EĞİTİMDE ÖLÇME ve DEĞERLENDİRME

Neşe GÜLER

5. Baskı

PEGEM AKADEMİ

egitim@hotmail.es

ÖNSÖZ

Bu kitabın amacı, öğretmenlere ve öğretmen adaylarına eğitimin en önemli öğelerinden biri olan ölçme ve değerlendirme ile ilgili temel bilgi ve becerileri kazanabilmelerine yardımcı olmaktır. Kitap, ölçme ve değerlendirme ile ilgili temel kavramları, ölçme araçlarında bulunması gereken nitelikleri, eğitimde kullanılan ölçme araç ve yöntemlerini kapsadığı gibi test geliştirme, ölçme sonuçları üzerinde istatistiksel işlemleri de içeren sekiz bölümden oluşmaktadır. Öğrenilenleri pekiştirmesi düşüncesiyle bölüm sonlarında alıştırma sorularına ve KPSS'de çıkmış sorulara yer verilmiştir. Sade bir anlatımla sunulmaya çalışan kitabın öğretmenlere ve öğretmen adaylarına yardımcı olabilmesi amaçlanmıştır.

Akademik hayatımda katkısı ve emeği bulunan değerli danışman hocalarım Prof. Dr. İ.Naci Cangül'e, Doç. Dr. Randall Penfield'a ve aynı zamanda kitabın ilk değerlendirilmesi aşamasında önerileriyle destekte bulunan Prof. Dr. Selahattin Gelbal'a teşekkürlerimi sunarım.

Her zaman yanımda olan, desteğiyle bana güç veren eşim Güçlen Güler'e, kitabın düzeltme aşamasındaki katkılarıyla, manevi olarak göstermiş oldukları ilgiyle yardımlarını benden esirgemeyen değerli meslektaşlarım Yrd. Doç. Dr. Bayram Çetin'e, Öğr. Gör. Gülşen Taşdelen'e ve Araş. Gör. Gülden Kaya'ya sevgilerimle...

Burada verilmeye çalışılan bilgiler öğrencilerim sayesinde çoğalmış ve gelişmiştir. Derslerde yönelttikleri sorularla beni yönlendiren ve bu kitabın oluşmasında bana cesaret veren öğrencilerime ben de iyi bir rehber olabilirsem ne mutlu bana...

Tüm öğrencilerime teşekkürlerimle,

Yrd. Doç. Dr. Neşe GÜLER

İÇİNDEKİLER

1. BÖLÜM
ÖLÇME VE DEĞERLENDİRMEYLE İLGİLİ TEMEL KAVRAMLAR

GİRİŞ 2
ÖLÇME VE DEĞERLENDİRMEYLE İLGİLİ TEMEL KAVRAMLAR 2
SABİT VE DEĞİŞKEN KAVRAMLARI 2
ÖLÇME VE TÜRLERİ 4
ÖLÇME TÜRLERİ 5
ÖLÇMEDE BİRİM 6
ÖLÇMEDE SIFIR NOKTASI 7
ÖLÇEK 8
ÖLÇEK TÜRLERİ VE ÖZELLİKLERİ 8
DEĞERLENDİRME 12
Ölçüte göre Değerlendirme Türleri 13
Amaca göre Değerlendirme Türleri 14
DEĞERLENDİRME SORULARI 15
KPSS'DE ÇIKMIŞ SORULAR 17

2. BÖLÜM
ÖLÇMEDE HATA

ÖLÇMEDE HATA 30
HATA TÜRLERİ 31
Hata Kaynakları 31
DEĞERLENDİRME SORULARI 32
KPSS'DE ÇIKMIŞ SORULAR 33

3. BÖLÜM
KORELASYON

KORELASYON 36
DEĞERLENDİRME SORULARI 41
KPSS'DE ÇIKMIŞ SORULAR 42

egitim@hotmail.es

4. BÖLÜM
ÖLÇME ARAÇLARINDA BULUNMASI GEREKEN NİTELİKLER

GÜVENİRLİK [0 , 1] 44
Güvenirlik Türleri ve Güvenirliği Kestirme Yöntemleri 46
Ölçme Sonuçlarının Standart Hatası 55
Ölçme Sonuçlarının Güven Aralığının Hesaplanması 56
GEÇERLİK [-1 , 1] 57
Kapsam Geçerliği 58
Yapı Geçerliği 58
Ölçüt Dayanaklı Geçerlik (Ölçüt Geçerliği) 60
Geçerliği Arttırmanın Yolları 61
Güvenirlik ile Geçerlik Arasındaki İlişki 62
KULLANIŞLILIK 64
DEĞERLENDİRME SORULARI 65
KPSS'DE ÇIKMIŞ SORULAR 67

5. BÖLÜM
EĞİTİMDE KULLANILAN ÖLÇME ARAÇ VE YÖNTEMLERİ

EĞİTİMDE KULLANILAN ÖLÇME ARAÇ VE YÖNTEMLERİ 80
SÖZLÜ SINAVLAR 80
Sözlü Sınav Nedir? 80
Sözlü Sınavlar Ne Zaman Tercih Edilebilir? 80
Sözlü Sınavlarda Nelere Dikkat Edilmelidir? 81
Sözlü Sınavların Avantajları Nelerdir? 81
Sözlü Sınavların Dezavantajları Nelerdir? 82
YAZILI SINAVLAR 82
Yazılı Sınav Nedir? 82
Yazılı Sınavlar Ne Zaman Tercih Edilebilir? 84
Yazılı Sınavlarda Nelere Dikkat Edilmelidir? 84
Yazılı Sınavların Avantajları Nelerdir? 85
Yazılı Sınavların Dezavantajları Nelerdir? 85
KISA CEVAPLI TESTLER 85
Kısa Cevaplı Test Nedir? 85
Kısa Cevaplı Testler Ne Zaman Tercih Edilebilir? 86
Kısa Cevaplı Testlerin Hazırlanmasında Nelere Dikkat Edilmelidir? 86
Kısa Cevaplı Testlerin Avantajları Nelerdir? 87
Kısa Cevaplı Testlerin Dezavantajları Nelerdir? 87

DOĞRU-YANLIŞ TESTLERİ (Sınıflama Gerektiren Testlerin En Çok Bilineni) 88
Doğru-Yanlış Test Nedir? 88
Doğru-Yanlış Testleri Ne Zaman Tercih Edilebilir? 88
Doğru-Yanlış Testlerin Hazırlanmasında Nelere Dikkat Edilmelidir? 88
Doğru-Yanlış Testlerin Avantajları Nelerdir? 88
Doğru-Yanlış Testlerin Dezavantajları Nelerdir? 89
ÇOKTAN SEÇMELİ TESTLER (Seçme Gerektiren Testler) 89
Çoktan Seçmeli Test Nedir? 89
Çoktan Seçmeli Testler Ne Zaman Tercih Edilebilir? 90
Çoktan Seçmeli Testlerin Hazırlanmasında Nelere Dikkat Edilmelidir? 90
Çoktan Seçmeli Testlerin Avantajları Nelerdir? 92
Çoktan Seçmeli Testlerin Dezavantajları Nelerdir? 92
Çoktan Seçmeli Madde Tipleri 92
PERFORMANS DEĞERLENDİRME 93
ÖĞRENCİ ÜRÜN DOSYASI (PORTFOLYO) 94
Portfolyo ile Değerlendirmenin Ögrenci Açısından Yararları 99
Portfolyo ile ile Değerlendirmenin Ögretmen Açısından Yararları 99
DERECELİ PUANLAMA ANAHTARI (RUBRİK) 99
Bütünsel (Holistik) Dereceli Puanlama Anahtarları 100
Analitik Dereceli Puanlama Anahtarları 101
DEĞERLENDİRME SORULARI 103
KPSS'DE ÇIKMIŞ SORULAR 105

6. BÖLÜM
TEST GELİŞTİRME SÜRECİ VE MADDE ANALİZİ

TEST GELİŞTİRME SÜRECİ VE MADDE ANALİZİ 126
TEST GELİŞTİRME SÜRECİNİN BASAMAKLARI 126
MADDE ANALİZİ 127
Basit Yöntem (alt – üst grup yöntemi) 129
Madde Güvenirliği (rx) 135
2. Yöntem 136
Test Parametrelerinin Hesaplanması 138
DEĞERLENDİRME SORULARI 141
KPSS'DE ÇIKMIŞ SORULAR 143

egitim@hotmail.es

7. BÖLÜM
TEST PUANLARI ÜZERİNDE İSTATİSTİKSEL İŞLEMLER

TEST PUANLARI ÜZERİNDE İSTATİSTİKSEL İŞLEMLERİ 150
Verilerin Düzenlenmesi 150
MERKEZİ YIĞILMA ÖLÇÜLERİ 154
MERKEZİ YIĞILMA ÖLÇÜLERİ ARASINDAKİ İLİŞKİ 158
MERKEZİ DAĞILIM ÖLÇÜLERİ 159
DEĞERLENDİRME SORULARI 165
KPSS'DE ÇIKMIŞ SORULAR 167

8. BÖLÜM
STANDART PUANLAR

STANDART PUANLAR 184
DEĞERLENDİRME SORULARI 187
KPSS'DE ÇIKMIŞ SORULAR 189

DEĞERLENDİRME SORULARI CEVAP ANAHTARI 191
KPSS'DE ÇIKMIŞ SORULAR CEVAP ANAHTARI 192

KAYNAKÇA 194
DİZİN 195

1 ÖLÇME VE DEĞERLENDİRMEYLE İLGİLİ TEMEL KAVRAMLAR

ANAHTAR KAVRAMLAR

Sabit, değişken ve türleri, ölçme, ölçmede birim, ölçmede sıfır noktası, ölçek ve türleri, değerlendirme, ölçüte göre değerlendirme türleri, amaca göre değerlendirme türleri

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Ölçme kavramını tanımlayabilecek,
- Ölçme türlerini tartışabilecek,
- Ölçmede kullanılan birimin özelliklerini açıklayabilecek,
- Ölçmede kullanılan birim türlerine örnekler verebilecek,
- Ölçmede tanımlanan iki farklı sıfır noktasının farkını açıklayabilecek,
- Ölçek türleri arasındaki farklılıkları tartışabilecek,
- Her bir ölçek türüyle ilgili örnekler verebilecek,
- Değerlendirmeyi tanımlayabilecek,
- Ölçme ile değerlendirme arasındaki farkı tartışabilecek,
- Ölçüte göre değerlendirme türlerini açıklayabilecek,
- Amaca göre değerlendirme türlerini tartışabileceksiniz.

