

RASCH MODELİ İLE ELDE EDİLEN YETENEK ÖLÇÜLERİNİN NİTELİKLERİ ÜZERİNDE BİR ÇALIŞMA

Yr. Doç. Dr. Giray BERBEROĞLU (*)

Giriş

Psikometrik testlerden elde edilen ham puanların karşılaştırılmaya elverişli hale getirilmesi amacı ile standart puanlara dönüştürülmesi test geliştirme süreci içerisinde sıkça karşılaşılan uygulamalardan biridir. Genel olarak ham puanlar üzerinde doğrusal (linear) ve doğrusal olmayan (nonlinear) standart puanlar olmak üzere iki türlü dönüştürme yapmak mümkündür (Anastasi, 1982). T ve z standart puanları gibi doğrusal dönüştürmelerde ham puanların çarpıklık, basıklık gibi dağılım özellikleri standart puanlarda aynen korunmakta, ancak doğrusal olmayan standart 9 ve normalleştirilmiş T gibi dönüştürmelerde ise bu özellikler değişmektedir. Şöyle ki, doğrusal olmayan dönüştürmeler sonucunda elde edilen standart puanlarda normal dağılım özellikleri sağlanmaktadır (Gronlund, 1976). Bu tür dönüştürmelerin temel amacı, aynı test ile dağılım özellikleri farklı gruplardan elde edilen puanların karşılaştırılmasına olanak sağlamaktır.

Son yıllarda uygulama alanı hızla genişlemekte olan yeni ölçme yaklaşımları (Item Response Theory) içerisinde ele alınan modellerde de test ile ölçülmekte olan nitelik, modellerin tanımladığı standart yetenek parametreleri ile açıklanmaya çalışılmaktadır. Bunlardan biri olan Rasch modeli (1960, 1966) bir maddeye doğru cevap verme olasılığını madde-nin güçlüğü (d_j) ve yetenek düzeyinin (b_j) bir fonksiyonu olarak tanımlamaktadır. Şöyle ki;

$$Pr(b_j) = \frac{1}{1 + \exp[-(b_j - d_i)]} \dots\dots\dots 1$$

(Baker, 1987)

Model, maddeleri ve yetenek düzeyini en çok olabilirlik fonksiyonlarını kullanarak ölçeklemekte, aynı ham puana sahip bireyler için bir yetenek tahmini vermektedir. Modelin söz konusu parametreleri (b_j ve d_j) ölçeklenirken farklı örneklem gruplarının kullanılması halinde madde ve yeterlik düzeyleri için aynı parametre tahminlerinin elde edileceği iddia edilmekte, yapılan bir çok araştırma da bu iddiaları doğrular

(*) ODTÜ Eğitim Fakültesi.

gözükmektedir. (Skaggs ve Lissitz, 1986). Ancak ham puanlar yerine Rasch yetenek tahminlerinin kullanılması halinde elde edilen dağılımların ne tür özellikler taşıdığı incelenmeye değer gözükmektedir.

Bu araştırmada 1986 Öğrenci Seçme Sınavı (ÖSS) sonuçları kullanılarak, Rasch modeli ile elde edilen yetenek tahminlerinin (b_j) farklı yeterlik düzeylerindeki ortalama, standart sapma, çarpıklık ve basıklık ölçüleri, ham puanlardan elde edilen ölçülerle karşılaştırılmıştır. Bu yolla, ham puanlar yerine Rasch yetenek tahminlerinin kullanılması halinde puan dağılımlarının göstereceği değişikliklerle birlikte, standart yetenek tahminlerinin ham puanlar üzerinde ne tür bir dönüştürme yaptığı incelenmeye çalışılacaktır.

Yöntem

Araştırmada, Ankara'da 1986 Öğrenci Seçme ve Yerleştirme Sınavının birinci basamağına katılan, liselerin son sınıfında okumakta olan adaylar kullanılmıştır. Dağılım özellikleri farklı okullar elde edebilmek amacı ile, daha önceki yıllarda elde edilen sınav sonuçlarına göre başarı sıraları belirlenen okullardan üst, orta ve alt yeterlik dilimleri oluşturulmuştur. Her bir dilim içinde örneklem sayısı 1000'in altına düşmeyecek şekilde, dağılımları birbirine benzeyen okullar birleştirilmiştir.

Araştırmada kullanılan okullar ve örneklem sayıları Tablo 1'de gösterilmiştir.

