

GENELLEŐTİRİLMİŐ AŐAMALI DOĐRUSAL MODEL İLE RASCH MODELİNİN PARAMETRELERİNİN DEĐİŐMEZLİĐİNİN KARŐILAŐTIRILMASI

Uzman Tlin ACAR

Parantez Eđitim AraŐtırma, totbicer@gmail.com

Dođ Dr. Hlya KELECİOĐLU

Hacettepe niversitesi, hulyaebb@hacattepe.edu.tr

zet: Bu alıŐmada, GenelleŐtirilmiŐ AŐamalı Dođrusal Model (GADM) ve Rasch Modeli ile kestirilen madde ve yetenek parametrelerinin deđiŐmezliđi incelenmiŐtir. Madde parametresinin deđiŐmezliđine cinsiyet ve yerleŐim yerine gre; yetenek parametresinin deđiŐmezliđine madde gclk ortalamaları eŐit ve farklı testlere gre bakılmıŐtır. AraŐtırmanın evreni Ortađretim Kurumlar Sınavı'na (OKS) 2006 yılında giren 789307 đrencidir. Bu evrenden tesadfi rnekleme yoluyla 10727 đrenci seilmiŐtir. OKS'nin Trke ve Sosyal Bilgiler alt testlerinden elde edilen veriler kullanılmıŐtır. Analiz sonularına gre iki model de madde parametrelerinin deđiŐmezliđini yksek dzeyde sađlamıŐtır. Yetenek parametrelerinin deđiŐmezliđi GADM ile tam olarak sađlanırken Rasch Modelde yksek dzeyde sađlanmıŐtır. İki model arasındaki uyum yksek bulunmuŐtur. Yetenek parametreleri arasındaki uyum, madde parametreleri arasındaki uyumdan daha yksektir.

Anahtar szckler: GADM, Rasch model, parametre deđiŐmezliđi

GİRİŐ

Son yıllarda zellikle geniŐ lekli test geliŐtirme ve deđerlendirme uygulamalarında Madde Tepki Kuramına (MTK) dayalı modellerden yararlanılmaktadır. MTK'nın test geliŐtirmede kullanılmasının en nemli gerekelerinden biri, madde ve yetenek parametrelerinin, verilerin elde edildiđi grubun zelliklerinden bađımsız olarak elde edilebilmesidir. Madde tepki kuramı, dođrusal olmayan regresyon modeline dayandıđından, test ve madde parametreleri rneklemden bađımsız kestirilebilmektedir (Breithaupt ve Zumbo, 2002). Model kuramsal olarak, parametrelerin cinsiyet, sosyo-ekonomik dzey gibi llen yetenekle ilgili olmayan zelliklerden etkilenmemesini sađlamaktadır. Bu zellik de parametrelerin deđiŐmezliđi olarak adlandırılmaktadır (Hambleton, Swaminathan, ve Rogers, 1991).

Parametre deęişmezlięi madde ve yetenek parametreleri boyutunda ele alınmaktadır. MTK'ya dayalı modellerle kestirilen madde özellikleri, bir cevaplayıcı evreninin özelliklerine baęlı deęildir. Bireylerin örtük özellik boyutundaki yerleri (yetenekleri), madde fonksiyonları bilinen maddelere verilen yanıtlara dayanarak kestirilebilir ve bireylerin yanıtladıęı maddeler farklı ölçme durumlarında elde edilse bile yetenek ölçüleri deęişmez. Madde ve yetenek parametrelerinin gruptan ve ölçme uygulamalarından baęımsız olarak kestirilmesinin test geliřtirmede, grupları karřılařtırmada ve deęerlendirme çalıřmalarında büyük avantajları vardır. Parametre deęişmezlięi, geniş ölçekli deęerlendirme çalıřmalarında, farklı sınıf düzeyinde elde edilen veriler arasında; aynı sınıf düzeyindeki farklı gruplar içinde; kültürlerarası karřılařtırmalarda ve deęişen madde fonksiyonunu incelemeye yansız bir karřılařtırma saęlamaktadır (Breithaupt ve Zumbo, 2002; Reise, Ainsworth ve Haviland, 2005).