İÇİNDEKİLER

Sabit ve değişken kavramları
Ölçme ve türleri
Ölçmede birim
Ölçmede sıfır noktası
Ölçek ve türleri
Değerlendirme
Çalışma soruları

BAŞLAMADAN ÖNCE

Sizce ölçme ve değerlendirme arasındaki farklılık nedir?

egitim@hotmail.es

GİRİŞ

Değerli okuyucular, öncelikle kitabımızın ana konusu olan "ölçme" ve daha sonra da "değerlendirme" kavramlarını tanımaya çalışacağız. Günlük hayatımızda da sıkça kullandığımız, hepimize tanıdık gelen bu kavramlara ilişkin bildiklerimizin ne kadarının doğru ne kadarının hatalı olduğunu anlayacak, bilmediğimiz ve bildikçe seveceğimiz yeni pek çok şey öğreneceğiz. Öğrendikçe bu dersi daha çok sevecek, sevdikçe de daha çok öğrenmek isteyeceğiz. İlk bölümde anlatılmaya çalışılan temel kavramlardan sonra diğer bölümlerle de daha detaylı olarak bu derse ilişkin bilgilerimizi genişletmeye çalışacağız...

ÖLÇME VE DEĞERLENDİRMEYLE İLGİLİ TEMEL KAVRAMLAR

- "Ölçme" dediğimizde aklınıza ne geliyor? Cetvel, terazi, bu sabah tartıda gördüğümüz kilomuz...
- En son ne zaman "ölçme" yaptığımızı bir düşünelim? Dün sabah boyumuzu ölçtük, okula giderken kaç dakikada okula ulaştığımızı hesapladık? Bugünün sıcaklığını öğrenmek için termometreye baktık...
- Peki yukarıdaki sorulara verilen cevapları da ek çokde buludurursak "ölçme" kavramını nasıl tanımlayabiliriz?

Ölçme, yaşamımızdaki ek çok şeyin sürekli değişmesinden ve insanoğlunun da bu değişimleri merak edip gözlemleme ihtiyacından ortaya çıkan bir süreçtir. Diyebiliriz ki ölçmenin gözlem kaynağı "değişkenler"dir. Burada öncelikle değişken ve değişkene ilişkin kavramları tanımlayalım:

SABİT VE DEĞİŞKEN KAVRAMLARI

Değişken: Kişiden kişiye, durumdan duruma, nesneden nesneye değişen özelliklere denir. Bir başka deyişle, iki ya da daha fazla değer alan özelliklere "değişken" adı verilir. İnsanların boy uzunlukları, zekâ düzeyleri, cinsiyet değişkene örnek olarak verilebilir.

Sabit: Kişiden kişiye, durumdan duruma, nesneden nesneye değişmeyen özelliklere denir. Sadece tek bir değer alabilen özelliklere sabit adı verilir. "Sabit" özellikler fen bilimlerinde sosyal bilimlerden daha fazla yer almaktadır. Örneğin; Einstein'ın enerji denklemindeki ($E=mc^2$) "c" değeri (c: ışık hızını göstermektedir) ve pi (π) sayısı birer sabittir. Bir araştırmacı sadece bir grup kız öğrencinin okumayı öğrenme hızı üzerinde bir araştırma yapıyorsa, bu çalışmada cinsiyet özelliği de (çalışmada hiç erkek öğrenci olmadığı için) bir sabit olmaktadır.

Değişkenler aldıkları değerlere göre farklı isimler almaktadır:

Nicel değişken: Bu tür değişkenler sayısal değişkendir. Diğer bir ifade ile sayılarla ifade edilebilebilirler. Örneğin; kişilerin boy uzunlukları, ağırlıkları, zekâ düzeyleri vb...

Nitel değişken: Bu tür değişkenler sıfatsal değişkendir. Sıfat ya da sembollerle ifade edilebilirler. Örneğin; cinsiyet, renkler, medeni durum vb... Bu özellikler sayılarla ifade edilse bile bu sayısal gösterimler matematiksel anlamda bir büyüklük, nicelik ifade etmezler. Örneğin; kız öğrencileri "1"erkek öğrencileri "2" ile göstermek, matematiksel olarak bir büyüklüğü ifade etmez ve bu sayılarla hiçbir işlem yapılamaz.

Sürekli değişken: Reel sayılarla eşleştirme yapılabilen değişkenlerdir. Daha açık olarak ifade edersek; iki değer arasında sonsuz sayıda başka değer yer alabiliyorsa bu tür değişkenlere sürekli değişken denir. Örneğin; uzunluk, ağırlık, zekâ, sıcaklık vb...

Süreksiz (kesikli) değişken: Tam sayılarla eşleştirme yapılabilen değişkenlerdir. Daha açık olarak ifade edersek; iki değer arasında ancak bir ya da sayılabilir çoklukta değer yer alabiliyorsa bu tür değişkenlere süreksiz değişken denir. Örneğin; cinsiyet, mezuniyet durumu, medeni durum vb...

Değişkenler ayrıca, bir araştırmada araştırmanın amacına göre değişebilen isimlerle de ifade edilebilir:

Bağımsız (etki eden) değişken: Araştırmacının kontrolü altında olan, düzenleyip, değiştirebildiği değişkendir. Daha açık olarak ifade edersek; araştırmada etkisinin olup olmadığı araştırılan değişkendir.

Bağımlı (etkilenen) değişken: Araştırmada diğer değişkene ya da değişkenlere (bağımsız değişkene) karşı tepkisinin araştırıldığı değişkenlerdir. Daha açık olarak ifade edersek; bağımsız değişkene bağlı olarak değişiminin araştırıldığı değişkendir.

Kontrol değişkeni: Araştırmacının araştırma ortamında sabit tuttuğu, araştırmaya etkisinin olmasının istenmediği için nötrlenmeye çalışılan değişkenlerdir.

Karıştırıcı (ara) değişken: Araştırmada doğrudan gözlenmeyen ancak bağımsız değişken ile bağımlı değişken arasındaki ilişkiyi etkileme olasılığı olan, kontrol altına alınamayan değişkendir. Bu tür değişkenler daha çok sosyal bilimlerde karşımıza çıkar. Fen bilimlerinde bu tür değişkenler laboratuar ortamında kontrol altında tutularak ortadan kaldırılabilirken, sosyal bilimlerde bu değişkenlerin çok fazla olması ve kontrolünün bazı durumlarda imkânsız olması, bu değişkenlerin araştırmada bulunmasını kaçınılmaz hale getirmektedir.

ÖRNEK: 8. Sınıf öğrencilerine buluş yoluyla öğrenme stratejisini uygulayan bir öğretmenin bu stratejinin öğrencilerinin başarıları üzerindeki etkisini araştırdığını düşünelim. Bu çalışmada;

- Buluş yoluyla öğrenme stratejisinin etkisinin olup olmadığı araştırıldığı için öğretmenin uyguladığı öğretme stratejisi bağımsız değişken,

egitim@hotmail.es

- Başarının üzerindeki etki araştırıldığı için öğrencilerin başarı düzeyi bağımlı değişken,
- Çalışmanın uygulandığı sınıf düzeyi sabit tutulduğu için 8. Sınıflar kontrol değişkeni
- Başarı düzeyine etki edebilecek, öğretmenin kontrol altında tutamadığı, öğrencilerin sosyo-ekonomik düzeyi, çalışma saatleri, derse karşı tutumları ve daha pek çok dış değişkende karıştırıcı değişkenler olmaktadır.

ÖLÇME VE TÜRLERİ

ÖLÇME: *Bir değişkenin (ölçülmek istenilen bir özelliğin) gözlenerek, gözlem sonucunun sayı ya da sembolle eşleştirilmesi sürecidir.* (Turgut, 1987).

- Bu süreç "ölçme" olarak tanımlanırken sürecin sonunda elde edilen sayı ya da sembole ölçme sonucu ya da "ölçüm" adı verilir.
- Ölçülen özelliğin hangi tür sayı ya da sembolle eşleştirileceğine karar verme işlemine de "ölçme kuralı" adı verilir.

