

DINA Model Parametreleri Kullanılarak Tahminlenen Madde Ayrıcılık İndekslerinin İncelenmesi

The Analyzing Item Discrimination Index Estimated by Using DINA Model Parameters

Tahsin Ođuz BAŐOKÇU*

Ege Üniversitesi

Öz

Bu arařtırmada, DINA model parametreleri kullanılarak elde edilen madde ayrıcılık indeksinin evrenden çekilen aynı büyüklükteki alt örneklemelerden ve örneklem büyüklüğünden nasıl etkilendiđi incelenmiřtir. DINA model, testi alan bireylerin bir maddeyi dođru yanıtlamak için gerekli olan niteliklere sahip olup olmadıklarını belirleyen bir örtük sınıf analizidir. Model, öđrenci yeteneđini test toplam puanından bađımsız olarak kategorik bir deđişken řeklinde hesaplamakta ve öđrencinin ait olduđu örtük sınıfı belirlemektedir. Bu nedenle modele iliřkin madde ayrıcılık parametresi geleneksel madde analizi yöntemlerinden farklı bir nitelik tařımaktadır. Arařtırmanın amacı, DINA modelin pratik kullanımlarını incelemektir. Bu amaca ulařmak için DINA modelin yapısı göz önüne alınarak geliřtirilen bir testte, modele iliřkin parametrelerle hesaplanan madde ayrıcılık indeksinin örneklem büyüklüğü ve deđişikliğinden nasıl etkilendiđi ve geleneksel yöntemler ve DINA modele ait sonuçlar arasındaki farklılıkların ve ortak noktaların ne olduđu ortaya koyulmaya çalıřılmıřtır.

Anahtar Sözcükler: Biliřsel Tanı Modelleri, DINA model, test geliřtirme, madde ayrıcılık indeksi.

Abstract

In this study, how to interpret the outcomes when using the DINA model parameters, which is one of the Cognitive Diagnostic Models for adaptation of the test development process, is studied. In this sense; DINA model which determines if the respondents have enough qualifications to answer the item correctly, provides to analyze the structure of the item difficulty parameter which is determining in test development process. The aim of this study is to explore the differences and common points of discrimination index of items in a developed test which are analyzed by traditional methods and the DINA model.

Keywords: Cognitive Diagnosis Models, DINA Model, Test development process, discrimination index

Summary

Purpose

The aim of the research is to examine the practical use of DINA model which is one of the BTM. Therefore, how to interpreted the results of DINA model when it is used in test development process studied in this research while focusing on the process of test related features besides determining the capabilities of the students, To realize this aim, the DINA model parameters compared with the item test correlation index which is used in traditional methods. As a result of these comparisons, it is learnt about the practical usage of the theories, by determining the similarities and differences between the theories.

* Öđ. Gör. Dr. T. Ođuz BAŐOKÇU, Ege Üniversitesi, Eđitim Fakültesi, Eđitimde Ölçme ve Deđerlendirme Anabilim Dalı. İzmir, tahsin.oguz.basokcu@ege.edu.tr

Results

In the study, two models for the material differences discrimination index are examined by using real application data. In the real application data's the sample size $N=403$, the number of attributes for DINA model $K=8$ and test length $J=40$. For comparing the methods, the real application data sample taken and analyzed in two different ways. Firstly, 5 sub-groups $N=100$ were taken from real application data. Then, 4 sub-group with different sample size $N=200$, $N=100$, $N=60$ and $N=20$ were randomly taken from data. At last the r_c and δ_j values for full data and sub-groups median's were compared. The Spearman's rho coefficient was calculated for r_{jx} values with full data and 5 sub-groups with $N=100$ was .84 and the same coefficient for DINA's δ_j values was .95. In the scope of this study, the lower level of discrimination index was accepted as 20 for r_c values. The r_c value calculated with bi-serial correlation technique for sub-groups and full data shows difference only in one item for the decision of the item discrimination. In DINA model, when the lowest level of discrimination index was accepted as .20, all the data and the sub-groups were coherent with each other.

According to the results; the incoherence of r values increase as the number of observation decrease. Especially in the samples which has low "n" number of 60 or 20 person, there has been many changes in the decision of item discrimination. For DINA model, it gives more coherent results to the change of the "n" number.

Conclusion

When the research results are analysed, it can be said that, the parameters the DINA model determines change less both against the change of the samples and the decrease in the sample size. DINA model produces more coherent values comparing to the traditional item discrimination for both situations.

Giriş

Madde analizi süreci, test geliştirme basamaklarının karar aşamasındaki en önemli uygulamalarından biridir. Madde analizi, madde istatistiklerinin hesaplanması, doğrudan teste konulabilecek maddelerin belirlenmesi ve bu maddeler üzerinde yapılabilecek düzeltme çalışmalarının ne doğrultuda olacağını saptanması, teste konulması mümkün olmayan maddelerin ayıklanması amacıyla yapılır (Turgut, Baykul, 2010).

DINA model, Bilişsel Tanı Modelleri (BTM) içinde en yaygın kullanılan analiz yöntemlerinin başında gelmektedir. BTM, testi alan bireylerin yeteneklerini maddeleri doğru cevaplamak için gerekli olan özelliklere sahip olup olmama durumları bakımından kategorik olarak belirleyen modellerdir. BTM geleneksel yöntemlerden farklı parametreler hesaplayarak hem testi alan bireylerin maddeleri doğru cevaplamak için gerekli özelliklere sahip olup olmadığını belirlemekte hem de madde güçlüğü oluşturulan koşulları daha açık ortaya koymaktadır.