Geniş ölçekli test uygulamalarında, deęişen madde fonksiyonunu belirlemek ve testi alan farklı grupları karřılařtırmak için son yıllarda kullanılmaya bařlanan dięer bir model de Ařamalı Doğrusal Model (ADM)'dir. ADM, özellikle eęitimde ařamalı olarak yer alan deęişkenlerin yapısına daha uygun bir çözüm üretmektedir. Eęitimde ölçmeye konu olan özellikler genellikle ařamalı bir yapı göstermektedir. Öğrenciler sınıf ve okul gibi ařamalı bir yapı içerisinde bulunurlar. Ařamalı yapı gösteren baęımsız deęişkenlere standart regresyon eřitlikleri uygulandıęında, baęımsızlık varsayımı ihlal edilmiř olur. Bu nedenle, ařamalı verilerin yapısına uygun olarak Ařamalı Doğrusal Modellemenin kullanılması önerilmektedir (Bryk ve Raudenbush 1992).

ARAřTIRMANIN AMACI

Bu çalıřmanın amacı, Türkiye'de geniş ölçekli test uygulamalarından biri olan Ortaöęretim Kurumlar Sınavı'nın Türkçe ve Sosyal Bilgiler alt testlerinin parametrelerinin deęişmezlięini Rasch Model ve GADM ile inceleyip sonuçlarını birbirleriyle karřılařtırmaktır.

Madde Tepki Kuramı

Madde Tepki Kuramında kiřinin örtük özellięi test maddelerine verdięi yanıtlardan kestirilebilir. Kuram, örtük özellikler ve madde karakteristik eęrisi kavramlarına dayanmaktadır. MTK'ya dayalı modeller yetenek düzeyi ve madde özelliklerinin bireyin maddeye verdięi tepkilerle olan iliřkisini tanımlar. Madde karakteristik eęrisi, bir maddeye verilen tepkinin olasılıęına baęlı olarak örtük özellikteki deęişimi gösterir (Embretson ve Reise, 2000, s. 40). Madde karakteristik fonksiyonunun aldıęı deęerler

örneklem grubuna bağlı değildir. Bu nedenle fonksiyonu tanımlayan parametreler örneklem grubuna göre değişmez (Lord, 1968).

Puanlaması 1 ve 0 biçiminde olan tek boyutlu testlerde en çok kullanılan MTK modelleri bir, iki ve üç parametrelili lojistik modellerdir. Bu modeller, kestirilen parametre sayısına göre adlandırılmışlardır. Bir parametrelili lojistik model ve onun özel hali olan Rasch Modelinde kestirilen b parametresi madde güçlüğü olarak adlandırılmaktadır. İki parametrelili lojistik modelde b madde güçlük parametresi ile birlikte kestirilen a parametresi madde ayırıcılık gücüdür. Üç parametrelili lojistik modelde a ve b parametresiyle birlikte şansla doğru cevabı bulma olasılığını gösteren c parametresi yer alır.

Lojistik modellerden Rasch Modelin genel denklemini aşağıdaki gibidir.

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\theta_s - \beta_i)}{1 + \exp(\theta_s - \beta_i)}$$

θ_s , bireyin yetenek düzeyini; β_i , i maddesinin güçlük indeksini göstermektedir. $P(X_{is} = 1 | \theta_s, \beta_i)$ ise, β_i güçlük düzeyindeki maddeyi, θ_s yetenek düzeyindeki bireyin doğru cevaplama olasılığını tanımlamaktadır (Embretson ve Reise, 2000, s. 50).

Aşamalı Doğrusal Model

Aşamalı Doğrusal Model, bireyin özellikleri gibi, grup üyelerinin özelliklerinin de içinde buldukları grubun bir fonksiyonu olarak açıklandığı doğrusal eşitlikler türeten bir modeldir. ADM, eğitimle ilgili problemler için de uygun çözümler üretir. Eğitimle ilgili pek çok değişken iç içe geçmiştir. Öğrenciler sınıfların içinde, sınıflar okulların içinde, okullar da sosyo-ekonomik düzeye göre farklı bölgelerde yerleşmişlerdir. Bu nedenle öğrenci başarısı ile ilgili değişkenler üzerinde yapılan çalışmalar çok düzeylidir (Arnold, 1992).