Örnekler:

- Bugün hava sıcaktır.
- Burada ölçülen özellik havanın sıcaklığıdır. Ölçme: Bugün hava sıcaktır. Ölçüm: Sıcak. Ölçme kuralı: Ölçülen özelliğin bir sıfat ile belirlenmesi.
- Bugün havanın sıcaklığı 30^0C'dir.
- Burada ölçülen özellik yine havanın sıcaklığıdır. Ölçme: Bugün havanın sıcaklığı 30^0C'dir. Ölçüm: 30^0C. Ölçme kuralı: Sıcaklığın ^{0}C cinsinden bir sayı ile belirlenmesi.
- Ayşe'nin boyu uzundur.
- Burada ölçülen özellik Ayşe'nin boy uzunluğudur. Ölçme: Ayşe'nin boyu uzundur. Ölçüm: Uzundur. Ölçme kuralı: Boy uzunluğunun bir sıfat olarak belirlenmesi.
- Ayşe'nin boyu 172 cm'dir.

Burada ölçülen özellik yine Ayşe'nin boy uzunluğudur. Ölçme: Ayşe'nin boyu 172 cm'dir. Ölçüm: 172 cm. Ölçme kuralı: boy uzunluğunun cm cinsinden bir sayı ile belirlenmesi.

Diğer örnekler:

- Sevda'nın matematik puanı 89'dur.
- Sınıfta 21 kız öğrenci vardır.
- Arabanın hızı 100km/saat'tir.
- Metin'in göz rengi mavidir.

Yukarıdaki örneklerden de anlaşılacağı üzere ölçme sonuçları bazen bir sayı bazen bir sıfat ile ifade edilebilmektedir. Ancak sayısal olarak belirtilen ölçme sonuçları çok daha net, anlaşılır ve nesnel olma özelliğine sahiptirler. Sayılarla ifade edilen

ölçme sonuçlarıyla iletişim çok daha kolay ve anlaşılır olmaktadır. Sembollerle ifade etmek (örn: Bugün hava sıcaktır.) subjektiftir, farklı kişiler farklı yorumlar yapabilir ("Bugün hava sıcak değil, soğuktur." Diyebilir). Ancak sayılarla ifade çok daha kesin ve anlaşılır olmaktadır (Bugün havanın sıcaklığı 30^0C'dir.).

ÖLÇME TÜRLERİ

Doğrudan (Temel) Ölçme	Dolaylı (Göstergeyle) Ölçme	Türetilmiş Ölçme
Ölçmek istenilen özellik doğrudan gözleniyorsa ya da bu özelliğin kendisiyle doğrudan ilişkili bir ölçme aracı ile ölçülebiliyorsa, bu tür ölçmelere doğrudan (temel) ölçme denir. *Doğrudan ölçme yapmak demek araç kullanmamak anlamına gelmez!* *Doğrudan ölçmede asıl olan; ölçülen özellik dışında bir başka özelliğin ölçme işlemine karışmamasıdır!* *Örneğin:* "Öğrencilerin cinsiyetlerini gözlemleyerek kız-erkek sembolleri ile eşleştirmek" doğrudan gözlemleyerek yapılan bir doğrudan ölçme işlemidir. "Bir masanın uzunluğunu cetvel ile ölçmek" masanın uzunluğunu, uzunluğu ölçen bir ölçme aracıyla, araya başka hiçbir özellik karışmaksızın ölçülebildiği için yine bir doğrudan ölçme işlemidir.	Bir özellik doğrudan gözlemlenemez ve ölçülemez, ancak bir başka özellik yardımıyla ölçülebilirse, bu tür ölçmelere dolaylı (göstergeyle) ölçme denir. *Dolaylı ölçmede ölçülmek istenilen özellik, başka bir özellik yardımıyla dolaylı olarak ölçülebilmektedir.* *Örneğin;* "Havanın sıcaklığını termometre ile ölçmek" dolaylı ölçmedir. Çünkü burada sıcaklık, termometrenin cam borusu içindeki civa ya da alkolun sıcaklıkla genleşerek boru içinde yükselmesiyle oluşan yüksekliğe bağlı olarak ölçülmektedir. "Öğrencinin matematik dersindeki başarısını ölçmek" de dolaylı ölçmedir. Matematik başarısı doğrudan gözlemlenemez, ancak var olduğu düşünülen matematik başarısı, öğrencinin bu durumunu ortaya çıkaracağı düşünülen sorulara verdiği cevaplarla gözlemlenmeye çalışılır ve böylece dolaylı olarak ölçülebilir. **NOT:** *Sosyal bilimlerde, bilişsel ve duyuşsal alanlarda yapılan ölçmeler dolaylı ölçmedir.*	Ölçülmek istenilen özellik kendisinden farklı iki ya da daha fazla özelliğin arasındaki matematiksel bir bağıntı (toplama çıkarma, çarpma ya da bölme işlemleri) yardımıyla ölçülebiliyorsa, bu ölçme işlemine türetilmiş ölçme adı verilir. *Örneğin;* "Hız= Yol / Zaman" Burada hız özelliği kendisinden farklı yol ve zaman gibi iki özelliğin matematiksel bir bağıntısı olan bölme işlemi yardımıyla ölçülebilmektedir. Yani hız ölçülürken türetilmiş ölçme işlemi yapılmaktadır. *Örnekler;* Yoğunluk= Kütle / Hacim Öğrencinin puanı=%70 sınavdan alınan puanlar+ %30 derse devam Bir sınıftaki öğrencilerin puanları ortalaması= Öğrencilerin puanları toplamı/ öğrenci sayısı vb…

Yukarıda da belirtildiği gibi ölçme sonuçlarının daha objektif ve kesin olabilmesi için sayılarla ifade edilmesi tercih edilir. Ölçme sonuçlarının sayılarla ifadesi için de genellikle ölçme araçlarından yararlanılır (cetvel, terazi, termometre, sınavlar vb…).

egitim@hotmail.es

Ölçme araçları bizlerin çok daha net ve doğru sonuçlara ulaşabilmemize yardımcı olurlar. Ölçme yaparken araç kullanmak, ölçmeye ilişkin iki kavramın daha bilinmesini gerektirir. Bu kavramlardan biri "ölçmede birim" kavramı diğeri ise "ölçmede sıfır noktası"dır.

ÖLÇMEDE BİRİM

Birim: Ölçmede kullanılan en küçük standart miktara verilen addır. Örnek olarak metre, gram, ^{0}C, puan verilebilir. Ölçmede iki tür birim vardır: doğal birim ve yapay birim.

Doğal birim: Doğada kendiliğinden var olan birimdir. İnsanlar tarafından birim olarak belirlenmesine gerek duyulmayan fiziksel olarak zaten var olan birimlerdir.

Tüm sayma işlemlerinde kullanılan birim doğal birimdir.

Örneğin; Sınıftaki öğrencilerin, bahçedeki ağaçların, masanın üstündeki kalemlerin veya bir sınav kâğıdındaki hataların sayısını ölçerken kullanılan birimler hep doğal birimdir. Bu ölçme işlemlerinde sayılan her bir kişi ya da nesnenin kendisi bir birimdir.

Yapay birim: Uzmanlar tarafından tanımlanmış, ölçmede pratikliği sağlamak için sonradan belirlenmiş standart miktarlara yapay birim adı verilir.

Örneğin; metre, santimetre, derece, kilogram, meridyen, vb...

Örneklerle de ifade edilen yapay birimler doğal olarak fark edilebilen, kendiliğinden var olan büyüklükler değildir. İnsanlar tarafından sonradan belirlenen standart miktarlara isimler verilerek oluşturulmuştur.

Ağırlık ölçerken hafif-ağır gibi ifadeler kullanmak yerine gram ya da kilogram, sıcaklık ölçerken sıcak-soğuk demek yerine ^{0}C ya da Kelvin cinsinden bir birim kullanmak ölçme sonuçlarının daha duyarlı ve güvenilir olmasına imkân tanır.

Ölçmede Birimin Özellikleri:

1. **Eşitlik:** Birimin her yerde ve kullanıldığı ölçme aracında daima aynı miktarı, büyüklüğü göstermesi özelliğine denir.

Örneğin; "1 cm" büyüklük, dünyanın her yerinde de, kullanıldığı cm bölmeli cetvelde de hep aynı miktarı gösterir. Hâlbuki bir uzunluğu karışımızla ölçsek, her defasında aynı büyüklüğü elde edemeyebiliriz ya da bir başkasının karışı bizim karışımızla aynı büyüklüğe sahip olmayacaktır. Bu sebeple "karış" bir birim olarak eşitlik ilkesine sahip değildir.

2. **Genellik:** Birimin olabildiğince geniş kitleler tarafından kullanılıyor ve anlaşılıyor olması özelliğidir.

Örneğin; Ağırlık ölçerken kg kullanırız. Hâlbuki "okka" kullanmış olsaydık, elde ettiğimiz sonuçtan pek çok kimse bir şey anlamamış olacaktı. Yani "okka" genellik özelliğine sahip bir birimdir diyemeyiz. Hâlbuki "kg" birimini pek çok kimse anlamakta ve kullanmaktadır. Bu da kg'ın genellik özelliğine sahip olan bir birim olduğunu gösterir.

3. **Uygunluk (Kullanışlılık):** Birimin, ölçülen özelliğin büyüklük ya da miktarına uygun olarak seçilebilmesi özelliğidir.

Örneğin; ağırlık olarak; kışlık kömür ihtiyacımız için kömür alırken "ton" birimini kullanırken, kuyumcudan altın alırken "gram" birimini kullanırız. Benzer şekilde; şehirlerarası yolların uzunluğunu ölçerken km, defterimizin uzunluğunu ölçerken cm kullanırız. Böylece birimin uygunluk (kullanışlılık) özelliğini kullanmış oluruz.

