.7	7.1
8	4.4	6.2	6.6	8	4.0	6.2	6.6
9	4.3	5.7	6.2	9	3.5	5.7	6.1
10	4.0	5.2	5.7	10	3.4	5.3	5.7
11	3.4	4.8	5.2	11	3.0	4.8	5.2
12	3.1	4.3	4.8	12	2.5	4.3	4.7
13	2.7	3.8	4.3	13	2.3	3.8	4.3
14	2.4	3.3	3.8	14	1.8	3.3	3.8
15	2.1	2.8	3.4	15	1.7	2.8	3.4
16	1.8	2.3	2.9	16	1.3	2.3	2.9
17	1.3	1.8	2.4	17	1.0	1.8	2.4
18	0.9	1.3	1.8	18	0.6	1.3	1.9
19	0.4	0.7	1.3	19	0.5	0.7	1.3
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	37.9*	10.1	11.1	1	40.7*	10.0	11.1
2	6.9	9.3	10.0	2	6.9	9.3	10.1
3	6.0	8.7	9.4	3	5.7	8.7	9.4
4	5.8	8.2	8.7	4	5.4	8.1	8.8
5	5.5	7.6	8.2	5	4.9	7.6	8.2
6	5.3	7.2	7.6	6	4.6	7.1	7.6
7	4.1	6.7	7.1	7	4.4	6.7	7.1
8	4.0	6.2	6.6	8	4.3	6.2	6.6
9	3.7	5.7	6.1	9	3.9	5.7	6.1
10	3.5	5.2	5.6	10	3.6	5.2	5.6
11	3.2	4.8	5.2	11	3.2	4.8	5.2
12	2.9	4.3	4.8	12	2.9	4.3	4.7
13	2.5	3.8	4.3	13	2.2	3.8	4.3
14	2.4	3.3	3.8	14	1.9	3.3	3.8
15	2.2	2.8	3.4	15	1.9	2.8	3.4
16	1.7	2.3	2.9	16	1.5	2.3	2.9
17	1.3	1.8	2.5	17	1.0	1.8	2.4
18	0.9	1.3	1.9	18	0.7	1.3	1.9
19	0.3	0.7	1.3	19	0.5	0.7	1.3
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance	Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	43.5*	10.1	11.1	1	41.1*	10.0	11.0
2	6.1	9.3	10.0	2	6.7	9.3	10.1
3	5.7	8.7	9.3	3	6.1	8.7	9.3
4	5.2	8.2	8.8	4	5.1	8.1	8.8
5	4.7	7.6	8.2	5	4.8	7.6	8.2
6	4.0	7.1	7.6	6	4.3	7.1	7.5
7	3.6	6.6	7.0	7	4.1	6.6	7.1
8	3.4	6.2	6.6	8	3.9	6.2	6.6
9	3.3	5.7	6.1	9	3.6	5.7	6.1
10	3.0	5.2	5.6	10	3.2	5.3	5.7
11	2.7	4.8	5.2	11	2.9	4.8	5.2
12	2.7	4.3	4.8	12	2.8	4.3	4.7
13	2.4	3.8	4.3	13	2.4	3.8	4.3
14	2.1	3.3	3.8	14	2.2	3.3	3.9
15	2.0	2.9	3.4	15	2.1	2.9	3.4
16	1.8	2.3	2.9	16	1.8	2.4	2.9
17	1.5	1.8	2.4	17	1.3	1.8	2.4
18	1.3	1.3	1.8	18	1.2	1.3	1.9
19	0.9	0.7	1.2	19	0.3	0.8	1.3
20	0.0	0.0	0.0	20	0.0	0.0	0.0