İç içe geçmiş veriler klasik doğrusal modellerle analiz edildiğinde gözlemlerin bağımsızlığı varsayımı ihlal edilir (Bryk ve Raudenbush, 1987). Eğer araştırmacı verilerinin aşamalı yapısını göz ardı ederse analiz sonuçları hatalı olacaktır. Örneğin, eğitimde sıkça öğrencilerin çeşitli alanlardaki başarı puanları incelenir. Öğrenciler sınıflardan ve sınıflar da içinde buldukları okullardan bağımsız olarak düşünülemez. Öğrenciler, aynı sınıfın içinde ortak bir öğretmeni, ortak bir öğretim ortamını ve ortak öğrenme deneyimlerini paylaşırlar. Bu ortak deneyimler nedeniyle, öğrenci başarı puanları, aynı sınıfın içindeki diğer öğrencilerin başarı puanlarıyla doğrudan ilişkilidir. Eğer iç içe geçmişlik göz ardı edilirse yapılan istatistiksel analizlerde düzeylerin

standart hataları kestirilemez. Çünkü, öğrenci ve sınıf düzeylerindeki değişkenlik istatistiksel olarak modele dahil edilmemiştir (Willms, 1999).

ADM, sonuç değişkeninin sürekli olduğu durum için geliştirilmiş bir modeldir ve pek çok varsayımı vardır. Sonuç değişkeninin sıralama veya sınıflama ölçeğinde olduğu durumlar için ise genelleştirilmiş aşamalı doğrusal model (GADM) geliştirilmiştir ve ADM'ye göre daha az varsayımı vardır (Kamata 1998; Williams, 2003; Tate, 2004; Subedi, 2005). GADM, son yıllarda değişen madde fonksiyonunu belirlemek amacı ile de kullanılmaya başlanmıştır (Chaimongkol, Huffer ve Kamata 2007; Vaughn 2006; Kamata, Chaimongkol, Genç ve Bilir 2005; Williams 2003; Luppescu 2002). Değişen madde fonksiyonu belirleme çalışmalarından yansız bir sonuç elde edebilmek için, uygulanan modellerin parametre değişmezliğini de sağlaması gerekmektedir.

ADM, yöntemi gereği (Koopmans,1998; Willms,1999),

- Bireysel düzeye bir regresyon eşitliği sağlar.
- Regresyon eşitliğinin parametrelerini grup üyeleri aracılığıyla değiştirir. Bireysel-düzey parametrelerinde değişimi açıklamak için grup-düzey değişkenlerini kullanır.
- Değişkenlere ilişkin ana etkilerin testlerine imkân verir.

ADM, düzey-1 ve düzey-2 olmak üzere çoklu düzey modellerini kapsayan bir istatistiksel model sağlar. Grup araştırmalarında düzey-1, bireyin düzeyini ve düzey-2 ise grup düzeyini temsil eder. Her grupta farklı regresyon doğrularının olduğu düşünüldüğünde, gözlem sayıları farklı birden fazla grup değişkeni ve birden fazla değişken ADM ile rahatlıkla modellenebilir (Gokiort ve Ricker, 2004).

Yapılan çalışmalar, GADM'de Rasch modeline göre madde güçlük ve yetenek parametrelerinin kestirilebildiğini; GADM ile Rasch modelin benzerliğini göstermiştir (Kamata, 2002; Kim, 2003). GADM de, madde puanlarının (sonuç değişkeninin) iki kategorili olduğu düzey 1 (madde düzeyi) ve düzey 2 (birey düzeyi) modelleri aşağıdaki eşitliklerdeki gibi kurulur (Kamata, 2002).

