

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME

BÖLÜM V

Test ve Madde İstatistikleri

Test Geliştirme

Test, bireylerin ölçme konusu olan özelliklerinin belirlenmesi amacıyla kullanılan ölçme araçlarına verilen genel bir isimdir. Test geliştirme sistematik olarak belli aşamalardan geçilerek gerçekleştirilir.

Test geliştirme aşamalarını şöyle sıralayabiliriz:

1. Amacın belirlenmesi
2. Kapsamı teşkil edecek davranışların belirlenmesi
3. Belirli tablosunun hazırlanması
4. Kullanılacak soru türünün belirlenmesi
5. Denemelik soruların yazılması
6. Soruların test düzenine konulması
7. Deneme uygulamasının yapılması
8. Madde analizi yapılarak maddelerin seçilmesi
9. Seçilmiş maddelerle testin oluşturulması
10. Test istatistiklerinin kestirilmesi

Test Geliştirme

Amacın belirlenmesi : Testin amacı ne?

- ✓ Tanıma ve Yerleştirmeye Yönelik Değerlendirme;
- ✓ Biçimlendirme - Yetiştirmeye Yönelik Değerlendirme (Formatif);
- ✓ Düzey (Seviye) Belirlemeye Yönelik Değerlendirme (Summatif)

Kapsamı teşkil edecek davranışların belirlenmesi

- ✓ Testin amacına bağlı olarak belirlenen test türüne göre **tüm davranışlar** (izleme testi) yada **kritik davranışlar** (başarı testi) alınır.

Belirli tablosunun hazırlanması

Kullanılacak soru türünün belirlenmesi

- ✓ Ölçülecek davranışların (kazanımların) bilişsel alanın hangi basamağında ise o tür sorular hazırlanır. (Bilgi, Kavrama, Uygulama, Analiz, Sentez, Değerlendirme)

Denemelik soruların yazılması

- ✓ Her bir davranışı ölçecek denemelik üç soru yazılır.

Test Geliştirme

Soruların test düzenine konulması

- ✓ Maddeler test formu içinde dağıtılır, test yönergesi yazılır, maddeler test formatına göre yazılır.

Deneme uygulamasının yapılması

- ✓ Asıl gruba benzer özellikler gösteren bir gruba deneme uygulaması yapılır.

Madde analizi yapılarak maddelerin seçilmesi

Seçilmiş maddelerle testin oluşturulması

- ✓ Madde analizi sonucunda seçilen maddeler ile nihai test oluşturulur.

Test istatistiklerinin kestirilmesi

- ✓ Nihai teste seçilen maddeler ile test istatistikleri hesaplanır.

Madde Analizi

Madde analizinde en çok kullanılan iki teknik;

1. **Madde puanları matrisi**ne dayalı olarak madde istatistiklerinin belirlenmesi,
2. **%27lik gruplar**a dayalı olarak madde istatistiklerinin belirlenmesidir.

➤ Deneme uygulamasındaki tüm cevaplar mı kullanılacak yoksa üst ve alt %27lik grupların cevapları mı kullanılacak sorusuna göre yöntemimiz değişir.

Madde Puanları Matrisi ile Analiz

Her bir öğrencinin sorulara verdikleri doğru ve yanlış yanıtları gösteren bir matris (1-0 matrisi) oluşturulur. (sıra öğrenciler, sütun sorular olacak şekilde)

Bu matrise bakarak madde istatistikleri hesaplanır.

	Madde 1	Madde 2	Madde 3	Madde 4	Madde 5	Toplam
Öğrenci-1	1	1	0	1	0	3
Öğrenci-2	0	1	1	1	0	3
Öğrenci-3	1	1	0	1	0	3
Öğrenci-4	0	1	0	1	0	2
Öğrenci-5	1	1	0	1	0	3
Öğrenci-6	1	0	0	1	0	2
Toplam	4	5	1	6	0	16

Madde Güçlük İndeksi (p_j)

- Her bir soru için hesaplanan madde güçlük indeksi soruların güçlük düzeyini gösterir. Soruyu doğru cevaplayanların sayısının toplam kişi sayısına oranı ile bulunur.

$$P_j = n(D) / n$$

- P_j değeri % şeklinde okunabilir: 'grubun yüzde kaç soruyu doğru cevapladı'
- Ne kadar çok öğrenci soruyu cevaplarsa soru o kadar kolay, ne kadar az öğrenci soruyu cevaplarsa soru o kadar zordur.
 - (Madde güçlük indeksi 1'e yaklaştıkça madde kolay, 0'a yaklaştıkça madde zor olarak yorumlanır.
- Öm:** Bir önceki madde puanları matrisinde madde güçlük indekslerini hesaplayın.

