



**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**



ÖLÇEKLERDE SAPTANAN MADDE İŞLEV FARKLILIĞININ KARMA RASCH MODELLERİ İLE İNCELENMESİ

Selcen YÜKSEL

**BIYOİSTATİSTİK ANABİLİM DALI
DOKTORA TEZİ**

**DANIŞMAN
Prof. Dr. Atilla Halil ELHAN**

2012- ANKARA

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇEKLERDE SAPTANAN MADDE İŞLEV FARKLILIĞININ
KARMA RASCH MODELLERİ İLE İNCELENMESİ**

Selcen YÜKSEL

**BİYOİSTATİSTİK ANABİLİM DALI
DOKTORA TEZİ**

**DANIŞMAN
Prof. Dr. Atilla Halil ELHAN**

2012- ANKARA

Ankara Üniversitesi Sağlık Bilimleri Enstitüsü
Biyoistatistik Doktora **Programı**
çerçevesinde yürütülmüş olan bu çalışma, aşağıdaki jüri tarafından
Doktora **Tezi** olarak kabul edilmiştir.

Tez Savunma Tarihi: 26/09/2012

Prof. Dr. Atilla Halil ELHAN
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Jüri Başkanı
Tez Danışmanı

Prof. Dr. Ergun KARAAĞAOĞLU
Hacettepe Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı

Prof. Dr. Şehim KUTLAY
Ankara Üniversitesi
Tıp Fakültesi
Fizik Tedavi ve Rehabilitasyon Anabilim Dalı

Doç. Dr. Erdem KARABULUT
Hacettepe Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Raportör

Doç. Dr. Yasemin Genç
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı

İÇİNDEKİLER

Kabul ve Onay	ii
İçindekiler	iii
Önsöz	v
Simgeler ve Kısaltmalar	vi
Şekiller	vii
Çizelgeler	viii
1. GİRİŞ	1
1.1. Araştırmanın Konusu	1
1.2. Teorik Yapı	4
1.2.1. Rasch Modelleri	4
1.2.1.1. İkili Rasch Modeli	4
1.2.1.2. Kısmi Kredi Modeli	4
1.2.1.3. Sıralı Sonuçlu Model	5
1.2.1.4. Rasch Modelinin Varsayımları	6
1.2.1.5. Madde İşlev Farklılığı (MİF)	7
1.2.1.6. MİF Saptama Yöntemleri	9
1.2.1.7. MİF Gösteren Maddelerin Değerlendirilmesi	10
1.2.2. Karma Rasch Model	11
1.2.2.1. Karma Dağılım Modelleri	12
1.2.2.2. Gizli Sınıf Analizi	13
1.2.2.3. İkili Karma Rasch Model	15
1.2.2.4. Çok Sonuçlu Karma Rasch Model	18
1.2.3. Gizli Sınıf Sayısının Belirlenmesi	19
2. GEREÇ ve YÖNTEM	21
2.1. Kullanılan Bilgisayar Programları Hakkında Genel Bilgi	21
2.2. Benzetim Çalışması	22
2.2.1. Benzetim koşullarının belirlenmesi	22
2.2.2. Verilerin Türetilmesi ve Türetilen Verilerin KKM ve KRM ile Değerlendirilmesi	24
2.2.2.1. Veri türetmede kullanılan istatistiksel model	24
2.2.2.2. Altın standart parametre değerlerinin türetilmesi	24
2.2.2.3. MİF gösteren maddelerin eşik zorluklarının hesaplanması ve her bir birey için yanıt deseninin türetilmesi	26
2.2.2.4. Türetilen veri setlerinin KKM ile analizi	27
2.2.2.5. Türetilen veri setlerinin KRM ile analizi	28
2.3. Benzetim Çalışmasının Sonuçlarının Değerlendirilmesi	28
2.3.1. Türetilen veri setlerinde KKM'nin MİF gösteren maddeleri doğru ve yanlış saptama oranlarının ve KRM'nin iki gizli sınıf saptama oranının değerlendirilmesi	28

2.3.2. KKM’de MİF analizinden ve iki gizli sınıflı KRM’den elde edilen θ kestirimlerinin ve standart hatalarının değerlendirilmesi	30
2.3.2.1. θ kestirimlerinin karşılaştırılması	31
2.3.2.2. θ kestirimlerinin standart hatalarının karşılaştırılması	32
3. BULGULAR	33
3.1. Benzetim Çalışması Bulguları	33
3.1.1. Türetilen veri setlerinde KKM’nin MİF gösteren maddeleri doğru saptama oranlarının ve KRM’nin iki gizli sınıf saptama oranının değerlendirilmesi	33
3.1.2. KKM’de MİF analizinden ve iki gizli sınıflı KRM’den elde edilen θ kestirimlerinin ve standart hatalarının değerlendirilmesi	37
3.1.2.1. θ kestirimlerinin karşılaştırılması	37
3.1.2.2. θ kestirimlerinin standart hatalarının karşılaştırılması	42
4. TARTIŞMA	45
5. SONUÇ ve ÖNERİLER	50
6. ÖZET	51
7. SUMMARY	52
8. KAYNAKLAR	53
9. EKLER	56
10. ÖZGEÇMİŞ	64

ÖNSÖZ

Bu çalışmada, çok sonuçlu maddeler içeren ölçeklerde saptanan madde işlev farklılığının incelenmesinde kullanılan klasik yöntemle Karma Rasch Modeli yönteminin karşılaştırılması üzerinde durulmuştur. Bu amaca yönelik farklı koşullarda ve küçük örneklem büyüklüğü altında bir benzetim çalışması yapılmıştır.

Aşağıda ismi geçen kişilere, çalışmam boyunca gösterdikleri ilgiden dolayı, içtenlikle teşekkür ederim.

Değerli danışman hocam Prof. Dr. Atilla Halil Elhan; doktora eğitimim ve tezim süresince benden ilginizi ve desteğinizi esirgemediniz. Tezim süresince karşılaştığım tüm sorunları çözecek bir danışmanım olduğunu bilmek bana güven ve cesaret verdi. Kişiliğinizi, bilimsel disiplininizi ve bilgi birikiminizi her zaman örnek aldım.

Değerli hocam Prof. Dr. Ergun KARAAĞAOĞLU; biyoistatistik üzerine çalışmaya başladığım günden itibaren her konuda yardımınızı, desteğinizi sürekli hissettim. Sizinle çalışma fırsatım olduğu için ve bilgi birikiminizden yararlanabildiğim için çok şanslıyım.

Tez izleme toplantılarına katılıp fikirlerini paylaşan tüm hocalarım; soru ve önerileriniz tezimin şekillenmesine oldukça katkı sağladı.

Değerli bölüm arkadaşlarım Yrd. Doç. Dr. Derya Öztuna, Dr. Beyza Doğanay Erdoğan ve Can Ateş; bilginiz, desteğiniz ve arkadaşlığınız sayesinde tez sürecim kolaylaştı ve keyifli hale geldi.

Biyoistatistik Anabilim Dalı'ndaki tüm hocalarım ve çalışma arkadaşlarım; desteğinizi, ilginizi sürekli hissettim.

Sevgili eşim Tolga, anneciğim, babacığım ve Burcu'm; bu tez sizlerin sayesinde tamamlandı.

Sevgili kızım İdil; bu süreçte seni çok ihmal etsem bile beni her gördüğünde gösterdiğin sevgi ile bana dünyanın en şanslı annesi olduğum hissini tattırdın.

SİMGELER ve KISALTMALAR

ABÖ:	Akaike bilgi ölçütü
BBÖ:	Bayes bilgi ölçütü
GSA:	Gizli sınıf analizi
HKO:	Hata kareleri ortalaması
İRM:	İkili Rasch modeli
KKM:	Kısmi kredi modeli
KRM:	Karma Rasch modeli
KDM:	Karma dağılım modeli
MC:	Monte Carlo
MİF:	Madde işlev farklılığı
MYT:	Madde yanıt teorisi
RM:	Rasch modeli
SKK:	Sınıf içi korelasyon katsayısı
SSM:	Sıralı sonuçlu model
θ :	Bireyin incelenen özellik düzeyi
β :	Maddenin zorluğu

ŞEKİLLER

	Sayfa
Şekil 3.1. MİF gösteren madde yüzdesi %10, MİF büyüklüğü 1 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri	35
Şekil 3.2. MİF gösteren madde yüzdesi %30, MİF büyüklüğü 1 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri	35
Şekil 3.3. MİF gösteren madde yüzdesi %10, MİF büyüklüğü 2 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri	36
Şekil 3.4. MİF gösteren madde yüzdesi %30, MİF büyüklüğü 2 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri	36
Şekil 3.5. Altın standart θ değerleri ile KKM ve KRM' den elde edilen θ değerleri arasındaki saçılım grafikleri	38
Şekil 3.6. KMM ve KRM' den elde edilen θ değerleri için yanlılık değerlerinin dağılımı	40
Şekil 3.7. KKM ve KRM' den elde edilen θ değerleri için HKO değerlerinin dağılımı	41
Şekil 3.8. KKM ve KRM' den elde edilen θ değerlerinin standart hatalarının dağılımları	43

ÇİZELGELER

		Sayfa
Çizelge 2.1.	Benzetim Koşulları	23
Çizelge 2.2.	Veri türetmede kullanılan kişilere ait gerçek parametre değerleri	25
Çizelge 2.3.	Veri türetmede kullanılan maddelere ait gerçek zorluk parametre değerleri	25
Çizelge 2.4.	Performans ölçütleri	31
Çizelge 3.1.	Her bir benzetim koşulundaki 200 MC tekrarda MİF gösteren maddelerin KKM ile doğru saptanma oranları	34
Çizelge 3.2.	KKM ve KRM'den elde edilen θ değerleri ile Altın Standart arasındaki SKK değerleri ve %95 güven sınırları	39
Çizelge 3.3.	KKM ve KRM'den elde edilen θ değerleri için yanlılık değerlerinin tanımlayıcı istatistikleri	40
Çizelge 3.4.	KKM ve KRM'den elde edilen θ değerleri için HKO değerlerinin tanımlayıcı istatistikleri	41
Çizelge 3.5.	KKM ve KRM'den elde edilen θ değerlerinin standart hatalarının tanımlayıcı istatistikleri	43
Çizelge 3.6.	KKM ve KRM'den elde edilen θ değerlerinin standart hatalarının ortalamalarına ilişkin %95 güven sınırları	44
Çizelge 3.7.	KKM ve KRM'den elde edilen θ değerlerinin standart hataları arasındaki SKK değeri ve %95 güven sınırları	44

1. GİRİŞ

1.1. Araştırmanın Konusu

Bilimsel çalışmalarda ölçme önemlidir. Ölçme, genel anlamıyla “herhangi bir niteliği gözlemek ve gözlem sonuçlarını sayılarla veya başka sembollerle ifade etmektir (Turgut, 1977). Ölçek, belirli bir özelliği ölçmek üzere geliştirilmiş, psikometrik özellikleri (geçerlik, güvenirlik, değişime duyarlılık) belirlenmiş ölçme aracı olarak tanımlanabilir. Ölçekler özellikle sağlık, eğitim ve sosyal bilimler alanlarında doğrudan ölçüm yapılamayan, gizli (latent) özelliklerin ölçülmesinde kullanılır. Zekâ, yetenek ve özürülük gibi özellikler soyut, gizli özelliklerdir. Ölçmenin amacı, belirlenen gizli özellikler açısından bireyleri sınıflandırarak belli kararlar vermek ve bu kararları genelleyeabilmektir. Benzer bireylerin maddelere benzer yanıtlar vermesi ölçmenin bu amacını kuvvetlendirmekte ve ölçeğin geçerliğini artırmaktadır. Ters bir durum söz konusu ise ölçülmek istenen özellik, bireylerin gözlenen gruplarından etkileniyor demektir. Bahsedilen bu gözlenen gruplara örnek olarak cinsiyet, eğitim durumu ve gelir durumu gibi doğrudan ölçülebilen özellikler verilebilir.

Bu çalışmada bir ölçeğin ölçmeyi amaçladığı özellik için (zeka ve yetenek gibi) **gizli özellik**, bu gizli özelliğe etki edebileceği düşünülen, doğrudan ölçülebilen özellik(ler) için (cinsiyet ve eğitim durumu gibi) **gözlenen grup(lar)**; doğrudan ölçülemeyen özellik(ler) için **gizli sınıf(lar)** tanımı kullanılacaktır.

Bireylerin gizli özelliklerinin ölçüldükten sonra sınıflandırılmalarını sağlayan modellere literatürde “Gizli Değişken Modelleri” denilmiştir. Bu Gizli Değişken Modelleri içerisinde, Temel Bileşenler Analizi, Faktör Analizi, Gizli Sınıf Analizi (GSA), Gizli Karakter Analizi, Gizli Profil Analizi ve Madde Yanıt Teorisi (MYT) gibi modeller bulunmaktadır (Çolak, 2003). Rasch Modeli (RM), tek parametrelili MYT’e benzer yapı göstermektedir. Rasch modelleri yaşam kalitesi, tutum ve özürülük düzeyi gibi gizli değişkenlerin değerlendirilmesinde kullanılmaktadır. Rasch analizinin sağlık, eğitim, psikoloji, pazarlama, ekonomi ve sosyoloji gibi farklı alanlarda kullanımı mevcuttur. Rasch modellerinin sağlık alanında ölçek geliştirmek amacıyla kullanımı standart bir yöntem haline gelmiştir. Bu nedenle, bu modellere ilişkin çalışmalar hızla artmaktadır (Erdoğan, 2012). Rasch Modelleri, incelenen veri

seti iki sonuçlu maddelerden oluşuyorsa İkili RM veya RM; çok sonuçlu maddelerden oluşuyorsa, maddelerin kategorilerinin özelliklerine bağlı olarak Kısmi Kredi Modeli (KKM) veya Sıralı Sonuçlu Model (SSM) olarak isimlendirilir. SSM ve RM, KKM'nin özel birer durumudur. Eşik değerleri arasındaki uzaklıklar her madde için aynı olduğunda KKM, SSM'ye; maddeler iki sonuçlu olduğunda ise RM'ye indirgenmektedir. Model tanımlarının detayı (1.2.1) alt başlığı altında verilmiştir. Bu çalışmada KKM kullanıldığı için metindeki tanımlar KKM'ye göre verilmiştir.

Rasch analizinde sıkça karşılaşılan sorunlardan biri maddelerin işleyişlerinin farklı olabilmesidir. Başka bir anlatımla, incelenen özellik düzeyi sabit tutulduğunda, veri setindeki bazı maddelere verilen yanıtlar, bireylere ait cinsiyet, yaş ve hastalık süresi gibi bazı faktörler açısından yanlılık gösterebilmektedir. Bu yanlılığa Rasch literatüründe madde işlev farklılığı (MİF) adı verilmektedir (Erdoğan, 2012).

Bu sorunu gidermek için öngörülen çözümlerden biri Karma Rasch Modelidir (KRM). RM ile yapılan analizlerde önceden belirlenen gözlenen gruplar açısından yanlılık gösteren maddeler belirlenir. Belirlenen bu maddeler ya ölçekten çıkartılır ya da ele alınan gözlenen grupların alt gruplarına göre madde parametreleri ayrı ayrı kestirilerek her birey için doğru yanıt olasılıkları hesaplanır. MİF gösteren maddelerin gözlenen grubun her bir alt grubuna göre ayrı parametre kestirimlerinin elde edilmesine **“KKM’de MİF analizi”** denir. Ancak bu yaklaşımın en büyük eksikliği, önceden belirlenen gözlenen gruplar üzerine odaklanması ve MİF'e neden olan gizli boyutu incelememesidir. KRM'de ise, bireyler önceden tanımlanmamış gizli sınıflara atanmaktadır. Bu sınıflarda bireylerin incelenen özellik (θ) düzeyleri açısından sınıf içi homojenlik, sınıflar arasında heterojenlik maksimumdur. Ayrıca KRM, her bir maddenin zorluğunun saptanan gizli sınıflarda farklı kestirilmesine olanak sağlamakta ve böylece MİF sorunu çözülmektedir. Dolayısıyla KRM'nin önceden saptanan gözlenen gruplar üzerine odaklanmadan bireyleri kendi aralarında heterojen gruplara ayırması ve her bir sınıfta madde parametre kestirimlerini sağlayarak MİF sorununu çözmesi, KKM'ye göre avantaj olarak gösterilmektedir (DeAyala et.al 2002).

KRM'nin literatürde öngörülen en büyük dezavantajı, model ile belirlenen gizli sınıfların gözlenen gruplar ile tanımlanabilme olasılığının düşük olmasıdır. Örneğin

KRM ile iki gizli sınıflı bir yapı oluşturulmuş ve veri setinde gözlenen grup olarak cinsiyet ele alınmışsa, KRM ile bulunan iki gizli sınıf ile cinsiyetin birebir örtüşme olasılığı, çoğu veri seti için düşüktür. KRM ile elde edilen sınıfları tanımlamaya yönelik olarak yapılmış birçok çalışma mevcuttur. Ancak bu çalışmaların büyük bir bölümü, KRM ile bulunan gizli sınıfları eldeki gözlenen gruplar ile tanımlamaya çalışmak yerine, doğrudan her bir gizli sınıftaki bireylerin yanıt desenlerini yorumlamanın daha etkin olacağını öngörmüşlerdir (Samuelson,2005; Cohen ve Bolt,2005; Cho, 2007; Bilir,2009).

MİF'in neden meydana geldiğine ilişkin olarak kullanılan iki farklı KRM yaklaşımı vardır. Bunlardan ilki, RM ile MİF gösteren maddelerin belirlenmesinden sonra, KRM ile MİF'in büyüklüğünün daha doğru kanıtının araştırmacıya sunulmasıdır. Mevcut yöntemler farklılık gösteren maddelerin sadece bir alt kümesini tanımlamakta ve MİF'in büyüklüğü için olması gereken değer altında kestirim yapmaktadır. Bu durum madde işlev farklılığına neden olan etkenleri tanımlamayı güçleştirmektedir. İkincisi ise, KRM ile grup üyeliklerinin kestirilmesi ve daha sonra eldeki gözlenen gruplar ile bu gizli sınıfların tanımlanmaya çalışılmasıdır. Bu durumda fazla sayıda MİF analizine gerek kalmadan MİF'in altında yatan neden hakkında bilgi alınmaktadır (Samuelson, 2005).

Bu çalışmada ilk olarak, benzetim çalışması ile MİF gösteren maddeler içeren veri setleri türetilmiştir. Türetilen bu veri setlerinde, KKM'de MİF analizi ile MİF gösteren maddelerin doğru saptanma oranları ve KRM ile gizli sınıf sayıları değerlendirilmiştir. İkinci olarak ise KRM ve KKM ile elde edilen θ kestirimleri ve bu kestirimlerin standart hataları karşılaştırılmıştır. Bu çalışmada KRM ve KKM ile elde edilen madde zorluk kestirimleri karşılaştırması yapılmamıştır. Çünkü KKM ile her bir maddeye ilişkin tek bir madde zorluk kestirimi elde edilirken KRM'de her bir gizli sınıf için madde zorluk kestirimleri elde edilmektedir. Literatürde yer alan ve bu çalışma ile benzer amaç güden çalışmalarda benzetim faktörleri de göz önüne alınarak, MİF analiz sonuçlarına etki edebileceği düşünülen faktörler benzetim çalışmasının koşulları olarak alınmıştır. Bu çalışmalarda örneklem büyüklüğü en az 500 olarak alınmıştır. Bu çalışmada ise beş yanıt kategorili maddeler için KRM ve KKM'nin karşılaştırması benzetim çalışması yapılarak değerlendirilmiştir. Ayrıca bu amaç küçük örneklem büyüklüğü altında tasarlanmıştır. Dolayısıyla bu çalışmanın

literatüre katkısı, belirli benzetim koşulları altında, KRM'nin çok sonuçlu maddeler ve küçük örneklemler için etkinliğinin değerlendirilmesi yönündedir.