TABLO 1
Örneklem Grupları

Örneklem Grubu	OKUL	n
ÜST	Atatürk Anadolu Lisesi	300
	Türk Eğitim Derneği Lisesi	<u>734</u>
	Toplam	1035
ORTA	Kocatepe Mimar Kemal Lisesi	482
	Bahçelievler Deneme Lisesi	441
	Bahçelievler Cumhuriyet Lisesi	<u>384</u>
	Toplam	1307
ALT	Abidinpaşa Lisesi	187
	Keçiören Fatih Sultan Mehmet Lisesi	315
	Keçiören Lisesi	316
	Gülveren Lisesi	<u>230</u>
	Toplam	1048
TOPLAM		3390

Araştırmanın analizlerinde 1986 Öğrenci Seçme Sınavının (ÖSS) alt testleri ayrı ayrı ele alınmış 33 maddeden oluşan Türkçe, 39 maddeden oluşan sosyal bilimler, 31 maddeden oluşan matematik, 41 maddeden oluşan fen bilimleri alt testlerinden elde edilen ham puan dağılımları ile.

ölçekleme sonucunda elde edilen yetenek ölçülerinin her bir adayın elde ettiği ham puanlara göre atanmasıyla elde edilen Rasch yetenek ölçüleri dağılımları; ortalama, standart sapma, çarpıklık ve basıklık ölçülerine göre karşılaştırılmıştır. Araştırmada Rasch yetenek ölçüleri RMAX (Powers, 1985) paket programı kullanılarak elde edilmiştir.

Bulgular

Her bir yetenek dilimi için, 1986 ÖSS alt testlerine göre elde edilen ham puan dağılımları Tablo 2'de verilmiştir.

TABLO 2

Örneklemler Gruplarında ÖSS Türkçe, Sosyal Bilimler Matematik ve Fen Bilimleri Alt Testleriyle Elde Edilen Puanların Ortalama ve Standart Sapmaları ile Çarpıklık ve Basıklık Ölçüleri

TEST	ÜST YARI		TÜM GRUP	
	ÜST	ORTA	ALT	
TÜRKÇE				
X	27,44	23,34	19,98	23,55
s	3,620	4,832	5,269	5,493
Çarpıklık	-1,543	-0,876	-0,328	-0,739
Basıklık	8,480	4,089	2,824	3,315
SOSYAL BİLİMLER				
X	24,68	20,52	17,35	20,81
s	6,230	5,955	5,806	6,651
Çarpıklık	-0,521	-0,156	0,157	-0,066
Basıklık	3,056	2,673	2,661	2,480
MATEMATİK				
X	19,58	13,61	7,58	13,57
s	6,449	6,896	5,379	7,879
Çarpıklık	-0,556	0,121	1,213	0,201
Basıklık	2,816	2,086	4,563	1,937
FEN BİLİMLERİ				
X	19,02	11,68	6,52	12,33
s	9,127	7,675	5,123	8,974
Çarpıklık	0,011	0,806	1,651	0,786
Basıklık	2,172	3,033	6,481	2,722

Simetrik normal dağılımda çarpıklık ölçüsünün 0, basıklık ölçüsünün ise 3 olduğu göz önünde bulundurulduğunda Tablo 2'deki bilgiler şöyle bir eğilime işaret eder niteliktedir: Türkçe alt puanı üst yeterlik diliminde negatif ivmeli ve sivri bir dağılım gösterirken alt yeterlik dilimlerine indikçe negatif ivme ile birlikte sivrilik azalmaktadır. Sosyal bilimler alt puanı dağılımında, üst yeterlik diliminden alt yeterlik dilimlerine indikçe ivme ölçüsü pozitif dönmekte ve basıklık artma eğilimi göstermektedir. Matematik alt puanında da durum sosyal bilimler alt puanına benzemektedir. Ancak burada çarpıklık pozitif yönde artarken sivrilik de artmaktadır. Fen bilimlerinde ise alt yeterlik dilimlerinde dağılım pozitif ivmeli hale gelirken sivrilik de artmaktadır.

Her bir yetenek dilimi için 1986 ÖSS alt testlerine göre elde edilen Rasch yetenek ölçüleri dağılımları Tablo 3'de verilmiştir.

Tablo 3 : Rasch Yetenek Ölçülerinin Gruplara Göre Elde Edilen Değerlerinin Ortalama, Standart Sapma İle Çarpıklık ve Basıklık Ölçüleri

TEST		ÜST YARI	ÜST YARI	ÜST YARI	TÜM
		ÜST	ORTA	ALT	GRUP
TÜRKÇE					
	X	2,188	1,194	0,587	1,265
	s	0,926	0,894	0,877	1,063
	Çarpıklık	-0,095	-0,152	0,093	0,015
	Basıklık	3,334	3,258	3,662	3,010
SOSYAL BİLİMLER					
	X	0,749	0,139	-0,273	0,187
	s	0,916	0,804	0,765	0,908
	Çarpıklık	-0,156	-0,032	0,001	0,090
	Basıklık	3,523	3,279	3,251	3,201
MATEMATİK					
	X	0,830	-0,383	-1,463	-0,380
	s	1,385	1,345	1,115	1,562
	Çarpıklık	-0,013	-0,076	0,335	0,089
	Basıklık	3,248	2,797	3,478	2,585
FEN BİLİMLERİ					
	X	-0,191	-1,197	-2,035	-1,132
	s	1,257	1,132	0,974	1,322
	Çarpıklık	-0,143	0,124	0,249	0,240
	Basıklık	3,294	2,861	3,272	2,791