Madde Güçlük İndeksi (p_j)

	Madde 1	Madde 2	Madde 3	Madde 4	Madde 5	Toplam
Öğrenci-1	1	1	0	1	0	3
Öğrenci-2	0	1	1	1	0	3
Öğrenci-3	1	1	0	1	0	3
Öğrenci-4	0	1	0	1	0	2
Öğrenci-5	1	1	0	1	0	3
Öğrenci-6	1	0	0	1	0	2
Toplam	4	5	1	6	0	16
P_j	4/6= 0.67	5/6=0.83	1/6=0.17	6/6=1	0/6=0	

- En kolay madde? En zor madde?

Madde Ayırtedicilik Gücü İndeksi (r_{jx})

Madde ayırtedicilik gücü indeksi, maddenin ilgili davranışa sahip olanla olmayana ne ölçüde ayırdığını gösterir.

$$r_{jx} = \frac{X_{jD} - X_x}{S_x} \cdot \sqrt{\frac{p}{1-p}}$$

- P_j : Madde ayırtedicilik gücü indeksi
- X_{jD} : j maddesini doğru cevaplayanların test puanı ortalaması
- X_x : Testin ortalaması
- S_x : Testin standart kayması
- p : madde güçlüğü

Madde Ayırtedicilik Gücü İndeksi (r_{jx})

- Madde ayırtedicilik indeksi, madde puanı ile test puanı arasında **korelasyon**dur. (-1) ile (+1) arasında değerler alır.
 - (+) işaretli değerler maddenin ölçtüğü niteliğin testin bütünüyle ölçülen nitelikle uyumlu olduğunu;
 - (-) işaretli değerler maddenin ölçtüğü değişkenin testin bütünüyle ölçülen nitelikte ters yönde ilişki içinde bulunan bir nitelik ölçtüğünü;
 - 0) ve civarındaki değerler de, maddenin ölçtüğü niteliğin testin bütünüyle ölçülen nitelik arasında ilişki bulunmadığını belirir.
- Bu nedenle (-) işaretli ve sıfır civarında değer alan maddeler teste konulmaz.

Madde Ayırtedicilik İndeksi	Madde Seçimi Kararı
0.19 ve daha düşük	Madde kesinlikle teste alınmamalı yada tamamen düzeltilmelidir.
0.20 ile 0.29 arasında	Gerekirse düzeltilecek teste alınabilir.
0.30 ile 0.39 arasında	Daha ayırt edici madde olmaması durumunda küçük düzeltmelerle kullanılabilir.
0.40 ve daha yüksek	Teste olduğu gibi alınabilir.

Madde standart sapması (S_j) ve Varyansı (S_j^2)

- Eğer testi oluşturan maddeler '1-0' şeklinde puanlanabiliyorsa her bir maddenin standart sapması S_j o maddenin madde güçlük indeksi (p_j) ile madde güçlük indeksinin 1'den farkının (q_j) çarpımının kareköküne eşittir.

- $S_j = \sqrt{p_j \cdot q_j}$
- Varyans standart sapmanın karesine eşittir ve aşağıdaki
- $S_j^2 = p_j \cdot q_j$

Yorum: Bir maddenin madde varyansı arttıkça, o madde ile ölçülmek istenen özellikler bakımından bireylerin farklılıklarının ortaya konulma gücü de artmış olur.

Soru: Madde varyansının en yüksek değeri kaç olur?

Madde standart sapması (S_j) ve Varyansı (S_j^2)

- Aşağıdaki tablodaki maddelerin madde varyanslarını hesaplayınız ve yorumlayınız

	Madde 1	Madde 2	Madde 3	Madde 4	Madde 5	Toplam
Öğrenci-1	1	1	0	1	0	3
Öğrenci-2	0	1	1	1	0	3
Öğrenci-3	1	1	0	1	0	3
Öğrenci-4	0	1	0	1	0	2
Öğrenci-5	1	1	0	1	0	3
Öğrenci-6	1	0	0	1	0	2
p_j	0.67	0.83	0.17	1	0	
q_j	0.33	0.17	0.83	0	1	
S_j^2	0.22	0.14	0.14	0	0	

KPSS-2008

Bir testteki maddelere ilişkin aşağıdaki değerler elde edilmiştir.

Maddeler	I	II	III	IV	V
Madde güçlüğü	0,1	0,15	0,5	0,6	0,9
Madde ayrıncılık gücü	0,08	-0,15	0,5	0,28	0,43

60. Yapılandırılacak bir testte bu maddeleri kullanma zorunluluğu olsaydı, en köklü değişikliğin hangi maddede yapılması gerekirdi?