1.2. Teorik Yapı

1.2.1. Rasch Modelleri

1.2.1.1. İkili Rasch Modeli

Danimarkalı matematikçi George Rasch tarafından 1960 yılında, iki sonuçlu (evet/hayır, var/yok, katılıyorum/katılmıyorum vb.) maddeler için geliştirilen İkili Rasch modelinde bireyin bir maddeye doğru yanıt verme olasılığı, bireyin θ düzeyi ile maddenin zorluğu (β) arasındaki farkın lojistik fonksiyonu olarak tanımlanır. İkili Rasch Modeline kısaca Rasch Model de denilmektedir (RM). İki sonuçlu maddeler, tek basamaklı maddeler olarak da adlandırılmaktadır. Eğer bireyler bu tek basamağı başarı ile gerçekleştirirlerse 1 puan, gerçekleştiremezlerse 0 puan almaktadırlar. Bu modelde madde zorluk parametresinin (β) yanı sıra bireyin θ düzeyi de kestirilmektedir. İncelenen özellik düzeyi θ_i ve madde zorluk düzeyi β_j olmak üzere i.bireyin j.inci maddeye x yanıtı verme olasılığı aşağıdaki gibi hesaplanmaktadır.

$$P(X = x_{ij} | \theta_i) = P_{ij} = \frac{\exp[x_{ij}(\theta_i - \beta_j)]}{1 + \exp(\theta_i - \beta_j)} \quad ,x=0, 1 \quad (1.1)$$

1.2.1.2. Kısmi Kredi Modeli

Çok sonuçlu maddeler için geliştirilmiş olan kısmi kredi modelinde (KKM), maddelere verilen yanıtların sadece doğru/yanlış olarak değil, doğruluk kavramını göreceli olarak geliştirip bir ya da daha fazla orta düzeye yayarak, θ düzeyi orta derecede olan kişilerin de kredilendirilmesi mümkün olmaktadır. Bu model Masters tarafından 1982 yılında çoklu aşama gerektiren ve çözümlene sürecinde farklı aşamaların tamamlanması durumunda kısmi kredi vermenin önemli olduğu veya Likert tipi maddelerde yanıt kategorileri arasındaki uzaklıkların maddeden maddeye

farklılık gösterdiği durumlar için geliştirilmiştir. Bu model ile i.nci kişinin, j.inci maddeden x puan alma olasılığı aşağıdaki gibi hesaplanabilmektedir.

$$P_{ijx} = \frac{\exp \sum_{k=0}^x (\theta_i - \beta_{jk})}{\sum_{k=0}^{m_i} \exp \sum_{t=0}^k (\theta_i - \beta_{jt})} \quad (1.2)$$

$$\beta_{j0} \equiv 0, \sum_{k=0}^0 (\theta_i - \beta_{jk}) = 0, \exp \sum_{k=0}^0 (\theta_i - \beta_{jk}) = 1$$

1.2.1.3. Sıralı Sonuçlu Model

Sıralı kategorileri aşağıdaki gibi olan bir davranış anketi için:

Kesinlikle katılmıyorum	Katılmıyorum	Katılıyorum	Kesinlikle katılıyorum
0	1	2	3

“Katılıyorum” kategorisini seçen bir bireyin; “Kesinlikle katılmıyorum” yerine “Katılmıyorum” kategorisini (yani ilk basamağı tamamladığı) ve aynı zamanda “Katılmıyorum” yerine de “Katılıyorum” kategorisini seçtiği (yani ikinci basamağı tamamladığı) düşünülür. Ancak “Katılıyorum” yerine “Kesinlikle katılıyorum” kategorisini seçmekte başarısız olduğu (yani üçüncü basamağı tamamlayamadığı) ifade edilebilir. Sıralı sonuçlu modelde (SSM), basamakların zorluk düzeylerinin maddeden maddeye değişmediği varsayılır. Bu varsayım KKM’de madde basamaklarını aşağıdaki gibi iki bölümden oluşacak şekilde yeniden yazarak gerçekleştirilebilir:

$$\beta_{jk} = \beta_j + \tau_k \quad (1.3)$$

β_j değeri, seçilen bir maddenin kategori eşik değerlerine göre ortalama zorluğu gösterirken, τ_k değeri her maddedeki k’ncü basamağın eşik değerini gösterir. KKM

için verilen (1.2) eşitliğinde β_{jk} yerine (1.3) eşitliği yazıldığında, SSM elde edilir. Bu model ile i. kişinin, j. maddeden x puan alma olasılığı aşağıdaki gibi hesaplanabilir.

$$P_{ijx} = \frac{\exp \sum_{k=0}^x (\theta_i - (\beta_j + \tau_k))}{\sum_{k=0}^{m_i} \exp \sum_{t=0}^k (\theta_i - (\beta_j + \tau_k))}, x=0,1,\dots,m_i \quad (1.4)$$

$$\tau_0 \equiv 0, \exp \sum_{k=0}^0 (\theta_i - (\beta_j + \tau_k)) = 1$$

1.2.1.4. Rasch Modelinin Varsayımları

Rasch modeli aşağıdaki varsayımlara sahiptir:

1. Ölçeği oluşturan maddelerin tümü tek bir özelliği ölçmeli yani ölçülen özellik tek boyutlu olmalıdır. Eğer tek boyutluluk sağlanmıyorsa doğru ölçüm yapılamaz.
2. Maddelere doğru yanıt verme olasılığı, θ düzeyi arttıkça monoton olarak artar.
3. Belirli bir θ düzeyinde, maddelere verilen yanıtlar birbirinden bağımsız olmalıdır.
4. Bireylere ve maddelere ait toplam puanlar sırasıyla kişi ve madde parametrelerinin yeterli istatistikleridir.

Yukarıda tanımlanan birinci varsayımın bozulmasına neden olan faktörlerden biri de maddelerin ölçeğin uygulandığı bireylerin oluşturduğu alt gruplarda işleyişinin farklı olmasıdır. MİF, aynı θ düzeyine sahip olan ancak farklı gruplardan gelen bireylerin bir maddeye yanıt verme olasılıklarının farklı olmasıdır. Gözlenen bir grup açısından MİF gösteren maddeler için öngörülen ortak çözüm, MİF'e neden olan maddelerin ölçekten çıkarılmasıdır. Ancak bu durum ölçekte klinik anlamlılığa sahip olan maddelerin kaybına neden olmakta ve madde kaybı o ölçeğin ölçmek istediği gizli özelliği tam olarak ölçememe sorununu beraberinde getirmektedir. Ölçekte gözlenen bir grup açısından (örn. cinsiyet) MİF gösteren maddelerin

ölçekten çıkarılması yerine bireylerin yanıt farklılıklarının altında yatan gizli yapıyı ortaya koymaya yönelik yeni istatistiksel modeller geliştirilmektedir. Bu modellerden biri KRM'dir. Bu çalışmada MİF gösteren veri setleri üzerinde KRM ve RM kestirimleri inceleneceği için MİF üzerinde detaylı açıklama aşağıdaki alt başlık altında verilmiştir.

1.2.1.5. Madde İşlev Farklılığı (MİF)

MİF, farklı gruptaki bireylerin θ düzeyleri sabit tutulduğunda, bu bireylerin madde(ler) üzerindeki başarı olasılıklarının veya olabirliklerinin farklı olmasıdır. Bu farklılık göz önüne alındıktan sonra bireyler ilgilenilen gizli özelliklerle eşleştirilirler (Clauser ve Mazor, 1998). MİF'in varlığı farklı gruptan bireylerin karşılaştırılmayacağını gösterir. MİF'in göz ardı edilmesi ise çıkarsamaların bireyler göz önüne alınarak yapıldığının göstergesidir. MİF, bazı veri setlerinde çok boyutluluk sinyalleri de verir (Ackerman, 1992). Sonuç olarak en temel düzeyde MİF'in varlığı, test sürecinde eşitlik ve doğruluk sorunları doğurmaktadır.

MİF önemli sorunlara neden olduğu için birçok kapsamlı çalışmaya konu olmuştur. MİF'in tanımlanma sürecinde çok fazla üzerinde durulmayan iki önemli problem vardır. Bu problemlerden ilki cinsiyet ve eğitim durumu gibi gözlenen gruplar ile ilgilidir. İkinci problem ise MİF'in neden olduğu bilgi kaybıdır.

Geleneksel MİF analizleri De Facto varsayımına dayalıdır. Bu varsayım gözlenen grubun alt gruplarındaki bireylerin grup içi benzer olup alt gruplar arasında bireylerin farklılaştığına yöneliktir (DeAyala et.al.,2002). Gerçekte bireye ilişkin faktörler kolayca tanımlanabilen gruplar olmasına rağmen bu gruplar genellikle homojen popülasyonları temsil etmezler. Gruplar içindeki bu değişkenliği gösteren bir örnek Amerika'daki İspanyol popülasyonudur. ABD nüfus sayım bürosuna göre, İspanyol kökenli bireylerin belirli bir kısmı kökenlerinin Meksikalı, Puerto Ricalı, Kübalı, Merkez veya Güney Amerikalı veya diğer İspanyol ırklarından biri olduğunu söylemektedir. İspanyol ırkından olan birinin herhangi bir ırka ait olacağı da unutulmamalıdır. Bu farklılık İspanyol sınıflamasının heterojen bir yapıda olduğunu göstermektedir. Aynı örnek farklı ırklar, soylar veya cinsiyet için de verilebilir. Bu homojenlik bozulmasına bağlı olarak bir gözlenen gruba göre maddeler MİF gösterirse bireylerin bir kısmının kendi grubundaki bireylerle benzer yanıtlar vermesi

beklenir. Cohen ve Bolt (2002) tarafından yapılan bir çalışmada maddelerin cinsiyete karşı MİF gösterdiği saptanmıştır. Bu maddeler gizli sınıf analizi ile tekrar incelendiğinde çalışmadaki kadınların yarısının, erkeklerin de %40'a yakınının ters grupta sınıflandığı görülmüştür. Bu örnek, gözlenen grupları kullanarak MİF'in nasıl yüzeysel kaldığının bir göstergesidir. Başka bir örnek olarak Hu ve Dorans'ın (1989) çalışması verilebilir. Hu ve Dorans, kadınların lehine olan maddelerin kaldırılması sonucu kadınların puanlarının biraz düştüğünü, erkeklerin puanlarının biraz yükseldiğini belirlemişlerdir. Araştırmacıları şaşırtan durum ise İspanyol ve Asyalı-Amerikalı erkeklerin puanlarının daha fazla artması olmuştur. Bu sonuçlar marjinaler üzerine odaklanan MİF analizindeki hatayı göstermektedir ve MİF analizindeki etkileşimlerin göz önünde bulundurulmasına dikkat çekmektedir. Dorans ve Holland'ın (1993) öngördüğü "Melting-Pot DIF" analizine göre her bir cinsiyet/etnik grup ile diğer bireylerin oluşturduğu popülasyon, madde fonksiyonları açısından karşılaştırılmaktadır (çeşitli ırk ve ulustan insanların kaynaştığı yer-the melting pot). Bu analizin daha az avantajlı olduğu saptanmıştır, çünkü gerekli olan analiz sayısı fazladır ve bu strateji ile daha geniş gruplarda MİF bulmak zordur. DeAyala, et.al (2002) çalışmasında bir gizli sınıfa ait olan siyahi bireylerin testteki üç maddeden etkilendiğini ancak diğer sınıftaki bireylerin etkilenmediğini saptamışlardır. Bu durumda öngörülen, maddelerin gözlenen gruplardan bağımsız MİF gösterdiğini saptamaktır. Aksi durumda gözlenen grubun tüm üyeleri MİF gösteren bir madde veya maddeler açısından yanlış olarak tanımlanacaktır.

MİF saptamada gözlenen grupları kullanmanın sakıncalı olmasının bir başka nedeni ise bu gözlenen grupların araştırmacının ilgilendiği durumla doğrudan ilgili olmayabileceğidir. Araştırmacılar cinsiyet ve ırk gibi gözlenen grupların gerçekte başka bir değişken için kestirici olabileceğini tartışmaktadır. Dorans ve Holland'a (1993) göre:

"Etnik grupların bazıları dezavantajlı atfedilirler ve odak grup tanımına uyarlar. Gerçekte bu grupların herhangi birinin içinde avantaj veya dezavantaja göre oldukça büyük bir değişkenlik olasılığı vardır. Belki de grup tanımımızı bu avantaj veya dezavantajı tanımlayarak ve ölçerek yapmalıyız."

Bazı maddelerin işleyişlerinin farklı olduğu sonucuna varıldıktan sonra, MİF'e neden olan kaynağı belirleme gereği doğar. MİF'e neden olan etkenlerin araştırıldığı

çalıřmalarda genellikle, daha önceki benzer çalıřmalarda MİF'e neden olan etkenler referans gösterilmiř ya da MİF'e neden olan gizli boyutu tanımlamaya yönelik yanıt modelleri arařtırılmıřtır. Birçok arařtırmacı bu yöntemlerin MİF saptamada sınırlı başarıya sahip olduđunu düşünmektedir. Var olan yöntemlerin neden yetersiz olduđuna dair en anlamlı cümlelerin Standards for Educational and Psychological Testing tarafından kurulduđu ifade edilebilir:

“MİF yöntemleri test kalitesini arttırmaya yönelik olsa bile, MİF içeren maddeleri tanımlayan yetersiz bir süreç vardır. Bu nedenle, bir testte bir gruptan diđer gruba işleyiř farklılıđı gösteren maddeler varsa bu farklılıđa neden olan asıl etkeni bulmak zordur”(1999).

Kelderman ve Macready (1990) MİF sorununa gizli sınıf analizi ile yaklařmanın avantajlı olduđunu ifade etmiřlerdir. Onlara göre: *“Gizli grup deđiřkeni kullanmak MİF'i herhangi bir özel deđiřkene veya deđiřken setine bađlamadan MİF deđerlendirmesi sađlar. Böylece MİF saptamasının daha tanımlayıcı olarak yapılması mümkün olabilir”.*

1.2.1.6. MİF Saptama Yöntemleri

Bazı MİF saptama yöntemleri ıkarsamalı istatistik tabanlıdırlar ve MYT modellerini kullanırlar. İki sonuçlu maddeler için birçok MİF analiz yöntemi vardır (Camilli, 2006; Clauser ve Mazor, 1998; Hills, 1990; Millsap ve Everson, 1993; Penfield ve Camilli, 2007). Bunların içinde Mantel-Haenszel test (Holland ve Thayer, 1988), standardizasyon istatistiđi (Dorans, 1989), SIBTEST istatistiđi (Shealy ve Stout, 1993), lojistik regresyon (Swaminathan ve Rogers, 1990), Lord'un ki-kare istatistiđi (Lord, 1977; Lord, 1980), Raju'nun alan ölçümleri (Raju, 1988; Raju, 1990) ve olabilirlik oran testi (Thissen et.al 1988; Thissen, et.al, 1993) sayılabilir. Bu yöntemlerden standardizasyon istatistiđi (Zwick ve Thayer, 1996), SIBTEST istatistiđi (Chang et.al, 1996), lojistik regresyon (French ve Miller, 1996) ve olabilirlik oran testi (Cohen, 1988) çok sonuçlu maddelere genellenebilir. Sadece çok sonuçlu maddeler için öngörülen MİF saptama yöntemleri arasında Mantel test (Zwick et.al, 1993), HW1 ve HW3 istatistiđi (Welch ve Hoover, 1993), lojistik ayırsama fonksiyon analizi (Miller ve Spray, 1993), Cox's β istatistiđi (Camilli ve Congdon, 1999) ve Liu-Agresti istatistiđi verilebilir. (Liu ve Agresti, 1996).

MİF gösteren maddelerin belirlenmesi için kullanılan en yaygın yaklaşım RM'de bulunan artıkların kullanılmasıdır. Ölçeği yanıtlayan n. Kişi, i. iki sonuçlu maddeye yanıt vermiş olsun. Bu yanıt için RM'den bulunan beklenen değer ve varyans da sırasıyla $E[y_{ni}]$ ve $V[y_{ni}]$ olsun. Bu yanıtla ilişkin standartlaştırılmış artık değeri aşağıdaki gibi olacaktır:

$$Z_{ni} = \frac{y_{ni} - E[y_{ni}]}{V[y_{ni}]} \quad (1.5)$$

Standartlaştırılmış artıklar bireylerin maddelere verdikleri bütün yanıtlar için hesaplanır. Daha sonra bireyler θ kestirimlerine göre, her sınıfta en az 50 birey olacak şekilde, G tane sınıf aralığına (class interval) bölünür. Her madde için standartlaştırılmış artık değerleri bağımlı değişken, bulunan G tane sınıf aralığı ve kişi faktörü bağımsız değişkenler olarak alınıp, iki yönlü varyans analizi gerçekleştirilir. Varyans analizi sonucunda iki çeşit MİF bulmak mümkündür. Birincisi θ düzeyi boyunca değişmeyen tekdüze (uniform) MİF, diğeri ise θ düzeyi boyunca farklılık gösteren tekdüze olmayan (non uniform) MİF'tir. Tekdüze MİF yalnızca kişi faktörünün varyans analizi sonucunda anlamlı bulunması, tekdüze olmayan MİF ise varyans analizi sonucunda kişi faktörü ve θ düzeyi sınıfları arasındaki etkileşim teriminin anlamlı bulunması ile ortaya konabilir (Erdoğan, 2012).

Bu çalışmada, RM ile elde edilen kestirimler üzerinde, MİF analizi için iki yönlü varyans analizi yöntemi kullanılacaktır. Benzetim çalışmasında ve gerçek veri seti üzerinde yapılan MİF analizinde tekdüze ve tekdüze olmayan MİF ayrımı yapılmamış, maddelerin sadece MİF gösterip göstermediği sorgulanmıştır.

1.2.1.7. MİF Gösteren Maddelerin Değerlendirilmesi

Uygulamada eğer bir madde RM ile yapılan analizlerde, önceden belirlenen gözlenen gruplar açısından yanlılık gösteriyorsa bu sorunu gidermek için öngörülen çözümlerden biri belirlenen bu maddeyi ölçekten çıkarmaktır. Diğeri bir çözüm ise MİF'e neden olan gözlenen grubun alt gruplarına göre madde parametresini ayrı ayrı

kestirerek her birey için doğru yanıt olasılıklarını bu düzeltme kullanarak hesaplamaktır. İkinci çözüme “**KKM’de MİF analizi**” denir.

MİF’i tanımlarken gözlenen grupları kullanmanın dışında öngörülen bir yöntem ise KRM’yi kullanmaktır. KRM ile bireyler, önceden tanımlanmamış gizli sınıflara atanırlar. Bu sınıflar kendi içinde homojen ve kendi aralarında heterojen yapıya sahiptir. Her bir gizli sınıfta madde parametreleri ayrı kestirilir ve maddelerin zorluk düzeylerinin sıralanışı da farklılaşabilmektedir. Bir madde için parametre kestirimi her bir gizli sınıfta ayrı ayrı yapıldığından dolayı bireyin dahil olduğu gözlenen gruptan bağımsız madde parametre kestirimleri elde edilmiş olur. Bu da bir madde için MİF sorununun ortadan kalktığını ifade etmektedir.

1.2.2. Karma Rasch Model

KRM ilk kez, iki sonuçlu maddeler için, 1990 yılında Rost tarafından tanımlanmıştır (Rost, 1990). Öngörülen modelin, RM’nin teorik gücü ile GSA’nın deneysel gücünü birleştirdiği ifade edilmiştir. Modelin, RM’nin ve GSA’nın varsayımlarını ortadan kaldırdığı öngörülmüştür. GSA’da her bir gizli sınıf içinde maddelerin zorluk sırasının aynı olması koşulu vardır. KRM ise farklı gizli sınıflarda madde zorluklarının sırasının da farklılaşmasına olanak sağlar. Dolayısıyla bu modelde farklı gizli sınıflara atanan bireyler için madde zorlukları da farklılaşabilecektir. KRM’de çok sonuçlu maddeler için parametre kestirimleri ise ilk kez von Davier ve Rost (1995) tarafından öngörülmüştür.

KRM, RM’nin kesikli karma dağılım modeline genelleştirilmesi ile elde edilir. Kısaca açıklanırsa, KRM, RM’nin tüm popülasyon için uyumlu olmayacağını ancak önceden bilinmeyen, bireylerin alt popülasyonları için uyumlu olacağını varsayar (Fischer ve Molenaar, 1995).

RM’den KRM’ye geçişi açıklamadan önce karma dağılım modellerinden bahsetmek gereklidir.

1.2.2.1. Karma Dağılım Modelleri

Karma dağılım modellerinin (KDM) ana varsayımı, bir gözlenen rastgele değişkenin (genellikle vektör değerli) tekdüze olasılık fonksiyonu şeklinde tanımlanamayıp koşullu olasılık fonksiyonları ile tanımlanmasıdır. Rastgele değişken kesikli ise kesikli KDM, sürekli ise sürekli KDM'den bahsedilir. KRM kesikli bir KDM'dir. Dolayısıyla bu bölümde kesikli karma dağılım modelleri üzerinde durulacaktır.

Karma değişkenin kesikli olması durumunda n madde sayısı olmak üzere, her bireyden elde edilen yanıt deseni $x=(x_1, \dots, x_n)$ 'in koşulsuz olasılığı olan $P(X=x)$ aşağıdaki eşitlik ile verilir:

$$P(X = x) = \sum_{c=1}^c P(c)P(X = x | \theta = c) \quad (1.6)$$

Burada c, kesikli karma değişkenin bileşenlerinin sayısını göstermektedir. Ağırlıklar olan P(c)'ler ise karma değişkenin c değerini alma olasılığıdır. Böyle bir kesikli KDM'de θ parametresi birey sınıflarını veya karma bileşenlerin sayısını gösterir. Bu nedenle, P(c) ağırlıkları sınıf büyüklüklerinin veya karma bileşenlerinin göreceli büyüklüklerini tanımlar. Bu durumda koşulsuz P(x) olasılığı, ağırlıklandırılmış toplamdır.

MYT modelleri genellikle KDM'dir; sürekli veya kesikli olabilmektedirler. MYT modelleri tarafından tanımlanan gözlenen rastgele değişkenler, bireylerin yanıt vektörleridir. θ karma değişkeni "gizli değişken" olarak bilinmektedir. Koşullu olasılıklar lojistik fonksiyon olarak tanımlanmaktadır ve örneğin RM için θ yetenek düzeyi bilinen bir bireyin x yanıt desenine sahip olma olasılığı aşağıdaki eşitlik ile verilir:

$$P(X = x | \theta) = \prod_{i=1}^k \frac{\exp[x_i(\theta - \beta_i)]}{1 + \exp(\theta - \beta_i)} \quad (1.7)$$

Bu eşitlikte k madde sayısı, β_i i.maddenin zorluk parametresidir. Bu durumda, RM'in sürekli KRM olduğu ifade edilebilir çünkü θ gerçek değerli gizli değişkendir.

$r = \sum_{i=1}^k x_i$ bir yanıt vektörünün toplam skorudur ve r , θ kestirimi için yeterli bir istatistiktir. Dolayısıyla her bireyi toplam skorlarına göre sınıflandırmak ile kestirilen $\hat{\theta}$ parametrelerine göre sınıflandırmak aynı sonucu verir ($\hat{\theta}$ 'ya göre sınıflandırmak, θ 'ya göre değil). Bu nedenle RM, θ 'dan bağımsız olacak şekilde yeniden parametrize edilerek kesikli KDM olarak yazılabilmektedir. RM'nin kesikli hali eşitlik (1.9)'da verilmiştir:

$$P(X = x) = P(r_x)P(X = x | r_x) \quad (1.8)$$

olmak üzere,

$$P(X = x | r_x) = \frac{\exp(\sum_{i=1}^k x_i \beta_i)}{\gamma_r(\exp(\beta))} \quad (1.9)$$

Burada $\gamma_r(\exp(\beta))$, $\exp(\beta_i)$ madde parametrelerinin r . sıra elementer simetrik fonksiyonlarıdır.