ÖLÇMEDE SIFIR NOKTASI

Bir ölçme aracıyla ölçme yapılırken bir başlangıç noktası alınır. Ölçmeye başlangıç olan bu noktaya ölçmede sıfır noktası adı verilir. İki tür sıfır noktası vardır: mutlak (gerçek) sıfır noktası ve göreceli (izafi, itibari) sıfır noktası.

1. **Mutlak (gerçek) sıfır noktası:** Ölçmeye başlangıç olarak seçilen sıfır noktası gerçek, mutlak yokluğu göstermektedir. Sıfır noktası ölçülen özelliğin gerçekten var olduğu noktadır ve bu nokta öncesi o özellik gerçekten yoktur, bulunmamaktadır.

Örneğin; boyumuzu ölçerken başlangıç olarak kabul ettiğimiz sıfır noktası boyumuzun başladığı noktadır, sıfır noktası öncesi boyumuz yoktur. Olmayan bir cismin ağırlığı da yoktur, bir cisim (nesne) var olmaya başladığı noktada ağırlığı da başlar, böylece ağırlıkta kullandığımız, sıfır dediğimiz nokta da gerçek yokluğu göstermekte olup, gerçek sıfır noktasını ifade eder.

2. **Göreceli (izafi, itibari) sıfır noktası:** Ölçmeye başlangıç olarak alınan sıfır noktası aslında o özelliğin gerçek yokluğundan başlamaz. Pratikte ölçmeye başlangıç olabilmesi için seçilen bir noktadır ve sıfır noktası olarak gösterilir. Kısacası göreceli sıfır noktası mutlak yokluğu göstermez.

Örneğin; Sıcaklığı termometre ile ölçerken kullandığımız "0" noktası yani 0^0C, sıcaklığın o noktada ve ondan öncesinde yok olduğunu göstermez. 0^0C demek "sıcaklık yok" anlamı taşımaz. Buradaki "0" noktası ölçmede pratik olarak bir başlangıcı göstermek üzere seçilmiş noktadır.

egitim@hotmail.es

ÖLÇEK

Ölçek kelimesi "birim", "derecelenmiş araç" gibi değişik anlamlarda kullanılmaktadır (Baykul, Gelbal ve Kelecioğlu, 2003). Ancak, aşağıda ifade edilen ölçek türleri incelendiğinde de görüleceği gibi biz ölçeği ölçme sonuçlarının matematiksel özellikleri olarak ele alacağız. Ölçme sonucu elde edilen sayı ya da sembollerin matematik özelliklerine göre sınıflandırılmasıyla ölçek türleri aşağıdaki şekliyle gösterilebilir:

ÖLÇEK TÜRLERİ VE ÖZELLİKLERİ

ÖLÇEK TÜRÜ	ÖZELLİKLERİ
Adlandırma (Sınıflandırma) Ölçeği	Nesneler belli bir özelliğe sahip olup olmamasına göre adlandırılmaktadır. Amaç sadece ölçme yapılan özellik açısından sınıflandırarak, ölçülen nesneleri (ya da bireyleri) birbirlerinden ayırmaktır. Örneğin; insanları cinsiyet özelliğine göre kadın-erkek, medeni hallerine göre evli-bekâr-dul şeklinde sınıflandırmak. Şehirlere plaka numarası vermek, futbolcuların sırt numaraları vb... sadece isim vermek adına yapılan ölçme işlemleridir. Ölçme sonuçlarının bazen sayılarla da gösterilebilmesine rağmen (Örneğin cinsiyet özelliğini ölçerken, kız öğrencileri "1" erkek öğrencileri "2" ile göstermek gibi) bu sonuçların sayısal anlamı yoktur, miktar belirtmez. Burada ölçmeden çok adlandırma yapılmaktadır. Özelliklere verilen sayıların tek fonksiyonu yalnızca sınıf, tip ya da kategori bildirmektir. Özelliklere verilen sayılar üzerinden hiçbir matematiksel işlem yapılamaz. İstatistiksel olarak da sadece frekans (bir özelliğin "kaç tane" olduğunun, "kaç kez tekrar ettiğinin" sayısı) ve mod (en çok tekrar sayısı ya da en büyük frekans değeri) hesaplanabilir. NOT: Frekans ve mod kavramlarıyla ilgili daha geniş bilgi 7. bölümde verilmiştir. Sayılarla hiçbir matematiksel işlem yapılamaz. Örneğin; kızlar "1", erkeklere "2" olarak tanımlanmışsa, "1+1=2" işlemi, burada anlamsız olmaktadır. Sınıflama ölçeğinde *sıfır noktası anlamsızdır ve birim yoktur.*
Sıralama Ölçeği	Bu ölçek türünde, nesneler belli bir özellik bakımından bir sıraya dizilmektedir. Nesneler bir özelliğe en az sahip olandan, en çok sahip olana (ya da en çok sahip olandan, en az sahip olana) doğru sıralanır ve bir sıra numarası verilir. Böylece her nesne bir "sayma sayısı" (1., 2., 3., vb....) ile eşleştirilmiş olur. Örneğin; sınıftaki öğrencileri boy sırasına dizmek, sınavdan alınan ham puanları büyükten küçüğe doğru sıralamak vb. Nesnelere verilen sayılar, nesnelerin belli bir özelliğe, birbirlerine göre daha az ya da daha çok sahip olma durumunu ifade etmektedir. Fakat sıraları gösteren sayılar arasındaki farklar, özelliğe sahip olma derecesini ve nesneler arasındaki farkın miktarını yansıtmaz. Bir başka deyişle, 1.sıradaki ile 2. Sıradaki özellik arasındaki farklılık 2. İle 3. Arasındaki farklılıkla aynıdır ya da farklıdır gibi bir bilgiyi göstermez. Sıralama ölçeğinde *sıfır noktası ve birim yoktur.* Sıralama ile elde edilen ölçme sonuçları ancak iki çeşit bilgi taşır: Sıralamada ayrı ayrı yer tutan iki birey, sıralama boyutunda birbirinden farklıdır. Büyüklük sırasında önce gelen bir birey, sonra gelen bireylerin hepsinden daha büyüktür. Örneğin; A>B ve B>C ise A>C'dir. Sıralama ölçeğiyle elde edilen sonuçlar arasında hiçbir matematiksel işlem yapılamaz. İstatistiksel olarak da sadece frekans, mod ve medyan (ortanca) hesaplanabilir.

egitim@hotmail.es

Eşit Aralıklı Ölçek

Eşit aralıklı ölçekte *göreceli bir sıfır noktası bulunmaktadır*.

Eşit aralıklı ölçekte *birim vardır*.

Ya iki noktası belirlenip bu iki nokta arası eşit aralıklarla bölünerek ya da bir noktası belirlenip bu noktadan itibaren tanımlanabilir bir birimle bölümlenerek geliştirilebilir.

Başlangıç (sıfır) noktası keyfi olarak saptanır ve bu noktanın sağa, sola, ileriye, geriye kaydırılması sonucu değiştirmez.

Tanımlanmış değişmez bir birim vardır. Ölçek, başlangıç noktasından itibaren her iki yana doğru bu birimle eşit olarak bölmelenmiştir.

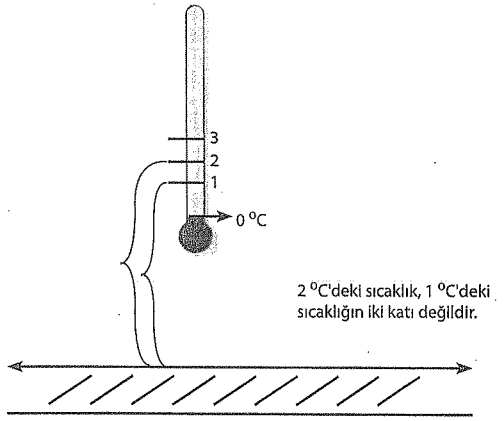
Bu ölçekle, nesnelerin belli bir özelliğe sahip oluşunun ve nesneler arasındaki bu özelliğe sahip oluş farkının miktarını da sayısal olarak belirtme olanağı vardır.

Ölçek eşit birimlerle bölmelenmiş olduğundan, birimler toplanıp çıkarılabilir. Örneğin, takvimler, sıcaklık derecesi vb.

Ölçekte sıfır, o noktada ölçülen özelliğin hiç bulunmadığı anlamına gelmez. Buradaki sıfır noktası doğal ya da mutlak sıfır değildir. Buradan alınan verilerle toplama ve çıkarma işlemi yapılabilir.

Sıfır noktası mutlak yokluğu göstermediği için, ölçülen özelliğin gerçek durumuna ilişkin oransal işlemler yani çarpma ve bölme yapmak mümkün değildir.

Örneğin; 20 °C'deki sıcaklık 10 °C'deki sıcaklığın iki katı değildir. Çünkü başlangıç noktası gerçek yokluğu göstermez.