A) I B) II C) III D) IV E) V

• Cevap: B

KPSS-2009

SORULAR

	1	2	3	4	5	6	7	8	9	10
A	1	1	1	0	0	1	1	0	1	0
B	1	1	0	0	1	1	1	0	0	1
C	0	1	0	0	1	1	1	1	1	1
D	1	1	1	0	1	0	1	1	1	0
E	0	0	1	0	0	0	1	1	1	0
F	1	1	0	0	1	1	1	1	1	1
G	1	1	0	0	1	1	1	0	0	0
H	1	1	0	0	1	1	1	0	0	1
I	1	1	1	0	0	0	1	0	0	1
J	0	0	0	0	0	1	1	1	0	0

59. Hangi sorular ölçülen özellikler açısından öğrencilerin bireysel farklılıklarını ortaya çıkarmada hiç etkili olmamıştır?

A) Yalnız 4. soru B) Yalnız 7. soru
C) Yalnız 8. soru D) 4. ve 7. sorular
E) 4., 7. ve 8. sorular

• Cevap: D

%27'lik gruplarla madde analizi

- %27lik gruplara dayalı olarak madde analizi tekniği işlemlerin basitliği nedeniyle en çok kullanılan tekniklerden biridir. Deneme grubunda en az 100, tercihen 200 cevaplayıcı olduğu durumlarda kullanılır.
- Deneme uygulaması yapıldıktan sonra grup puanlarına göre sıralanır ve en başarılı %27'lik üst grup ile, en az başarılı %27'lik alt grup karşılaştırılır.
- Alt ve üst grupların sorulara verdikleri yanıtlar kullanılarak madde istatistikleri hesaplanır.

%27'lik gruplarla madde güçlük indeksi(p_j)

- Madde güçlük indeksi soruyu doğru cevaplayanların sayısının, toplam kişi sayısına oranıdır. Bu yöntemde madde güçlük indeksi;

$$P_j = \frac{n(dü) + n(da)}{N}$$

☆ % 27 (Üst Grup)

(Öğrenci notlarının % 46'sı İşleme tabi tutulmaz)

☆ % 27 (Alt Grup)

P_j = Maddeyi güçlük indeksi
n (dü) = Maddeyi üst grupta doğru cevaplayanlar
n (da) = Maddeyi alt grupta doğru cevaplayanlar
N = Üst ve alt gruptaki öğrenci sayısı

%27'lik gruplarla madde güçlük indeksi(p_j)

Soru: Aşağıdaki tablodaki verilerden faydalanarak, %27lik alt ve üst gruplara göre madde güçlük indeksini hesaplayınız.

	A cevabını veren öğrenci sayısı	B cevabını veren öğrenci sayısı	C cevabını veren öğrenci sayısı (Doğru cevap)	D cevabını veren öğrenci sayısı	E cevabını veren öğrenci sayısı	Toplam
Üst Grup	15	20	30	20	15	100
Alt Grup	20	17	15	20	28	100

$$P_j = \frac{n(dü) + n(da)}{N}$$

• $P_j = 30 + 15 / 200 = 0,23$
(nasıl bir madde?)

%27'lik gruplarla madde ayrıncılık gücü indeksi(r_{jx})

- Madde ayrıncılık indeksi ölçülen davranışa sahip olanlarla (üst grup) bu davranışa sahip olmayanları (alt grup) ayırt etmede kullanıldığı için üst gruptaki doğru cevap sayısından alt gruptaki doğru cevaplayanların sayısını çıkarmaya dayalı bir eşitlikle hesaplanır.

$$r_{jx} = \frac{n(dü) - n(da)}{n}$$

☆ % 27 (Üst Grup)

(Öğrenci notlarının % 46'sı İşleme tabi tutulmaz)

☆ % 27 (Alt Grup)

r_{jx} = Maddeyi ayırt edicilik gücü
n (dü) = Maddeyi üst grupta doğru cevaplayanlar
n (da) = Maddeyi alt grupta doğru cevaplayanlar
n = Üst ya da alt gruptan herhangi birisinin eleman sayısı (Her iki grupta da eleman sayısı eşittir)

%27'lik gruplarla madde ayırıcılık gücü indeksi (r_{jx})

- Soru:** Aşağıdaki tablodaki verilerden faydalanarak, %27'lik alt ve üst gruplara göre madde ayırıcılık gücü indeksini hesaplayınız.