RM'nin bu şekilde kesikli KDM olarak yazılmasının avantajı karma değişken olan θ 'dan bağımsız olduğu için θ dağılımı hakkında kısıtlayıcı herhangi bir varsayımın olmamasıdır. Kesikli RM'nin ve GSA'nın teorik yapıları kullanılarak KRM parametreleri elde edilir. Dolayısıyla bu noktada GSA'dan bahsetmek gereklidir.

1.2.2.2. Gizli Sınıf Analizi

Gizli Sınıf Analizi (GSA), iki veya daha fazla kategorik değişken arasındaki ilişkiyi, değişkenlerin çapraz sınıflaması ile açıklayarak çok boyutlu, kategorik gizli değişkenleri tanımlayan, ilgili alt grupların belirlenmesini sağlayan istatistiksel bir yöntemdir. GSA'nın amacı, gözlenen değişkenlerden ya da ölçümlerden gizli değişken olarak adlandırılan tek bir ölçüt elde etmektir. GSA'da iki tür değişken

tanımlanır: Doğrudan gözlenen değişkenler ile doğrudan gözlenemeyen değişkenler. GSA'da model parametreleri, koşullu yanıt olasılıklarının en çok olabilirlik kestirimleridir. Koşullu olasılıklar ise her bir gizli sınıf için, o sınıfın bir üyesinin açıklayıcı değişkene öngörülen değeri vermesi olasılığıdır. Her bir sınıf içerisinde toplamları 1'i veren, kestirilen grup üyelikleri olasılıkları da GSA'nın model parametreleridir. En çok olabilirlik kestirimi, adımsal olarak en olası gözlenen sonuçları verecek model parametrelerini hesaplar. Ayrıca, en çok olabilirlik kestirim yöntemi ile bireylerin gizli sınıflara ataması yapılır. GSA'da sınıf sayısı arttıkça model parametreleri de artmaktadır.

GSA'da temel adımlardan birincisi, var olan veri setinden model parametrelerinin kestirilmesidir. Daha sonra optimal sınıf sayısının belirlendiği model seçimi aşaması gelir. Model seçimi yapıldıktan sonra seçilen modelin uyum değerlendirmesi yapılır ve model açıklaması ile aşamalar son bulur.

GSA'da ilk temel varsayım, popülasyonun bağımsız ve homojen örneklemelerden oluştuğudur. İkinci temel varsayım ise yerel bağımsızlık varsayımıdır. Bu varsayım gözlenen değişkenler arasındaki ilişkinin bu değişkenlerin gizli sınıflamayla olan ilişkileri ile açıklanabilmesidir. Buna bağlı olarak, verilen bir gizli örnekte, tüm gözlenen değişkenler istatistiksel olarak bağımsız olacaktır. GSA'nın merkezini koşullu bağımsızlık oluşturur. Bu ise her bir gizli sınıf için gözlenen değişkenlerin bağımsızlığı demektir. GSA işlemsel açıdan kolay ve hazır yazılımlarla çözümlenebilen bir yöntemdir. Diğer yöntemlerden ayrılan yanı ise yöntemin gizli çok değişkenli normal dağılım varsayımının olmayışdır (Çolak, 2003).

GSA'da madde yanıt olasılıkları gizli sınıf üyelik fonksiyonu olarak ortaya konulmuştur. Gizli sınıf modellerinin uygulanmasındaki en önemli varsayımlardan birisi, yerel bağımsızlıktır. Yerel bağımsızlık ise, sınıf üyelikleri üzerinde koşullu olarak madde yanıtlarının istatistiksel olarak bağımsız olmaları demektir. Bu kural sınıf içi homojenliğini garanti etmek için gerekli bir varsayımdır.

Gizli sınıf modelleri, madde yanıtları için en basit kesikli KDM olarak tanımlanabilir.

$$P(X = x) = \sum_{c=1}^C \pi_c \prod_{i=1}^k \pi_{ic}^{x_i} (1 - \pi_{ic})^{1-x_i} \quad (1.10)$$

$\pi_c = P(c)$, c.sınıfın ortaya çıkma olasılığıdır ve bu KDM'de karma oranı olarak bilinir. Diğer bölümü ise koşullu yanıt olasılıkları olan $\pi_{x|c} = P(x_i = 1 | c)$ olasılıklarıdır.

1.2.2.3. İkili Karma Rasch Model

Bir kesikli karma dağılımda koşullu olasılıkların RM ile tanımlandığı varsayılarak (ki bunların kendileri sürekli karmadır) aşağıdaki eşitlikler elde edilir:

$$P(X = x) = \sum_{c=1}^c \pi_c P(X = x | c) \quad (1.11)$$

Burada $P(X=x)$, bir bireyin x yanıt desenine sahip olma olasılığıdır. c , kesikli karma değişkenin bileşenlerinin sayısını göstermektedir. π_c ise GSA'da tanımlanan sınıf büyüklüğü parametresidir. $\pi_{x|c} = P(x | c)$ 'dir. Bu olasılık c .sınıfta olduğu bilinen bir bireyin x yanıt desenine sahip olma olasılığıdır. θ sürekli değişkeni her biri c sınıfında r . sıra γ_{rc} elementer simetrik fonksiyonların tanıtılması ile modelde devre dışı bırakılabilir. Ayrıca, koşullu yanıt olasılıkları aşağıdaki gibi yazılabilir:

$$P(X = x | c) = \pi_{x|c} \frac{\exp\left(\sum_{i=1}^k x_i \beta_{ic}\right)}{\gamma_r(\exp(\beta_c))} \quad (1.12)$$

Bu model KRM olarak bilinmektedir (Rost, 1990). Bu modelde $\pi_{x|c}$ gizli skor olasılıklarıdır ve β_{ic} ise sınıfa özel madde parametreleridir. Skor olasılıkları gizli olarak adlandırılır çünkü doğrudan gözlenen frekanslardan elde edilemezler. Bunun yerine gizli sınıf yapısının bir parçasıdır ve bu nedenle gözlenen skor olasılıklarının bileşenlerine bağlıdır.

$$P(x) = \sum_{c=1}^c \pi_c \pi_{x|c} \quad (1.13)$$

Bileşen veya sınıf sayısının (c 'nin) bir model parametresi olmadığına dikkat etmek gerekmektedir; sınıf sayısı bir önsel olarak varsayılmalıdır ya da farklı sınıf sayıları altında kurulan modellerin uyumu değerlendirilmelidir. Parametrelerin normalleştirme koşulları aşağıdaki gibidir. Sınıf büyüklüklerinin toplamı, iki uç skordan 1'in çıkarılması ile elde edilir:

$$\sum_{c=1}^c \pi_c = 1 - P(r = 0) - P(r = k) \quad (1.14)$$

Her bir sınıf içinde madde parametrelerinin normalleştirilmesi ise şöyledir:

$$\sum_{i=1}^k \beta_{ic} = 0 \quad (1.15)$$

KRM'de parametre kestirimi Expectation-Maximization (EM) algoritması kullanılarak yapılır. EM algoritmasındaki adımların içerikleri aşağıda tanımlanmıştır.

E.adım

E. adımda her bir sınıf için beklenen yanıt frekansları, model parametrelerinin başlangıç kestirimleri (veya başlama değerlerini) temel alınarak kestirilir.

$$\hat{n}(x, c) = n(x) \frac{\pi_c P(x | c)}{\sum_{c=1}^c \pi_c P(x | c)} \quad (1.16)$$

Burada $n(x)$ $X=x$ vektörünün gözlenen frekansını ve $n(x,c)$ ise c .sınıf için $X=x$ vektörünün gözlenen frekansını gösterir. Koşullu yanıt olasılıkları $P(x|c)$ KRM ile tanımlanmıştır.

M.adım

M. adımda gözlenen yanıt frekanslarının bu oranları, π_c , $\pi_{x|c}$, β_{ic} kestirimlerini elde etme amaçlı kullanılır. Bu parametreler, her bir gizli sınıf için, c .sınıfta log-olabilirlik fonksiyonunu en büyükleyerek ayrı ayrı kestirilir:

$$\ln L_c = \sum_x \hat{n}(x | c) \left[\ln \pi_{x|c} + \sum_{i=1}^k x_i \beta_{ic} - \ln[\gamma_r(\exp \beta_c)] \right] \quad (1.17)$$

Bu eşitliğin β_{ic} 'ye göre ilk kısmi türevi alınıp 0'a eşitlenirse, c sınıfında madde parametreleri için kestirim eşitliği elde edilir:

$$\hat{\beta}_{ic} = \ln \frac{n_{ic}}{\sum_{r=0}^k m_{rc} \gamma_{r-1}^{(i)} / \gamma_r} \quad (1.18)$$

Burada n_{ic} , c.sınıfta i. maddeye 1 yanıtı veren birey sayısının başlangıç kestirimidir. m_{rc} ise c. sınıftaki r skorlu bireylerin sayısının kestirimidir. Hem n_{ic} hem de m_{rc} M. adımda elde edilen $n(x,c)$ 'den elde edilir. $\gamma_{r-1}^{(i)}$ ise i.madde hariç tüm $\exp(\beta_i)$ parametrelerinin r-1. sıra elementer simetrik fonksiyonlarıdır.

Bu süreçte tek bir iterasyon yeterlidir, çünkü bu süreç EM algoritmasının her bir M. adımında uygulanır. Bu nedenle EM süreci boyunca maksimum olabilirlik çözümüne yakınsanır. Skor olasılıkları ve sınıf büyüklüklerinin kestiricileri açıkça şöyle elde edilir:

$$\hat{\pi}_{x|c} = \frac{m_{rc}}{n_c} \text{ ve } \pi_c = \frac{n_c}{N} \quad (1.19)$$

n_c , c. sınıftaki birey sayısını gösterir ve $n(x,c)$ 'ye bağlıdır.

Gizli özellik parametresi olan θ , bu eşitliklerde bulunmaz çünkü bunların yerini, θ 'ların yeterli istatistikleri olan r_c 'ler almıştır. r_c 'ler, maddelerin ve skor parametrelerinin sonsal kestirimleri kullanılarak kestirilirler. Bu ise sınıf içi olabilirliği en büyüklenerek yapılabilir:

$$\ln L_c = \sum_v \ln P(x_v | c) = \sum_v r_v \theta_{vc} + \sum_i n_{ic} \beta_{ic} - \sum_v \sum_i \ln(1 + \exp(\theta_{vc} + \beta_{ic})) \quad (1.20)$$

Bu eşitlik, bilinmeyen özellik parametresi θ_{vc} 'yi baz alır. θ_{vc} , v. bireyinin r_v skoruna bağlıdır. c. sınıfta i. maddeye yanıt veren birey sayısı n_{ic} olmak üzere, n_{ic} 'nin bilinmesi gerekmez. (1.20)'nin θ_{vc} 'ye göre diferansiyeli alınırsa bu terim yok edilir. Kestirim eşitliği,

$$r_v = \sum_{i=1}^k \frac{\exp(\theta_{vc} + \beta_{ic})}{1 + \exp(\theta_{vc} + \beta_{ic})} \quad (1.21)$$

ile verilir. Bu eşitlik iteratif olarak çözülebilir.

Dolayısıyla model (1.12)'nin sınıfa bağımlı incelenen özellik parametreleri, kestirim sürecinin ikinci adımında kestirilebilir ve bu kestirimler ilk adımda elde edilen koşullu madde parametrelerini (β_{ic}) kullanmaktadır.

1.2.2.4. Çok Sonuçlu Karma Rasch Model

(1.2) ile tanımlanan KKM, karma dağılım modeline genelleştirilebilir (Rost,1991). c.sınıfta olduğu bilinen i.bireyin j.maddeden x puanı alma olasılığı aşağıdaki eşitlik ile verilir:

$$P_{ijx|c} = \sum_{c=1}^c \pi_c \frac{\exp \sum_{k=0}^x (\theta_{ic} - \beta_{jkc})}{\sum_{k=0}^{m_i} \exp \sum_{t=0}^k (\theta_i - \beta_{jtc})}, \quad x=0,1,\dots,m_i \quad (1.22)$$

Madde parametreleri ve incelenen özellik parametreleri sınıfa özgüdür ve her bir gizli sınıf için kestirilir. KKM'deki madde parametreleri gibi çok sonuçlu KRM'de de eşik parametreleri bir normalleştirme koşulu içerir; yani bir gizli sınıf içinde eşik parametrelerinin toplamı sıfırdır. Ayrıca π_c ile gösterilen her bir sınıfa atanma olasılıklarının toplamı bire eşittir.

$$\sum_c \pi_c = 1 \quad (1.23)$$

Çok sonuçlu KRM için parametre kestirimi iki sonuçlu KRM için tanımlanan E-M algoritması kullanılarak yapılır.

1.2.3. Gizli Sınıf Sayısının Belirlenmesi

İki sonuçlu ve çok sonuçlu KRM’de modelin öngördüğü gizli sınıf sayısını belirlemeye yönelik bilgi ölçütlerinden yararlanılır. Bilgi ölçütleri “-2log olabirlik” değerine dayanmaktadır. Model parametre sayısı ve örneklem büyüklüğü gibi faktörlerin eşit olması durumunda log olabirlikleri eşit olan iki modelden parametre sayısı en az olan model, en iyi model olarak ifade edilir. Literatürde en sık kullanılan bilgi ölçütleri Akaike ve Bayes Bilgi Ölçütleridir.

Akaike Bilgi Ölçütü (ABÖ) bilgi kuramından çıkarılmış kavramlara dayanan, model uyumunda kullanılan bir ölçüttür. ABÖ, aynı veriye uydurulmuş iki veya daha fazla modelin karşılaştırılmasında, olabirliğinde en küçük azalma beklenen modeli destekler. ABÖ için eşitlik aşağıda verilmiştir:

$$\log_e^{(L_h)} = \sum_{s=1}^{2^y} n_s \log_e^{[P_h(y_s)]} \text{ olmak üzere,}$$

$$ABÖ_h = -2\log_e^{(L_h)} + 2m_h \quad (1.24)$$

Burada m_h , h.modelin veriye uydurulmasında kestirilen bağımsız parametre sayısını ifade etmektedir. $P_h(y_s)$ ise h. modelin en çok olabirlik kestirimine dayanan yanıt vektörünün olasılık değeridir. n_s ise örneklem büyüklüğüdür.

Literatürde ABÖ doğrudan örneklem büyüklüğünü içermediği için asimptotik tutarlılık açısından eleştirilmektedir. Bu dezavantaja alternatif olarak Schwarz (1978) asimptotik tutarlı bir ölçüt olarak Bayes Bilgi Ölçütü’nü (BBÖ) geliştirmiştir. BBÖ için eşitlik aşağıda verilmiştir:

$$BB\ddot{O}_h = -2\log_e^{L_h} + \log_e^N \times m_h \quad (1.25)$$

BBÖ'de model seçimi, ABÖ'de olduğu gibidir. Fakat BBÖ, ABÖ'ye göre daha az karmaşık, daha az parametre içeren modelleri seçmeye eğilimlidir. Preinerstorfer ve Formann (2011) tarafından yapılan çalışmada KRM için BBÖ'nün daha iyi sonuç verdiği benzetim çalışması ile kanıtlanmıştır. Dolayısıyla bu çalışmada da KRM ile gizli sınıfları belirlerken BBÖ'ü kullanılmıştır.

Bu çalışmanın amacı, çok sonuçlu maddeler içeren ölçeklerde saptanan madde işlev farklılığının incelenmesinde kullanılan klasik yöntemle Karma Rasch Modeli yönteminin karşılaştırılmasıdır. Bu amaca yönelik farklı koşullarda ve küçük örneklem büyüklüğü altında bir benzetim çalışması yapılmıştır.

2. GEREÇ YÖNTEM

2.1. Kullanılan Bilgisayar Programları Hakkında Genel Bilgi

KRM 1990 yılında Rost tarafından geliştirildikten sonra, birçok çalışmada MIF'i tanımlamak için kullanılmıştır. KRM'ye yönelik olarak geliştirilen bilgisayar programlarının hepsi 2000'li yıllarda geliştirilmiştir. Bu programlardan en sık kullanılanı ise 2001 yılında Matthias von Davier tarafından yazılan WINMIRA adlı paket programdır (von Davier,2001). Bu program dışında 2011 yılında Preinerstorfer tarafından R Kütüphanesi olarak geliştirilen *mRm* de kullanılmaktadır. Ancak bu kütüphane sadece iki sonuçlu maddeler için parametre kestirimlerini sağlamaktadır. Ayrıca 2011 yılında Frick tarafından KRM parametre kestirimlerini sağlayan *psychomix* adlı R Kütüphanesi oluşturulmuştur. Fakat bu kütüphane ile θ kestirim sonuçlarını kullanıcıya sunan fonksiyon tanımlanamadığından kullanımının çok etkin olmadığı ifade edilmektedir.

Bu çalışma kapsamında veri türetilmesinde ve KKM sonuçlarının elde edilmesinde R 2.15.0 ve RUMMss programı, KRM sonuçları için de WINMIRA paket programı kullanılmıştır. R, istatistiksel hesaplamalar yapmak ve grafikler çizmek için yazılım ortamı sağlayan; UNIX, Windows ve MacOS ortamlarında çalıştırılabilen açık kaynak kodlu bir programdır (R Development Core Team, 2011). İstatistik alanında yapılan en son gelişmeler, genellikle R kütüphanesi olarak uygulamaya sunulmaktadır. Ayrıca geliştirilen hazır kütüphane ve fonksiyonların yanı sıra, özel olarak kodlama yapabilmek için de uygundur. R programı son zamanlarda uluslararası akademik platformda yaygın olarak kullanılan bir dil olmuştur. WINMIRA ise iki sonuçlu ve çok sonuçlu kategorik veriler için farklı kesikli KDM'leri analiz etmeye yönelik yazılmıştır. Bu program iki ve çok sonuçlu maddeler için RM, GSA, KRM ve Hibrid Model içerir. Literatürde kullanıcıya KRM'de çok sonuçlu maddeler için tüm parametre kestirimlerini sağlayan tek yazılımdır. RUMMss (Marais ve Andrich, 2007) programı ile madde eşik değerleri türetilmiştir. Bu program RM için veri setleri türetmeyi sağlar.

R ve WINMIRA θ kestirimleri için koşullu en çok olabilirlik kestirimi kullanmaktadır. Dolayısıyla bu programların kullanılması, KKM ve KRM kestirimlerinin karşılaştırılmasında herhangi bir olumsuzluk yaratmamıştır.

2.2. Benzetim Çalışması

2.2.1. Benzetim koşullarının belirlenmesi

Benzetim çalışması için, literatürde yapılan bazı karma model çalışmaları (Samuelsen, 2005; De Ayala et al., 2002 ;Cho, 2007) dikkate alınarak benzetim koşulları oluşturulmuştur. Bu çalışmalarda ölçme aracındaki madde sayısı, örneklem büyüklüğü, MİF gösteren madde yüzdesi ve MİF büyüklüğünün benzetim faktörleri olarak kullanıldığı belirlenmiştir. Bu faktörlerin benzetim sonuçlarına ilişkin özet bilgi aşağıda verilmiştir.

Ölçme Aracında Yer Alan Madde Sayısı ve Maddelerin Özellikleri:

Samuelsen (2005) ve De Ayala et al. (2002) 20 madde ile benzetim çalışması yapmışlardır. Bazı çalışmalar 15 ve 40 madde olarak karma dağılım modelleri üzerinde çalışmışlardır (Cho, 2007). Bu çalışmada ise Samuelsen (2005) ve De Ayala et al. (2002) çalışmaları ile karşılaştırılabilir olması açısından, 20 madde ile benzetim çalışması tasarlanmıştır. Ayrıca bahsedilen çalışmaların tümü iki sonuçlu maddeler için tasarlanmıştır. Sağlık alanında kullanılan ölçeklerin çoğu çok sonuçlu maddeler içerir. Dolayısıyla bu çalışmada beş yanıt kategorili maddeler için benzetim sonuçları elde edilmiştir.

Örneklem Büyüklüğü:

Cho (2007), karma modeller için minimum örneklem büyüklüğünün 1000 olması gerektiğini ifade etmiştir. Samuelsen (2005), örneklem büyüklüğünü 500 ve 2000 olarak benzetim çalışması tasarlamıştır. Cho (2007) ise 8000 birey ile benzetim çalışması yapmıştır. Bu çalışmalardan elde edilen sonuçlar, eğitim alanı için genelleştirilmiştir. Sağlık alanında büyük örneklemlemler ile çalışmak her zaman mümkün olmadığı için bu çalışmada da 300 bireyin yanıt desenleri türetilmiştir. Literatürde küçük örneklemlemler için KKM ve KRM kestirimlerinin performansını değerlendiren çalışma yoktur. Bu çalışmada, gerçekçi olması için, KKM ve KRM'nin 300 birey içeren ve MİF sorunu içeren veri setleri üzerinde performansları sorgulanmıştır.

MİF Gösteren Madde Yüzdesi:

Cho (2007) çalışmasında toplam madde sayısının %10'u ve %30'unun MİF gösterdiği varsayımı altında çalışmasını tasarlamıştır. Samuelsen (2005) bu oranları; %10, %30 ve %50 olarak ele almıştır.

Sağlık alanındaki uygulamalarda maddelerin %30'undan fazlasının MİF gösterme olasılığı düşük olduğu için bu çalışmada sadece %10 ve %30 oranları dikkate alınarak çalışma yürütülmüştür.