DINA Model

BTM alanında geliştirilen birçok model bulunmaktadır. Bu modellerin çoğu pratikte kullanım alanı olmasa da süreç içinde BTM'nin gelişim basamaklarını oluşturmaktadır. Sympson (1978) tarafından geliştirilen Denkleştirici olmayan Çok Boyutlu MTK Modeli (Noncompensatory MIRT model), Whitely(1980) ve Embretson (1997) tarafından geliştirilen Çok Bileşenli Örtük Özellik Modeli (Multicomponent latent trait model), Embretson (1985, 1997) tarafından geliştirilen Genel Bileşenli Örtük Özellik Modeli (General component latent trait model) bu modellerden birkaçıdır.

BTM içinde çok boyutlu örtük sınıf modelleri bir dönem araştırmacılar tarafından yoğun olarak çalışılmıştır. Özellikle Maris (1999) farklı özellikleri olan birkaç çok kategorili örtük sınıf modeli önermiştir. Ayırıcı Çok Kategorili Örtük Sınıf Modeli (Disjunctive MCLCM) ve Denkleştirici Çok Kategorili Örtük Sınıf Modeli (Compensatory MCLCM) bu modeller

arasındadır. Bunun yanında von Davier (2006), von Davier (2005) ve von Davier ve Yamamoto (2004) tarafından da yeteneğin sürekli kabul edildiği denkleştirici MTK modellerine benzer ama yetenek parametresinin kategorik olarak alındığı denkleştirici çok kategorili örtük sınıf modelleri geliştirmiştir.

BTM içinde daha yaygın olarak araştırılan, pratik uygulamalarda kullanılan ve yazılım desteği olan modeller de bulunmaktadır. Bu modellerden en önemlileri, Tatsuoka ve Tatsuoka'nın (1982) Rule Space Modeli, Hartz (2002) tarafından geliştirilen Reparametrize Birleşik Model, Gitomer ve Yamamoto'nun (1991) HYBRID modeli ve Haertel'in (1984, 1990) Sınırlandırılmış Örtük Sınıf Modeli'dir.

Haertel (1989) tarafından geliştirilen DINA model, ikili yetenek modellerine benzeyen bir örtük sınıf analizidir. Bu anlamda DINA model, Madde Tepki Kuramı'nın (MTK) bir uzantısıdır. Bununla birlikte DINA model, MTK'dan ayrı olarak öğrencilerin farklı büyüklükte sürekli biçimde dağılmış yetenekleri olduğunu varsaymamakta, daha çok öğrencileri kesin olarak belirlenmiş az sayıda farklı örtük sınıflara ayırmaktadır (Haertel, 1989).

DINA model kısaca şu şekilde tanımlanabilir: X_{ij} nin i cevaplayıcısının j maddesine verdiği yanıt olduğu farz edilirse, $i=1, \dots, I$ ve $j=1, \dots, J$ olur. Cevaplayıcının ikili özellik vektörü $\alpha_i = \{\alpha_{ik}\}$ şeklinde gösterilirse, $k=1, \dots, K$ için cevaplayıcının k inci elemanının 1 olduğu durum k özelliğine sahip olduğunu, 0 olduğu durum ise k özelliğine sahip olmadığını göstermektedir (de la Torre 2009a). Özellik olarak adlandırılan terim genel olarak nitelik, beceri, yetenek, bilginin sunumu, bilişsel süreç olarak tanımlanabilir (Tatsuoka, 1995). Çoğu BTM cevaplama özelliğinin belirlendiği $J \times K$ şeklinde oluşturulan ve 1-0 şeklinde kodlanan bir Q matrisi üzerinden hesaplama yapar (Embretson, 1984; Tatsuoka, 1985). Q matrisinde sütunlar özellikleri ve satırlar maddeleri gösterir. Matriste q_{jk} hücresi j maddesini doğru cevaplamak için k özelliğine sahip olunması gerekip gerekmediğini işaret eder.

DINA model cevaplayıcıları her özellik için iki boyutta sınıflar. Birinci sınıf "yokluk sınıfı" yani belirlenen özelliğe sahip olmayanların oluşturduğu sınıf, diğeri ise "tam sınıf" yani belirlenen özelliğe sahip olanların oluşturduğu sınıftır. Görüldüğü gibi DINA model özellik parametresini sürekli değil, kategorik bir değişken olarak tanımlar. DINA modelde Q matrisinde tanımlanan k tane özellik için 2^k tane örtük sınıf mevcuttur. Modelde her özellik için cevaplayıcının o özelliğe sahip olup olmama durumu hesaplanır. Bu şekilde cevaplayıcının özelliklerin toplamının oluşturduğu 2^k örtük sınıftan hangisine dahil olduğunu belirleyen Alfa (α) parametresi belirlenir. Örneğin α_1, α_2 ve α_3 olarak belirlen 3 özelliikle temsil edilen 4 maddelik bir teste ilişkin Q matrisi aşağıdaki şekilde belirlenmiş olsun:

Tablo 1.

Örnek Q Matris

Maddeler	α_1	α_2	α_3
1	1	0	0
2	1	1	0
3	1	1	1
4	0	1	1

Yukarıdaki Q matris, birinci maddenin doğru cevaplanması için sadece α_1 'e sahip olmanın yeterli olduğunu, ikinci maddenin doğru cevaplanması için ise hem α_1 hem de α_2 'ye sahip olunması gerektiğini göstermektedir. Örnek Q matrisine göre yapılan analiz sonucunda cevaplayıcılar 2^3 örtük sınıfa yerleştirilirler. Bu örtük sınıflar (000), (100), (010), (001), (110), (101), (011) ve (111) şeklinde belirlenir. Örtük sınıflar öğrencinin hangi özelliklere sahip olup hangilerine sahip olmadıklarını tam olarak ifade eder. Yukarıdaki örnekte belirlenen örtük sınıflardan (000)'a dahil olan öğrenci hiçbir özelliğe sahip değildir. Benzer şekilde (100) örtük sınıfında sadece α_1 'e sahip olan öğrenciler yer alırken, (011) sınıfında ise α_1 'e sahip olmayıp α_2 ve α_3 'e sahip olan öğrenciler

bulunmaktadır. Bu durum BTM'nin tanı işlevini açıklamaktadır. BTM ile geliştirilmiş ve analiz edilmiş bir testte, bir toplam puan ya da toplam alt ölçek puanları yerine, sınavı alan her bir bireyin sınavın kapsamındaki becerilerden hangilerine sahip olduğu ve hangileri konusunda eksiklikleri bulunduğu belirlendiği bir ölçme yapılır. Bu anlamda Bilişsel Tanı Modeli'yle geliştirilmiş testler sadece değerlendirme sürecine değil aynı zamanda her bir öğrencinin eğitim ihtiyacını belirleme konusuna da hizmet eder (Cheng Y., Chang H.2007).