* Advised number of dimensions: 1

* Advised number of dimensions: 1

EK-B: Gerçek Veriye Ait Paralel Analiz Sonucu

No	Açıklanan Varyans Yüzdeleri		
	Gerçek Veri	Tesadüfi Veri	Tesadüfi Verinin 95. Yüzdelik Dilimi
1	30.60	6.0	6.60
2	5.60	5.60	6.10
3	4.30	5.30	5.70
4	3.80	5.10	5.40
5	3.60	4.80	5.10
6	3.30	4.60	4.90
7	3.10	4.50	4.70
8	3.10	4.30	4.50
9	2.90	4.20	4.40
10	2.90	4.00	4.20
11	2.80	3.90	4.00
12	2.70	3.70	3.90
13	2.50	3.60	3.70
14	2.50	3.40	3.60
15	2.40	3.30	3.40
16	2.30	3.20	3.30
17	2.20	3.00	3.10
18	2.00	2.70	3.00
19	2.00	2.70	2.90
20	1.80	2.60	2.80
21	1.70	2.40	2.60
22	1.70	2.30	2.50
23	1.60	2.10	2.30
24	1.50	2.00	2.20
25	1.30	1.90	2.10
26	1.20	1.70	1.90
27	1.10	1.50	1.80
28	1.00	1.40	1.60
29	0.90	1.20	1.50
30	0.80	1.00	1.30
31	0.60	0.80	1.10
32	0.20	0.60	0.90
33	0.00	0.40	0.60
34	0.00	0.00	0.00

EK-C: M-ST Analizi için R Programı Kodu

```

library("MplusAutomation")

for (j in 1:100) {

  setwd("D:\\mimic analizi_v2\\Kosul6")

  dmldata_j <- read.table(paste0("dmldata_rep", j, ".txt"))

  setwd("D:\\mimic analizi_v2\\Kosul6\\MIMIC-ST")

  dir.create(paste0("Rep", j))

  setwd(paste0("D:\\mimic analizi_v2\\Kosul6\\MIMIC-ST\\Rep", j))

  dmldata <- as.data.frame(dmldata_j)

  names(dmldata) <- c("grup", paste("madde", 1:20, sep = ""))

  for (i in 1:20) {

    dmldata.mimic <- mplusObject(TITLE = "MIMIC DMF ANALIZI",

                                rdata = dmldata,

                                VARIABLE = "Categorical are madde1-madde20;",

                                usevariables = c("grup", paste0("madde", 1:20)),

                                ANALYSIS = "Estimator is WLSMV;",

                                MODEL = paste("f by madde1-madde20;

                                                f on grup;

                                                madde", i, " on grup;", sep = ""),

                                OUTPUT = "Modindices (all)")

    mplusModeler(dmldata.mimic, dataout = "veri.dat",

                  modelout = paste("rep", j, ",madde", i, ".inp", sep = ""), run = 1L,

```

```
check = FALSE, varwarnings = TRUE,  
  
Mplus_command = "Mplus")  
  
}  
  
}
```



EK-Ç: Alt Problem 1'e İlişkin Faktöriyel ANOVA Tablosu

Etki Kaynağı	Serbestlik Derecesi	F	p	Kısmi η^2
Yöntem	1	22.07	.00*	.33
Örn. Büyüklüğü	2	14.33	.00*	.39
O/R Oranı	1	7.42	.00*	.14
DMF Oranı	1	99.25	.00*	.69
DMF Büyüklüğü	2	22.89	.00*	.50
Yöntem*Örn. Büyüklüğü	2	.80	.46	.03
Yöntem*O/R Oranı	1	4.12	.05	.08
Yöntem*DMF Oranı	1	1.78	.19	.04
Yöntem*DMF Büyüklüğü	2	2.09	.14	.09
Örn. Büyüklüğü*O/R Oranı	2	1.61	.21	.07
Örn. Büyüklüğü*DMF Oranı	2	3.65	.03*	.14
Örn. Büyüklüğü*DMF Büyüklüğü	4	.45	.77	.04
O/R Oranı*DMF Oranı	1	4.12	.05	.08
O/R Oranı*DMF Büyüklüğü	2	1.72	.19	.07
DMF Oranı*DMF Büyüklüğü	2	8.23	.00*	.27