	A cevabını veren öğrenci sayısı	B cevabını veren öğrenci sayısı	C cevabını veren öğrenci sayısı	D cevabını veren öğrenci sayısı	E cevabını veren öğrenci sayısı	Toplam
Üst Grup	15	20	30	20	15	100
Alt Grup	20	17	15	20	28	100

$$r_{jx} = \frac{n(d_{ii}) - n(d_{ia})}{n}$$

- $r_{jx} = 30 - 15/100 = \underline{0,15}$ (nasıl bir madde?)

Madde güvenilirlik indeksi (r_j)

- Madde ayırt edicilik indeksi ile madde standart sapmasının çarpımına eşittir.

$$r_j = r_{jx} * S_j$$

Soru: Bir önceki örnek için madde güvenilirliği kaçtır?

- (madde ayırt ediciliği 0,15 ve madde güçlük indeksi 0,23 idi)
- $r_j = 0,15 * \sqrt{0,23 * 0,77}$
- $r_j = 0,06$

Madde seçimi nasıl yapılmalıdır?

1. **Madde güçlük indeksi kontrol edilmelidir.** Öncelikle testin hangi amaçla yapıldığı belirlenmelidir.

- Testin amacı **öğrenci başarısını belirlemek** ise test maddelerinin çoğu **orta güçlükte** olmalı, yani sıra hem kolay hem de zor maddeler teste yer almalıdır. (Madde güçlük indeksi 0,20 ile 0,80 arasındaki maddeler uygundur.)
- Testin amacı **seçim yapmak** ise ve çok sayıda başvuran varsa, madde güçlük indekslerinin olabildiğince **düşük olması**, maddelerin zor olması beklenir.

Madde seçimi nasıl yapılmalıdır?

2. **Madde ayırt edicilik indeksi kontrol edilmelidir.** 0,40'tan yüksek ve olabildiğince 1'e yakın maddeler teste alınmalıdır.

Madde ayırt edicilik düzeyi yeterli çıkmadığında, madde kökünün değiştirilmesi yerine çeldiriciler üzerinden düzeltme yapılır. Zayıf olan çeldirici anlamca doğru yanıtla yakınlaştırılır. Bu işlem aynı zamanda soruyu zorlaştırır.

	A	B	C (Doğru cevap)	D	E	Toplam
Üst Grup	15	20	30	20	15	100
Alt Grup	20	17	15	20	28	100

Madde seçimi nasıl yapılmalıdır?

Çeldiriciler:

- Doğru cevap seçeneği dışındaki seçeneklere çeldirici denir.
- Üst grupta doğru cevaba giden öğrenci sayısı alt grupta doğru cevaba giden öğrenci sayısından fazla olmalıdır.
- Üst grupta çeldiricilere giden öğrenci sayısı alt grupta çeldiricilere giden öğrenci sayısından az olmalıdır.
- Çeldiricilere giden öğrenci sayısı dengeli olmalıdır.
- Hiçbir öğrenci tarafından işaretlenmeyen çeldiriciler yeniden düzenlenmelidir.

Madde seçimi nasıl yapılmalıdır?

3. **Madde standart sapması ve varyansı kontrol edilmelidir.**

Bir maddenin varyansı/standart sapması arttıkça o madde ile ölçülmek istenen özellik bakımından **bireyler arasındaki farklılıkların ortaya konulma gücü** artar. Bu nedenle test maddelerini seçerken olabildiğince varyansı yüksek maddeler seçilmelidir. (varyansı 0,25'e yakın maddeler seçilmelidir)

Madde seçimi nasıl yapılmalıdır?

- 4. Madde güvenilirliği kontrol edilmelidir. Maddelerin güvenilirliği ne kadar yüksekse, bu maddelerden oluşan testin güvenilirliği de o kadar yüksek olur. Bu nedenle teste madde seçilirken olabildiğince **madde ayırt ediciliği** ve **madde standart sapması** (dolayısıyla **madde güçlüğü**) büyük maddeler seçilmelidir.

Madde istatistiklerinden test istatistiklerinin kestirilmesi

Madde analizi ile hesaplanan madde istatistiklerine dayalı olarak testin bütünün bazı istatistikleri hesaplanabilir.

Bunlar:

- Testin aritmetik ortalaması
- Testin ortalama güçlüğü
- Test puanlarının standart sapması ve varyansı

Madde istatistiklerinden test istatistiklerinin kestirilmesi

• Testin Aritmetik Ortalaması:

Testteki tüm soruların madde güçlük indekslerinin toplamı, testin aritmetik ortalamasına eşittir.