DINA modelin diğer BTM'den en önemli farkı ise bir madde için gerekli olan özelliklerden sadece birine bile sahip olmayan cevaplayıcıyı yokluk sınıfında görmesidir. Başka bir deyişle sadece bir maddeyi doğru cevaplamak için gerekli özelliklerin tamamına sahip olan cevaplayıcının maddeyi doğru cevaplama olasılığı 1'dir. Aşağıda bütün özelliklere sahip olan bireyin maddeyi doğru cevaplama olasılığının fonksiyonu verilmiştir:

$$P [Y_{ij} = 1 | \eta_{ij}, s_j, g_j] = (1 - s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}}$$

P , aranan bütün becerilere sahip olan öğrencinin maddeyi doğru cevaplama olasılığıdır. η_{ij} , α tarafından belirlenen örtük cevaplama ve i 'inci konunun niteliği ve q_j 'nin vektörüdür. Q matrisinin j 'inci maddesine tekabül eden sıra şu şekilde gösterilebilir:

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}$$

DINA modelin diğer BTM'lere göre en önemli avantajı, gerek uygulama gerekse yorumlama süreçlerinin daha az karmaşık olmasıdır. Bununla birlikte DINA model basitliğine rağmen yüksek düzeyde model-data uyumu verdiği ve bazı modifikasyonlarla modelin farklı stratejilere kolaylıkla uyum sağladığı, de la Torre ve Douglas (2004, 2008) tarafından göstermiştir. Aynı zamanda de la Torre çalışmalarında daha yeni ve daha karmaşık olan BTM'lerle DINA model sonuçları arasında yüksek düzeyde uyum olduğunu belirtmiştir (de la Torre, 2008b, 2009b; de la Torre & Liu, 2008). Bu gerekçeler genel olarak BTM arasında DINA modelin öneminin bir göstergesi olduğu gibi bu çalışmada DINA modelin seçilmesinin nedenini oluşturmaktadır.

DINA Model Parametreleri

DINA model cevaplayıcıların sahip olduğu özellikleri belirlerken aynı zamanda tahmin (guess) g ve kaydırma (slip) s parametreleri olarak adlandırılan iki parametre daha hesaplamaktadır.

$$s_j = P [Y_{ij} = 0 | \eta_{ij} = 1] s_j = P [Y_{ij} = 0 | \eta_{ij} = 1] \text{ ve} \\ g_j = P [Y_{ij} = 1 | \eta_{ij} = 0] g_j = P [Y_{ij} = 1 | \eta_{ij} = 0],$$

s_j , örtük özelliğe sahip bireyin j maddesine yanlış cevap verme olasılığını gösteren durumu (yanlış pozitif olasılık) ve g_j ise örtük sınıfa sahip olmayan bireyin doğru cevap verme olasılığı durumunu (doğru pozitif olasılık) ifade eder. s_j parametresi "slip = kaydırma" anlamına gelir ve bu parametre ne kadar düşük olursa, aranan özelliklere sahip bireylerin doğru cevap verme olasılığı o kadar artar.

g_j parametresine "guessing" tahmin parametresi de denir. Maris (1999) alternatif olarak g_j parametresini, zihinsel tahmin yeteneğini başarıyla kullanmak olarak da açıklar. Görüldüğü gibi DINA model g parametresi MTK 3 parametrelilikte görülen c "şans" parametresinden biraz farklıdır. DINA modelde hesaplanan g sadece bir maddeyi tahminle doğru cevaplamak değil, aynı zamanda Q matris tarafından tanımlanamamış bir özelliği kullanarak maddeyi doğru cevaplamak anlamına da gelmektedir.

DINA modelde s ve g parametreleri madde düzeyinde ortaya çıkmaktadır. Her madde popülasyonu iki sınıfa böler ve aynı sınıfa düşen öğrencilerin o maddeye doğru cevap verme olasılıkları eşittir. Öğrencilerin sınav performansı, testte yoklanan niteliğin vektörünün tam bir göstergesi değildir. Bu nedenle olasılık temelinde bir model ancak s ve g 'nin olasılığını görmeye

izin verir. “s” durumu, öğrencinin maddede aranan niteliğe sahip olmasına rağmen alt görevi doğru cevaplayamadığı ya da maddeyi doğru cevaplayamadığında gerçekleşir. “g” ise öğrencinin maddede aranan özelliklerden birine ya da birkaçına sahip olmamasına rağmen alt görevleri tamamlaması veya maddeyi doğru cevaplama durumu (de la Torre & Douglas, 2008).