Düzye 1 modeli: $\eta_{ij} = \log\left(\frac{P_{ij}}{1-P_{ij}}\right) = B_{0j} + B_{1j}X_{1j} + B_{2j}X_{2j} + \dots + B_{(k-1)j}X_{(k-1)j}$

Düzye 2 modeli:

$$B_{0j} = y_{00} + u_{0j}$$

$$B_{1j} = y_{10}$$

...

$$B_{(k-1)j} = y_{(k-1)0}$$

Denklemdaki P_{ij} , j. kişinin i. maddeyi doğru cevaplandırma olasılığı ve X_{1j} , j. kişi için i. yapay değişkendir. Gözlem i. madde üzerinde olduğu zaman yapay değişken 1 değerini alır, aksi takdirde 0 değerini alır. β_{0j} : kesme noktasıdır. β_{ij} , $i=1$ 'den $(k-1)$ 'e kadar X_{ij} ilişkili bir katsayıdır. u_{0j} , β_{0j} 'nin tesadüfî bileşenidir. u_{0j} , dağılımın ortalaması 0 ve varyansı τ olan normal dağılım gösterir. Düzey 1 ve 2 modelleri, i. madde j. kişi için birleştirildiğinde i. maddenin j.kişi için doğru cevaplandırılma olasılığı aşağıdaki gibidir.

$$P_{ij} = \frac{1}{1 + \exp\{-[u_{0j} - (-\gamma_{i0} - \gamma_{00})]\}}$$

Buradaki u_{0j} , j. kişinin yetenek parametresi ve $(-\gamma_{i0} - \gamma_{00})$ işlemi de i. maddenin güçlük parametresidir. Dolayısıyla GADM'de kestiricisiz denklemlerden elde edilen model, 1 parametrelî Rasch modeline eşittir (Kamata,2002). Bu gerekçe ile bu çalışmada, Rasch Model ve GADM'nin madde ve yetenek parametrelerinin değişmezliğini sağlamada benzerlik gösterip göstermediği incelenmiştir.

Problem

Madde ve yetenek parametrelerinin değişmezliğini sağlamada Rasch Modeli ve GADM arasında benzerlik var mıdır?

Alt Problemler

1. GADM ile kestirilen madde güçlük ve yetenek parametrelerinin değişmezliği nasıldır?
2. Rasch modeli ile kestirilen madde güçlük ve yetenek parametrelerinin değişmezliği nasıldır?
3. GADM ve Rasch model ile kestirilen madde güçlük ve yetenek parametrelerinin değişmezlikleri arasındaki ilişki ne düzeydedir?

YÖNTEM

Araştırmanın türü

Bu araştırma Rasch modeli ile GADM'yi madde ve yetenek parametrelerinin değişmezliği bakımından karşılaştıran betimsel bir çalışmadır.

Evren ve örneklem

Araştırmanın evrenini 2006 OKS'ye giren 798307 öğrenci; örneklemine de bu evrenden tesadüfî örnekleme yöntemi ile seçilen 10727 öğrenci oluşturmuştur.

Verilerin analizi

Araştırmanın verileri, 2006 yılı OKS Türkçe ve Sosyal Bilgiler alt testlerine verilen yanıtlardan oluşmaktadır. Madde parametrelerinin değişmezliğini incelemek için, Türkçe ve Sosyal Bilgiler alt testleri cinsiyete göre iki ve mezun olunan okulun bulunduğu yerleşim yerine göre üç gruba ayrılmıştır. Mezun olunan okulun bulunduğu yerleşim yeri, coğrafi bölgelerin cari fiyatlarla kişi başına düşen gayri safi yurtiçi hasıla (YTL) verilerinden yararlanarak sınıflandırılmıştır (TÜİK, 2001). Buna göre kuzey doğu, orta doğu, orta ve güneydoğu Anadolu ile doğu Karadeniz bölgesindeki okullar birinci (Yer1); batı Karadeniz, Akdeniz, batı Anadolu ve batı Marmara bölgesindeki okullar ikinci (Yer2); Ege, doğu Marmara ve İstanbul bölgesindeki okullar, üçüncü bölge (Yer3) olarak ele alınmıştır. Cinsiyet için iki, yerleşim yeri için üç gruptan madde güçlük parametreleri kestirilmiştir.