$p < .05$

EK-D: Alt Problem 2'ye İlişkin Faktöriyel ANOVA Tablosu

Etki Kaynağı	Serbestlik Derecesi	F	p	Kısmi η^2
Yöntem	1	122.05	.00*	.73
Örn. Büyüklüğü	2	34.54	.00*	.61
O/R Oranı	1	9.78	.00*	.18
DMF Oranı	1	188.11	.00*	.81
DMF Büyüklüğü	2	68.43	.00*	.75
Yöntem*Örn. Büyüklüğü	2	7.64	.00*	.25
Yöntem*O/R Oranı	1	.04	.84	.00
Yöntem*DMF Oranı	1	2.64	.11	.06
Yöntem*DMF Büyüklüğü	2	7.73	.00*	.26
Örn. Büyüklüğü*O/R Oranı	2	1.75	.19	.07
Örn. Büyüklüğü*DMF Oranı	2	16.80	.00*	.43
Örn. Büyüklüğü*DMF Büyüklüğü	4	5.89	.70	.34
O/R Oranı*DMF Oranı	1	4.17	.05	.09
O/R Oranı*DMF Büyüklüğü	2	1.88	.16	.08
DMF Oranı*DMF Büyüklüğü	2	32.86	.00*	.59

$p < .05$



EK-E: Alt Problem 3'e İlişkin Faktöriyel ANOVA Tablosu

Etki Kaynağı	Serbestlik Derecesi	F	p	Kısmi η^2
Yöntem	1	81.06	.00*	.64
Örn. Büyüklüğü	2	232.45	.00*	.91
O/R Oranı	1	39.84	.00*	.47
DMF Oranı	1	341.18	.00*	.88
DMF Büyüklüğü	2	568.19	.00*	.96
Yöntem*Örn. Büyüklüğü	2	.33	.73	.01
Yöntem*O/R Oranı	1	.31	.58	.01
Yöntem*DMF Oranı	1	23.30	.00*	.34
Yöntem*DMF Büyüklüğü	2	.45	.64	.02
Örn. Büyüklüğü*O/R Oranı	2	.46	.64	.02
Örn. Büyüklüğü*DMF Oranı	2	.73	.49	.03
Örn. Büyüklüğü*DMF Büyüklüğü	4	6.69	.00*	.37
O/R Oranı*DMF Oranı	1	.31	.58	.01
O/R Oranı*DMF Büyüklüğü	2	.22	.80	.01
DMF Oranı*DMF Büyüklüğü	2	3.79	.03*	.14

$p < .05$



EK-F: Alt Problem 4'e İlişkin Faktöriyel ANOVA Tablosu

Etki Kaynağı	Serbestlik Derecesi	F	p	Kısmi η^2
Yöntem	1	55.24	.00*	.55
Örn. Büyüklüğü	2	206.62	.00*	.90
O/R Oranı	1	27.30	.00*	.38
DMF Oranı	1	205.74	.00*	.82
DMF Büyüklüğü	2	370.26	.00*	.94
Yöntem*Örn. Büyüklüğü	2	.09	.92	.00
Yöntem*O/R Oranı	1	.02	.89	.00
Yöntem*DMF Oranı	1	4.25	.05	.09
Yöntem*DMF Büyüklüğü	2	.31	.74	.01
Örn. Büyüklüğü*O/R Oranı	2	.03	.97	.00
Örn. Büyüklüğü*DMF Oranı	2	2.39	.10	.10
Örn. Büyüklüğü*DMF Büyüklüğü	4	7.22	.00*	.39
O/R Oranı*DMF Oranı	1	.16	.69	.00
O/R Oranı*DMF Büyüklüğü	2	.53	.59	.02
DMF Oranı*DMF Büyüklüğü	2	8.63	.00*	.28