DINA model öğrencilerin maddeyi doğru cevaplama için gerekli becerilere sahip olup olmama durumlarını belirlemek noktasında oldukça başarılıdır. Bununla birlikte belirlediği s ve g parametreleri yoluyla madde güçlüğü ve ayırt ediciliği şeklinde yorumlanabilecek bazı indekslerin hesaplanmasına da olanak sağlamaktadır. De la Torre (2008a) maddenin kalitesini belirleyen ayırt edicilik indeksini belirlemek amacıyla hem s hem de g parametrelerinin birlikte kullanıldığı bir hesaplama yöntemi önermiştir. De la Torre’ye göre δ_j (delta), ayırt edicilik indeksi olarak $1 - s_j - g_j$ formülüyle hesaplanabilmektedir. Örtük özelliğe göre öğrencileri mükemmel düzeyde ayırt eden bir madde için $\delta_j = 1$ olmalıdır ve bu değer “sıfır” a yaklaştıkça maddenin ayırt edicilik gücü düşmektedir. Modelde s ve g parametreleri 0 ile 1 arasında değer alabilmektedir. Bu nedenle madde ayırtıcılığını ifade eden δ_j katsayısı $-1 \geq \delta_j \leq 1$ arasında değerler almaktadır.

DINA modelde hesaplanan δ_j parametresi konusunda de la Torre (2008a) tam bir ölçütün olmadığını belirtmiştir. Daha yüksek δ_j değeri daha ayırıcı bir maddeyi işaret etmektedir. Bununla birlikte modelin yeni oluşu ve bu konuda yapılan çalışmaların sınırlı sayıda olması, henüz böyle bir ölçütün belirlenememesi durumunu doğurmuştur.

Araştırmanın Amacı

Araştırma, BTM’den biri olan DINA modelin pratik kullanımını incelemek amacı taşımaktadır. DINA model kullanılarak yapılan araştırmalar genel olarak simülasyon düzeyinde kalmaktadır. Bu çalışmada gerçek uygulama verisi kullanılarak modelin gerçek koşullarda nasıl sonuçlar ürettiği incelenmiştir. Araştırmada evrenden çekilen aynı büyüklükteki alt örneklerde ve örneklem büyüklüğü değiştiğinde DINA modele ilişkin madde ayırtıcılık indeksindeki farklılaşmaların betimlenmesi amaçlanmıştır. Evrenden çekilen farklı örneklerde madde ayırtıcılık indeksindeki değişimi gözlemek amacıyla gerçek uygulama verisine ait %25’lik alt örnekler belirlenmiştir. 403 kişilik gerçek uygulama verisinden 100 kişilik 5 alt örneğe ait DINA modele ilişkin madde ayırtıcılık indeksi (δ_j) hesaplanmıştır. Hesaplanan bu değerler incelenerek tüm veri için hesaplanan δ_j değerleri ile aynı büyüklükteki alt örnekler için hesaplanan δ_j değerlerinin değişimi belirlemeye çalışılmıştır. Bununla birlikte araştırmada örneklem büyüklüğü değiştiğinde de δ_j değerindeki farklılaşmalar belirlenmeye çalışılmıştır. Aynı zamanda her bir örneklem için hesaplanan madde test korelasyon değerleriyle DINA modelle belirlenen δ_j parametresinin farklılıkları ve bezerlikleri de araştırılmıştır. DINA modelde δ_j parametresi modelin ürettiği iki parametre olan g ve s parametrelerine dayandığından, δ_j parametresinin alt örneklerden ve örneklem büyüklüğünden etkilenme düzeyi g ve s parametrelerinin de bu değişkenlerden etkilenip etkilenmediğinin bir göstergesi olacaktır. Bu nedenle araştırmada tek bir parametre incelenerek modelin bütün parametrelerinin alt örneklem ve örneklem büyüklüğünden nasıl etkilendiği incelenmeye çalışılmıştır. Çalışmada, DINA modelin öğrencilerin yeteneklerinin belirlenmesi amacıyla kullanımı dışında, test geliştirme ve teste ilişkin özelliklerin tanımlaması sürecinde kullanılması durumunda sonuçların nasıl yorumlanabileceği araştırılmıştır. Araştırmada bu amaçlara ulaşmak için aşağıdaki soruya yanıt aranmıştır:

DINA modele göre belirlenen madde ayırtıcılık gücü parametresi, aynı evrenden çekilen farklı örneklerden ve örneklem büyüklüğünden ne derece etkilenmektedir?

Yöntem

Araştırmanın Türü

Araştırma, test geliştirme sürecinde kullanılan madde ayırtıcılık gücünün madde test korelasyonu ve DINA modelle elde edilen sonuçların incelenmesine dayandığından betimsel

bir araştırmadır. Bu incelemeler sonucunda kuramların benzerlikleri ve farklılıklarına ilişkin durumlar saptanarak kuramların pratik kullanımlarına ilişkin bilgiler elde edilmiştir.

Çalışma Grubu

Araştırmada analizler gerçek veriler kullanılarak gerçekleştirilmiştir. Gerçek verilerin toplandığı grup Ege Üniversitesi Eğitim, Edebiyat ve Fen Fakültesi bünyesinde okuyan 403 öğrenciden oluşmuştur.

Ölçme Aracı

Araştırmada gerçek veriyi elde etmek için kullanılan ölçme aracı “Eğitimde Ölçme ve Değerlendirme” dersi için hazırlanan 40 maddeden oluşmaktadır. Ölçme aracında yer alan 40 madde, “ölçmenin temel kavramları”, “ölçme aracında bulunması gereken nitelikler” ve “değerlendirme ve değerlendirme türleri” konularını kapsamaktadır. Ölçme aracına ilişkin betimsel istatistikler Tablo 2’de verilmiştir.

Tablo 2.