Yetenek parametrelerinin değişmezliği, Rasch Modeli ile kestirilen b parametrelerine göre iki ayrı durumda incelenmiştir. Birinci durumda madde güçlük indeksi ortalamaları birbirine eşit olan iki ayrı test oluşturulmuştur. İkinci durumda maddeler kolay ve zor olarak ayrılarak iki test elde edilmiştir. Türkçe ve Sosyal Bilgiler alt testlerine yanıt veren bireylerin, yukarıdaki dört ayrı durum için yetenek parametreleri kestirilmiştir. Tablo 1’de alt testlere göre b değerlerinin cinsiyet, yerleşim yeri ve yetenek gruplarının alt testlere göre istatistikleri verilmiştir.

Tablo 1. Gruplara göre b değerlerinin ortalama ve standart sapmaları

Gruplar	k	Türkçe		Sosyal B.	
		Ort.	Ss	Ort.	Ss
Tüm Grup (n=10727)	25	0.75	0.90	0.54	0.57
Kız (n=5423)	25	0.65	0.92	0.52	0.64
Erkek (n=5304)	25	0.86	0.89	0.55	0.52
Yer1 (n=2294)	25	0.82	0.97	0.59	0.55
Yer2 (n=4711)	25	0.73	0.89	0.52	0.58
Yer3 (n=3722)	25	0.74	0.89	0.53	0.58
Test1 (n=10727)	13	0.73	1.08	0.54	0.62
Test2 (n=10727)	12	0.73	0.73	0.54	0.55
Kolay Maddeler (n=10727)	12	1.44	0.56	0.94	0.35
Zor Maddeler (n=10727)	13	0.08	0.64	0.16	0.49

Her grup için Rasch Modeli ve GADM ile b parametreleri ve yetenek kestirimleri elde edilmiştir. Parametreler arasındaki benzerlik düzeyi için Pearson korelasyon katsayısı kullanılmıştır. Verilerin analizinde BILOG-MG, HLM5 ve SPSS15.0 programlarından yararlanılmıştır.

BULGULAR

GADM ile kestirilen madde güçlük ve yetenek parametrelerinin değişmezliği

OKS-Türkçe ve Sosyal Bilgiler alt test maddelerinin GADM yöntemi ile kestirilen madde güçlüğü değerlerinin farklı örneklemlerdeki (kız-erkek ve yer1, yer2, yer3) kestirimleri arasındaki ilişki Tablo 2’de gösterilmiştir.

Tablo 2: GADM ile kestirilen madde güçlükleri arasındaki korelasyonlar

Alt testler	Kız-Erkek	Yer 1-Yer2	Yer 1-Yer3	Yer2-Yer 3
Türkçe	0.993	0.992	0.994	0.998
Sosyal Bilgiler	0.970	0.991	0.990	0.999

Türkçe ve Sosyal Bilgiler alt test maddelerinin GADM yöntemi ile farklı örneklemlerden kestirilen madde güçlükleri arasındaki korelasyonların 0.970 ile 0.999 ($p < 0.001$) arasında değiştiği bulunmuştur. Farklı alt örneklemlerden kestirilen madde güçlükleri arasında pozitif yönlü ve korelasyon miktarı bakımından kuvvetli düzeyde ilişkilerin olduğu söylenebilir.

Türkçe ve Sosyal Bilgiler alt testleri için madde güçlük ortalamaları eşit olan (Test1 ve Test2) ve zor ve kolay maddelerden oluşan testlerden GADM ile elde edilen yetenek kestirimleri arasındaki korelasyonlar Tablo 3’te verilmiştir.

Tablo 3: GADM ile kestirilen yetenek parametreleri arasındaki korelasyonlar

Alt testler	Test1-Test2	Kolay-Zor
Türkçe	1.00	1.00
Sosyal Bilgiler	1.00	1.00

GADM ile madde güçlük ortalamaları eşit olan testlerden ve kolay ve zor maddelerden oluşturulan testlerden elde edilen yetenek kestirimleri arasındaki ilişki 1.00 bulunmuştur. GADM ile aynı bireylerin, madde güçlükleri benzer ve farklı olan maddelere göre kestirilen yetenek parametrelerinde değişme olmamıştır.