Oran Ölçeği

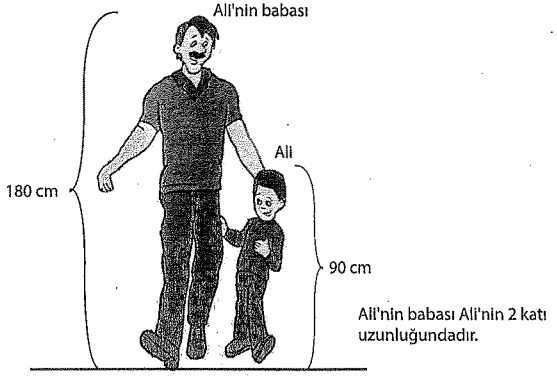
Eşit oranlı ölçekte *mutlak (gerçek) sıfır noktası bulunmaktadır*.

Eşit oranlı ölçekte *birim vardır*.

Bir değişkenin iki gerçek değerinin birbirine oranının, bu değerlere karşılık olan iki ölçme sonucunun birbirine oranının eşit olması durumudur.

Ölçme sonuçları üzerinde her türlü matematiksel işlem yapılabilir.

Örneğin; 180 cm. uzunluğundaki bir kimse 90 cm. uzunluğundaki bir kimsenin iki katı uzunluktadır diyebiliriz. Çünkü boy ölçerken kullandığımız uzunluk ölçüsünün sıfır noktası gerçek yokluğu göstermektedir.



NOT: *"Ölçek ve türleri" konusuyla ilgili daha detaylı bilgi için kaynakçada yer alan (Turgut, 1987) ve diğer ölçme ve değerlendirme kitaplarından yararlanabilirsiniz.*

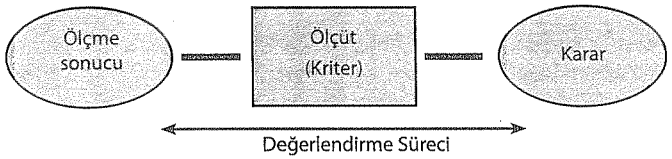
egitim@hotmail.es

I. **Yukarıdaki ölçek türleri en az bilgi veren (en basit) ölçekten en çok bilgi veren (en üst düzey) ölçeğe göre sıralanmıştır. Üst düzeydeki bir ölçek türünde bulunan bir ölçme sonucunu alt düzeydeki ölçeğe taşımak mümkündür. Ancak alt düzeydeki bir ölçek türünde bulunan bir ölçme sonucunu üst düzeydeki ölçeğe taşımak ya imkânsız ya da çok zordur. Örneğin; Sınıflandırma ölçeğinde olan cinsiyet özelliğini ele alalım. Cinsiyet özelliği bir çokluk, büyüklük belirtmediği için sıralanamaz, o halde sıralama ölçeğinde olamaz. Cinsiyet özelliğinin birimi ve sıfır noktası da anlamsız olduğu için eşit aralıklı veya eşit oranlı ölçek türünde olması da söz konusu olamaz. Benzer durum sınıflandırm ölçeğinde yer alan tüm diğer özellikler (medeni durum, mezun olunan bölüm vb.) için de geçerlidir.**

II. **Bilişsel ve duyuşsal davranışlara ilişkin yapılan ölçme sonuçları en fazla eşit aralıklı ölçek düzeyindedirler.**

DEĞERLENDİRME

Günlük konuşma dilinde her ne kadar ölçme ile değerlendirme birbiri yerine kullanılıyor olsa da birbirlerinden çok farklı kavramlardır. Ölçme kavramını yukarıda açıkladık. Değerlendirme ise ölçme sonuçlarının bir ölçüt (kriter) ile karşılaştırılarak sonuca varma sürecini içerir. Tanımdan da anlaşılacağı üzere ölçme, değerlendirme sürecinin ilk ve olmazsa olmaz basamağıdır. Ancak değerlendirmenin yapılabilmesi için bir de "ölçüt"e ihtiyaç vardır. Bu ölçüt ile karşılaştırılan ölçme sonucu ya da sonuçlarına göre karara varılarak değerlendirme işlemi gerçekleştirilmiş olur. Böylece değerlendirme kısaca üç basamaktan oluşmaktadır diyebiliriz: 1. Ölçme, 2. Ölçüt ve 3. Karar verme, sonuca varma.



Ölçme işlemi hem subjektif hem de objektif olabilir. Ancak ölçmenin objektif olması istenir. Bu nedenle ölçme araçları kullanılarak elde edilen ve sayılarla ifade edilen ölçme sonuçları tercih edilir. Değerlendirme ise subjektif bir süreçtir. Ölçütü belirleyen kişi veya kişilere göre değerlendirme sonucu değişebilir. Değerlendirmenin de olabildiğince objektif yapılabilmesi için ölçütlerin çok dikkatli ve doğru olacak şekilde belirlenmesi çok önemlidir.

Değerlendirme, kullanılan ölçüte göre ve kullanılış amacına göre çeşitlilik göstermektedir. Şimdi bu türleri tanımlamaya çalışalım:

Ölçüte Göre Değerlendirme Türleri

Mutlak Değerlendirme	Bağıl Değerlendirme
- Mutlak değerlendirmede, mutlak ölçüt adı verilen önceden belirlenmiş bir kriter kullanılır.	- Bağıl değerlendirmede, bağıl ölçüt adı verilen ölçme sonuçlarına göre sonradan belirlenmiş bir kriter kullanılır.
- Mutlak ölçütle yapılan değerlendirmede, her bir bireyin (öğrencinin) değerlendirme sonucu, içinde bulunduğu sınıf, topluluk, grup vb.den bağımsızdır.	- Bağıl ölçütle yapılan değerlendirmede, her bir bireyin (öğrencinin) değerlendirme sonucu, içinde bulunduğu sınıf, topluluk, grup vb.den etkilenir.
- "Her koyun kendi bacağından asılır." Deyimi bu değerlendirme için uygun olacaktır.	- Her öğrenci sınıftaki diğer arkadaşlarının da başarılarını önemsemek durumundadır.
- Örneğin; Bir sınavdaki 50 sorudan 30'unu doğru yanıtlamış olmak, sınavdan 100 üzerinden en az 65 almak ya da 10 hedef-davranışın 7'sini başarmış olmak gibi bireyin sadece kendi başarısıyla değerlendirileceği durumlar, mutlak değerlendirme kapsamına girer. Eğer 50 sorudan 30'unu doğru yanıtlamış olmanın başarı sayılacağı bir sınavda bir öğrenci, örn; Can, 35 soruyu doğru yanıtlamışsa, sınavın uygulandığı sınıftaki diğer arkadaşlarının tümü 50 soruyu da doğru cevaplasa ya da tümü hiçbir soruyu da doğru bilmemiş olsa Can bu sınav için belirlenen ölçüte göre daima başarılı sayılacaktır.	- Örneğin; Bir sınav sonucunda, sınıftaki öğrencilerin aldıkları puanların aritmetik ortalamasının üstünde puan almış olmanın başarının ölçütü olması ya da tüm öğrenciler içinde sınavdan en yüksek puan alan ilk üç öğrencinin arasına girmenin başarı sayılması gibi durumlarda öğrencinin sadece kendi başarısı değil içinde bulunduğu grubun başarı durumu da başarılı sayılabilmesi için önem taşımaktadır.
- Mutlak değerlendirmenin yapıldığı sınıf ortamlarında, ders çalışırken, öğrenciler arası bilgi alışverişinde bulunmak, yardımlaşmak daha yaygındır (Elbette sınav ortamında değil ☺).	- Bağıl değerlendirmenin yapıldığı sınıf ortamları öğrencileri daha çok çalışmaya sevk eder. - Ancak bu tür değerlendirmenin yapılması, öğrenciler arası rekabeti arttıracağı gibi, öğrencilerin arasındaki dayanışmanın azalmasına sebep olcaktır.

egitim@hotmail.es

Amaca Göre Değerlendirme Türleri

Tanıma- Yerleştirmeye Dönük Değerlendirme	Biçimlendirme ve Yetiştirmeye Dönük Değerlendirme (Formatif Değerlendirme)	Değer Biçmeye Dönük Değerlendirme (Summatif Değerlendirme)
Eğitim sürecinin başında yapılır. Amaç, öğrenciyi tanımak ve onu uygun olan eğitim programına yerleştirmektir. Bir bakıma öğrencinin hazırbulunuşluk düzeyini belirlemeyi amaçlar. Bu değerlendirmede amaç öğrencinin aldığı nota göre en doğru programa (ya da sınıfa) yerleşmesini sağlamak olduğu için ASIL AMAÇ öğrenciye NOT vermek DEĞİLDİR. Her öğrencinin kendisine uygun olan programa yerleştirilmesi, zaman, para ve emek açısından ekonomi sağlar. Bu değerlendirme türüne; seviye tespit sınavları, muafiyet sınavları vb. örnek olarak verilebilir.	Eğitim süreci devam ederken, belirli periyotlarda yapılır. Amaç, öğrencinin eksiklerini, hatalarını belirleyerek bunları düzeltmek, öğrenciye geri bildirimde bulunmaktır. Bir bakıma eğitim süreci bitmeden, belirli aralıklarla eğitimin hedefine ulaşıyor olup olmadığı hakkında fikir sahibi olmak için yapılır. Böylelikle eksiklikler ya da yapılan yanlışların birikerek başa çıkılmaz bir hal alması engellenmiş olur. Eğitimde, istenen hedefe doğru daha sağlam adımlarla ilerleme sağlanır. Bu değerlendirmede amaç öğrencinin aldığı nota göre geri bildirimde bulunmak olduğu için ASIL AMAÇ öğrenciye NOT vermek DEĞİLDİR. Bu değerlendirme türüne; quizler, ünite testleri, tarama testleri örnek olarak verilebilir.	Eğitim sürecinin sonunda yapılır. Amaç, öğrenciye tamamlanan eğitim sonucundaki seviyesini, ulaştığı noktayı bildirmektir. Bir bakıma öğrenciye eğitim süreci sonunda aldığı nota göre bir değer biçilmiş olur. Bu değerlendirmede ASIL AMAÇ öğrenciye NOT VERMEKTİR. Çünkü artık hedeflenen eğitim süreci sona ermiştir. Bu değerlendirme türüne; bitirme sınavları, uzmanlık sınavları, sertifika sınavları, finaller örnek olarak verilebilir.