Gerçek Uygulama Verisi Betimsel İstatistikler

N	403
Ortalama	23,71
Standart Hata	.26
Ortanca	24
Mod	21
Standart Sapma	5,28
Basıklık	-0,24
Çarpıklık	.18
Ranj	28
En Büyük	11
En Küçük	39
Güvenirlilik (Alfa)	.85

Verilerin Analizi

Araştırmada klasik test kuramı için madde parametrelerinin belirlenmesinde ITEMAN paket programı kullanılmıştır. DINA model parametrelerinin kestirimi için ise OX EDIT programı altında çalışan kodlardan yararlanılmıştır. Baykul (2000), maddelerin 1 ve 0 olarak (iki kategorili) puanlandığı durumlarda, süreksiz yapay ikilem olan madde puanları ile sürekli olan test puanları arasında madde ayırt edicilik indeksi olarak Pearson momentler çarpımı korelasyon katsayısının özel hali olan nokta-çift serili ya da çift serili korelasyon katsayıları kullanılabileceğini belirtmiştir. Bu nedenle geleneksel yöntemle madde ayırtıcılıkları için çift serili korelasyon tekniği kullanılmıştır. İki veri içinde maddelerin genel olarak orta güçlük düzeyinde olmalarından dolayı madde ayırtıcılık indeksinin hesaplanmasında çift serili korelasyon katsayısının kullanılması daha uygundur görülmüştür (Baykul, 2000).

Araştırmada kullanılan ölçme aracında yer alan maddelerle özellikler arasındaki ilişkiyi tanımlayan Q matrisi, uzman görüşlerine başvurularak hazırlanmıştır. Bu süreçte öncelikle uzman görüşleri sonucunda 40 maddeyi tanımlayan 8 özellik belirlenmiştir. Bu 8 özellik konu alanı içindeki alt bölümleri temsil etmektedir. Sonraki aşamada uzmanların her maddenin doğru cevaplanması için öğrencilerin hangi özelliklere sahip olması gerektiğini belirlemeleri istenmiştir. Bu doğrultuda her uzman madde özellik ilişkisini gösteren bir Q matrisi hazırlamıştır. Uzmanların hazırladıkları Q matrislerinde karşılaştırılarak her madde için uzman uyumu göz önüne alınarak analizde kullanılacak Q matrisi belirlenmiştir. Bu süreç sonucunda belirlenen Q matrisi Tablo 3’te verilmiştir.

Tablo 3.
Gerçek Uygulama Verisine Ait Q Matrisi

Madde	Özellikler								Madde	Özellikler							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
1	0	0	1	0	0	0	0	0	21	0	0	0	0	0	0	0	1
2	1	0	0	0	0	0	0	0	22	0	0	0	0	1	0	0	0
3	0	0	0	0	0	0	0	1	23	0	1	0	0	0	1	0	0
4	0	0	1	0	0	0	0	0	24	0	0	0	0	0	1	0	0
5	1	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	1
6	1	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	1
7	1	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	1
8	0	0	0	0	1	0	0	0	28	0	0	0	0	0	1	0	0
9	0	0	0	0	0	1	0	0	29	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	1	30	0	0	0	0	0	0	0	1
11	0	0	0	1	0	0	0	0	31	0	0	0	0	0	1	0	0
12	0	1	0	0	0	1	0	0	32	0	1	0	0	0	0	0	0
13	0	0	0	0	0	1	0	0	33	0	1	0	0	0	0	1	1
14	0	0	0	1	1	0	0	0	34	0	0	0	0	0	0	1	0
15	0	0	1	0	0	0	0	0	35	0	0	0	0	0	0	1	0
16	0	0	1	0	0	0	0	0	36	1	0	0	0	0	0	0	0
17	1	0	0	0	0	0	0	0	37	0	0	0	0	0	0	1	0
18	0	0	0	0	1	0	0	0	38	0	0	0	0	0	0	1	1
19	0	0	0	0	0	0	0	1	39	0	0	0	0	0	0	1	0
20	0	0	0	1	0	0	0	0	40	0	0	0	0	0	0	0	1

Gerçek uygulama verileri için iki farklı yöntemle örneklem çekilerek analiz yapılmıştır. Öncelikle ölçme aracı uygulanan 403 kişilik grubun cevap örüntüsünden 100'er kişilik 5 farklı örneklem tam seçkisiz yöntemle çekilmiştir. İkinci uygulamada ise aynı veriden sırayla 200, 100, 60 ve 20 kişilik örneklem çekilerek madde ayırıcılık değerleri hesaplanmıştır.

Belirlenen grupların test cevaplarından yola çıkarak maddelerin ayırıcılık değerleri ITEMAN ve OX EDIT programları kullanılarak hesaplanmıştır.

Bulgular

Örneklem Değişikliğine İlişkin Bulgular: Gerçek uygulama verisinden tam (yerine koyma yöntemiyle) seçkisiz yöntemle çekilen 100 kişilik 5 alt gruba ait 40 madde için hesaplanan r_{ζ} ve δ_j değerleri Tablo 4'te verilmiştir. Yorumu kolaylaştırmak ve tabloyu daha anlaşılır kılmak amacıyla 5 alt örneklem için belirlenen katsayıların ortanca değerine tabloda yer verilmiştir.

Tablo 4.

Gerçek Uygulama Verisi ve 100 kişilik Alt Gruplara Ait Madde Ayırıcılık Değerleri

Madde	GELENEKSEL		DINA		Madde	GELENEKSEL		DINA	
	Tam Veri r_{ζ}	5 Örneklem Ortanca r_{ζ}	Tam Veri δ_j	5 Örneklem Ortanca δ_j		Tam Veri r_{ζ}	5 Örneklem Ortanca r_{ζ}	Tam Veri δ_j	5 Örneklem Ortanca δ_j
1	.40	.36	.22	.21	21	.37	.37	.53	.54
2	.42	.39	.17	.13	22	.47	.47	.80	.98
3	.29*	.26	.08*	.09	23	.36	.44	.30	.23
4	.27*	.17	-.07*	-.07	24	.47	.42	.32	.24
5	.12	.08	.09	.05	25	.29	.24	.68	.76
6	.60	.59	.75	.69	26	.14	.14	-.37	-.30
7	.28	.29	.23	.22	27	.48*	.51	-.06*	-.06
8	.39	.39	.19	.18	28	.44	.41	.39	.49
9	.36	.30	.13	.03	29	.31	.22	.25	.26
10	.26	.23	.52	.60	30	.42	.33	.22	.25
11	.33	.34	.16	.08	31	.42	.27	.33	.27
12	.36	.35	.26	.24	32	.48	.57	.31	.67
13	.48	.50	.40	.46	33	.43	.45	.34	.40
14	.21	.24	.16	.11	34	.44	.48	.26	.22
15	.45	.45	.28	.25	35	.34	.24	.45	.30
16	.53	.60	.59	.98	36	.40	.43	.25	.23
17	.28	.39	.24	.23	37	.29	.40	.30	.42
18	.25*	.27	-.02*	-.01	38	.22	.23	.45	.30
19	.42	.34	.27	.24	39	.40	.32	.30	.39
20	.38	.32	.39	.98	40	.37*	.38	.02*	.01