$p < .05$



EK-G: Çalışma Grubu 1 için DMF Analiz Sonuçları

Madde No	M-ST Beta Estimate	p	M-AÇM Beta Estimate	p
1	-0.05	0.58	-0.01	0.91
2	-0.16	0.09	-0.12	0.29
3	0.20	0.01*	0.23	0.03*
4	0.16	0.18	0.19	0.15
5	0.09	0.23	0.13	0.22
6	-0.16	0.07	-0.12	0.29
7	0.01	0.88	0.04	0.65
8	-0.42	0.00*	-0.37	0.01*
9	0.34	0.01*	0.36	0.01*
10	-0.04	0.69	-	-
11	0.18	0.06	0.21	0.05
12	0.04	0.78	0.07	0.59
13	0.11	0.16	0.14	0.12
14	-0.30	0.01*	-0.25	0.03*
15	-0.17	0.04*	-0.13	0.15
16	-0.20	0.02*	-0.16	0.13
17	-0.02	0.82	0.03	0.82
18	-0.02	0.81	-	-
19	-0.08	0.44	-0.03	0.82
20	0.07	0.40	0.10	0.31
21	-0.03	0.72	0.01	0.90
22	0.43	0.00*	0.46	0.00*
23	-0.07	0.38	-0.03	0.76
24	0.26	0.01*	0.29	0.00*
25	0.07	0.57	-	-
26	0.31	0.00*	0.33	0.00*
27	-0.06	0.47	-	-
28	-0.17	0.04*	-0.12	0.24
29	0.07	0.51	0.10	0.39
30	-0.08	0.32	-0.04	0.72
31	-0.25	0.01*	-0.20	0.12
32	-0.00	0.96	0.04	0.71
33	0.13	0.28	0.17	0.23
34	0.13	0.11	0.16	0.10

$p < .05$,

- ile gösterilen maddeler çapa olarak kullanıldığından beta kestirimi yapılmamıştır.

EK-H: Çalışma Grubu 2 için DMF Analiz Sonuçları

Madde No	M-ST Beta Estimate	p	M-AÇM Beta Estimate	p
1	-0.07	0.39	-0.09	0.37
2	-0.13	0.17	-0.14	0.20
3	0.20	0.01*	0.17	1.00
4	0.13	0.28	0.11	0.38
5	0.04	0.63	0.01	0.94
6	-0.31	0.01*	-0.32	0.01*
7	0.05	0.50	0.03	0.75
8	-0.35	0.00*	-0.37	0.00*
9	0.22	0.09	0.21	0.11
10	-0.01	0.91	-	-
11	0.15	0.13	0.12	0.25
12	-0.06	0.65	-0.08	0.56
13	0.07	0.36	0.06	0.54
14	-0.34	0.00*	-0.35	0.00*
15	-0.21	0.01*	-0.22	0.02*
16	-0.09	0.30	-0.11	0.31
17	-0.05	0.55	-0.07	0.50
18	0.03	0.69	-	-
19	-0.05	0.65	-0.08	0.57
20	0.16	0.05	0.14	0.17
21	-0.01	0.90	-0.03	0.73
22	0.46	0.00*	0.42	0.00*
23	-0.04	0.63	-0.06	0.57
24	0.24	0.00*	0.21	0.03*
25	0.11	0.38	-	-
26	0.36	0.00*	0.34	0.00*
27	-0.02	0.86	-	-
28	-0.15	0.05	-0.17	0.08
29	0.05	0.64	0.03	0.82
30	-0.08	0.30	-0.11	0.33
31	-0.21	0.03*	-0.23	0.07
32	-0.01	0.86	-0.04	0.70
33	0.06	0.61	0.04	0.78
34	0.18	0.02*	0.16	0.09

$p < .05$,

- ile gösterilen maddeler çapa olarak kullanıldığından beta kestirimi yapılmamıştır.