Burada yer alan değerlendirme türleri dışında eğitim programının değerlendirilmesi, öğretim yönteminin değerlendirilmesi vb. farklı değerlendirme türleri de bulunmaktadır.

NOT: *"Değerlendirme" konusuyla ilgili daha detaylı bilgi için kaynakçada yer alan (Özçelik, 1998) ve diğer ölçme ve değerlendirme kitaplarından yararlanabilirsiniz.*

DEĞERLENDİRME SORULARI

1. Aşağıdakilerden hangisi bir ölçme işlemi ifade eder?

A) Basketbol takımına seçmeler için, boy uzunluğunda bir alt sınırın belirlenmesi
B) Sınıf ortalamasının altında ve üstünde kalan puanların belirlenmesi
C) 25 soruluk final sınavında 15 doğru cevap veremeyen öğrencilerin başarısız sayılması
D) Doğru-yanlış testi için cevap anahtarının hazırlanması
E) Çoktan seçmeli bir testte öğrencilerin doğru yanıt sayılarının belirlenmesi

2. Aşağıdakilerden hangisi dolaylı ölçmeye bir örnektir?

A) Masanın üzerindeki kitapların sayısı
B) Öğrencilerin fen bilgisi dersindeki başarıları
C) Bir çuval patatesin ağırlığı
D) Sınıfın yüksekliği
E) Öğrencilerin boy uzunlukları

3. Aşağıdakilerden hangisi doğrudan ölçülebilir?

A) Kardeş sayısı
B) Zeka puanı
C) Tutum düzeyi
D) Akademik başarı
E) Kaygı düzeyi

4. Aşağıdakilerden hangisinde mutlak sıfır noktası vardır?

A) Bugün hava sıcaklığı 29 °C'dir.
B) Ağrı Dağı'nın rakımı 5174 metredir.
C) Sevgi'nin matematik sınavından aldığı puan 65'tir.
D) İstanbul'un fetih tarihi 1453'tür.
E) Ece'nin tarih sınavındaki doğru sayısı 15'tir.

5. Bir ölçüm yapılmış ve sonuç 25 birim bulunmuştur.

Bu sonuç 50 birimin gerçekten yarısını gösteriyorsa, bu birimin ölçek düzeyi aşağıdakilerden hangisidir?

A) Oranlı
B) Eşit aralık
C) Sıralama
D) Sınıflama
E) Sabit

egitim@hotmail.es

6.

I. Ahmet'in geçme notu pekiyidir.
II. Ayşe'nin sınavdan aldığı en yüksek puan 70'tir.
III. Geçme notu 60 olup Ali 70 alarak başarılı olmuştur.
IV. Hasan sınavdan 45 puan aldı.
V. Narin sınavdan 80 aldı ve geçti.

Yukarıdaki ifadeler aşağıdaki seçeneklerden hangisinde doğru olarak tanımlanmıştır?

	Ölçme	Değerlendirme
A)	I ve II	III,IV ve V
B)	I, II ve III	IV ve V
C)	II ve IV	I, III ve V
D)	III ve IV	I, II ve V
E)	IV ve V	I, II ve III

7. **Öğrencilerin matematik sınavından aldıkları ham puanlar hangi ölçek türündedir?**

A) Sınıflama
B) Sıralama
C) Eşit Aralıklı
D) Eşit Oranlı
E) Başarı ölçeği

8. Bir matematik öğretmeni dersine girdiği 8. sınıflara öğrenci merkezli öğretim yöntemini uygulamaktadır. Kullandığı bu yöntemin öğrencilerinin başarısına etkisini merak ederek bir araştırma yapmak ister.

Öğretmenin araştırmasında kullanacağı bağımlı değişken aşağıdakilerden hangisidir?

A) Matematik başarısı
B) 8. sınıf
C) Matematik öğretmeni
D) Öğrenci merkezli öğretim yöntemi
E) Öğrenciler

9. **Aşağıda verilen değerlendirme örneklerinin hangisinde kullanılan ölçüt bağıldır?**

A) Desten başarılı sayılabilmek için 100 puan üzerinden en az 65 almış olmak
B) Sınıfı geçebilmek için sınıf ortalamasının 5 puan üzerinde puan almış olmak
C) Pilot olabilmek için en fazla 180 cm. boy uzunluğuna sahip olmak
D) Başarılı sayılabilmek için sınav soruların en az %80'ini doğru cevaplamış olmak
E) Doktoraya kabul edilebilmek için yabancı dil sınavından en az 75 puan almış olmak

10.

I. Ayşe'nin fizik dersinden başarılı sayılabilmesi için alması için en az 65 puan alması gerekmektedir.
II. Ayşe'nin fizik sınavı puanı 85'tir.
III. Bu durumda Ayşe fizik dersinde başarılı olmuştur.

Aşağıdaki seçeneklerden hangisinde yukarıdaki ifadeler <u>sırasıyla</u> doğru olarak tanımlanmıştır?

A) Ölçme- Değerlendirme- Bağıl ölçüt
B) Mutlak ölçüt –Değerlendirme –Ölçme
C) Mutlak ölçüt - Ölçme–Değerlenirme
D) Bağıl ölçüt - Ölçme – Değerlendirme
E) Ölçme- Bağıl ölçüt - Değerlendirme

KPSS'DE ÇIKMIŞ SORULAR

1.

I. Herhangi bir özelliği, o özelliğe uygun bir araçla karşılaştırarak sonucu sayı veya sembollerle belirtmek
II. Bir gözlem sonucunu uygun bir ölçütle karşılaştırarak, bu sonuca ilişkin bir karara ya da yargıya ulaşmak

Yukarıda tanımlanan iki kavram, aşağıdakilerin hangisinde birlikte verilmiştir? (2001)

	I	II
A)	Değerlendirme	Ölçme
B)	Ölçme	Ölçek
C)	Değişken	Değerlendirme
D)	Ölçme	Değerlendirme
E)	Ölçek	Ölçme

2. **Aşağıdakilerden hangisi, ölçmenin yapıldığı öğrenci grubuna bağlı değerlendirme için bir ölçüt <u>değildir</u>? (2001)**

A) Öğrenilmesi beklenen kapsamın en az % 60'ını bilme
B) Sınıfta ortalama başarı düzeyi gösteren öğrenciden daha başarılı olma
C) Sınıfta en yüksek başarı gösteren ilk beş öğrencinin arasına girme
D) Sınıftaki öğrencilerin % 70'inden daha başarılı olma
E) Sınıf ortalamasının bir standart sapma üzerinde puan alma

3. **Aşağıdakilerden hangisi bir ölçme işlemidir? (2002)**

A) Bir yazılı yoklamadaki cevaplar için puanlama anahtarının hazırlanması
B) Sözlü yoklamada bir öğrencinin verdiği cevapların doğruluk derecesinin bir puanla belirtilmesi
C) Sınıfın bir sınavdaki başarı ortalamasının bulunması
D) Öğrencinin bir dersten başarılı sayılıp sayılmayacağına karar verilmesi
E) Bir sınavda başarılı sayılabilmek için gerekli olan en küçük puanın belirlenmesi

egitim@hotmail.es

4. **Bir öğretmen, değerlendirmede şu amaçlara hizmet etmeyi düşünmektedir:**

I. Öğrencilerin birbirlerine göre başarı düzeylerini belirleyerek ileri ve geri öğrenciler için ek önlemler alma

II. Ders süresince öğrencilerin öğrenme eksikliklerini belirleyerek bunları, tamamlayıcı etkinliklerle giderme

III. Dönem sonunda öğrencilerin bu dersten alacakları notu belirlerken sınıf ortalamasını dikkate alma

Bu öğretmen, I, II ve III numaralı amaçları için hangi tipte birer değerlendirme yapmalıdır? (2002)

	I. amaç	II. amaç	III. amaç
A)	Bağıl	Mutlak	Bağıl
B)	Mutlak	Mutlak	Bağıl
C)	Bağıl	Bağıl	Mutlak
D)	Mutlak	Bağıl	Bağıl
E)	Bağıl	Mutlak	Mutlak

5. **Bir öğrencinin aşağıdaki özelliklerden hangisi üzerinde yapılacak ölçme işlemi dolaylı ölçmeye örnektir? (2002)**

A) Akademik başarısı
B) Sahip olduğu kitap sayısı
C) Devam ettiği ders sayısı
D) Arkadaş sayısı
E) Boyunun uzunluğu

6. **Aşağıdaki ölçme sonuçlarından hangisi** "oranlı ölçek" **ile elde edilmiş olabilecek ölçülere bir örnektir? (2002)**

A) Anne babaların eğitim düzeyleri
B) Okulların üniversiteye giriş sınavı sonuçlarına göre sıraları
C) Akademik başarı ve okula yönelik tutum arasındaki ilişki
D) Öğrencilerin boyları
E) Sınıfların genel başarı düzeyleri

7. **Aşağıdakilerden hangisi seçme amacına yönelik bir testin kullanılmasını gerektirir? (2002)**

A) Öğrencilerin bir konudaki güçlü ve zayıf yönlerinin ortaya çıkarılması
B) Sıklıkla karşılaşılan yanlış öğrenmelerin belirlenmesi
C) İngilizcesi en ileri düzeyde olan üç kişinin belirlenmesi
D) Bir ünite işlendikten sonra, hedeflere ulaşılma derecesinin belirlenmesi
E) Bir programda hedeflerin ne kadarına ulaşıldığının belirlenmesi

8. **Aşağıdakilerden hangisi "ölçme" işlemine örnek <u>değildir</u>? (2003)**

A) Hangi puanı alanların derste başarılı sayılacağını gösteren bir sınır belirleme
B) Öğrencileri kısadan uzuna doğru dizerek onların boy sıralarını belirleme
C) Öğrencilerin projelerini 1 ile 10 arasında derecelendirme
D) Öğrencilerin yazılarındaki yazım yanlışlarını sayma
E) Bir sınavda, verilen cevaplara bakarak öğrencilerin kaç puan alacaklarını belirleme

9.