(*) olan değerler, DINA ve geleneksel yöntemde birbiriyle uyuşmayan maddeleri göstermektedir.

❖ Altı çizili değerler aynı model tam veriye uyuşmayan değerleri göstermektedir.

Tablo 4'te tüm veriden çekilen 5 alt örneklem için belirlenen δ_j değerlerinin ortancası ile tüm veriye ilişkin δ_j değerleri karşılaştırılmıştır. Sıra farklar korelasyonuyla iki δ_j değeri arasındaki ilişkiye bakıldığında, tüm veri ile alt gruplar arasında .95 düzeyinde bir ilişki gözlenmiştir. Aynı ilişki madde test korelasyonu yöntemi için .84 düzeyinde hesaplanmıştır. İki yöntemde de tüm veri ve alt örneklem için hesaplanan ayırıcılık değerlerinin birbiriyle benzer ve yüksek düzeyde ilişkiyi gösterdiği görülmektedir.

Çalışmada daha önce değinildiği gibi DINA model ayırıcılık parametresinin nasıl bir ölçütle sınırlandırılacağına ilişkin herhangi bir araştırma bulunmamaktadır. Bununla birlikte de la Torre (2008a) DINA model için δ_j değerinin sıfıra yakın olmasının ve (-) değer almasının maddenin yeterli ayırt ediciliğe sahip olmadığını bir göstergesi olduğunu belirtmiştir.

İki yöntemin tüm veri için hesapladığı ayırıcılıkla alt örneklem arasındaki uyuma bakıldığında, 4. madde için geleneksel yöntemde bir farklılık göze çarpmaktadır. Tablo 4'te altı çizili olarak ifade edilen bu madde için tüm veride r_{jx} .27 olarak hesaplanırken alt grup r_{jx} değerlerinin ortancası .17 değerinde kalmıştır. Geleneksel yöntemde .20 r_{jx} değeri ayırıcılık indeksinin alt sınır olarak kabul edilirse, 4. maddeye ilişkin ayırıcılık kararında tüm veri ve alt gruplar arasında farklılık olduğu görülmektedir. Maddenin ayırıcılığına ilişkin kararlar konusunda ise DINA model daha tutarlı sonuçlar vermiştir. Yine 4. maddeye bakıldığında DINA model hem tüm veri için hem de alt gruplar ortancası göz önüne alındığında maddenin düşük ayırıcılığa sahip olduğu kararını verdiği görülmektedir.

Örneklem Büyüklüğü Değişikliğine İlişkin Bulgular: DINA model δ_j parametresinin örneklem büyüklüğündeki değişiklikten nasıl etkilendiğini belirlemek amacıyla 403 kişiden oluşan gerçek uygulama verisi içinden tam (yerine koyma yöntemiyle) seçkisiz yöntemle 200, 100, 60 ve 20 kişilik gruplar çekilmiş ve bu örneklemere ilişkin r_{jx} ve δ_j değerleri hesaplanmıştır. Gruplara ilişkin değerler Tablo 5'te verilmiştir.

Tablo 5.

Gerçek Uygulama Verisi Örneklem Büyüklüğü Değişikliğine İlişkin Madde Ayırıcılık Değerleri