MİF büyüklüğü: Eğer bir madde gözlenen bir grup açısından MİF gösteriyorsa, bu gözlenen grubun alt kategorileri için kestirilen madde zorlukları arasındaki mutlak fark MİF büyüklüğü olarak tanımlanmaktadır. Samuelson (2005) benzetim çalışmasında MİF büyüklüğünü 0.4, 0.8, ve 1.2 almıştır. Cho (2007) ise bu farkı 0,4; 0,6; 0,8; 1 ve 1,2 olarak alıp çalışma sonuçlarını değerlendirmiştir. Bu benzetim çalışmasında MİF büyüklüğü 1 ve 2 olarak ele alınmıştır.

Sonuç olarak, bu çalışmanın yukarıda bahsedilen çalışmalardan farkı MİF gösteren maddeler içeren veri setlerinde, maddelerin çok sonuçlu yanıt kategorisine sahip olması durumunda, küçük örneklem için KKM ve KRM'nin kestirim performanslarının değerlendirilmesidir. Çalışmada alınan benzetim faktörleri ölçme aracındaki madde sayısı (20), örneklem büyüklüğü (300), MİF gösteren madde yüzdesi (%10; %30) ve MİF büyüklüğü (1; 2) dür. Bu faktörlerin kombinasyonlarından elde edilen dört farklı koşul için veri türetilmiştir (Çizelge 2.1). Her bir koşul için 200 Monte Carlo (MC) tekrarı yapılmıştır.

Çizelge 2.1. Benzetim Koşulları

Ölçme Aracındaki Madde Sayısı	MİF gösteren madde yüzdesi	MİF büyüklüğü	Örneklem Büyüklüğü	MİF gösteren maddeler
20	10%	1	300	5;15
20	30%	1	300	2;5;8;12;15;18
20	10%	2	300	5;15
20	30%	2	300	2;5;8;12;15;18

2.2.2. Verilerin Türetilmesi ve Türetilen Verilerin KKM ve KRM ile Değerlendirilmesi

2.2.2.1. Veri türetmede kullanılan istatistiksel model

SSM ve RM, KKM'nin özel birer durumudur. Eşik değerleri arasındaki uzaklıklar her madde için aynı olduğunda KKM, SSM'ye; maddeler iki sonuçlu olduğunda ise RM'ye indirgenmektedir. Bu nedenle bu tez kapsamında veri türetmek için kullanılacak model, Master'ın (1982) KKM'si olarak seçilmiştir. KKM'ye göre, herhangi bir i . bireyin j . maddeye x ($x=0,1,\dots,m_j$) yanıtını verme olasılığı eşitlik (1.2)'de tanımlanmıştır. Rastgele veri türetme işlemi bu olasılık modeline göre gerçekleştirilmiştir.

2.2.2.2. Altın standart parametre değerlerinin türetilmesi

Bu çalışmanın ana amacı, MİF gösteren maddeler içeren veri setlerinde KKM ve KRM'nin performanslarının değerlendirilmesidir. Dolayısıyla MİF gösteren maddeler içeren veri setlerinin türetilmesi amaçlanmıştır. Bu amaca yönelik, MİF'e neden olan gözlenen grubun iki alt gruptan oluştuğu, bu alt grupların birinde ilk 150 bireyin, diğer alt grupta ise geri kalan 150 bireyin olduğu varsayılmıştır. Başka bir deyişle gözlenen grup %50 / %50 oranlarında iki alt gruba ayrılmıştır. MİF önceki bölümlerde, gözlenen grubun alt gruplarında maddelerin zorluklarının farklı olması olarak tanımlanmıştı. Dolayısıyla MİF'in tanımının sağlandığı veri seti türetmek için ilk gruba yönelik madde zorlukları türetilmiş, ikinci gruba yönelik de birinci grup için türetilen madde zorluklarına belirli büyüklükler eklenerek yeni madde zorlukları türetilmiştir. Sonuç olarak maddelerin ikinci gruptaki bireyler için daha zor olduğu söylenebilecektir.

Bu amaca yönelik benzetim çalışmasında öncelikle, altın standart olarak tanımlanan, birey parametreleri (θ) ve madde zorluk parametreleri (β) türetilmiştir. θ parametresi, birinci gruptaki 150 birey için θ düzeyleri, $[-4,+4]$ aralığında, eşit aralıklarla artma koşulu ile belirlenmiştir. Birinci birey incelenen özellik açısından en kötü, sonuncu birey de incelenen özellik açısından en iyi olacak şekilde ilk 150 birey için θ değerleri türetilmiştir. İkinci gruptaki 150 bireyin θ düzeyleri birinci gruptaki

bireylerle aynı alınmıştır (Çizelge 2.2.). 20 madde için madde zorlukları (β), [-4,+4] aralığında, eşit aralıklarla artma koşulu ile belirlenmiştir. Birinci soru en kolay, son soru en zor olacak şekilde 20 madde için, eşik değerleri ise RUMMss (Marais, 2003) paket programı kullanılarak oluşturulmuştur. Eşik zorlukları, eşik değerlerinin madde zorluğuna eklenmesi ile elde edilmiştir (Çizelge 2.3).

Çizelge 2.2. Veri türetmede kullanılan kişilere ait gerçek parametre değerleri

Kişi	1	2	...	149	150	151	152	...	299	300
0	-4	-3,95	...	3,95	4,00	-4	-3,95	...	3,95	4,00

Çizelge 2.3. Veri türetmede kullanılan maddelere ait gerçek zorluk parametre değerleri

Madde	1.eşik zorluğu	2.eşik zorluğu	3.eşik zorluğu	4.eşik zorluğu	Madde zorluğu
1	-5,378	-4,056	-3,559	-3,007	-4,0
2	-4,945	-3,880	-3,343	-2,149	-3,579
3	-5,010	-3,536	-3,497	-0,588	-3,158
4	-4,730	-3,280	-1,629	-1,309	-2,737
5	-3,129	-2,745	-2,272	-1,118	-2,316
6	-2,914	-2,209	-1,913	-0,543	-1,894
7	-3,505	-3,160	-0,877	1,646	-1,474
8	-1,676	-1,444	-0,995	-0,098	-1,053
9	-2,985	-0,510	-0,342	1,309	-0,632
10	-1,613	-1,076	-0,475	2,320	-0,211
11	-1,641	-0,128	0,577	2,036	0,211
12	-1,106	-0,065	1,241	2,459	0,632
13	-0,004	1,149	1,341	1,726	1,053
14	0,058	1,556	1,724	2,557	1,473
15	0,209	2,067	2,492	2,812	1,895
16	0,271	1,129	2,490	5,374	2,316
17	1,486	1,757	3,576	4,129	2,737
18	0,793	1,986	4,662	5,191	3,158
19	0,599	3,547	5,045	5,125	3,579
20	2,262	2,406	5,506	5,827	4,0

2.2.2.3. MİF gösteren maddelerin eşik zorluklarının hesaplanması ve her bir birey için yanıt deseninin türetilmesi

MİF'e neden olan değişkenin iki sınıflı olması öngörülmüştür. Daha açık bir ifadeyle örneklem büyüklüğü 300 iken ilk 150 bireyin birinci grubu, ikinci 150 bireyin ise ikinci grubu oluşturduğu ve bu gruplarda maddelerin işleyişinin farklı olduğu öngörülmüştür. Bu amaca yönelik olarak, örneklem büyüklüğünün yarısı için Çizelge 2.3.'de verilen eşik zorlukları alınarak, diğer yarısı için de bu eşik zorluklarına belirlenen MİF büyüklüğü eklenerek veri türetilmiştir. Veri türetilirken *datagen(.)* fonksiyonundan yararlanılmıştır (Erdoğan, 2012). Kullanılan datagen isimli fonksiyonun argümanları aşağıdaki gibidir:

n_g : her bir gruptaki birey sayısı

n_i : madde sayısı

t: eşik değeri sayısı

beta: eşik zorluklarını içeren [eşik sayısı×madde sayısı] boyutlu matris

theta: θ düzeylerini içeren kişi sayısı uzunluğunda vektör.

Örneğin Çizelge 2.1'de verilen benzetim koşullarından birincisi göz önüne alınsın. Bu benzetim koşulu için altın standart θ ve β değerleri kullanılarak 300 birey için 20 maddelik yanıt deseni türetilirken aşağıdaki aşamalar tamamlanmıştır.

Adım 1. datagen(.) fonksiyonu aracılığı ile birinci gruptaki 150 birey için Çizelge 2.3.'de elde edilen eşik zorlukları kullanılarak beş kategorili yanıt desenleri R 2.15.0 programı kullanılarak oluşturulmuştur.

Adım 2. 5 ve 15.madde için elde edilen madde zorluklarına 1 eklenmiştir. Çünkü benzetim koşulunda bu maddelerin işleyişinin ikinci sınıfta farklı olduğu öngörülmüştür. datagen(.) fonksiyonu aracılığı ile ikinci gruptaki 150 birey için bu yeni madde zorluklarıyla eşik değerleri toplanarak yeni eşik zorlukları elde edilmiştir. Bu yeni eşik zorlukları kullanılarak beş kategorili yanıt desenleri R 2.15.0 programı kullanılarak oluşturulmuştur.

Şimdiye kadar tanımlanan aşamalar tamamlanırken birinci ve ikinci gruptaki bireylerin θ düzeyleri benzer alınmıştır. Dolayısıyla benzer θ düzeyine sahip olan bu gruplarda 5. ve 15.maddenin zorluğu farklılaştırılarak, MİF tanımı da sağlanmıştır.

Adım 3. Son olarak, Adım 1 ve Adım 2’de edilen veri setleri birleştirilerek 300 birey içeren veri seti elde edilmiştir. Dolayısıyla 300 birey için minimum yanıt 0, maksimum yanıt ise eşik sayısı olacak şekilde [kişi sayısı×madde sayısı] boyutlu veri seti türetilmiştir.

2.2.2.4. Türetilen veri setlerinin KKM ile analizi

KKM için R programındaki *eRm* kütüphanesi (Mair ve Hatzinger, 2007) kullanılmıştır. *eRm* kütüphanesi, ikili RM, KKM, SSM ve MYT modellerinden doğrusal lojistik test modeli, doğrusal SSM ve doğrusal KKM için geliştirilmiştir. Madde ve birey parametreleri, bu kütüphane yardımıyla kısmi kredi modeli kullanılarak kestirilmiştir. *eRm* kütüphanesinde, madde parametrelerinin kestiriminde koşullu en çok olabilirlik kestirimi kullanılırken, kişi parametrelerinin kestiriminde ise klasik en çok olabilirlik kestirimi kullanılmaktadır. Koşullu en çok olabilirlik yaklaşımında, madde parametreleri kişilere ait θ düzeyi dağılımından bağımsız olarak kestirilir (Eşitlik 2.2). İkili Rasch modeli için koşullu olabilirlik fonksiyonu aşağıdaki gibidir:

$$L = \frac{\exp(-\sum_j \beta_j s_j)}{\prod_r \sum_{x_j} \exp(-\sum_j x_j \beta_j)} \quad (2.1)$$

Bu eşitlikte s_j j. maddeye ait toplam skoru, x_j j. maddeye ait yanıt desenini, r ise kişilere ait toplam skorları göstermektedir. Rasch modellerinde kişilere ait toplam skorlar (r) kişi θ düzeyi kestirimlerinin yeterli istatistikleridir. Bu nedenle, kişilere ait parametreler koşullu olabilirlik kestiriminde yer almaz. Madde parametrelerinin kestirimi, kişilere ait toplam skorlara koşullu olarak yapılır. Daha sonra bulunan madde parametre kestirimleri, bileşik olabilirlik fonksiyonunda kullanılarak (Eşitlik 2.2) kişi parametre kestirimleri yapılır. İkili RM için bileşik olabilirlik fonksiyonu aşağıdaki gibidir:

$$L = \frac{\exp(\sum_i \theta_i r_i) \exp(-\sum_j \beta_j s_j)}{\prod_i \prod_j (1 + \exp(\theta_i - \beta_j))} \quad (2.2)$$

Burada β_j 'ler yerine, koşullu en çok olabilirlik kestirimleri yazılır. Toplam skoru en düşük olanlar ya da ölçme aracından olabilecek en yüksek skoru alanlar için kestirimler yapılamaz.

2.2.2.5. Türetilen veri setlerinin KRM ile analizi

Karma Rasch Analizi için WINMIRA programı kullanılmıştır. Programa öncelikle R ile türetilen veri seti tanıtılmış, sonra bu veri setine bir gizli sınıflı model (RM) uydurulmuştur. Bir sınıflı modelden BBÖ elde edilmiştir. Daha sonra aynı veri setine iki gizli sınıflı model uydurulmuştur ve yine iki gizli sınıflı modelden BBÖ elde edilmiştir. BBÖ, (1.39) eşitliği ile bir önceki bölümde verilmiştir. Bir gizli sınıflı ve iki gizli sınıflı modelden elde edilen BBÖ değerleri karşılaştırılıp uygun sınıf sayısı belirlenmiştir. BBÖ değeri en küçük olan model veriye en uygun model olarak tanımlanır. Benzetim çalışmasında ele alınan her bir benzetim koşulundaki 200 tekrar için bir ve iki sınıflı modeller oluşturulmuştur. Bu noktada KRM sonuçlarından beklenen, MİF'e neden olan grubun iki düzeyi olduğu için KRM'nin de iki gizli sınıflı yapıyı ortaya koyabilmesidir.

2.3. Benzetim Çalışmasının Sonuçlarının Değerlendirilmesi

KKM ve KRM kestirimleri kullanılarak yapılan MİF değerlendirmesi, iki farklı alt başlık altında ele alınmıştır.

2.3.1. Türetilen veri setlerinde KKM'nin MİF gösteren maddeleri doğru ve yanlış saptama oranlarının ve KRM'nin iki gizli sınıf saptama oranının değerlendirilmesi

Bu alt başlık ile türetilen veri setlerinde MİF gösterdiği bilinen maddelerin KKM ile doğru saptanma oranlarının belirlenmesi ve KRM ile de bu veri setlerinde iki gizli sınıflı yapının saptanıp saptanmadığının sorgulanması amaçlanmıştır. Böylece her iki

modelin de MİF'e neden olan etkenin saptanması üzerine performansları değerlendirilmiştir.

Bu amaca yönelik olarak, KKM'den elde edilen standartlaştırılmış artıklar bağımlı değişken, türetilen iki düzeyli grup faktörü (ilk 150 birey birinci gruba, ikinci 150 birey ikinci gruba oluşturmak üzere) ve altı sınıf aralığına bölünmüş θ düzeyi de bağımsız değişkenler olarak alınarak, iki yönlü varyans analizi uygulanmıştır. θ düzeyini sınıf aralıklarına bölmek için *classInt* kütüphanesi kullanılmıştır. MİF analizi için *rasch(.)* fonksiyonu kullanılmıştır (Erdoğan, 2012). Bu fonksiyonun argümanları aşağıdaki gibidir:

data: [kişi sayısı×(madde sayısı+grup faktörü)] boyutlu veri matrisi.

t: maddelere ilişkin eşik değeri sayısı

ns: her bir gruptaki kişi sayısı

ni: madde sayısı

ci: sınıf aralığı sayısı, bu çalışmada 6 olarak seçilmiştir.

Varyans analizi sonucunda, etkileşim terimi anlamlı bulunursa tekdüze olmayan MİF olduğu, yalnızca grup faktörü anlamlı bulunursa tekdüze MİF olduğu, etkileşim ve grup faktörü anlamlı bulunmazsa MİF olmadığı sonucuna varılmıştır. MİF gösteren maddeleri doğru saptama oranları değerlendirilirken, tekdüze olan veya olmayan MİF ayrımına bakmaksızın sadece maddelerde MİF saptanıp saptanmadığı üzerine odaklanılmıştır. Her bir benzetim koşulunda 20 madde için 200 MC tekrarı yapıldığından dolayı 4000 maddede KKM ile MİF saptanıp saptanmadığı sorgulanmıştır. MİF gösteren madde yüzdesine (%10, %30) göre önceden belirlenen MİF gösteren ve göstermeyen maddeler altın standart olarak ele alınmıştır. KKM ile elde edilen sonuçlar altın standart değerler ile çapraz tablo kullanılarak karşılaştırılmıştır. Aşağıdaki tanımlar kullanılarak KKM'den elde edilen sonuçlar yorumlanmıştır.

1. Doğru pozitif oran (DPO): Altın standart olarak MİF gösteren maddelerde KKM ile de MİF saptanan madde oranı

2. Yanlış negatif oran (YNO): 1- DPO

3.Dođru negatif oran (DNO): Altın standart olarak MİF göstermeyen maddelerde KKM ile de MİF saptanmayan oranı

4. Yanlıř pozitif oran (YPO): 1-DNO

5.Dođru sonuç oranı (DSO): Altın standart olarak MİF gösteren maddelerde KKM ile de MİF saptanan maddelerin ve altın standart olarak MİF göstermeyen maddelerde KKM ile de MİF saptanmayan maddelerin sayısının toplamının toplam madde sayısına oranı

Benzetim alıřması iin nsel olarak MİF'e neden olan deđiřkenin iki sınıflı olması ngrlmřtr. Dolayısıyla maddelerin iřleyiřini deđiřtiren faktr iki sınıflıdır ve bu iki sınıf belli bir zellik bakımından birbirinden farklı bireylerden oluřtuđu iin veri setinin homojenliđini bozmaktadır. Benzetim alıřması sonucunda da KRM ile bu iki gizli sınıfın saptanması beklenmektedir. Her bir benzetim kořulu iin tretilen veri setleri KRM ile deđerlendirilip BB ile iki gizli sınıfın saptanıp saptanmadıđı sorgulanmıřtır. KRM, KKM gibi her bir maddenin gzlenen bir grup aısından MİF gsterip gstermediđi hakkında bilgi vermez. KRM ile elde edilen veri setinin kendi aralarında heterojen ka tane gizli sınıf ierdiđi sorgulanır. Oluřturulan bu gizli sınıflarda maddelerin farklı iřlediđi dolayısıyla bu gizli sınıflarda madde zorluklarının sırasının farklı olduđu varsayılır. Eđer kendi aralarında heterojen olan bu gizli sınıflar iin her bir gruba iliřkin madde zorlukları elde ediliyorsa MİF sorunu da zlmř demektir. Sonu olarak birinci stratejide KKM ile elde edilen dođru MİF saptanma oranları ile KRM ile elde edilen sınıf sayıları deđerlendirilmiřtir.

2.3.2. KKM'de MİF analizinden ve iki gizli sınıflı KRM'den elde edilen θ kestirimlerinin ve standart hatalarının deđerlendirilmesi

Bu alt bařlıđın amacı ise KKM ve KRM'nin θ kestirimlerini ve θ kestirimlerinin standart hatalarını karřılařtırmaktır. Uygulamada KKM ile bazı maddelerde grup faktr aısından MİF saptandıđı zaman grubun her bir kategorisi iin ilgili maddelerde ayrı ayrı madde parametre kestirimleri yapılmaktadır. Bu iřleme "**KKM'de MİF analizi**" denilir. Bu nedenle, alıřmamızda da MİF gsteren maddeler iin iki dzeyli grup faktr aısından madde parametre kestirimleri ayrı

ayrı elde edilerek bu kestirimler aracılığı ile θ kestirimleri ve standart hataları elde edilecek ve iki gizli sınıflı KRM'den elde edilenlerle karşılaştırılacaktır.

2.3.2.1. θ kestirimlerinin karşılaştırılması

KKM ile MİF analizi sonucunda ve KRM ile iki gizli sınıf olması durumunda elde edilen θ değerlerinin 200 MC tekrarı üzerinden ortalaması alınmıştır. Her bir benzetim koşulu için, altın standart olarak belirlenen θ değerleri ile iki modelden elde edilen θ kestirimleri arasındaki saçılım grafikleri çizilmiştir. Saçılım grafiklerine 45 derecelik eğime sahip bir doğru eklenerek, saçılım grafiğindeki θ ' ların bu doğrudan ne kadar saptığı gözlenmiştir.

Ayrıca, KKM ve KRM'den elde edilen θ kestirimlerinin karşılaştırılması için yanlılık ve hata kareler ortalaması ölçütlerinden yararlanılmıştır. Geliştirilmiş performans ölçüm değerlerinin bazıları doğruluğu (accuracy), bazıları etkinliği (efficiency) ve bazıları da hem doğruluğu hem de etkinliği bir arada bulundurur. Doğruluk, yanlılık miktarını; etkinlik ise 1/varyasyon miktarını yansıtmaktadır. Yansız bir kestirimin etkinliği çok düşük olabilir ya da etkin bir kestirim doğru değerden çok uzakta olabilir. Bu nedenle kullanılacak olan yanlılık bir doğruluk ölçütü iken, hata kareleri ortalaması (HKO) hem doğruluk hem de etkinliği bir arada bulandıran karma bir ölçüttür (Çizelge 2.4). Bu çalışma kapsamında bu ölçütler, altın standart parametre değerleri ile KKM ve KRM parametre kestirimlerinin kıyaslanmasından elde edilmiştir.

Çizelge2.4. Performans ölçütleri

Yanlılık=	$\hat{\theta} - \theta$
Hata Kareleri Ortalaması (HKO)=	$(\hat{\theta} - \theta)^2 + \text{Var}(\hat{\theta})$
θ	Altın standart parametre değeri
$\hat{\theta} = \frac{1}{200} \sum_{i=1}^N \hat{\theta}_i$	RM veya KRM'den elde edilen parametre değerinin 200 benzetim tekrarı üzerinden ortalama kestirimi
$\text{Var}(\hat{\theta}) = \frac{1}{200} \sum_{i=1}^{10} (\hat{\theta}_i - \hat{\theta})^2$	200 benzetim tekrarından bulunan deneysel hata

Her bir benzetim koşulu için KKM ve KRM'den elde edilen yanlılık ve HKO dağılımı kutu-çizgi grafiklerinden yararlanılarak değerlendirilmiştir. Ayrıca bu değerler için tanımlayıcı istatistikler verilmiştir. KKM ve KRM'den elde edilen θ değerlerinin uyumunu incelemek için sınıf içi korelasyon katsayısından (SKK) yararlanılmıştır.