EK-I: Çalışma Grubu 3 için DMF Analiz Sonuçları

Madde No	M-ST Beta Estimate	p	M-AÇM Beta Estimate	p
1	-0.12	0.29	-0.07	0.62
2	-0.07	0.57	-0.03	0.87
3	0.09	0.38	0.14	0.31
4	-0.07	0.66	-	-
5	0.12	0.27	0.17	0.23
6	-0.17	0.17	-0.11	0.45
7	0.04	0.72	-	-
8	-0.34	0.01*	-0.28	0.06
9	0.35	0.04*	0.39	0.03*
10	0.06	0.69	0.11	0.53
11	0.16	0.24	0.20	0.17
12	0.08	0.63	0.13	0.49
13	-0.10	0.37	-	-
14	-0.31	0.01*	-0.26	0.09
15	-0.25	0.03*	-0.21	0.11
16	-0.02	0.87	0.03	0.80
17	-0.01	0.92	0.05	0.74
18	0.11	0.35	0.15	0.28
19	-0.12	0.44	-0.04	0.84
20	0.27	0.02*	0.31	0.02*
21	0.03	0.78	0.09	0.55
22	0.57	0.00*	0.61	0.00*
23	-0.01	0.95	0.04	0.75
24	0.12	0.25	0.17	0.20
25	0.22	0.21	0.26	0.17
26	0.32	0.00*	0.35	0.01*
27	-0.06	0.65	-	-
28	-0.18	0.10	-0.11	0.45
29	-0.05	0.72	-0.01	0.98
30	-0.19	0.09	-0.13	0.40
31	-0.44	0.00*	-0.36	0.04*
32	-0.10	0.34	-0.03	0.83
33	0.07	0.68	0.12	0.53
34	0.23	0.03*	0.27	0.04*

$p < .05$,

- ile gösterilen maddeler çapa olarak kullanıldığından beta kestirimi yapılmamıştır.

EK-İ: Çalışma Grubu 4 için DMF Analiz Sonuçları

Madde No	M-ST Beta Estimate	p	M-AÇM Beta Estimate	p
1	-0.11	0.39	0.00	1.00
2	-0.07	0.62	0.06	0.74
3	0.31	0.00*	0.43	0.01*
4	0.18	0.29	0.27	0.18
5	0.07	0.52	0.21	0.22
6	-0.17	0.23	-0.06	0.72
7	0.12	0.29	0.23	0.15
8	-0.37	0.02*	-0.25	0.18
9	-0.11	0.63	-	-
10	0.21	0.17	0.32	0.09
11	0.25	0.08	0.36	0.04*
12	-0.28	0.23	-0.15	0.56
13	0.21	0.09	0.29	0.05
14	-0.31	0.02*	-0.17	0.29
15	-0.10	0.38	-	-
16	-0.18	0.15	-0.06	0.72
17	0.05	0.70	0.19	0.35
18	0.05	0.67	0.15	0.34
19	-0.20	0.17	-0.04	0.85
20	-0.11	0.40	0.00	1.00
21	-0.04	0.72	0.07	0.69
22	0.21	0.14	0.32	0.09
23	0.00	0.99	0.11	0.46
24	0.27	0.02*	0.39	0.02*
25	-0.09	0.64	-	-
26	0.45	0.00*	0.52	0.00*
27	-0.19	0.15	-0.11	0.44
28	-0.28	0.02*	-0.14	0.43
29	-0.05	0.78	-	-
30	-0.06	0.65	0.07	0.69
31	-0.28	0.05	-0.14	0.47
32	-0.05	0.68	0.10	0.58
33	0.26	0.17	0.36	0.10
34	0.12	0.34	0.21	0.20

$p < .05$,

- ile gösterilen maddeler çapa olarak kullanıldığından beta kestirimi yapılmamıştır.