Madde	Geleneksel					DINA				
	Tam Veri r_c	200 r_c	100 r_c	60 r_c	20 r_c	Tam Veri δ_j	200 δ_j	100 δ_j	60 δ_j	20 δ_j
1	.40	.27	.56	.48	<u>.09</u>	.22	.22	.20	.19	.18
2	.42	.38	<u>.19</u>	.26	.33	.17	.15	.13	.73	.60
3	.29	.26	.22	.27	<u>.12</u>	.08	-.09	.09	-.08	<u>.66</u>
4	.27	<u>.18</u>	<u>.17</u>	<u>.11</u>	.44	-.07	-.02	-.04	-.07	.00
5	.12	.09	<u>.20</u>	-.34	-.09	.09	.09	.08	<u>.29</u>	<u>.20</u>
6	.60	.62	.54	.55	<u>.19</u>	.75	.86	.98	.32	.31
7	.28	.27	.39	<u>.06</u>	<u>.22</u>	.23	<u>.13</u>	<u>.11</u>	.42	.89
8	.39	.29	.53	.48	<u>.01</u>	.19	.13	.18	.99	.99
9	.36	.32	<u>.19</u>	.33	.53	.13	<u>-.23</u>	<u>.03</u>	<u>.00</u>	.15
10	.26	.35	.23	.36	<u>.16</u>	.52	.47	.30	.27	-.06
11	.33	.32	.38	.41	<u>-.21</u>	.16	.12	.23	<u>-.01</u>	.21
12	.36	.34	.39	.59	.41	.26	.18	<u>-.02</u>	.17	.33
13	.48	.51	.54	.60	.42	.40	.17	.81	.79	.38
14	.21	<u>.18</u>	.26	<u>-.04</u>	<u>-.19</u>	.16	<u>.03</u>	.15	<u>.06</u>	.18
15	.45	.49	.45	.42	.54	.28	.29	.17	.23	.25
16	.53	.47	.48	.66	.45	.59	.99	.98	.99	.99
17	.28	.36	.27	<u>.17</u>	<u>.16</u>	.24	.28	.22	.28	.20
18	.25	<u>.16</u>	.27	<u>.18</u>	.23	-.02	-.04	-.05	-.08	<u>.12</u>
19	.42	.44	.33	.64	.80	.27	.22	.20	.22	.23
20	.38	.43	.32	.62	.45	.39	.99	.98	.99	.99
21	.37	.39	.25	.38	.47	.53	.44	.48	.47	.40
22	.47	.47	.42	.41	<u>.05</u>	.80	.99	.98	.21	.26
23	.36	.46	.43	.42	.55	.30	<u>-.28</u>	.21	.21	.21
24	.47	.39	.55	<u>.18</u>	<u>-.03</u>	.32	.79	.24	.63	.50
25	.29	.37	.28	.54	.55	.68	.61	.81	.96	.23
26	.14	.03	.15	<u>.46</u>	-.20	-.37	-.27	-.27	-.14	.03
27	.48	.53	.68	.57	<u>.03</u>	-.06	-.26	.05	-.08	.07
28	.44	.48	.41	.72	<u>.11</u>	.39	.26	.49	.17	.30
29	.31	.30	.35	.32	<u>.13</u>	.25	.47	.26	.16	.29
30	.42	.28	.28	.57	<u>.16</u>	.22	<u>-.10</u>	.15	<u>-.15</u>	<u>-.06</u>
31	.42	.43	.20	.53	.72	.33	.25	.27	.14	.50
32	.48	.45	.57	.48	.26	.31	.53	.69	.71	.99
33	.43	.51	.39	.54	.48	.34	.50	.27	.51	.58
34	.44	.34	.48	<u>.08</u>	.53	.26	.27	.25	.22	.08
35	.34	.29	.43	.49	<u>-.08</u>	.45	.15	<u>.08</u>	.27	.21
36	.40	.41	.58	<u>-.16</u>	.33	.25	.20	.17	.28	.20
37	.29	.29	.45	.54	.45	.30	.33	.98	.26	.67
38	.22	.24	<u>.07</u>	<u>.18</u>	.60	.45	.42	.30	.30	.20
39	.40	.33	.26	.34	<u>.15</u>	.30	.58	<u>.02</u>	.99	.63
40	.37	.31	.33	<u>.12</u>	.33	.02	-.07	.01	-.22	<u>.17</u>

(*) olan değerler, DINA ve geleneksel yöntemde birbiriyle uyuşmayan maddeleri göstermektedir.

❖ Altı çizili değerler, tam veriye uyuşmayan değerleri göstermektedir.

İki yöntemle hesaplanan madde ayırıcılık indeksi değerlerinin örneklem büyüklüğünden etkilenme durumlarının belirlenmesi amacıyla her bir yöntemin tüm veriye ilişkin değerleriyle

alt guruplara ait değerleri arasındaki ilişki, sıra farklar korelasyonu ile incelenmiştir. İki yöntemle ilişkin sonuçlar Tablo 6’da verilmiştir.

Tablo 6.
Tam Veri ile Alt Örneklem Arasındaki Korelasyon

N	Tam Veri r_c	Tam Veri δ_j
200	0,80**	0,80**
100	0,65**	0,80**
60	0,51**	0,60**
20	0,16	0,42**

Tablo 6’da görüldüğü üzere geleneksel yöntemde ve DINA modelde tam veri ile farklı n sayısındaki alt örneklem arasındaki ilişkinin azaldığı gözlenmektedir. Bununla birlikte DINA model için tam veri ile farklı büyüklüklerdeki alt örneklemde ait δ_j değerleri arasındaki ilişkinin yüksek ve istatistiksel olarak anlamlı olduğu görülmektedir. Madde test korelasyonu yöntemine ilişkin sonuçlar incelendiğinde, sadece 20 kişilik alt örneklem için belirlenen r_{jx} değerleri ile tam veri için belirlenen değer arasındaki ilişki istatistiksel olarak anlamlı çıkmamıştır.

Tartışma

Araştırma sonuçlarına göre DINA model, hem evrenden aynı büyüklükte farklı örneklem çekildiğinde hem de örneklemdeki eleman sayısı azaldığında geleneksel yöntemde göre daha az değişen ayırıcılık değerleri hesaplamaktadır.

Araştırma sonuçlarında ortaya çıkan bir diğer bulgu ise geleneksel yöntem ve DINA model tarafından aynı maddeler için farklı ayırıcılık kararlarının verilmesidir. DINA modelde madde ayırıcılık indeksi için belirlenmiş bir alt sınır bulunmamaktadır. Fakat de la Torre (2008a) “0” değerine yakın ve (-) değer alan maddelerin yeterli ayırıcılığa sahip olmadıkları yönünde karar verilebileceğini belirtmiştir. Araştırmada uygulanan testte yer alan 40 maddenin 5 tanesi madde test korelasyonu yöntemine göre yeterli ayırıcılıkta bulunurken DINA model parametrelerine göre aynı maddeler için belirlenen ayırıcılık indeksleri “0”a çok yakın ya da (-) değer almıştır. Bununla birlikte araştırmada elde edilen sonuçlar incelendiğinde, DINA model için elde edilen δ_j değerlerinin r_{jx} değerlerinden nispeten düşük olduğu gözlenmiştir. Konuyla ilgili gelecekte yapılacak çalışmaların DINA modele ilişkin δ_j parametresinin madde geçerliği olarak yorumlanabilecek alt sınırının belirlenmesi yönünde yarar vardır.