2.3.2.2. θ kestirimlerinin standart hatalarının karşılaştırılması

Her bir benzetim koşulu için, KKM ile MİF analizi sonucunda ve KRM ile iki gizli sınıf olması durumunda elde edilen θ değerlerinin standart hatalarının 200 MC tekrarı üzerinden ortalaması alınmıştır. KKM ve KRM'den elde edilen θ standart hatalarının dağılımı kutu-çizgi grafiklerinden yararlanılarak değerlendirilmiştir. Ayrıca, standart hatalar için tanımlayıcı istatistikler verilmiştir.

3. BULGULAR

3.1. Benzetim Çalışması Bulguları

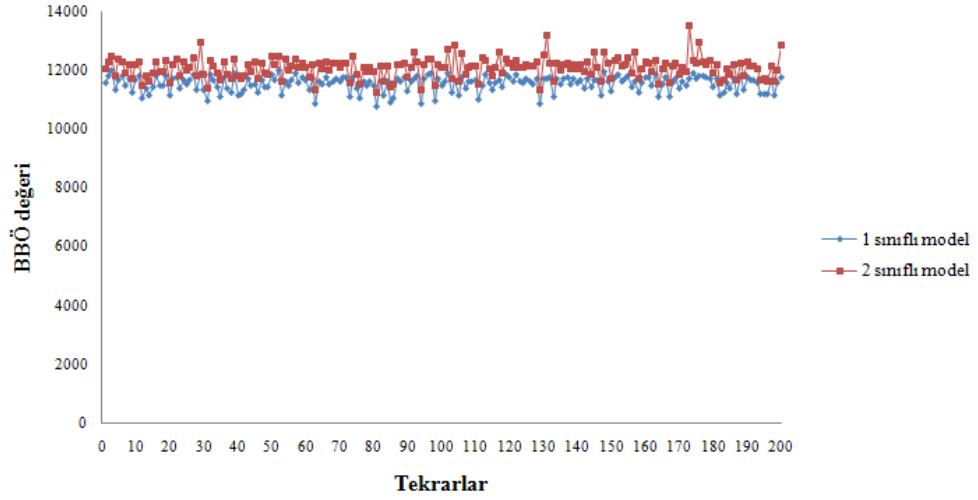
3.1.1. Türetilen veri setlerinde KKM'nin MİF gösteren maddeleri doğru saptama oranlarının ve KRM'nin iki gizli sınıf saptama oranının değerlendirilmesi

20 madde içerisinde MİF gösteren maddelerin KKM ile doğru ve yanlış saptanma oranları Çizelge 3.1'de verilmiştir. MİF büyüklüğünün 1 ya da 2 olmasından bağımsız olarak, MİF gösteren madde sayısı %10 iken en yüksek DPO, DNO ve DSO değerlerinin elde edildiği görülmektedir. En düşük DPO, DNO ve DSO değerleri ise MİF gösteren madde sayısı %30 olduğu zaman elde edilmiştir. Çizelge 3.1.'e bakıldığında, her iki MİF gösteren madde yüzdesi için de MİF büyüklüğünün 1'den 2'ye çıkmasının DPO oranını arttırdığı söylenebilir. Aynı koşullar altında, DNO oranları beklendiği gibi azalmıştır. Dolayısıyla MİF gösteren madde sayısından bağımsız, MİF büyüklüğü arttıkça KKM ile MİF saptanan bir maddenin yanlış pozitif olma olasılığının yüksek olduğu göz önünde bulundurulmalıdır. Benzer şekilde MİF gösteren madde sayısından bağımsız, MİF büyüklüğü azaldıkça KKM ile MİF saptanmayan bir maddenin yanlış negatif olma olasılığının yüksek olduğu göz ardı edilmemelidir.

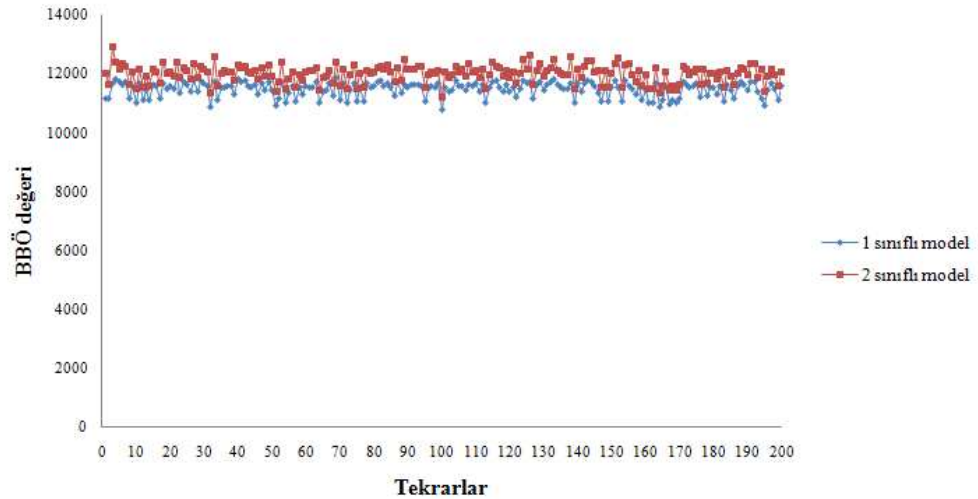
Çizelge 3.1. Her bir benzetim koşulundaki 200 MC tekrarda MİF gösteren maddelerin KKM ile doğru saptanma oranları

MİF gösteren madde yüzdesi	MİF büyüklüğü	Altın Standart		
		KKM	MİF gösteren	MİF göstermeyen
% 10 (2 madde)	1			
		MİF gösteren	353	13
		MİF göstermeyen	47	3587
		TOPLAM	400	3600
		DPO=0,882	DNO=0,996	DSO=0,985
MİF gösteren madde yüzdesi	MİF büyüklüğü	Altın Standart		
		KKM	MİF gösteren	MİF göstermeyen
% 30 (6 madde)	1			
		MİF gösteren	623	105
		MİF göstermeyen	577	2695
		TOPLAM	1200	2800
		DPO=0,519	DNO=0,962	DSO=0,829
MİF gösteren madde yüzdesi	MİF büyüklüğü	Altın Standart		
		KKM	MİF gösteren	MİF göstermeyen
% 10 (2 madde)	2			
		MİF gösteren	400	49
		MİF göstermeyen	0	3551
		TOPLAM	400	3600
		DPO=1	DNO=0,986	DSO=0,987
MİF gösteren madde yüzdesi	MİF büyüklüğü	Altın Standart		
		KKM	MİF gösteren	MİF göstermeyen
% 30 (6 madde)	2			
		MİF gösteren	1188	1121
		MİF göstermeyen	12	1679
		TOPLAM	1200	2800
		DPO=0,99	DNO=0,599	DSO=0,716

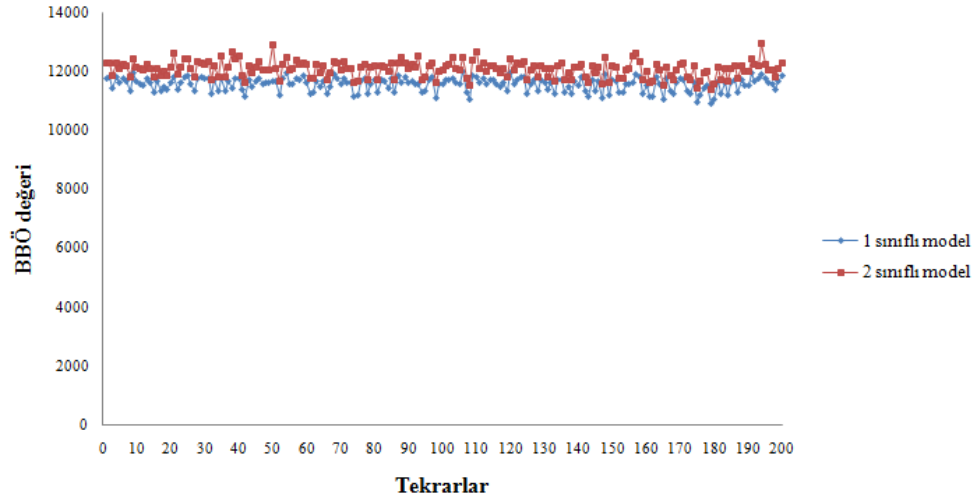
Türetilen veri setlerinde KRM ile saptanan gizli sınıf sayıları Şekil 3.1, Şekil 3.2, Şekil 3.3 ve Şekil 3.4’de verilmiştir. Her bir benzetim koşulu için KRM kestirimleri ile hesaplanan BBÖ’ler ile tek gizli sınıf öngörülmüştür. Türetilen veri setlerinde belirlenen benzetim koşulları altında, iki kategorili gözlenen grupta belirli maddelerin işleyişlerinin farklı olduğu bilindiği için KRM kestirimlerine yönelik elde edilen BBÖ değerlerinin bu grubu iki gizli sınıfla tanımlayabilmesi beklenmekteydi.



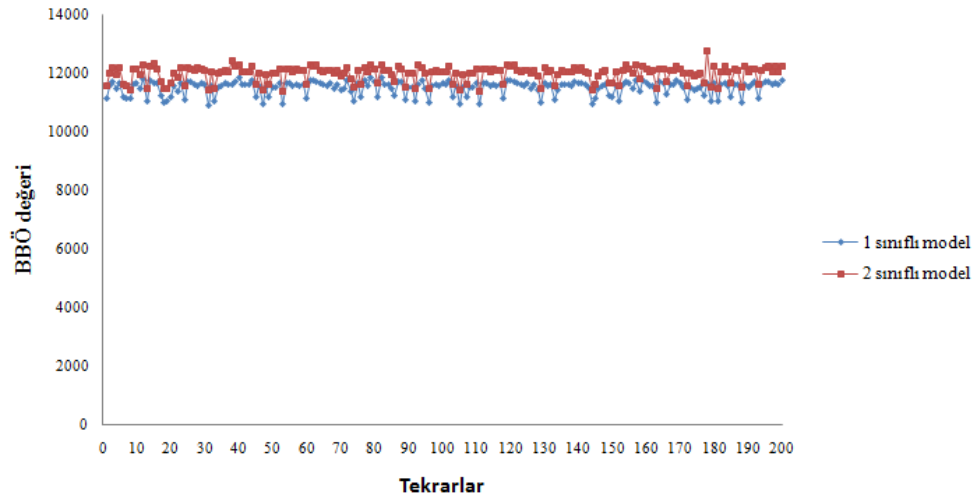
Şekil 3.1. MİF gösteren madde yüzdesi %10, MİF büyüklüğü 1 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri



Şekil 3.2. MİF gösteren madde yüzdesi %30, MİF büyüklüğü 1 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri



Şekil 3.3. MİF gösteren madde yüzdesi %10, MİF büyüklüğü 2 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri



Şekil 3.4. MİF gösteren madde yüzdesi %30, MİF büyüklüğü 2 olduğu durumda KRM ile 1 sınıflı model ve 2 sınıflı model için elde edilen BBÖ değerleri

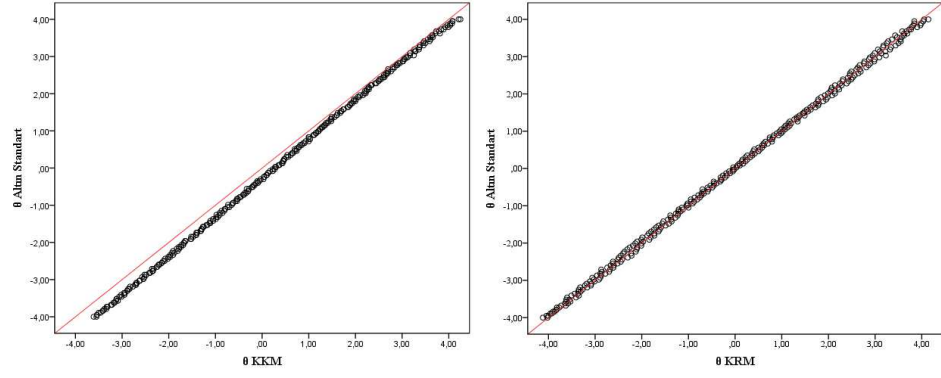
3.1.2. KKM’de MİF analizinden ve iki gizli sınıflı KRM’den elde edilen θ kestirimlerinin ve standart hatalarının değerlendirilmesi

3.1.2.1. θ kestirimlerinin karşılaştırılması

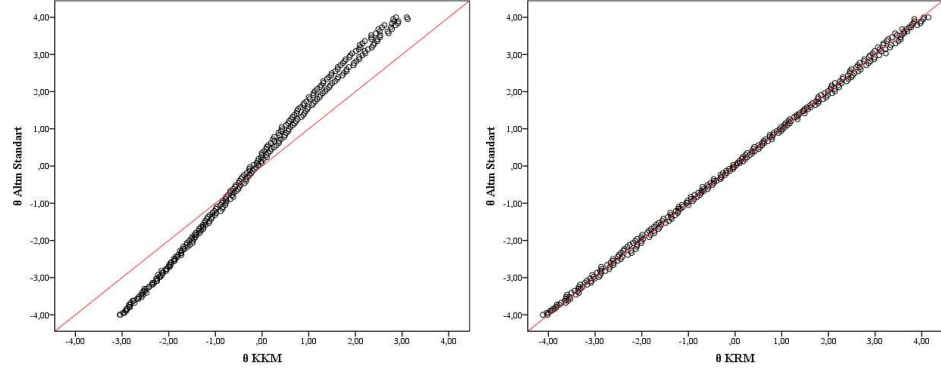
Bu bölümde, KKM’de MİF analizi ile elde edilen θ kestirimleri ve KRM ile iki gizli sınıf olması durumunda elde edilen θ kestirimleri karşılaştırılmıştır. Ayrıca, hesaplanan yanlışlık ve HKO değerlerinin, 300 kişi parametresi üzerinden gösterdiği dağılımlara ilişkin grafikler ve bu değerlere ilişkin tanımlayıcı istatistikler verilmiştir.

Şekil 3.5.’de MİF gösteren madde yüzdesi %10 ve %30 iken MİF büyüklüğünün 1 veya 2 olması durumunda KRM’ den elde edilen θ kestirimleri ile altın standart θ kestirimlerinin oldukça benzer olduğu söylenebilir. Ancak aynı MİF büyüklüğü altında, MİF gösteren madde yüzdesinin %30 olduğu durumda %10 olduğu duruma göre KKM ile elde edilen θ kestirimlerinin 45 derecelik doğrudan daha fazla saptığı görülmektedir. Ayrıca, aynı MİF gösteren madde yüzdesi altında, MİF büyüklüğünün 2 olduğu durumda 1 olduğu duruma göre, KKM ile elde edilen θ kestirimlerinin 45 derecelik doğrudan daha fazla saptığı ifade edilebilir. Sonuç olarak, tüm benzetim koşulları için KRM’ den elde edilen θ kestirimlerinin KKM’den elde edilenlere göre altın standart θ kestirimlerine daha benzer olduğu belirlenmiştir.

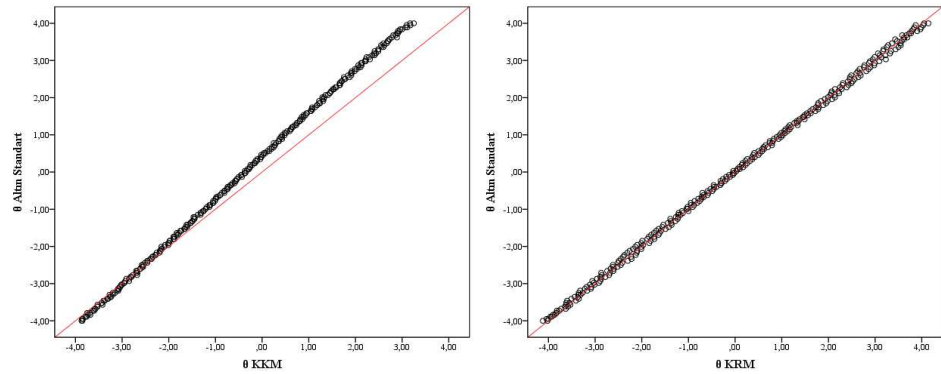
Çizelge 3.2. KKM ve KRM’den elde edilen θ değerleri ile altın standart arasındaki SKK değerleri incelendiğinde KRM’den elde edilen SKK değerlerinin KKM’ye göre yüksek olduğu söylenebilir.



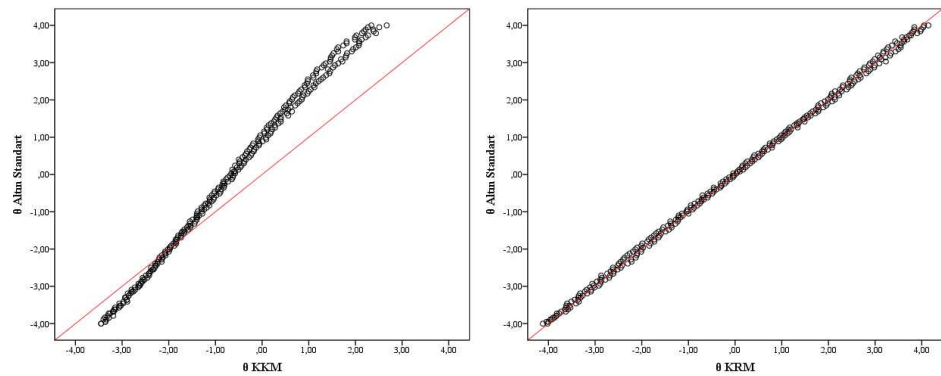
MİF gösteren madde yüzdesi %10; MİF büyüklüğü 1



MİF gösteren madde yüzdesi %30; MİF büyüklüğü 1



MİF gösteren madde yüzdesi %10; MİF büyüklüğü 2



MİF gösteren madde yüzdesi %30; MİF büyüklüğü 2

Şekil 3.5. Altın standart θ değerleri ile KKM ve KRM' den elde edilen θ değerleri arasındaki saçılım grafikleri

Çizelge 3.2.KKM ve KRM'den elde edilen θ değerleri ile Altın Standart arasındaki SKK değerleri ve %95 güven sınırları

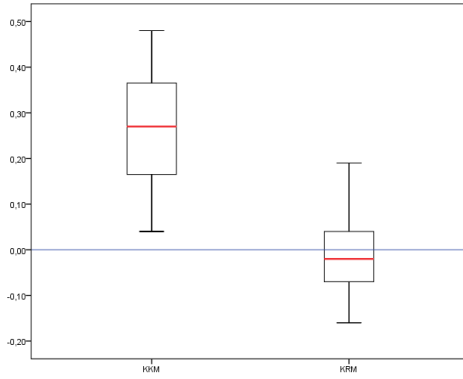
Benzetim Koşulları	KKM'den elde edilen θ değerleri ile Altın Standart arasındaki SKK (%95 güven sınırı)	KRM'den elde edilen θ değerleri ile Altın Standart arasındaki SKK (%95 güven sınırı)
MİF gösteren madde yüzdesi % 10 MİF büyüklüğü 1	0,992 (0,990-0,994)	1,0 (0,999-1,0)
MİF gösteren madde yüzdesi % 30 MİF büyüklüğü 1	0,944 (0,930-0,955)	1,0 (0,999-1,0)
MİF gösteren madde yüzdesi % 10 MİF büyüklüğü 2	0,975 (0,969-0,980)	1,0 (0,999-1,0)
MİF gösteren madde yüzdesi % 30 MİF büyüklüğü 2	0,896 (0,871-0,916)	1,0 (0,999-1,0)

Gerçek kişi parametre değerleri (θ) ile KKM ve KRM ile bulunan parametre kestirimleri ($\hat{\theta}$) arasındaki yanlılık değerlerinin, 300 kişi parametresi için dağılımlarına bakılacak olursa, bütün benzetim koşulları için, KRM sonucunda bulunan kestirimlerdeki yanlılığın KKM'ye göre sifira daha yakın olduğu görülmektedir (Şekil 3.6.). KKM ile elde edilen en büyük yanlılık değeri MİF gösteren madde yüzdesinin %30 ve MİF büyüklüğünün 2 olduğu durumdur. KRM ile elde edilen en büyük yanlılık değeri ise MİF gösteren madde yüzdesinin % 10 ve MİF büyüklüğünün 1 olduğu durumdur.

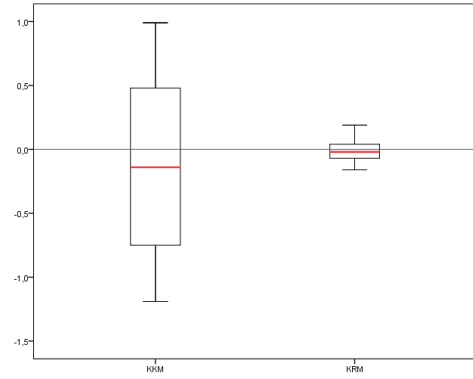
HKO bakımından θ kestirimleri değerlendirildiğinde bütün benzetim koşulları için, KRM sonucunda bulunan kestirimlerdeki HKO'ların KKM'ye göre daha düşük olduğu görülmektedir (Şekil 3.7.). KKM ile en büyük HKO değeri MİF gösteren madde yüzdesinin %30 ve MİF büyüklüğünün 2 olduğu durumda elde edilmiştir. KRM ile elde edilen en büyük HKO değeri ise MİF gösteren madde yüzdesinin % 10 ve MİF büyüklüğünün 1 olduğu durumdur.