	Madde 1	Madde 2	Madde 3	Madde 4	Madde 5	Toplam
Öğrenci-1	1	1	0	1	0	3
Öğrenci-2	0	1	1	1	0	3
Öğrenci-3	1	1	0	1	0	3
Öğrenci-4	0	1	0	1	0	2
Öğrenci-5	1	1	0	1	0	3
Öğrenci-6	1	0	0	1	0	2
Toplam	4	5	1	6	0	16
P _j	4/6=0.67	5/6=0.83	1/6=0.17	6/6=1	0/6=0	

$$X_{\text{ort}} = \sum p$$

$$X = 3+3+3+2+3+2/6 = 2,67$$

Veya

$$0.67+0.83+0.17+1+0 = 2.67$$

Madde istatistiklerinden test istatistiklerinin kestirilmesi

• Testin Ortalama Güçlüğü:

Testteki soruların güçlük indekslerinin toplamının testteki soru sayısına bölümüyle testin ortalama güçlüğü bulunur.

$$P_{\text{ort}} = \sum P / K \text{ veya}$$

Testin güçlük indeksleri toplamı testin aritmetik ortalamasına eşit olduğu için

$$P_{\text{ort}} = X_{\text{ort}} / K$$

Örn: 100 sorudan oluşan bir testin ortalamasının 50 olması isteniyorsa madde güçlükleri kaç olmalıdır?

$$P_{\text{ort}} = 50/100 = 0.50$$

Madde istatistiklerinden test istatistiklerinin kestirilmesi

• Test Puanlarının Standart Sapması ve Varyansı:

Test puanlarının standart sapmasının hesaplanması için madde güvenilirliği kullanılır. Testin standart sapması, testteki tüm maddelerin güvenilirliklerinin toplamına eşittir.

$$S_x = \sum r_j$$

KPSS 2004

Aşağıdaki sorulardan hangisi madde analizi ile elde edilecek bilgilerden yararlanılarak **cevaplanamaz**?

A. Testin zorluk derecesi amaca uygun mudur?

B. Testteki sorularda yoklanan davranışlar , kazandırılması amaçlanan davranışlarla aynı mıdır?

C. Testteki maddelerin çeldiricileri beklenene yakın biçimde çalışıyor mu?

D. Test, ölçülmek istenen nitelikleri ne derece bir tutarlılıkla ölçmektedir?

E. Test çok öğrenmiş ve az öğrenmiş olan öğrencileri birbirinden ayırabiliyor mu?

KPSS Ölçme Sorularının İncelemesi Grup Projesi

- KPSS-2009
- Madde No : 45
- Madde Kökü:
- Aşağıdaki değerlendirme süreçlerinin hangisinde bağıl ölçüt kullanılmıştır?
- **Hangi konu kapsamındadır?**
- Ölçme ve Değerlendirme, Bağıl ölçüt konusu
- **Bloom Taksonomisinin bilişsel alanın hangi basamağına aittir?**
- Kavrama basamağına aittir. Burada öğrenci ilk önce bağıl ölçütü bilmesi gerekecek, onunla ilgili özellikleri hatırlayacak en son olarak da bu örnekler arasından bağıl ölçüte en uygun olanı bulacak.

KPSS Ölçme Sorularının İncelemesi Grup Projesi

- **Çeldirciler ve Doğru yanıt;**
- **Çeldirici: A) Mezun olmak için diploma notunun 4 üzerinden en az 2 olması**
- Öğrencinin mezun olmak için diploma notunun 4 üzerinden 2 olması mutlak bir ölçüttür. Önceden belirlenen bir kriter vardır. Ama bağıl ölçütte ise sonradan belirlenen bir ölçüt vardır. Bağıl ölçüt, bir grup içerisindeki bireyleri birbirleri ya da kendisi ile karşılaştırmak, sıralamak ve içerisinde seçim yapmak için kullanılır.
- **Doğru Yanıt; B) Ortalamanın 10 puan üzerinde alanların sınıfı geçmesi**
- **Doğru olan yanıt B seçeneğidir.** Bağıl ölçüt, bir grup içerisindeki bireyleri birbirleri ya da kendisi ile karşılaştırmak, sıralamak ve içerisinde seçim yapmak için kullanılır. A,C,D ve E seçeneklerinde seçim için alınan ölçütlerde karşılaştırmanın yapılmadığı koşulunun belirlendiği görülmektedir. B seçeneğinde ise ortalama, değerlendirme için ölçüt olarak kullanılmıştır.
- **Çeldirici: C) ...**
- **Çeldirici: D) ...**
- **Çeldirici: E) ...**