Rasch Modeli ile kestirilen madde güçlük ve yetenek parametrelerinin değişmezliği

OKS-Türkçe ve Sosyal Bilgiler alt test maddelerinin Rasch modeli ile kestirilen madde güçlüğü değerleri arasındaki ilişki Tablo 4’de gösterilmiştir.

Tablo 4: Rasch yöntemi ile kestirilen madde güçlükleri arasındaki korelasyonlar

Alt testler	Kız-Erkek	Yer 1-Yer2	Yer 1-Yer3	Yer2-Yer 3
Türkçe	0.993	0.992	0.994	0.998
Sosyal Bilgiler	0.970	0.990	0.990	0.997

Türkçe ve Sosyal Bilgiler alt test maddelerinin Rasch yöntemi ile farklı örneklemelerden kestirilen madde güçlükleri arasındaki korelasyonların 0.970 ile 0.999 ($p<0.001$) arasında değiştiği görülmektedir. Farklı örneklemelerden kestirilen madde güçlükleri arasında pozitif yönlü ve yüksek ilişkilerin olduğu görülmektedir.

Türkçe ve Sosyal Bilgiler alt testleri için madde güçlük ortalamaları eşit olan ve (Test1 ve Test2) zor ve kolay maddelerden oluşan testlerden Rasch Modeli ile elde edilen yetenek kestirimleri arasındaki korelasyonlar Tablo 5’te verilmiştir.

Tablo 5: Rasch Modeli ile kestirilen yetenek parametreleri arasındaki korelasyonlar

Alt testler	Test1-Test2	Kolay-Zor
Türkçe	0.747	0.621
Sosyal Bilgiler	0.748	0.748

Rasch Modeli ile kestirilen yetenek parametreleri arasındaki korelasyonlar 0.621 ile 0.748 arasında değişmektedir ($p<0.01$). Madde güçlük düzeyleri eşit olan testler için her iki alt testte ait değişmezlik yüksek bulunmuştur. Kolay ve zor maddelerden oluşan testlerin değişmezliği Türkçe alt testi için orta, Sosyal Bilgiler alt testi için yüksek düzeydedir.

GADM ve Rasch Model ile kestirilen madde güçlük ve yetenek parametrelerinin değişmezlikleri arasındaki ilişki ne düzeydedir?

Farklı örneklem büyüklüklerinde GADM ve Rasch Modeli ile kestirilen Türkçe ve Sosyal Bilgiler alt test maddelerinin madde güçlükleri arasındaki ilişki Tablo 6’da gösterilmiştir.

Tablo 6: GADM ve Rasch Modeli ile kestirilen madde güçlükleri arasındaki ilişkiler

Alt testler	Kız	Erkek	Yer 1	Yer2	Yer 3	Tüm grup
Türkçe	1.00	1.00	1.00	1.00	1.00	1.00
Sosyal Bilgiler	1.00	1.00	1.00	1.00	1.00	1.00

GADM ve Rasch modelleri ile farklı örneklemler üzerinde kestirilen madde güçlük parametreleri arasındaki korelasyon katsayılarının tümü Türkçe ve Sosyal Bilgiler alt test maddeleri için 1.00 bulunmuştur. GADM ve Rasch Modeliyle hesaplanan madde güçlük parametreleri arasında tam bir doğrusal ilişkinin olduğu söylenebilir.

GADM ve Rasch Model ile kestirilen yetenek parametrelerinin değişmezlikleri arasında ilişki ise Tablo 7’de gösterilmiştir.

Tablo 7: GADM ve Rasch Modeli ile kestirilen yetenek parametreleri arasındaki ilişkiler

Alt testler	Test1	Test2	Kolay	Zor	Tüm Grup
Türkçe	0.868	0.872	0.842	0.940	0.999
Sosyal Bilgiler	0.934	0.932	0.915	0.944	0.999

GADM ve Rasch modeli ile madde güçlük ortalamaları eşit ve farklı olan testlere göre elde edilen yetenek kestirimleri arasındaki ilişkiler 0.842 ile 0.944 arasında değişmektedir ($p<0.01$). Tablo 7’deki korelasyon katsayıları incelendiğinde Türkçe alt testindeki yetenek parametrelerinin değişmezliğinin Sosyal Bilgiler alt testine göre daha düşük olduğu görülmektedir.