Çizelge 3.3.KKM ve KRM'den elde edilen θ değerleri için yanlılık değerlerinin tanımlayıcı istatistikleri

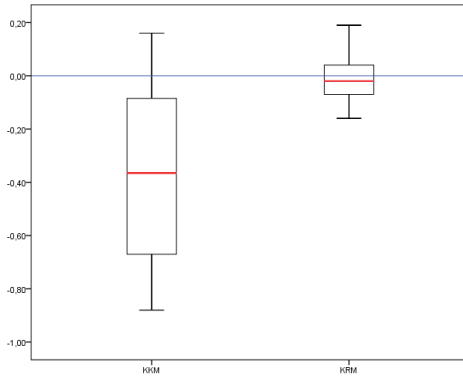
Benzetim Koşulları	Ortalama(standart-sapma)		Ortanca(min;maks)	
	KKM	KRM	KKM	KRM
MİF gösteren madde yüzdesi % 10				
MİF büyüklüğü 1	0,265 ±0,11	-0,021±0,07	0,27	-0,015
MİF gösteren madde yüzdesi % 30				
MİF büyüklüğü 1	-0,122±0,67	-0,016±0,05	-0,14	-0,018
MİF gösteren madde yüzdesi % 10				
MİF büyüklüğü 2	-0,37±0,31	-0,012±0,03	-0,37	-0,012
MİF gösteren madde yüzdesi % 30				
MİF büyüklüğü 2	-0,61±0,71	-0,013±0,032	-0,65	-0,015
			(-1,78;0,59)	(-0,14;0,21)



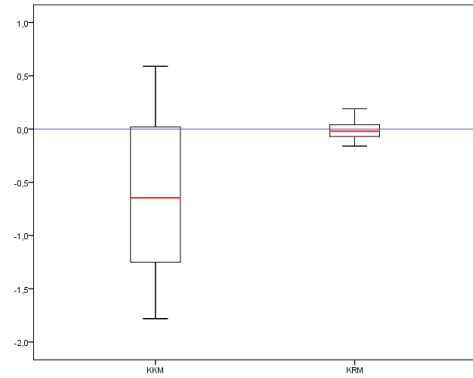
MİF gösteren madde yüzdesi % 10
MİF büyüklüğü 1



MİF gösteren madde yüzdesi % 30
MİF büyüklüğü 1



MİF gösteren madde yüzdesi % 10
MİF büyüklüğü 2

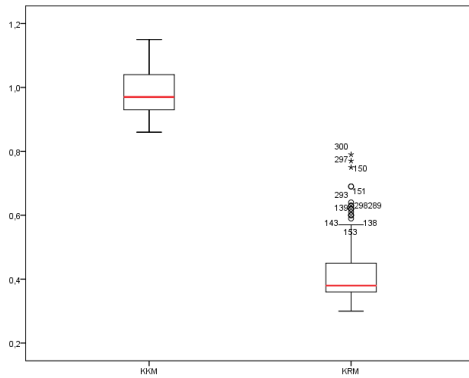


MİF gösteren madde yüzdesi % 30
MİF büyüklüğü 2

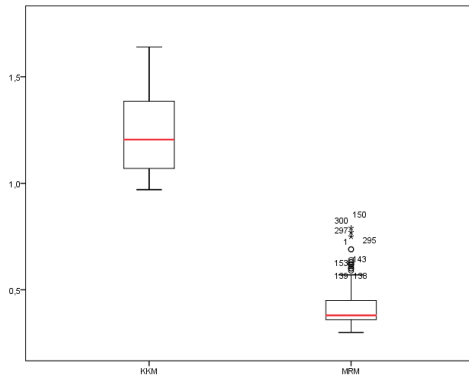
Şekil 3.6. KMM ve KRM' den elde edilen θ değerleri için yanlılık değerlerinin dağılımı

Çizelge 3.4. KKM ve KRM'den elde edilen θ değerleri için HKO değerlerinin tanımlayıcı istatistikleri

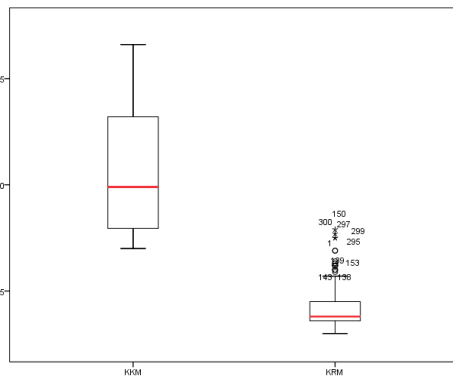
Benzetim Koşulları	Ortalama \pm standart sapma		Ortanca(min;maks)	
	KKM	KRM	KKM	KRM
MİF gösteren madde yüzdesi % 10	0,99 \pm 0,07	0,43 \pm 0,08	0,97	0,38
MİF büyüklüğü 1			(0,86-1,15)	(0,30-0,79)
MİF gösteren madde yüzdesi % 30	1,23 \pm 0,17	0,40 \pm 0,08	1,20	0,37
MİF büyüklüğü 1			(0,97;1,64)	(0,28;0,75)
MİF gösteren madde yüzdesi % 10	1,06 \pm 0,28	0,42 \pm 0,09	0,99	0,35
MİF büyüklüğü 2			(0,70;1,66)	(0,29;0,74)
MİF gösteren madde yüzdesi % 30	1,43 \pm 0,23	0,41 \pm 0,09	1,36	0,36
MİF büyüklüğü 2			(1,14;1,99)	(0,31;0,79)



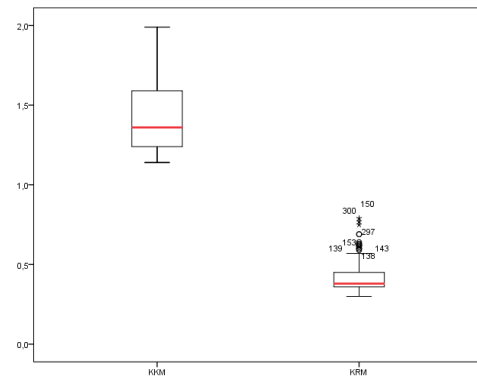
MİF gösteren madde yüzdesi % 10
MİF büyüklüğü 1



MİF gösteren madde yüzdesi % 30
MİF büyüklüğü 1



MİF gösteren madde yüzdesi % 10
MİF büyüklüğü 2



MİF gösteren madde yüzdesi % 30
MİF büyüklüğü 2

Şekil 3.7. KKM ve KRM'den elde edilen θ değerleri için HKO değerlerinin dağılımı

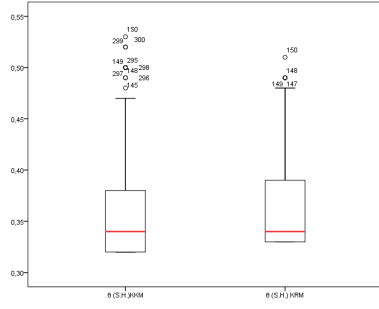
3.1.2.2. θ kestirimlerinin standart hatalarının karşılaştırılması

Bu bölümde, KKM'de MİF analizinden elde edilen θ kestirimlerinin standart hataları ve KRM'de iki gizli sınıf olması durumunda elde edilen θ kestirimlerinin standart hataları karşılaştırılmıştır.

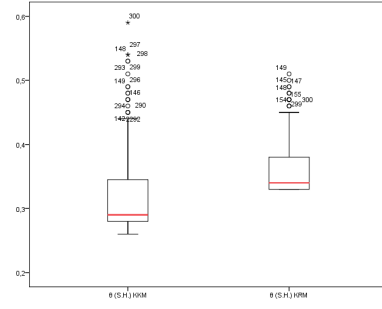
Şekil 3.8. ve Çizelge 3.4.'de görüldüğü gibi, MİF gösteren madde yüzdesinin %10 ve MİF büyüklüğünün 1 olduğu durumda KKM ve KRM ile elde edilen θ değerlerinin standart hatalarının ortancası eşit bulunmuştur. Bu benzetim koşulu dışındaki diğer tüm koşullar için, Çizelge 3.5.'de standart hataların ortalamalarına ilişkin %95 güven aralıkları da incelendiğinde, KKM ve KRM'den elde edilen standart hata değerlerinin benzer olduğu söylenebilir. KKM ve KRM ile en büyük standart hata ortalama değeri MİF gösteren madde yüzdesinin %10 ve MİF büyüklüğünün 1 olduğu durumda elde edilmiştir.

Çizelge 3.5.'de KKM ve KRM'den elde edilen θ değerlerinin standart hata ortalamalarının %95 güven sınırı incelendiğinde elde edilen güven sınırlarının uzunluklarının oldukça benzer olduğu görülmektedir.

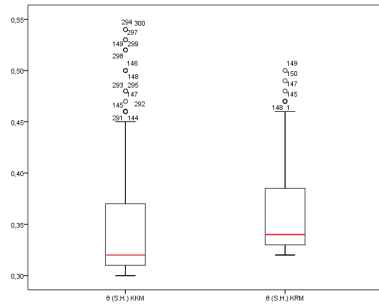
Çizelge 3.6.'de KKM ve KRM'den elde edilen θ değerlerinin standart hataları arasındaki SSK değerleri incelendiğinde, en yüksek SSK değerinin MİF gösteren madde yüzdesinin %10, MİF büyüklüğünün 1 olduğu durum için elde edildiği görülmektedir. En düşük SSK değeri ise MİF gösteren madde yüzdesinin %30, MİF büyüklüğünün 2 olduğu durum için elde edilmiştir. Dikkat çeken bir başka nokta ise MİF büyüklüğünden bağımsız, MİF gösteren madde yüzdesinin %30 olduğu durum için SSK değerinin oldukça düşük olmasıdır.



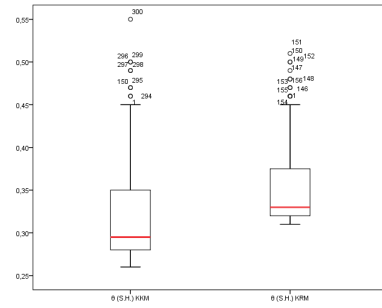
MİF gösteren madde yüzdesi %10
MİF büyüklüğü 1



MİF gösteren madde yüzdesi %30
MİF büyüklüğü 1



MİF gösteren madde yüzdesi %10
MİF büyüklüğü 2



MİF gösteren madde yüzdesi %30
MİF büyüklüğü 2

Şekil 3.8. KKM ve KRM' den elde edilen θ değerlerinin standart hatalarının dağılımları

Çizelge 3.5. KKM ve KRM' den elde edilen θ değerlerinin standart hatalarının tanımlayıcı istatistikleri

Benzetim Koşulları	Ortalama \pm standart sapma		Ortanca(min;maks)	
	<i>KKM</i>	<i>KRM</i>	<i>KKM</i>	<i>KRM</i>
MİF gösteren madde yüzdesi %10	0,36 \pm 0,05	0,37 \pm 0,04	0,34	0,34
MİF büyüklüğü 1			(0,32;0,53)	(0,33;0,51)
MİF gösteren madde yüzdesi %30	0,32 \pm 0,06	0,36 \pm 0,04	0,29	0,34
MİF büyüklüğü 1			(0,26;0,59)	(0,32;0,53)
MİF gösteren madde yüzdesi %10	0,35 \pm 0,05	0,36 \pm 0,04	0,32	0,35
MİF büyüklüğü 2			(0,30;0,54)	(0,32;0,50)
MİF gösteren madde yüzdesi %30	0,32 \pm 0,06	0,35 \pm 0,05	0,30	0,34
MİF büyüklüğü 2			(0,26;0,55)	(0,31;0,51)

Çizelge 3.6. KKM ve KRM'den elde edilen θ değerlerinin standart hatalarının ortalamalarına ilişkin %95 güven sınırları

Benzetim Koşulları	θ değerlerinin standart hatalarının ortalamalarına ilişkin %95 güven sınırı	
	<i>KKM</i>	<i>KRM</i>
MİF gösteren madde yüzdesi % 10	0,3522-0,363	0,3606-0,3702
MİF büyüklüğü 1		
MİF gösteren madde yüzdesi % 30	0,3143-0,3286	0,3586-0,3684
MİF büyüklüğü 1		
MİF gösteren madde yüzdesi % 10	0,3401-0,3524	0,3552-0,3646
MİF büyüklüğü 2		
MİF gösteren madde yüzdesi % 30	0,3123-0,3251	0,3447-0,3549
MİF büyüklüğü 2		

Çizelge 3.7. KKM ve KRM'den elde edilen θ değerlerinin standart hataları arasındaki SKK değerleri ve %95 güven aralıkları

Benzetim Koşulları	KKM ve KRM'den elde edilen θ değerlerinin standart hataları arasındaki SKK değeri (%95 güven sınırı)
MİF gösteren madde yüzdesi % 10	0,957 (0,946-0,965)
MİF büyüklüğü 1	
MİF gösteren madde yüzdesi % 30	0,625 (0,551-0,689)
MİF büyüklüğü 1	
MİF gösteren madde yüzdesi % 10	0,848 (0,813-0,877)
MİF büyüklüğü 2	
MİF gösteren madde yüzdesi % 30	0,620 (0,545-0,685)
MİF büyüklüğü 2	

4. TARTIŞMA

4.1. Sonuçların Özeti ve Literatürle Karşılaştırılması

KRM üzerine bugüne kadar yapılan tüm benzetim çalışmaları iki sonuçlu maddeler için tasarlanmışlardır. Bu çalışmaların çok büyük bir kısmı eğitim alanına yönelik tasarlandığı için, büyük örneklem ele alınarak sonuçlar değerlendirilmiştir. Literatürde küçük örneklem altında, çok sonuçlu maddeler içeren veri setlerinde KRM'nin performansını değerlendiren bir çalışma bulunmamaktadır.

De Ayala et.al (2002) yaptıkları çalışmada farklı MİF saptama yöntemleri kullanmışlar ve bu yöntemlerin MİF saptama açısından doğruluklarını değerlendirmişlerdir. Bu yöntemlerin içinde KRM de bulunmaktadır. KRM'nin önceden saptanan gözlenen gruplar üzerine odaklanmadan bireyleri kendi aralarında heterojen gruplara ayırmasının ve her bir sınıfta madde parametre kestirimlerini sağlayarak MİF sorununu çözmesinin, gözlenen grupları referans alan yöntemlere göre avantaj olduğunu ifade etmişlerdir. Bu çalışmada iki sonuçlu maddeler üzerinden yorum yapılmıştır.

Samuelsen (2005) yaptığı çalışmada, iki sonuçlu maddeler için, MİF gösteren madde sayısının ve MİF büyüklüğünün arttığı durumda MİF içeren maddelerin doğru tanımlanma oranlarını incelemiştir. Çalışmada özellikle MİF'e neden olan gözlenen grup üzerine odaklanılmış ve bu grubu oluşturan alt grupların büyüklükleri de göz önüne alınarak benzetim çalışması yapılmıştır. Aynı çalışmada sadece KRM ile belirlenen sınıf sayısı altında elde edilen θ kestirimlerinin dağılımları incelenmiştir. Ayrıca bu çalışmada küçük örneklem için KRM'den elde edilen sonuçların etkin olmadığı öngörülmüştür. Veri setindeki gözlenen grupların her zaman gizli sınıfların bir tanımlayıcısı olamayacağı üzerinde de durulmuştur. KRM ile elde edilen gizli sınıfları gözlenen gruplar ile tanımlamaya çalışmak yerine, doğrudan her bir gizli sınıftaki bireylerin yanıt desenlerini yorumlamanın daha etkin olacağı öngörülmüştür.

Cho (2007) yaptığı çalışmada iki sonuçlu maddeler için çok düzeyli Karma Madde Yanıt Modelinin MİF'i tanımlama üzerine etkisini incelemiştir. Ayrıca Karma Madde Yanıt Modelinden elde edilen gizli sınıfların gözlenen gruplar ile

tanımlanabilirliği sorgulanmıştır. Bu çalışmada da Samuelsen (2005) ile benzer sonuçlar elde edilmiştir.

Sağlık alanında büyük örneklemeler ile çalışmak her zaman mümkün olmamaktadır. Ayrıca sağlık alanında kullanılan ölçekler genellikle çok sonuçlu maddelerden oluşmaktadır. Bunun sebebi, çok sonuçlu maddelerin iki sonuçlu maddelere göre daha fazla bilgi sağlaması ve çok sonuçlu maddelerden elde edilen θ kestirimlerinin standart hatalarının iki sonuçlu maddelerden elde edilene göre daha düşük olmasıdır. Sağlık alanı için yorumlanabilecek sonuçlar elde etmeye yönelik bu çalışmada örneklem büyüklüğü 300 alınarak ve ölçme aracındaki maddeler çok sonuçlu olacak şekilde belirlenerek benzetim çalışması kurgusu yapılmıştır. Bu çalışmada veri setinin iki gizli sınıf içerdiği varsayımı altında KRM ile parametre kestirimleri elde edilmiştir.

KRM ile veri setinin tek sınıflı olduğu saptandığında bu veri setinin RM'ye uyum gösterdiği söylenir. Başka bir deyişle, çok sonuçlu maddeler için tek sınıflı KRM KKM'ye denktir. Tek sınıflı KRM (ya da KKM) veri setindeki bireylerin homojen olduğunu varsayar. KRM ile iki veya daha fazla gizli sınıf saptanırsa veri setinde kendi içinde homojen ve kendi aralarında heterojen bireylerden oluşan iki veya daha fazla sınıf olduğu ifade edilir. Türetilen veri setlerinin MİF gösteren maddeler içermesi, gözlenen bir grup açısından bireylerin maddelere farklı yanıt vermesi anlamına gelmektedir. Dolayısıyla bireyler MİF gösteren maddelere verdikleri yanıtlar açısından heterojen bir yapı oluştururlar. Benzetim çalışmasında türetilen veri setlerinin hepsi belirlenen yüzde ve büyüklüklerde MİF gösteren maddeler içermektedir. Dolayısıyla bu veri setlerinin tamamının bireyler açısından heterojen olduğu söylenebilir. Fakat bu heterojenliğin MİF gösteren madde yüzdesinin %10, MİF büyüklüğünün 1 olduğu durumda daha az, MİF gösteren madde yüzdesinin %30, MİF büyüklüğünün 2 olduğu durumda daha fazla olması beklenmektedir.

MİF gösteren maddelerin KKM ile doğru saptanma oranları incelendiğinde MİF büyüklüğünün 1 ya da 2 olmasından bağımsız olarak, MİF gösteren madde yüzdesi %10 iken en yüksek DPO, DNO ve DSO değerlerinin elde edildiği görülmektedir. En düşük DPO, DNO ve DSO değerleri ise MİF gösteren madde yüzdesi %30 olduğu zaman elde edilmiştir. Bu sonuç yukarıda tanımlanan homojen

ve heterojen veri setlerine yönelik tanım ile uyumludur. Tamamen heterojen olmayan gruplarda KKM ile daha iyi sonuçlar elde edilmiştir. Ayrıca MİF gösteren madde yüzdesinin KKM sonuçları üzerine etkisinin oldukça yüksek olduğu görülmektedir.

Samuelsen (2005) iki sonuçlu maddeler için küçük örneklem büyüklüğünde KRM'den elde edilen sonuçların etkin olmadığını öngörmüştür. Bu çalışmada çok sonuçlu maddeler ele alındığı ve çok sonuçlu maddeler üzerinde iki sonuçlu maddelere göre daha fazla parametre kestirimi elde edildiği için çok sonuçlu maddelerde daha büyük örneklem üzerinde çalışmak gerektiği açıktır. Her bir benzetim koşulu için türetilen veri setlerinde KRM ile tek gizli sınıf saptanmıştır. KRM ile beklenen iki gizli sınıflı yapının ortaya koyulamamasının nedeninin benzetim çalışmasının küçük örneklem altında tasarlanmış olduğu düşünülmektedir. Dolayısıyla çalışmamızdan elde edilen bu sonuç, Samuelsen'in (2005) çalışması ile uyumludur.

θ parametreleri açısından KKM ve KRM'den elde edilen yanlılık incelendiğinde KRM'den elde edilen yanlılık değerlerinin sıfıra çok yakın olduğu görülmektedir. Ayrıca iki modelden elde edilen HKO değerlerine bakıldığında da KRM'den elde edilen HKO değerlerinin KKM'ye göre daha düşük olduğu belirlenmiştir.. Dolayısıyla her bir benzetim koşulu için KRM'den elde edilen θ parametre kestirimlerinin KKM'ye göre daha iyi olduğu söylenebilir. KKM ve KRM θ parametrelerinin standart hataları açısından karşılaştırıldığında genel olarak KKM ve KRM'den edilen standart hataların benzer olduğu görülmektedir. Bu sonuç da heterojen gruplarda KRM ile daha iyi θ kestirimlerinin elde edildiğini göstermektedir.

θ değerleri ile altın standart arasındaki SKK değerleri incelendiğinde KRM'den elde edilen SKK değerlerinin KKM'ye göre yüksek olduğu görülmektedir. Eğer bir veri seti homojen ise, KKM'nin veriye uyum göstereceği bilinmektedir. MİF gösteren madde yüzdesinin ve MİF büyüklüğünün düşük olduğu durumda veri setinin homojene yakın olduğu söylenebilir. Dolayısıyla türetilen veri setlerinde MİF gösteren madde yüzdesinin %10, MİF büyüklüğünün 1 olduğu durumda KKM'den elde edilen SKK değerinin yüksek olması beklenen bir durumdur. KKM için en düşük SKK değeri ise MİF gösteren madde yüzdesinin %30, MİF büyüklüğünün 2 olduğu durum için elde edilmiştir. Bu koşullarda veri setinin heterojen olduğu

söylenbilir ve KKM ve KRM'den elde edilen θ deęerleri ile altın standart arasındaki SKK deęerleri incelendięinde KRM'den elde edilen SKK deęerlerinin daha yüksek olduęu görölmektedir. Bu sonuç, heterojen veri setlerinden KRM ile daha iyi parametre kestirimleri elde edildięinin göstergesidir.