Tablo 3'teki bilgileri Tablo 2'deki ham puanlarla ilgili dağılım ölçüleriyle karşılaştırıldığında, Rasch yetenek ölçüleriyle ilgili dağılımların birbirine daha çok benzediği görülmektedir. Standart sapmalar 1'e yaklaşırken, çarpıklık ve basıklık ölçütleri de normal dağılımla ilgili değerlere yaklaşmaktadır. Üst gruptan alt gruplara doğru gidildikçe ham puanların çarpıklık ve basıklık ölçülerinde gözlenen gelişme, Rasch yetenek tahminlerinde kaybolmaktadır.

Sonuç

Bu çalışmada farklı yeterlik dilimlerinden 1986 ÖSS alt testlerinde elde edilen ham puan dağılımları, aynı gruplarda ilgili ham puanlara Rasch yetenek ölçülerinin atanması ile elde edilen standart puanların dağılımlarıyla karşılaştırılmıştır.

Tablo 2 ve 3'teki bilgiler Rasch yetenek ölçülerinin normalleştirilmiş T ve standart 9 ölçeğinde olduğu gibi, doğrusal olmayan bir transformasyon sonucunda normal dağılıma yakın değerler verdiğini gösterir niteliktedir. Bu dağılımlarda standart sapma bir'e yaklaşırken çarpıklık ve basıklık ölçütleri normal dağılıma yakın değerler vermektedir. Ancak burada dikkati çeken bir nokta gruplar arasında aynı testten elde edilen ham puan ortalamaları farklarının, yetenek ölçülerinde de korunuyor olmasıdır. Bu özellikler, Rasch yetenek ölçülerinin dağılımları ve yeterlik düzeyleri farklı grupların karşılaştırılmasını anlamlı kılacak nitelikte olduğunu göstermektedir. Şöyle ki; Rasch yetenek ölçüsü bilinen bir bireyin ölçülen yeterlik boyutundaki düzeyi grubun ham puan dağılımlarından etkilenmeden belirlenebileceği gibi, yetenek düzeyleri farklı grupların ortalamalarına göre, bireyin söz konusu grup içerisinde hangi düzeyde olduğu da söylenebilecektir. Bu bulgular Rasch yetenek tahminlerinin kullanılan örneklem grubundaki yetenek dağılımından bağımsız olarak tahmin edilebileceğini, böylece bir testin ulaşılabilen herhangi bir grupta ölçeklenebileceğini de gösterir niteliktedir (Wright, 1967).

Söz konusu modelin farklı nitelikteki testler ve gruplar için incelenmesinde yarar görülmektedir.

Kaynaklar

- Anastasi, A. (1982) **Psychological Testing** New York: MacMillan Publishing Co., Inc.
- Baker, B.F. (1987) "Methodology Review: Item Parameter Estimation Under the One-Two and Three-Parameter Logistic Models" **Applied Psychological Measurement**. 11; 2; 111-141.
- Berberoğlu, G. (1988) Seçme Amacıyla Kullanılan Testlerde Rasch Modelinin Katkıları. Yayınlanmamış Doktora Tezi. (Hacettepe Üniversitesi).

- Gronlund, N. (1976) **Measurement and Evaluation in Teaching**. New York. Macmillan Publishing Co., Inc.
- Hambleton, R.K., Swaminathan, H. (1985) **Item Response Theory: Principles and Applications**. Boston: Kluwer. Nijhoff Publishing.
- Powers, S. (1985) "A Microcomputer Program that Performs a Rasch Calibration of Test Items". **Educational and Psychological Measurement**. 45; 1; 1987-188.
- Rasch, G. (1960) **Probabilistic Models for Some Intelligence and Attainment Tests**. Copenhagen: Danish Institute for Educational Research. Bölüm V.
- Rasch, G. (1966) "An Item Analysis Which Takes Individual Differences Into Account" **British Journal of Mathematical and Statistical Psychology** 19.1:49-57.
- Skaggs, G., Lissitz., W.R. (1986) "IRT Test Equating: Relevant Issues and a Review of Recent Research". **Review of Educational Research**. 56: 495-529.
- Wright, B.D. (1967) "Sample free Test Calibration and Person Measurement". Princeton: Educational Testing Service.

Bugünden Yarına Ortaöğretimimiz

- Bildiriler, Panel, Tartışmalar

TED Bilim Dizisi Yayını