SONUÇ VE ÖNERİLER

Sonuçlar

Bu çalışmada elde edilen bulgular, GADM ve Rasch Modeli ile yapılan kestirimlerin genel olarak madde ve yetenek parametrelerinin değişmezliğini sağladığını göstermiştir. Bu durum Türkçe ve Sosyal Bilgiler alt testlerine göre farklılık göstermemiştir.

İki modelde de cinsiyet ve yerleşim yeri değişkenlerine göre elde edilen madde güçlükleri arasında elde edilen korelasyonların en küçük değeri 0.97’dir. Korelasyon değerlerinin çoğu 1.00’e yaklaşmaktadır. Bu durum, iki modelin de madde güçlük parametresinin değişmezliğini sağladığını göstermektedir.

GADM ile madde güçlükleri eşit olan testlerden ve madde güçlük ortalamaları farklı olan testlerden elde edilen yetenek parametreleri arasındaki korelasyon 1.00’dir. Rasch modeli ile aynı testler için elde edilen yetenek kestirimleri arasındaki korelasyonlar ise 0.621 ile 0.747 arasında değişmektedir.

İki model ile elde edilen yetenek parametreleri arasındaki korelasyonlar madde güçlükleri eşit olan testlerde 0.868 ile 0.934 arasında değişmektedir. Kolay ve zor

maddelerden elde edilen yetenek parametreleri arasındaki korelasyon katsayıları ise 0.842 ile 0.944 arasındadır.

Araştırmanın bulguları, GADM ve Rasch modelin parametre değişmezliğini yüksek düzeyde sağladığını göstermektedir. GADM yetenek parametrelerinin değişmezliğini tam olarak sağlarken, Rasch Modeli Türkçe testinde kolay-zor maddeler arasında orta düzeyde, diğer kestirimlerde yüksek düzeyde sağlamıştır. Rasch modeli ile elde edilen yetenek kestirimleri arasında madde güçlük ortalamaları eşit olan testlerde Türkçe ve Sosyal Bilgiler alt testleri için birbirine yakın korelasyonlar elde edilmiştir. Kolay-zor maddeler için Sosyal Bilgiler alt testi için elde edilen yetenek kestirimleri arasındaki korelasyon Türkçe alt testinden daha yüksektir.

Sonuç olarak, GADM ile madde ve yetenek parametrelerinin değişmezliği çok yüksek düzeyde sağlanırken; Rasch Modeli ile madde parametrelerinin değişmezliği çok yüksek, yetenek parametrelerinin değişmezliği yüksek düzeyde sağlanmıştır. Bu araştırmanın sonuçlarına göre GADM ve Rasch Model, parametre değişmezliği için uygulanabilecek modeller olarak ele alınabilir.

Öneriler

Bu çalışmada model denklemlerinin birbirine eşit olduğu gösterilmiş olan GADM ve Rasch Modeli (Kamata, 2002) ile elde edilen madde ve yetenek parametrelerinin değişmezlikleri incelenmiş ve iki modelin de büyük ölçüde parametre değişmezliğini sağladığı görülmüştür. Parametre değişmezliğinin sağlanması, değişen madde fonksiyonunu belirleme çalışmalarında ve farklı gruplardan elde edilen test sonuçlarını karşılaştırmada büyük önem taşımaktadır. Bu nedenle bu iki modelin değişen madde fonksiyonunu belirleme konusundaki benzerlik düzeyleri incelenebilir.

Bu çalışmada, GADM ile MTK'ya dayalı modellerden Rasch modeli karşılaştırılmıştır. Parametre değişmezliği çalışması, yetenek parametreleri üzerinde iki ve üç parametrelili modelleri de içine alacak şekilde genişletilebilir.

Araştırmada kullanılan veriler, ikili kategorili madde puanlarıdır. Benzer bir karşılaştırma madde puanlarının çoklu puanlandığı testlerle yapılabilir. MTK'ya dayalı çok kategorili puanlama modelleri ile GADM ile elde edilen parametrelerin değişmezliği karşılaştırılabilir.