I. İş deneyimi sekiz yılın üstünde olanlar işe alınmıştır.
II. Ahmet, güreşten önce tartılmış; 130 kg gelmiştir.
III. Canan, ortalamanın üstünde puan alarak sınıfını geçmiştir.
IV. Pelin'in LES puanı 68'dir.

Yukarıdakilerin hangilerinde "değerlendirme" **işlemi yapılmıştır? (2003)**

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

10. **Aşağıdakilerden hangisi** "mutlak değerlendirme" **ye bir örnektir? (2003)**

A) Bir işletmeye, sınavda en yüksek başarıyı gösteren beş kişiyi almak
B) Ortalamanın bir standart sapma üstünde puan alanları başarılı kabul etmek
C) Yüksek lisans programında en başarılı üç kişiyi doktora programına almak
D) Sınıftaki başarı sıralamasında en geride kalan 10 kişiyi yetiştirme kursuna almak
E) Sınavda 65 ya da üstünde puan alanları başarılı kabul etmek

11. **Eğitimdeki ölçme ve değerlendirmeler genellikle şu amaçlara yöneliktir:**

I. Öğrenme eksiklerini belirleme
II. Öğrenme düzeyini saptama
III. Öğrencileri tanıma

Bu üç amaçla yapılacak ölçme ve değerlendirmelerin zamanı, aşağıdakilerin hangisinde doğru olarak verilmiştir? (2003)

	Dönemin (Dersin) Başı	Ünitelerin (Konuların) Sonu	Dönemin (Dersin) Sonu
A)	III	I	II
B)	I	III	II
C)	II	I	III
D)	I	II	III
E)	III	II	I

egitim@hotmail.es

12.

I. Elmas sınıfını "pekiyi" ile geçti.
II. Burç, resim sınavından 10 üzerinden 8 puan aldı.
III. Sınavda en yüksek puanı Erdem elde etti.
IV. Demet 160'ın üzerinde puan aldığı için sınavı kazandı.

ifadeleri, "ölçme" ve "değerlendirme" kavramlarını örneklendirme bakımından, aşağıdakilerin hangisinde doğru olarak gruplanmıştır? (2004)

	Ölçme	Değerlendirme
A)	I, III	II, IV
B)	I, II	III, IV
C)	II, III	I, IV
D)	II, IV	I, III
E)	II	I, III, IV

13. **Bir öğrencinin başarısını değerlendirmede aşağıdakilerden hangisinin dikkate alınması, bağıl değerlendirmeye örnek oluşturmaz? (2004)**

A) Programın hedefleriyle tutarlı öğrenmelerin
B) Standartlaştırılmış bir testin normlarının
C) Yetenek düzeylerine dayalı beklentilerin
D) Ön-test ve son-test puanları arasındaki farkın
E) Sınavdan elde edilen puanların dağılımının

14. Bir ölçme işleminde elde edilen ölçülerden biri 60, diğeri 30'dur.

Bu ölçülerden 60 ile ifade edilen büyüklük gerçekten de 30 ile ifade edilenin iki katı ise bu ölçme işleminde aşağıdaki ölçek türlerinden hangisi kullanılmış olabilir? (2004)

A) Sınıflama türü
B) Sıralama türü
C) Standart tür
D) Eşit aralıklı tür
E) Eşit oranlı tür

15. **Öğrencinin notunu belirlemede sınıf ortalamasını ölçüt alan bir öğretmen, ne tür bir değerlendirme yapmaktadır? (2005)**

A) Tanımaya yönelik
B) Biçimlendirici
C) Düzey belirleyici
D) Mutlak
E) Bağıl

16. **Aşağıdakilerden hangisi, değerlendirme sonunda ulaşılan kararın isabetsiz olmasının nedenlerinden biri olamaz? (2005)**

A) Ölçütün uygun olmaması
B) Öğrenme düzeyinin düşük olması
C) İşlemlerde yanlışlık yapılması
D) Gözlemlerin yanlı olması
E) Ölçme hatasının büyük olması

17. Eşit kollu terazinin bir kefesine nesneyi, diğer kefesine de gramları koyarak nesnenin kütlesi ölçülebilir. Bu işlem doğrudan ölçmeye örnektir. Bireylerin bilgi, zekâ, kaygı, tutum gibi özellikleri ise onlara sorulan sorulara verdikleri tepkiler gözlenerek ölçülür. Bu işlem de dolaylı ölçmeye bir örnektir.

Bu bilgilere göre, dolaylı ölçmeyi, doğrudan ölçmeden ayıran özellik aşağıdakilerden hangisidir? (2005)

A) Uygun olmayan koşullarda ölçme yapılması
B) Özelliklerin ayrıntılarına önem verilmesi
C) Basit ölçme araçlarından yararlanılması
D) Doğrudan karşılaştırma imkânı olmaması
E) Ölçmenin belirli aralıklarla yapılabilmesi

18.

I. Matematik sorularının %75'ini, fizik sorularının da %60'ını doğru cevaplayan öğrenciler bu dersleri başarmış sayılır.
II. İlk iki sorudan alınabilecek en yüksek puan 25'tir.
III. Sınavda 65'in altında puan alan öğrenciler dersi tekrar ederler.
IV. Türkçe ve matematik derslerinden başarısız olan öğrenciler bir sonraki döneme devam edemezler.

Yukarıdakilerden hangileri birer değerlendirme ölçütünü ifade etmektedir? (2006)

A) Yalnız I B) Yalnız II C) II ve III
D) III ve IV E) I, III ve IV

19. **Aşağıdaki öğrenci özelliklerinden hangisi doğrudan ölçülebilir? (2006)**

A) Zekâ puanı
B) Boy uzunluğu
C) Derse karşı tutum
D) Yetenek
E) Ders başarısı

egitim@hotmail.es

20. İş birliğine dayalı bir öğrenme uygulamasında gruplar öğrencilerin soyadlarının baş harfleri kullanılarak oluşturulmuştur.

Bu uygulamadaki gruplara ayırma, bir ölçme işlemi olarak kabul edilirse kullanılan ölçek türü aşağıdakilerden hangisidir? (2006)

A) Sıralama
B) Eşit oranlı
C) Sınıflama
D) Dereceleme
E) Eşit aralıklı

21. Yabancı dil sınavında Duru 80, Berke 40 puan almıştır.

Duru'nun yabancı dil bilgisinin Berke'nin yabancı dil bilgisinin iki katı kadar olduğunu öne süren bir kişi, bu ölçme sonuçlarının aşağıdaki ölçek türlerinin hangisiyle elde edildiğini kabul etmektedir? (2006)

A) Eşit oranlı
B) Sıralama
C) Dereceleme
D) Sınıflama
E) Eşit aralıklı

22. **Aşağıdakilerden hangisi bağıl değerlendirmede kullanılan bir ölçüt değildir? (2006)**

A) Sınıftaki diğer öğrencilerin başarısı
B) İzlenen programın hedefleri
C) Öğrencinin yeteneği
D) Ülke çapındaki normlar
E) Öğrencinin sınıf içindeki başarı sırası

23. Bir öğretmen fen bilgisi öğretiminde performans değerlendirme yaklaşımını kullanmanın öğrencilerin bilimsel yöntem süreç becerilerine etkisini incelemektedir.

Bu araştırmadaki bağımlı değişken aşağıdakilerden hangisidir? (2006)

A) Bilimsel yöntem süreç becerileri
B) Fen bilgisi dersi
C) Öğretmen
D) Öğrenciler
E) Performans değerlendirme

24. **Aşağıdaki örneklerin hangisinde sözü edilen ölçme işleminde bağıl (keyfî, itibari) sıfır söz konusu değildir? (2007)**

A) Ankara ili Çankaya ilçesinin deniz seviyesinden yüksekliği 870 metredir.
B) Öğrencilerin felsefe dersine karşı tutum puanları ortalaması 207 dir.
C) Tarih sorularından en az dört tanesine doğru cevap veremeyenler sıfır puan almıştır.
D) Bugün hava sıcaklığı 15 °C dir.
E) Bu kitabın ilk 4 sayfasında kitabın kapsamı tanıtılmaktadır.