DINA modelle hesaplanan madde ayırıcılık değeri, modelin g ve s parametrelerinin bir uzantısıdır. Madde güçlük değerinin belirlenmesini sağlayan bu iki parametre, madde güçlüğüne yapıyı özelliklerin ilişkili olduğu maddeler kapsamında değerlendirir. Çalışmada geleneksel yöntem olarak kullanılan madde test korelasyonu, her bir maddenin testin bütünüyle olan ilişkisini açıklamaktadır. DINA model parametreleri ise maddenin özelliğe sahip olan öğrenciler tarafından yanlış ve özelliğe sahip olmayan öğrenciler tarafından doğru cevaplanma durumunu dikkate alarak hesaplama yapmaktadır. Bu şekilde bakıldığında DINA model maddeyi doğru cevaplamak için gerekli olan özellik kapsamında bir çeşit üst grup ve alt grubun hata miktarını belirleyerek artı kalan oranı madde geçerliği olarak tanımlamaktadır. Bu anlamda iki model arasındaki farkların olası nedenlerinin daha detaylı incelenmesi gerekmektedir.

Kaynakça

- Baykul, Y.(2000) *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: ÖSYM Yayınları.
- Cheng Y.& Chang H.(2007). The Modified Maximum Global Discrimination Index Method for Cognitive Diagnostic Computerized Adaptive Testing. Presented at the CAT and Cognitive Structure Paper Session, June 7
- de la Torre, J. (2008a). An empirically-based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, 45, 343–362.
- de la Torre, J. (2008b). The generalized DINA model. Uluslar arası Psychometric Society toplantısında sunulmuştur, Temmuz, Durham, NH.
- de la Torre, J. (2009a). DINA Model and Parameter Estimation: A Didactic. *Journal of Educational and Behavioral Statistics* March, Vol. 34, No. 1, ss. 115–130.
- de la Torre, J. (2009b). A cognitive diagnosis model for cognitively-based multiple-choice options. *Applied Psychological Measurement*, 33, 163–183.
- de la Torre, J. & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*. V69, 3, s. 333-353.
- de la Torre, J. & Douglas, J. (2008). Model Evaluation and Multiple Strategies in Cognitive Diagnosis: An Analysis of Fraction Subtraction Data. *Psychometrika*. V73, N3, s. 595-624.
- de la Torre, J. & Lee, Y.S. (2010). A note on Invariance of the DINA Model Parameters, *Journal of Educational Measurement*, Bahar 2010, Sayı 47, No:1, s: 115-127.
- de la Torre, J. Hong, Y. & Deng W. (2010). Factors Affecting the Item Parameter Estimation and Classification Accuracy of the DINA Model. *Journal of Educational Measurement*, Yaz 2010, Sayı 47, No:2, s: 227-249.
- de la Torre, J., & Liu, Y. (2008). A cognitive diagnosis model for continuous response. *National Council on Measurement in Education toplantısında sunulmuştur*, Mart, New York, NY.
- Embretson, S. (1984). A general latent trait model for response processes. *Psychometrika*, 49, 175–186.
- Embretson, S.E. (1985). *Multicomponent latent trait models for test design*. In: Embretson, S.E. (Ed.), *Test Design: Developments in Psychology and Psychometrics*. Academic Press, New York, ss. 195–218.
- Embretson, S.E. (1997). Multicomponent response models. In: van der Linden, W.J., Hambleton, R.L. (Eds.), *Handbook of Modern Item Response Theory*. Springer, New York, pp. 305–321.
- Gitomer, D.H., Yamamoto, K. (1991). Performance modeling that integrates latent trait and class theory. *Journal of Educational Measurement* 28, 173–189.
- Haertel, E.H. (1984). An application of latent class models to assessment data. *Applied Psychological Measurement* 8, 333–346.
- Haertel, E.H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26, 333-352.
- Haertel, E.H. (1990). Continuous and discrete latent structure models of item response data. *Psychometrika* 55, 477–494.
- Hartz, S.M. (2002). A Bayesian framework for the Unified Model for assessing cognitive abilities: Blending theory with practicality. Yayınlanmamış doktora tezi, University of Illinois, Champaign, IL
- Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, 64, 187-212.
- Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, 64, 187-212.

- Sympson, J.B. (1977). A model for testing with multidimensional items. In: Weiss, D.J. (Ed.), *Proceedings of the 1977 Computerized Adaptive Testing Conference*. University of Minnesota, Department of Psychology, Psychometric Methods Program, Minneapolis, ss. 82–88.
- Tatsuoka, K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. *Journal of Educational Statistics*, 12, 55–73.
- Tatsuoka, K. (1995). *Architecture of knowledge structures and cognitive diagnosis: A statistical pattern recognition and classification approach*. In P.D. Nichols, S. F.
- Tatsuoka, K. Linn, R.L. (1982). Indices for detecting unusual patterns: Links between two general approaches and potential applications. *Applied Psychological Measurement* 7, 81-96.
- Turgut, M. F. & Baykul Y. 2010. *Eğitimde Ölçme ve Değerlendirme*, Pegem Akademi, 2.baskı, Ankara.
- von Davier, M. (2005). A general diagnostic model applied to language testing data. *ETS Araştırma Raporu: RR-05-16*, Educational Testing Service, Princeton, NJ.
- von Davier, M., Yamamoto, K. (2004). A class of models for cognitive diagnosis. *ETS Spearman Konferansında bildiri olarak sunulmuştur*, The Inn at Penn, Philadelphia, PA, Invitational Conference, Kasım.
- Wenmin, Z. (2006). Detecting Differential Item Functioning Using the DINA Model. The University of North Carolina at Greensboro. Yayınlanmamış doktora tezi. Greensboro
- Whitely, S.E. (1980). Multicomponent latent trait models for ability tests. *Psychometrika* 45, 479–494. Yazarın şu anki ismi Embretson, S.E.