KKM ve KRM'den elde edilen θ deęerlerinin standart hataları arasındaki SKK deęerleri incelendięinde en yüksek SKK deęerinin MİF gösteren madde yüzdesi %10, MİF büyüklüęü 1 olduęu durum için elde edildięi görölebilir. Bu koşullar altında verinin homojen olduęu düşünülürse elde edilen sonuç θ kestirimleri için elde edilen sonuç ile uyumludur. θ deęerlerinin standart hataları için dikkat çeken durum MİF büyüklüęünden bağımsız MİF gösteren madde yüzdesi %30 iken iki modelden elde edilen SKK deęerlerinin oldukça düşük olmasıdır. Bu sonuç, MİF gösteren madde yüzdesinin kestirimler üzerine etkisinin bir başka kanıtıdır.

Sonuç olarak, var olan MİF'i belirlemede küçük örneklem ve çok sonuçlu maddeler altında KKM ve KRM'nin başarılı olamadıęı söylenbilir. Çünkü KKM ile hiçbir benzetim koşulu için DSO %100'e ulaşmamıştır. Ayrıca KRM ile de var olan iki sınıf saptanamamıştır. Genel olarak θ kestirimlerinde KRM'nin KKM'den daha iyi sonuç verdięi görölmektedir. θ kestirimlerinin standart hatalarının da benzer olduęu görölmüştür. Dolayısıyla heterojen veri setlerinde KRM'den elde edilen parametre kestirimlerinin daha iyi olduęu söylenbilir.

4.2. Çalışmanın Kısıtlılıkları

Bu çalışma kapsamında yapılan benzetim çalışması, Intel(R) Core™ 2 CPU, Intel(R) Core™ I5 özelliklerindeki üç bilgisayarda paralel olarak gerçekleştirilmiştir. R'da verilerin türetilmesi, türetilen verilerden çalışma ile ilgili gerekli parametrelerin elde edilmesi yaklaşık 1,5 ay almıştır. WINMIRA programının iterasyonda yavaş olması bu süreyi uzatmıştır. Tüm veri türetme işlemlerinin R'da düzenlenmesi ile bu süreç hızlandırılarak her bir benzetim koşulu için tekrarlar artırılabilir. Fakat bu düzenleme daha ileri düzey bir yazılım bilgisi gerektirir. Bu çalışmada benzetim çalışması içerięinde iki faktör deęiştirilerek benzetim senaryoları oluşturulmuştur. Benzetim faktörleri artırılarak sonuçlar farklı senaryolar altında deęerlendirilebilir. MİF'e neden olan gözlenen grubun iki alt

gruptan oluřtuđu, bu alt grupların birinde ilk 150 bireyin, diđer alt grupta ise geri kalan 150 bireyin olduđu varsayılmıřtır. Bařka bir deyiřle gözlenen grup %50 / %50 oranlarında iki alt gruba ayrılmıřtır. Bu oranlar da benzetim alıřmasında deđiřen bir faktör olarak alınabilir. Ayrıca farklı örneklem büyüklükleri, madde sayıları ve eřik deđeri sayılarının da sonuçlar üzerindeki etkisi incelenebilir. Benzetim alıřmasında MİF gösterdiđi öngörülen tüm maddelerin zorluklarının bir grupta diđerine göre daha yüksek olduđu varsayılmıřtır. MİF gösterdiđi öngörülen bu maddelerin bazılarının bir grupta, geri kalanların diđer grupta daha zor oldukları varsayımı altında benzetim alıřması tekrarlanabilir.

5. SONUÇ ve ÖNERİLER

Rasch analizinin temel varsayımlarından biri ölçeği oluşturan maddelerin tümünün tek bir özelliği ölçmesi gerekliliğidir. Yani ölçek tek boyutlu olmalıdır. Eğer tek boyutluluk sağlanmıyorsa doğru ölçüm yapılamaz. Bu varsayımın bozulmasına neden olan faktörlerden biri de maddelerin ölçeğin uygulandığı bireylerin oluşturduğu alt gruplarda işleyişinin farklı olmasıdır, yani maddelerin MİF göstermesidir. MİF göstermeyen maddelerden oluşan bir ölçeğin geçerliğinin yüksek olması beklenir.

Bu çalışmada farklı yüzdelerde ve büyüklüklerde MİF gösteren maddeler içeren veri setleri türetilmiş ve bu veri setlerinde KKM ve KRM'nin parametre kestirimleri üzerine performansları karşılaştırılmıştır. Türetilen veri setlerinde MİF'e neden olan gözlenen grup, iki alt gruptan oluşan bir değişken olarak tanımlanmıştır. Dolayısıyla türetilen veri setlerinin farklı özellikteki iki gruptan oluşmaları ve heterojen oldukları söylenebilir. MİF gösteren madde yüzdesinin ve büyüklüğünün değişmesine bağlı olarak KKM'den elde edilen DSO değerlerinin değişmesi ve KRM ile tek gizli sınıf elde edilmesi nedeniyle, küçük örneklem altında iki modelin de MİF tanımlamada başarılı olamadıkları söylenebilir. Tüm benzetim koşulları için KRM'den elde edilen θ kestirimleri KKM'den elde edilenlere göre altın standart değerlere daha yakın çıkmıştır. İki modelden elde edilen θ kestirimlerinin standart hataları da benzerdir. Sonuç olarak heterojen bireylerden oluşan veri setlerinde KRM'nin KKM'ye göre daha iyi parametre kestirimleri sağladığı söylenebilir. KRM ile MİF'e neden olan gözlenen grup hakkında önsel bilgiye gerek yoktur. Dolayısıyla klinik araştırmalar tasarlanırken MİF kaynağı olabilecek önsel gözlenen faktörler tanımlanamıyorsa KRM ile veri setinin iki veya daha fazla gizli sınıf içerip içermediği sorgulanabilir. Sonuç olarak yeterli örneklem büyüklüğü altında, homojen veri setlerinde KKM'den elde edilen sonuçlara güvenilebilir. Fakat verinin homojenliği hakkında net bir bilgi yoksa veri KRM ile tekrar değerlendirilip homojenliği bozan gizli sınıflar tanımlanmalıdır.

ÖZET

Ölçeklerde Saptanan Madde İşlev Farklılığının Karma Rasch Analizi ile İncelenmesi

Bu çalışmanın amacı, madde işlev farklılığı (MİF) gösteren maddeler içeren veri setlerinde Karma Rasch modelinin (KRM) MİF'e neden olan grubu belirlemedeki ve kişi parametre kestirimleri üzerindeki başarısını değerlendirmektir. Literatürde küçük örneklem büyüklüğüne sahip ve çok sonuçlu maddeler içeren veri setlerinde KRM'nin performansını değerlendiren çalışma yoktur. Bu çalışmanın sonuçları küçük örneklem büyüklüğü ve çok sonuçlu maddeler için genellenmiştir.

Çalışmanın amacına yönelik, altın standart olarak MİF gösteren maddeler ve kişi parametreleri tanımlanmıştır. Yapılan benzetim çalışmasında altın standart değerler kullanılarak, farklı yüzde ve büyüklüklerde MİF gösteren maddeler içeren veri setleri türetilmiştir. Türetilen bu veri setleri sayesinde maddelere verilen yanıtlar açısından bireyler arasında farklılık yaratıldığı için, veri setlerinin heterojen olduğu söylenebilir. Türetilen veri setlerinde kısmi kredi modelinin (KKM) doğru saptadığı MİF gösteren madde yüzdeleri değerlendirilmiş, KRM ile gizli sınıf sayıları saptanmıştır. Ayrıca her iki modelden elde edilen kişi parametre kestirimleri altın standart değerler ile karşılaştırılmıştır.

KKM ile MİF'i doğru saptama oranı sadece MİF gösteren madde yüzdesinin %10 ve MİF büyüklüğünün 1 olduğu durumda yüksek elde edilmiştir. KRM ile tüm veri setlerinde tek sınıflı yapı saptanmıştır. Dolayısıyla iki modelin de var olan MİF'i belirlemede tam olarak başarılı olamadıkları söylenebilir. Tüm benzetim koşulları için KRM'den elde edilen kişi parametre değerlerinin altın standart parametre değerlerine oldukça yakın olduğu görülmüştür. Veri setinin heterojenliği arttığında, KKM'den elde edilen kişi parametre değerlerinin altın standart parametre değerlerinden saptığı gözlenmiştir. Parametre kestirimleri bakımından KRM ile elde edilen yanlılık ve hata kareleri ortalaması değerlerinin KKM'ye göre daha düşük olduğu saptanmıştır. İki modelden elde edilen kişi parametrelerinin standart hataları benzerdir.

Sonuç olarak, heterojen veri setlerinde KRM ile elde edilen kişi parametre kestirimlerinin altın standart değerlere oldukça yakın olduğu ve KKM'ye göre daha iyi kestirimler elde edildiği söylenebilir. Literatürde iki sonuçlu maddeler içeren ve küçük örnekleme sahip veri setlerinde KRM'nin etkin sonuç vermediği ifade edilmiştir. Bu çalışma ile çok sonuçlu maddeler içeren ve küçük örnekleme sahip veri setlerinde KRM'nin MİF'e neden olan etkeni tanımlamada başarılı olamadığı kanıtlanmıştır.

Anahtar Kelimeler: gizli sınıf, Karma Rasch modelleri, kısmi kredi modeli, küçük örneklem, madde işlev farklılığı

SUMMARY

Analyzing Differential Item Functioning by Mixed Rasch Models which Stated in Scales

The aim of this study was to investigate the effectiveness of Mixed Rasch Model (MRM) on detecting the group that cause differential item functioning (DIF) and on person estimates while the data set known to be contain DIF. There is no study at the literature about how the MRM performs while the sample is small and items are polytomous. In this study, the results were generalized for small sample size and polytomous items.

For this aim, gold standard items that include DIF and gold standard person parameters were defined. Using these gold standard values, the simulation study was performed. In this simulation study, items indicate different DIF percents and magnitudes. So that simulated data sets were heterogeneous with respect to person. True detection rates on gold standard DIF items were evaluated by Partial Credit Model (PCM). Number of latent classes was evaluated by MRM. Also person parameters got from both PCM and MRM were compared with gold standard values.

True detection rate is high only for DIF percent is %10 and DIF magnitude is 1 or 2. For all simulated data set, one latent class was found by MRM. So one can conclude that both models were not successful on detected DIF. Person parameters were the same with gold standard parameters for MRM. While the data set became heterogeneous, PCM parameter estimates deviated from gold standard values. Also bias and mean square error for MRM was lower than PCM. Standard errors of person parameters were same for MRM and PCM.

As a result, for heterogeneous data sets the MRM person estimates were similar with the gold standard person parameters and better estimates were obtained comparing with PCM. In the literature it is stated that MRM performs bad while the sample is small and items are dichotomous. In this study, it is verified that MRM performs bad to detect the group that cause DIF for data sets include small sample size and polytomous items.

Keywords: differential item functioning, latent class, Mixed Rasch Model, partial credit model, small sample

KAYNAKLAR

- ACKERMAN, T.A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, **29**(1), 67-91.
- BİLİR, M. K. (2009). Mixture Item Response Theory-Mimic Model: Simultaneous Estimation Of Differential Item Functioning For Manifest Groups And Latent Classes. Doctor of Philosophy. Department of Educational Psychology and Learning Systems.
- CAMILLI, G., CONGDON, P. (1999). Application of a method of estimating DIF for polytomous test items. *Journal of Educational and Behavioral Statistics*, **4**, 323-341.
- CAMILLI, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.). Westport, CT: Praeger Publishers.
- CHANG, H., MAZZEO, J., ROUSSOS, L. A. (1996). Detecting DIF for polytomously scored items: an adaptation of the SIBTEST procedure. *Journal of Educational Measurement*, **33**(3), 333-353.
- CHO, S.J. (2007). A Multilevel Mixture IRT Model for DIF Analysis. Doctor of Philosophy. Graduate Faculty of The University of Georgia.
- CLAUSER, B. E., MAZOR, K. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*, **17**(1), 31-44.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- COHEN, A.S., BOLT, D.M. (2002). A mixture model analysis of differential item functioning. Paper presented at the annual meeting of the American Educational Research Associations, New Orleans.
- COHEN, A.S, BOLT D.M. (2005). A Mixture Model Analysis of Differential Item Functioning. *Journal of Educational Measurement*, Vol. **42**, No. 2, pp. 133-148.
- ÇOLAK, M. (2003). Gizli Sınıf (Latent Class) Analizi ve Tanı Testlerinin Değerlendirmesi. Doktora tezi, Sağlık Bilimleri Enstitüsü, Hacettepe Üniversitesi.
- DeAYALA, R.J., KIM, SEOCK-HO, STAPLETON, L.M., DAYTON, C.M. (2002). Differential item functioning: A mixture distribution conceptualization. *International Journal of Testing*, **2**(3&4), 243-276.
- DORANS, N. J. (1989). Two new approaches to assessing differential item functioning: Standardization and the Mantel-Haenszel method. *Applied Measurement in Education*, **2**, 217-233.
- DORANS, N.J., HOLLAND, P.W. (1993). DIF Detection and Description: Mantel-Haenszel and Standardization. In P.W.
- ERDOĞAN, B. (2012). Çoklu Atama Yöntemlerinin Rasch Modelleri İçin Performansının Benzetim Çalışması İle İncelenmesi. Doktora tezi, Sağlık Bilimleri Enstitüsü, Ankara Üniversitesi.
- FISCHER HG, MOLENAAR (1995). *Rasch Models. Foundations, Recent Developments and Applications*. Springer; New York.
- FRENCH, A. W., MİLLER, T. R. (1996). Logistic regression and its use in detecting differential item functioning in polytomous items. *Journal of Educational Measurement*, **33**(3), 315-332.

- HILLS, J. (1990). Screening for potentially biased items in testing programs. *Educational Measurement: Issues and Practice*, **8**, 5-11.
- HU, P.G., DORANS, N.J. (1989). The Effects of Deleting Differentially Functioning Items on Equating Functions and Reported Score Distributions. Princeton, NJ: Educational Testing Service.
- HOLLAND, P.W., THAYER, D.T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H.I. Braun (Eds.), *Test Validity*. Hillsdale, N.J.: Erlbaum.
- KELDERMAN, H., MACREADY, G.B. (1990). The use of loglinear models for assessing differential item functioning across manifest and latent examinee groups. *Journal of Educational Measurement*, **27**, 307-327.
- LIU, I., AGRESTI, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics*, **52**, 1223-1234.
- LORD, F. M. (1977). A study of item bias using item characteristic curve theory. In Y. H. Poortings (Ed.), *Basic problems in cross-cultural psychology* (pp. 19-29). Amsterdam: Swets and Zeitlinger.
- LORD, F. M. (1980). *Applications Of Item Response Theory To Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- MARAIS I, ANDRICH D. RUMMSS. Rasch unidimensional measurement models simulation studies software. Perth: The University of Western Australia; 2007.
- MILLER, T. R., SPRAY, J. A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement*, **30**(2), 107-122.
- MILLSAP, R. E., EVERSON, H. T. (1993). Methodology review: statistical approaches for assessing measurement bias. *Applied Psychological Measurement*, **17**(4), 297-334.
- PENFIELD, R. D., CAMILLI, G. (2007). Differential item functioning and item bias. In C. R. Rao, & S. Sinharay (Eds.), *Psychometrics*. Amsterdam: Elsevier.
- RAJU, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, **53**, 495-502.
- RAJU, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, **14**, 197-207.
- R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, A Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, Erişim: [<http://www.R-project.org>].
- ROST, J. (1990). Rasch Models in Latent Classes: An Integration of Two Approaches to Item Analysis. *Applied Psychological Measurement*, **14**(3), 271-282.
- SAMUELSEN, K. (2005). Examining Differential Item Functioning From A Latent Class Perspective. Doctor of Philosophy. Faculty of the Graduate School of the University of Maryland, College Park.
- SHEALY, R., STOUT, W. (1993). A model-based 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.

- Standards for Educational and Psychological Testing (1999). Washington, DC: American Educational Research Association, American Psychological Association & National Council on Measurement in Education.
- SWAMINATHAN, H., ROGERS, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, **27**(4), 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. I. Braun (Eds.), *Test validity* (pp. 147-169). Hillsdale, NJ: Erlbaum.
- THISSEN, D., STEINBERG, L., WAINER, H. (1993). Detecting of differential item functioning using the parameters of item response models. In P. W.
- TURGUT, M.F. (1977). *Eğitimde Ölçme ve Değerlendirme Metodları*. Ankara: Nüve Matbaası.
- VON DAVIER, M. (2001). WINMIRA 2001: Software and user manual. Erişim:[<http://winmira.von-davier.de/>]
- VON DAVIER, M., ROST, J., (1995). *Rasch Models. Foundations, Recent Developments and Applications*. Springer; New York. Chapter 20: Polytomous Mixed Rasch Model. Page:371-380
- WELCH, C. J., HOOVER, H. D. (1993). Procedures for extending item bias detection techniques to polytomously scored items. *Applied Measurement in Education*, **6**, 1-19.
- ZWICK, R., DONOGHUE, J. R., GRIMA, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, **30**(3), 233-251.
- ZWICK, R., THAYER, D. T. (1996). Evaluating the magnitude of differential item functioning in polytomous items. *Journal of Educational and Behavioral Statistics*, **21**, 187-201.

EKLER

EK-1

Veri türetmede ve 2.3.1. alt başlığını değerlendirmede kullanılan fonksiyon.

```
library(norm)
library(mice)
library(eRm)
library(classInt)

ni=20 # number of items
t=4 # number of thresholds
ns=150 # number of sample for each group
b.min=-4 #minimum item difficulty
b.max=4 # maximum item difficulty
theta.min=-4
theta.max=4
theta=seq(theta.min,theta.max,((theta.max-theta.min)/(ns-1)))
beta1=seq(b.min,b.max,((b.max-b.min)/(ni-1)))

t1=c(-1.378,-0.056,0.441,0.993)
t2=c(-1.366,-0.301,0.236,1.430)
t3=c(-1.852,-0.378,-0.339,2.570)
t4=c(-1.993,-0.543,1.108,1.428)
t5=c(-0.813,-0.429,0.044,1.198)
t6=c(-1.019,-0.314,-0.018,1.352)
t7=c(-2.031,-1.686,0.597,3.120)
t8=c(-0.623,-0.391,0.058,0.955)
t9=c(-2.353,0.122,0.290,1.941)
t10=c(-1.402,-0.865,-0.264,2.531)
t11=c(-1.852,-0.339,0.366,1.825)
t12=c(-1.738,-0.697,0.609,1.827)
t13=c(-1.057,0.096,0.288,0.673)
t14=c(-1.416,0.082,0.250,1.083)
```

```

t15=c(-1.686,0.172,0.597,0.917)
t16=c(-2.045,-1.187,0.174,3.058)
t17=c(-1.251,-0.980,0.839,1.392)
t18=c(-2.365,-1.172,1.504,2.033)
t19=c(-2.980,-0.032,1.466,1.546)
t20=c(-1.738,-1.594,1.506,1.827)

th=rbind(t1,t2,t3,t4,t5,t6,t7,t8,t9,t10,t11,t12,t13,t14,t15,t16,t17,t18,t19,t20)

beta2=beta1
beta2[c(5,15)]=beta2[c(5,15)]+1
beta1=beta1+th
beta2=beta2+th

datagen<-function(ns,ni,t,beta,theta){
## calculate probabilities & info

  resp=matrix(0,ns,ni)
  prob=array(0,c(ns,(t+1),ni))
  for (i in 1:ni){
    ncat=rep(t+1,ni)
    beta_vec=unlist(beta[i,])
    beta_vec=c(0,beta_vec)
    num=matrix(0,ns,ncat)
    tsum=0
    den=rep(0,ns)
    cc=0
    for(k in 1:max(ncat)){
      tsum=tsum+beta_vec[k]
      num[,k]=exp(cc*theta-tsum)
      den=den+num[,k]
      cc=cc+1
    }
  }
}

```

```

for (k in 1:max(ncat)){
  prob[,k,i]=num[,k]/den
}

}

random=matrix(runif(ns*ni),ns,ni)
for (i in 1:ni){
  sump=numeric(ns)
  for (k in 1:t){
    sump=sump+prob[,k,i]
    resp[,i]=ifelse(random[,i]>sump, resp[,i]+1, resp[,i]+0)
  }
}
return(resp)
}

#####MERGED DATA#####

g1=datagen(ns,ni,t,beta1,theta)
g2=datagen(ns,ni,t,beta2,theta)
total<-rbind(g1,g2)
write.table(total,file="1.txt",sep="",row.names=FALSE,col.names=FALSE)

#####PARTIAL CREDIT MODEL#####

model=PCM(total)

beta_rasch=thresholds(model)
thr=matrix(beta_rasch$threspar,ncol=2,byrow=T)
loc=(cbind(thr[,1])+cbind(thr[,2]))/2
p.s<-person.parameter(model)
extreme=matrix(p.s$pers.ex,nrow=length(p.s$pers.ex),ncol=1)
extreme=cbind(extreme)