EK-J: Araştırma Etik Komisyon İzin Muafiyeti Formu

	Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması/Araştırma Etik Komisyon İzin Muafiyeti Formu	F46
27 / 05 / 2024		
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Ana Bilim Dalı Başkanlığına		
Tez/Araştırma Başlığı	Değişen Madde Fonksiyonunun MIMIC Yöntemleriyle Farklı Koşullar Altında İncelenmesi	
Yukarıda başlığı/konusu verilen tez/araştırma çalışmam,		
1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır. 2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmamasını gerektirmektedir. 3. Beden bütünlüğüne veya ruh sağlığına müdahale içermemektedir. 4. Anket, ölçek (test), mülakat, odak grup çalışması, gözlem, deney, görüşme gibi teknikler kullanılarak katılımcılardan veri toplanmasını gerektiren nitel ya da nicel yaklaşımlarla yürütülen araştırmalar niteliğinde değildir. 5. Diğer kişi ve kurumlardan temin edilen veri kullanımını (kitap, belge vs.) gerektirmektedir. Ancak bu kullanım, diğer kişi ve kurumların izin verdiği ölçüde Kişisel Bilgilerin Korunması Kanuna riayet edilerek gerçekleştirilecektir.		
Çalışmada kullanacağım veriler: (X) Kamusal erişime açık (buraya yazınız): Çalışmada gerçek veri olarak PISA 2012 verileri kullanılmıştır. Araştırmada kullanılan veriler kamusal erişime açık olup OECD'nin resmi Web sayfası üzerinden ulaşılmıştır. Veriler PISA 2012 uygulamasına ait https://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm adresinden indirilmiştir. () Özel izin ve onaya tabi (buraya yazınız): (X) Üretilmiş veri (buraya yazınız): Çalışmada araştırmacı tarafından R yazılımı kullanılarak üretilen simülasyon verisi kullanılmıştır. () Diğer (buraya yazınız):		
Yükseköğretim Kurumları Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre çalışmamın yürütülebilmesi için herhangi bir Etik Komisyondan/Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.		
Gereğini saygılarımla arz ederim.		
Pelin TAMCI (Araştırmacı Adı Soyadı, İmza)		
Araştırmacı Bilgileri		
Adı Soyadı	Pelin TAMCI	
Öğrenci İse No	N18147352	
Ana Bilim Dalı	Eğitim Bilimleri Ana Bilim Dalı	
Programı	Eğitimde Ölçme ve Değerlendirme	
Statüsü	☐ Yüksek Lisans ☒ Doktora ☐ Bütünleşik Dr. ☐ Diğer	
Danışman Görüşü ve Onayı*		
Çalışmada simülasyon veri ve PISA 2012 verileri kullanılmış olup, kişiler üzerinde bir uygulama yapılmamıştır		
Prof. Dr. Hülya KELECİOĞLU (İmza) (Danışmanın İmzası, Adı ve Soyadı)		
*Tez ve tezden üretilen yayınlarda gereklil		
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Beytepe Yerleşkesi, 06800 – Çankaya / ANKARA Telefon: 0(312) 207 85 72 Belgegeçer: 0(312) 207 85 00 e-Ağ: http://ebe.hacettepe.edu.tr/ e-Posta: ebe@hacettepe.edu.tr		

EK-K: Etik Beyanı

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

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

beyan ederim.

...../...../.....

(İmza)

Ad SOYADI

EK-L: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu

27/08/2024

HACETTEPE ÜNİVERSİTESİ
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Ana Bilim Dalı Başkanlığına,

Tez Başlığı: Değişen Madde Fonksiyonunun MIMIC Yöntemleriyle Farklı Koşullar Altında İncelenmesi

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

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
27/08/2024	75	92573	25/06/2024	%11	2438996725

Uygulanan filtreler:

1. Kaynaklar hariç
2. Alıntılar dâhil
3. 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

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

Ad Soyadı: Pelin TAMCI

Öğrenci No.: N18147352

Ana Bilim Dalı: Eğitim Bilimleri Ana Bilim Dalı

İmza

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: ☐ Y.Lisans ☒ Doktora ☐ Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. Hülya KELECİOĞLU

EK-M: Thesis/Dissertation Originality Report

27/08/2024

HACETTEPE UNIVERSITY
Graduate School of Educational Sciences
To The Department of Educational Sciences

Thesis Title: Examination of Differential Item Functioning Under Different Conditions Using MIMIC Methods

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

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
27/08/2024	75	92573	25/06/2024	%11	2438996725

Filtering options applied:

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

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

I respectfully submit this for approval.

Name Lastname: Pelin TAMCI

Student No.: N18147352

Department: Educational Sciences

Program: Educational Measurement and Evaluation

Status: ☐ Masters ☒ Ph.D. ☐ Integrated Ph.D.

Signature

ADVISOR APPROVAL

APPROVED
Prof. Dr. Hülya KELECİOĞLU

EK-N: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

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

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

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- ☐ Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- ☐ Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir. ⁽²⁾
- ☐ Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

..... / /
(imza)

Öğrencinin Adı SOYADI

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tez erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

*Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