KAYNAKLAR

- Arnold, C. (1992). An introduction to hierarchical linear models. *Measurement and Evaluation in Counseling and Development*, 25(2), 58-90.
- Breithaupt, K. ve Zumbo, B. D. (2002). Sample Invariance of the Structural Equation Model and the Item Response Model: A Case Study. *Structural Equation Modeling*, 9(3), 390-412.
- Bryk, A. S. ve Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change, *Psychological Bulletin*, 101(1), 147-158.
- Chaimongkol S., Huffer , F.W. ve Kamata, A. (2007). An explanatory differential item functioning (DIF) model by the WinBug 1.4. *Songklanakarın Journal of. Science Technology*, 29(2), 449-458.
- Embretson, S. E. ve Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates, Publishers, New Jersey.
- Gokiert, R.J. ve K. L. Ricker (2004). Gender differential item functioning on the WISC-II: Analysis of the Canadian standardization sample. *Centre for Research in Applied Measurement and Evaluation* (http://www.education.ualberta.ca/educ/psych/crame/files/Gender_DIF_WISCIIL.pdf).
- Hambleton, R. K., Swaminathan, H. ve Rogers, J. H. (1991). *Fundamentals of item response theory*, Sage Publications, Boston.
- Kamata, A. (1998). *Some generalizations of the Rasch model: An application of the hierarchical generalized linear model*. Yayınlanmamış doktora tezi, Michigan State University.
- Kamata, A. (2002). Procedure to perform item response analysis by hierarchical generalized linear model. *Paper presented at the annual meeting of the American Educational Research Association*, April, 2002; New Orleans.
- Kamata, A., Chaimongkol S, Genc, S. ve Bilir, E. (2005). Random-Effect Differential Item Functioning Across Group Unites by the Hierarchical Generalized Linear Model. *Paper presented at the annual meeting of the American Educational Research Association*, Montreal, Canada.
- Koopmans, M. (1998). Application of HLM to the evaluation of the effects of basic skills programs on math achievement. *Paper presented at the annual meeting of the American Educational Research Association*, SanDiego, CA, April 13-17.

- Kim, W. (2003). *Development of a differential item functioning procedure using the hierarchical generalized linear model: A comparison study with logistic regression procedure*. Yayınlanmamış doktora tezi, The Pennsylvania State University.
- Luppescu, S. (2002). DIF detection in HLM. *Paper presented at the Annual Meeting of the American Educational Research Association*, New Orleans.
- Lord, F. M. ve Novick, M. R.(1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.
- Reise, S. P., Ainsworth, A. T. ve Haviland, M. G. (2005). Item response theory: Fundamentals, applications, and promise in psychological research. *Current Directions In Psychological Science*, 14(2), 95-101.
- Subedi, R. (2005). *A demonstration of the three-level hierarchical generalized linear model applied to educational research*. Yayınlanmamış doktora tezi, The Florida State University.
- Tate, R. (2004). Interpreting hierarchical linear and hierarchical generalized linear models with slopes as outcomes. *The Journal of Experimental Education*, 73, 71-95.
- TÜİK (2001). [http://www.tuik.gov.tr/BolgeselIstatistik/ tabloOlustur.do](http://www.tuik.gov.tr/BolgeselIstatistik/tabloOlustur.do). (Erişim: Haziran 2007).
- Vaughn, K. V. (2006). *A hierarchical generalized linear model of random differential item functioning for polytomous items: A bayesian Multilevel approach*. Yayınlanmamış doktora tezi, The Florida State University.
- Williams, N. J. (2003), *Item and person parameter estimation using hierarchical generalized linear models and polytomous item response theory models*. Yayınlanmamış doktora tezi, Texas University, Austin.
- Willms, J.D. (1999). Basic Concepts in Hierarchical Linear Modeling with Applications for Policy Analysis. In G. Cizek (Ed.), *Handbook of educational policy*. San Diego, CA: Academic Press. (<http://www.unb.ca/crisp/pdf/9806land.pdf>: Erişim 10.02.2008)