```

```

id<-1:nrow(total)
if (dim(extreme)[1]!=0){
  id_y<-id[-extreme[,1]]
  pers<-cbind(id_y,p.s$theta.table[[1]])
  pers.se<-cbind(id_y,unlist(p.s$se.theta))

  pers.extrm=cbind(extreme[,1],NA)
  pers.extrm.se=cbind(extreme[,1],NA)
  pers=rbind(pers,pers.extrm)
  pers.se=rbind(pers.se,pers.extrm.se)

  pers=pers[order(pers[,1]),]
  pers.se=pers.se[order(pers.se[,1]),]

}else{
  id_y<-id
  pers<-cbind(id_y,p.s$theta.table[[1]])
  pers.se<-cbind(id_y,unlist(p.s$se.theta))

  pers=pers[order(pers[,1]),]
  pers.se=pers.se[order(pers.se[,1]),]

}

#####DIF#####

grp1<-cbind(0,1:ns)
grp2<-cbind(1,1:ns)
a<-cbind(grp1[,1])
b<-cbind(grp2[,1])
grup=rbind(a,b)
difgrup<-cbind(grup,total)
grup<-difgrup[,1]

```

```

data<-difgrup[,2:(ni+1)]
ci<-((ns*2)/50)
if (dim(extreme)[1]!=0){
    z_res<-residuals(p.s)
    manifest<-factor(grup[-extreme])
    CInt<-classIntervals(var=pers[-extreme,2],n=ci,style="equal")
    theta_gr<-findCols(CInt)
    theta_gr<-as.factor(theta_gr)

} else {

    z_res<-residuals(p.s)
    manifest<-factor(grup)
    CInt<-classIntervals(var=pers[,2],n=ci,style="equal")
    theta_gr<-findCols(CInt)
    theta_gr<-as.factor(theta_gr)

}

dif<-matrix(NA,nrow=ni,ncol=2)
for(i in 1:ni){
    anov<-aov(z_res[,i]~theta_gr+manifest+theta_gr*manifest)
    anov.test<-summary(anov)
    dif[i,]<-anov.test[[1]][,5][2:3]
}

cv<- 0.05/(ni*3)
dif2<- numeric(ni)
dif2[dif[,1]>=cv & dif[,2]>=cv]<-0
dif2[dif[,1]<cv & dif[,2]>=cv]<-1
dif2[dif[,1]>=cv & dif[,2]<cv]<-2
dif2[dif[,1]<cv & dif[,2]<cv]<-2
dif3<-cbind(1:ni,dif2)

```

EK-2

Veri türetmede ve 2.3.2. alt başlığını değerlendirmede kullanılan fonksiyon.

```
library(norm)
library(mice)
library(eRm)
library(classInt)

ni=20 # number of items
t=4 # number of thresholds
ns=150 # number of sample for each group
b.min=-4 #minimum item difficulty
b.max=4 # maximum item difficulty
theta.min=-4
theta.max=4
theta=seq(theta.min,theta.max,((theta.max-theta.min)/(ns-1)))
beta1=seq(b.min,b.max,((b.max-b.min)/(ni-1)))

t1=c(-1.378,-0.056,0.441,0.993)
t2=c(-1.366,-0.301,0.236,1.430)
t3=c(-1.852,-0.378,-0.339,2.570)
t4=c(-1.993,-0.543,1.108,1.428)
t5=c(-0.813,-0.429,0.044,1.198)
t6=c(-1.019,-0.314,-0.018,1.352)
t7=c(-2.031,-1.686,0.597,3.120)
t8=c(-0.623,-0.391,0.058,0.955)
t9=c(-2.353,0.122,0.290,1.941)
t10=c(-1.402,-0.865,-0.264,2.531)
t11=c(-1.852,-0.339,0.366,1.825)
t12=c(-1.738,-0.697,0.609,1.827)
t13=c(-1.057,0.096,0.288,0.673)
t14=c(-1.416,0.082,0.250,1.083)
t15=c(-1.686,0.172,0.597,0.917)
t16=c(-2.045,-1.187,0.174,3.058)
t17=c(-1.251,-0.980,0.839,1.392)
```

```

t18=c(-2.365,-1.172,1.504,2.033)
t19=c(-2.980,-0.032,1.466,1.546)
t20=c(-1.738,-1.594,1.506,1.827)

th=rbind(t1,t2,t3,t4,t5,t6,t7,t8,t9,t10,t11,t12,t13,t14,t15,t16,t17,t18,t19,t20)

beta2=beta1
beta2[c(5,15)]=beta2[c(5,15)]+1
beta1=beta1+th
beta2=beta2+th

total<-read.fwf("1.txt", width=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1))

out <- total
for (i in 1:length(which((beta1-beta2)[,1]!=0))){
out[1:ns,which((beta1-beta2)[,1]!=0)[i]]<-paste(NA,out[1:ns,which((beta1-
beta2)[,1]!=0)[i]],sep="&")
out[(ns+1):(2*ns),which((beta1-beta2)[,1]!=0)[i]]<-
paste(out[(ns+1):(2*ns),which((beta1-beta2)[,1]!=0)[i]],NA,sep="&")
}
outlist <- apply(out, 2, strsplit, split="&")
total=do.call( "cbind.data.frame", lapply( outlist, do.call, what="rbind" ) )
total=as.matrix(total)
total=apply(total,2,as.numeric)

#####PARTIAL CREDIT MODEL#####

model=PCM(total)

beta_rasch=thresholds(model)
thr=matrix(beta_rasch$threspar,ncol=2,byrow=T)
loc=(cbind(thr[,1])+cbind(thr[,2]))/2
p.s<-person.parameter(model)
extreme=matrix(p.s$pers.ex,nrow=length(p.s$pers.ex),ncol=1)

```

```

extreme=cbind(extreme)

id<-1:nrow(total)
if (dim(extreme)[1]!=0){
  id_y<-id[-extreme[,1]]
  pers<-cbind(id_y,p.s$theta.table[[1]])
  pers.se<-cbind(id_y,unlist(p.s$se.theta))

  pers.extrm=cbind(extreme[,1],NA)
  pers.extrm.se=cbind(extreme[,1],NA)
  pers=rbind(pers,pers.extrm)
  pers.se=rbind(pers.se,pers.extrm.se)

  pers=pers[order(pers[,1]),]
  pers.se=pers.se[order(pers.se[,1]),]

}else{
  id_y<-id
  pers<-cbind(id_y,p.s$theta.table[[1]])
  pers.se<-cbind(id_y,unlist(p.s$se.theta))

  pers=pers[order(pers[,1]),]
  pers.se=pers.se[order(pers.se[,1]),]

}

pers_1<-cbind(pers[,2])
pers.se_1<-cbind(pers.se[,2])

```


ÖZGEÇMİŞ

1- Bireysel Bilgiler

Adı: SELCEN
Soyadı: YÜKSEL
Doğum yeri ve tarihi: ELAZIĞ-13.06.1981
Uyruğu: T.C.
Medeni durumu: EVLİ
İletişim adresi ve telefonu: Ankara Üniv. Tıp Fak. Biyoistatistik AD, 06100,
Sıhhiye, Ankara. 0312 – 5958175.
e-mail adresi: selcenpehlivan@yahoo.com

2- Eğitim

2007-Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı (Doktora)
2004-2007 Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı
(Yüksek Lisans)
2000-2004 Fırat Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü (Lisans)

Yabancı Dil: İngilizce

3- Ünvanları

Uzman (2007)

Tez Konusu: Tarama Testlerinin Etkinliklerinin Karşılaştırılması İçin
Örneklem Genişliği Saptaması Ve Tarama Testleri Yardımı İle Prevalans Kestirimi

4- Mesleki Deneyimi

2007- Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı
(Araştırma Görevlisi)

2005-2007 Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Bilgisayar
Eğitimi Bölümü (Araştırma Görevlisi)

5- Üye Olduğu Bilimsel Kuruluşlar

- International Biometric Society - Eastern Mediterranean Region (IBS/EMR)
- Biyoistatistik Derneği
- International Association of Computer Science and Information Technology (IACSIT)Tıp Bilişimi Derneği
- International Young Bernoulli Statistician Group
- American Statistical Association (ASA) (2009-)

6 – Bilimsel İlgi Alanları

Yurtdışı Yayınlar:

1.Ozen T, Kamburoğlu K, Cebeci AR, **Yüksel SP**, Paksoy CS. “*Interpretation of chemically created periapical lesions using 2 different dental cone-beam computerized tomography units, an intraoral digital sensor, and conventional film*”Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2009 Mar;107(3):426–32.

2.Kamburoğlu K, Kiliç C, Ozen T, **Yüksel SP**.”*Measurements of mandibular canal region obtained by cone-beam computed tomography: a cadaveric study.*”Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2009 Feb;107(2):e34–42.

3. Kamburoğlu K, Murat S, **Yüksel SP**, Cebeci AR, Horasan S.”*Detection of vertical root fracture using cone-beam computerized tomography: an in vitro assessment.*” Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2010 Feb;109(2):e74–81.

4. Kamburoğlu K, Murat S, **Yüksel SP**. “*The effects of digital image enhancement on the detection of vertical root fracture.*” Dent Traumatol. 2010 Feb;26(1):47-51.

5. Kamburoglu K, Senel B, **Yüksel SP**, Ozen T.” *A comparison of the diagnostic accuracy of in vivo and in vitro photostimulable phosphor digital images in the detection of occlusal caries lesions.*” Dentomaxillofac Radiol. 2010 Jan;39(1):17–22.

6. Kamburoğlu K, Murat S, **Yüksel SP**, Cebeci AR, Paksoy CS.” *Occlusal caries detection by using a cone-beam CT with different voxel resolutions and a*

digital intraoral sensor”Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2010 May;109(5):e63-e69.

7. Kilic C, Kamburoglu K, **Yuksel SP**, Ozen T.” *An Assessment of the Relationship between the Maxillary Sinus Floor and the Maxillary Posterior Teeth Root Tips Using Dental Cone-beam Computerized Tomography.*” Eur J Dent. 2010 Oct;4(4):462-7.

8. Senel B, Kamburoglu K, Uçok O, **Yuksel SP**, Ozen T, Avsever H.” *Diagnostic accuracy of different imaging modalities in detection of proximal caries.*” Dentomaxillofac Radiol. 2010 Dec;39(8):501-11.

9. Bodur A., Turgut Z., Uraz A., Koç E., Karaduman B., Altok K., **Pehlivan S.**, Bal B. “Kronik Hemodiyaliz Hastalarında Periodontal Sağlık Durumunun Değerlendirilmesi” GÜ Diş Hek Fak Derg. 27(1): 29-35, 2010

10. Kamburoğlu K, Kurşun S, **Yuksel S**, Oztaş B.” *Observer ability to detect ex vivo simulated internal or external cervical root resorption.*” J Endod. 2011 Feb;37(2):168-75.

11. Ceylan C., Ceylan T., Odabaş Ö., **Yuksel S.**, Doğan S., Yiğman M. “*Evaluation of the role of digital rectal examination and transrectal ultrasonography in diagnosis of prostate cancer in Turkish men*”JCAM 2011 March; DOI: 10.4328/JCAM.610

12. Ceylan C., Ceylan T., Odabaş Ö., **Yuksel S.**, Doğan S., Yiğman M. “*Correlation of serum free/total prostatespecific antigen (Psa) levels with histological findings in the Turkish men with prostatic disease after the first biopsy*” International Journal of Medicine (IJM) Vol. 1(1) pp. 01-07 March, 2011

13.Kamburoğlu K, Murat S, Kolsuz E, Kurt H, **Yuksel S**, Paksoy C.”*Comparative assessment of subjective image quality of cross-sectional cone-beam computed tomography scans.*” J Oral Sci. 2011 Dec;53(4):501-8.

14. Kamburoglu K, **Yuksel S.**”*A comparative study of the accuracy and reliability of multidetector CT and cone beam CT in the assessment of dental implant site dimensions.*” Dentomaxillofac Radiol. 2011 Oct;40(7):466-7

15. Bavbek S, Yılmaz I, Celik G, Aydın O, Erkeköl FO, Orman A, Kurt E, Ediger D, Dursun B, Abadoğlu O, Ozşeker F, Akkaya E, Karakış GP, Canbakan S, **Yuksel S**, Mısırlıgil Z. “*Prevalence of aspirin-exacerbated respiratory disease in*

patients with asthma in Turkey: A cross-sectional survey.” Allergol Immunopathol (Madr). 2011 Aug 31

16. Ceylan C., Ceylan T., Odabaş Ö., **Yüksel S.**, Doğan S., Yiğman M.” *Incidental detection of asymptomatic chronic prostatitis and benign prostatic hyperplasia in patients with high serum PSA levels: a preliminary retrospective study*” Turkish Journal of Urology 2011;37(2):86-90

17. Idil A., Ozen M., Atak N, Elhan A.,**Pehlivan S.** “*Validity and reliability study of Turkish version on low vision with quality of life questionnaire*” Int J. Ophthalmol 2011;4(5):546-551

18. Kamburoglu K, Kolsuz E, Murat S, **Yüksel S**, Ozen T.”*Proximal caries detection accuracy using intraoral bitewing radiography, extraoral bitewing radiography and panoramic radiography.*” Dentomaxillofac Radiol. 2012 Sep;41(6):450-9.

19. Murat S, Kamburoğlu K, Isayev A, Kurşun S, **Yüksel S.** “*Visibility of Artificial Buccal Recurrent Caries Under Restorations Using Different Radiographic Techniques.*” Oper Dent. 2012 Aug 23.

20. Karagol BS, Orun UA, Zenciroglu A, Yuksel SP, Okumus N, Karademir S. “*The diameter of coronary arteries in healthy newborns at birth, 1 and 6 months of ages*” J Matern Fetal Neonatal Med. 2012 Aug 27.

Yurtiçi ve Yurtdışı Sözlü Bildiriler:

1.Pehlivan S, Karaağaoğlu E.”*Tarama Testlerinin Eşleşmiş ve Eşleşmemiş Araştırma Tasarımları için Örneklem Genişliği Saptaması.*”IX.Ulusal Biyoistatistik Kongresi 5-9 Eylül 2006-Zonguldak

2.Pehlivan S, İncetaş O, Aydın S.”*Türkiye’de İmpakt Faktörünün Akademik Değerlendirmede ve Üniversitelerin Bilimsel Sıralamasında Kullanımı.*” 3.Ulusal Tıp Bilişimi Kongresi. 16-19 Kasım 2006-Antalya

3.Pehlivan S.”*Sample Size Determination for Paired and Unpaired Study Designs of Screening Tests.*” 56th Session of the International Statistical Institute. 22-29 August 2007-Portugal-Lisbon

4. Tüfekçi A, Çakır H, İncetaş O, **Pehlivan S**, Tanyeri U. "Intranet Tabanlı Bir Sınav Sistemi: Gerekçe ve Tasarım" The Proceedings of 7th International

Educational Technology Conference, IETC 2007, May 3-5, 2007, Near East University - Cyprus 2007.

5. Tüfekçi A, Çakır H, İncetaş O, **Pehlivan S**, Tanyeri U. "İnternet Tabanlı Yarı Çevrimiçi Bir Sınav Sistemi Örneği" 1. Uluslararası Bilgisayar ve Öğretim Teknolojileri Sempozyumu, May 16-18, 2007, S. 867- 875, Çanakkale

6. Varol N, Özkal P, Gümüş Akay G, **Yüksel S**, Köroğlu R, Küçük Ö, İbiş E, Aras G, Sunguroğlu A. "Kültüre Edilmiş İnsan Lenfosit Hücreleri Üzerine "Samarium"un Olası Genotoksik Etkisi." Uluslar arası Katılımlı VIII.Ulusal Tıbbi Genetik Kongresi 06-09 Mayıs 2008 Çanakkale.

7. Genç Y, **Yüksel S**, Demir S, Tanyeri U, Ateş C, Öztuna D. "Kümelenmiş Verilerde Mann-Whitney U Testinin Uygulanması" XI. Ulusal Biyoistatistik Kongresi 27–30 Mayıs 2008 Malatya

8. Doğanay B, Genç Y, **Yüksel S**, Gültekin S. "İşlem Karakteristiği Eğrilerinde Ortak Değişken Etkisinin İncelenmesi: Lehmann Ailesi Yaklaşımı." XI. Ulusal Biyoistatistik Kongresi 27–30 Mayıs 2008 Malatya

9. Uraz A, Tacoy G, Karched M, Turgut Z, **Yüksel S**, Doğan B, Bodur A, Asikainen S. "Systemic Immunoinflammatory Markers in Periodontitis Patients with Acute Myocardial Infarction." UADR 86th Genera Session July 2 2008

10. Doğan B, Uraz A, Oscarsson J, Granström E, Turgut Z, **Yüksel S**, Takoy G, Bodur A, Asikainen S. "Immunoreactivity to bacteria-associated molecular patterns (BAMPs) in periodontitis" EROPERİO 2009 Stocholm

11. **Yüksel S**, Elhan A.H, Doğanay B, Küçükdeveci A.A., Kutlay S, Tennant A. "Determination of a Diagnostic Cut-Point for Mini Mental State Examination in a Normal Population in Turkey" 5.th Conference of the IBS-EMR 10-14 May 2009

12. Genç Y., Öztuna D., Yüksel S, Ateş C. "An Overview of Closed Form Methods for the Analysis of Clustered Data". Applied Statistics 2009 International Conference Sept. 20 – 23, 2009 Ribno (Bled), Slovenia

13. Bavbek S, Yılmaz İ, Çelik G, Aydın Ö, Erkeköl F Ö, Aktaş A, Kurt E., Ediger D, Dursun B, Abadoğlu Ö, Özşeker F, Akkaya E, Karakış G, Canbakan S, Yüksel S, Mısırlıgil Z. "Türkiye’de Erişkin Astımlı Hastalarda Aspirin Duyarlılığı Prevelansı". XVII. Ulusal Allerji ve Klinik İmmünoloji Kongresi 3-7 Kasım 2009 Antalya (Çok merkezli bildirimler dalında birincilik ödülü aldı)

14. Yılmaz İ, Seçil D, Yüksel S, Mısırlıgil Z, Demirel Y.S, Bavbek S. “Mevsimsel alerjik rinitli (MAR) hastalarda bronşial remodelling’in indükte balgamda vascular endothelial growth factor (VEGF) ölçümü ile değerlendirilmesi ve bunun bronş aşırı duyarlılığı ile ilişkisi” Türk Toraks Derneği 13.Yıllık Kongresi. 5-9 Mayıs 2010

15.Yüksel S., Elhan AH., Oztuna D., Kucukdeveci AA., Kutlay S. “Fitting a Mixed Rasch Model to Nottingham Health Profile in Turkish Population.” International Conference on Probabilistic Models for Measurement in Education, Psychology, Social Science and Health, 13-16 June 2010, Copenhagen, Denmark.

Poster Sunumları:

1.Yüksel S.Sample Size Calculation for Screen-Positive Designs of Screening Tests. XXIV. International Biometrics Conference. 13–18 July 2008, University College Dublin Ireland

2. Yılmaz İ, Seçil D, **Yüksel S**, Mısırlıgil Z, Demirel Y.S, Bavbek S. Evaluation of vascular remodeling in induced sputum and relationship to bronchial hyperreactivity in patients with seasonal allergic rhinitis. 29th EAACI Congress 5-9 June 2010 London (Poster dalında birincilik ödülü kazandı)

3. Yılmaz İ, Bayraktar N, Ceyhan K, Derya S, **Yüksel S**, Mısırlıgil Z, Demirel Y.S, Bavbek S. “Perrenial alerjik rinitli hastalarda indükte balgamda vasküler endothelial growth factor ve Endostatin düzeyi ve bunun metakolin bronş duyarlılığı ile ilişkisi” Türkiye Ulusal Alerjik Hastalıklar ve Klinik İmmunolojide Tedavi Kongresi 01-04 Kasım 2011 Antalya (“Erişkin Solunum Yolu Allerjileri kategorisinde sunulan posterler arasında birincilik ödülü kazandı)

4. Çiftçi Ö., Güloğlu D. , Çipe F. , Polat M. , **Yüksel S.** , Köse K. , Doğu F. , İkinciogulları A. “Atopik dermatitli çocuklarda T reglatuvar (Treg) hücre düzeyi” Türkiye Ulusal Alerjik Hastalıklar ve Klinik İmmunolojide Tedavi Kongresi 01-04 Kasım 2011 Antalya

Projeler:

1. Mevsimsel ve Yıl Boyu Alerjik Riniti Olan Olgularda Bronşial Remodelling’in İndükte Balgamda Vascular Endothelial Growth Factor ve

Endostatin Ölçümü ile Değerlendirilmesi ve Bunun Bronş Aşırı Duyarlılığı ile İlişkisi. Ankara Üniversitesi Bilimsel Araştırma Projesi. Proje Yürütücüsü: Sevim Bavbek Yardımcı Araştırmacılar: İnsu Yılmaz, Derya Seçil, Koray Ceyhan, Nilüfer Bayraktar, **Selcen Yüksel**, Zeynep Mısırlıgil, Yavuz Selim Demirel Proje Numarası : 09B3330019 Proje Başlangıç Tarihi: Mayıs 2009 Proje Bitiş Tarihi: Ağustos 2011

Seminerler:

1. Tarama testleri için prevalans tahmini. 13 Nisan 2004 Yer: Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik A.D.

2. Tarama testlerini karşılaştırmada kullanılan araştırma tasarımları için örneklem genişliği saptaması 13 Aralık 2005 Yer: Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik A.D.

3. Tarama testlerini karşılaştırmada kullanılan tarama-pozitif araştırma tasarımları için örneklem büyüklüğü saptaması.5 Şubat 2008.Yer: Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik A.D.

4. Tanı Testlerinin Doğruluğunu Belirlemede RASCH Analizi Olasılık Kestirimiyle Klasik Yöntem Kestirimlerinin Karşılaştırılması.6 Ocak 2009. Yer: Ankara Üniversitesi Tıp Fakültesi Biyoistatistik A.D.

5. Referans tanı testine karşı yeni tanı testi:3 olası rol ve bu roller için kullanılan istatistiksel yöntemler. 20 Nisan 2010. Yer: Ankara Üniversitesi Tıp Fakültesi Biyoistatistik A.D.