25. Değerlendirme yapılırken, ölçüt ile ölçme sonuçlarının aynı birimle ifade edilmesi gerekir.

Bunun en önemli amacı aşağıdakilerden hangisidir? (2007)

A) Ölçme aracının geçerliğini artırmak
B) Karşılaştırmada doğru karar verilmesini sağlamak
C) Ölçme aracının güvenirliğini artırmak
D) Sonuçların anlaşılmasını kolaylaştırmak
E) Ölçütün uygun olmasını sağlamak

26. **Aşağıdaki durumların hangisinde kriter referanslı (ölçüt dayanaklı) değerlendirme, norm referanslı (norm dayanaklı) değerlendirmeden daha uygun olur? (2007)**

A) Öğrencilerin mezun edilip edilmemelerine karar verilmesinde
B) Kontenjanları sınırlı olan yükseköğretim programlarına öğrenci seçiminde
C) Öğrencilerin hazırladıkları fen bilimleri projelerinden en iyisinin belirlenmesinde
D) Okulda dereceye girecek öğrencilerin belirlenmesinde
E) Bir sınıf içindeki en başarısız öğrencilerin belirlenmesinde

27. **Durum değerlendirme ya da düzey belirleme denen ölçme ve buna bağlı değerlendirme sürecinde temel amaç aşağıdakilerden hangisidir? (2008)**

A) Öğretim programında belirtilen kazanımların gerçekleşme derecesini belirlemek
B) Öğrencilerin yaşadıkları öğrenme güçlüklerini ortaya çıkarmak
C) Öğrencilerin sosyal, psikolojik ve ekonomik durumlarını belirlemek
D) Öğrencileri belli özelliklere göre sıralayıp ilgili programlara yerleştirmek
E) Öğrencilerin kavramsal gelişim düzeylerini ve kavram yanılgılarını tespit etmek

28. Matematik dersinde bilgisayar programlarından yararlanmanın öğrencilerin matematik başarısına etkisini ölçmeyi planlayan bir öğretmen, bir sınıfında geleneksel öğretim yöntemini kullanırken diğer bir sınıfında ise işlenen her konudan sonra bilgisayar programıyla alıştırmalar yaptırmaktadır.

Bu araştırmadaki bağımsız değişken aşağıdakilerden hangisidir? (2008)

A) Öğretmen özellikleri
B) Matematik notları
C) Öğretim yöntemleri
D) Sınıf mevcutları
E) Öğrenci nitelikleri

egitim@hotmail.es

29. **Aşağıdakilerden hangisi ölçülürken dolaylı ölçme yapılır? (2008)**

A) Park yerindeki araçların sayısı
B) Bilgisayar masasının yüksekliği
C) Bir sepet elmanın kütlesi
D) Öğrencilerin dört işlem becerileri
E) Koridorun uzunluğu

30. Değerlendirme, ölçme sonuçlarının bir ölçütle karşılaştırılarak bir yargıya varılması işidir.

Aşağıdakilerden hangisi "mutlak ölçüt"ün kullanıldığı değerlendirmeye örnektir? (2008)

A) Bilgi yarışmasında en az puanla elenen kişi olma
B) Yüksek atlamada en iyi dereceyi yapma
C) Sınıfta başarı açısından ilk yüzde yirmilik grupta olma
D) 100 sorudan 50'sini doğru cevaplama
E) Grup ortalamasının üzerinde performans gösterme

31. Arif Öğretmen, öğretim yılının başında, öğrencilerinin matematik dersine karşı tutumlarının nasıl olduğunu belirlemek ve elde edilen sonuçlardan hareketle öğretim uygulamalarına yön vermek istemektedir. Bu amaçla 10 maddeden oluşan bir tutum ölçeğini öğrencilerine uygulamış ve onların matematik dersine karşı tutumlarını ölçmüştür.

Arif Öğretmen bu uygulamasında hangi ölçme ve değerlendirme türlerini kullanmaktadır? (2009)

	Ölçme	Değerlendirme
A)	Dolaylı	Tanılayıcı
B)	Doğrudan	Biçimlendirici
C)	Türetilmiş	Tanılayıcı
D)	Dolaylı	Düzey belirleyici
E)	Doğrudan	Düzey belirleyici

32. **Aşağıdaki değerlendirme süreçlerinin hangisinde bağıl ölçüt kullanılmıştır? (2009)**

A) Mezun olmak için diploma notunun 4 üzerinden en az 2 olması
B) Ortalamanın 10 puan üzerinde alanların sınıfı geçmesi
C) Askere ağırlığı 150 kg'nin altındaki erkeklerin alınması
D) Soruların % 80'ini doğru cevaplayanların bilgi yarışması elemelerine katılması
E) Yüksek lisansa girebilmek için yabancı dil sınavından en az 70 puan alınması

33. Bu yıl yükseköğretim programına alınan 120 öğrenci, çoktan seçmeli bir yabancı dil testine alınacaktır. Bu testten en yüksek puanı alan 40 öğrenci öğrenimlerine birinci sınıftan başlayacak; diğer 80 öğrenci ise hazırlık sınıflarında bir yıl yabancı dil eğitimi alacaklardır.

Bu süreçte, ölçme değerlendirme ilkelerine aykırı olan durum aşağıdakilerden hangisidir? (2009)

A) Puanlamada objektifliğin sağlanamaması
B) Ölçme aracı olarak çoktan seçmeli testin kullanılması
C) Testin puanlanmasında şans başarısından arıtma yöntemi uygulanmaması
D) Birinci sınıfa seçilen öğrenci sayısının az olması
E) Ölçüt dayanaklı ölçme yapılması gerekirken norm dayanaklı ölçme yapılması

34. **Aşağıdaki işlemlerin hangisi ölçme sonuçlarının bir ölçütle karşılaştırılarak bir yargıya varma sürecine örnektir? (2010-2)**

A) Bilim şenliği jüri üyelerinin seçimi için kullanılacak özelliklerin belirlenmesi
B) Performans görevlerinin puanlarına göre sıralanması
C) Bilgi yarışmasında okulu temsil edecek öğrencilerin seçilmesi
D) Komposizyon sınavı için dereceli puanlama anahtarı oluşturulması
E) Sözlü sınavların genel ortalamadaki ağırlığının belirlenmesi

35. Bingöl'de bir kış günü saat 12.00'de hava sıcaklığı -15°C ölçülürken aynı gün ve saatte Eskişehir'de sıcaklık -5°C olarak ölçülmüştür. Son nüfus sayımlarında Bingöl'ün nüfusu yaklaşık 250 000, Eskişehir'in nüfusu yaklaşık 750 000 olarak tespit edilmiştir. Bingöl'ün yüz ölçümü 8277 km2, Eskişehir'in yüz ölçümü ise 13 925 km2 olarak ölçülmüştür.

Bu ölçümler ve kullanılan ölçekler düşünüldüğünde, aşağıdaki ifadelerden hangisi yanlıştır? (2010-2)

A) Bingöl'de hava Eskişehir'in üç katı soğuktur.
B) Eskişehir'de sıcaklık Bingöl'den 10°C daha yüksektir.
C) Eskişehir'in nüfusu Bingöl'ün nüfusunun yaklaşık üç katıdır.
D) Bingöl'ün nüfusu Eskişehir'in nüfusundan 500 000 daha azdır.
E) Eskişehir'de birim alana düşen kişi sayısı Bingöl'den daha fazladır.

36. Bir araştırmacı özel öğretim yöntemleri dersinde bir sınıfta öğrenci merkezli yöntem ve teknikleri, diğer sınıfta ise öğretmen merkezli yöntem ve teknikleri uygulamış ve her iki sınıftaki öğrencilerin uygulama öncesi ve sonrasındaki başarı ve tutum puanları arasındaki farkları incelemiştir.

Bu araştırmadaki bağımsız değişken aşağıdakilerden hangisidir? (2010-2)

A) Öğrencilerin hazırbulunuşluk düzeyleri
B) Başlangıçtaki başarı ve tutum puanları
C) Başarı ve tutum puanları arasındaki farklar
D) Uygulanan öğretim yöntem ve teknikleri
E) Öğrencilerin sosyo-ekonomik düzeyleri

egitim@hotmail.es

37. Bir Kuzey Avrupa ülkesinde hemşirelik lisans programı mezunlarının çalışmaya başlayabilmeleri için yeterli düzeyde bilgi ve beceri sahibi olduklarını çoktan seçmeli bir testle göstermeleri gerekmektedir. Bu testte adayların başarılı sayılmaları için doğru cevaplamaları gereken soru sayıları meslek standartları doğrultusunda belirlenir ve yeterlik düzeyi için bir kesme puanı oluşturulur. Testten başarılı olan adaylar, çalışmaya başlamak için gerekli olan sertifikalarını almaya hak kazanırlar.

Yapılan değerlendirmenin amacına ve ölçütüne göre türü aşağıdakilerin hangisinde doğru olarak verilmiştir? (2011)

	Amacına göre	Ölçütüne göre
A)	Biçimlendirme ve yetiştirme	Mutlak
B)	Tanıma ve yerleştirme	Bağıl
C)	Değer biçme	Bağıl
D)	Tanıma ve yerleştirme	Mutlak
E)	Değer biçme	Mutlak

38. VE 39. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

İnsan nefesinden, kandaki alkol düzeyini ölçen alkolmetrelerin kullanılmadığı bir ülkede, trafik polisleri alkol kontrolünü "yedi adım testi" ile yapmaktadır. Bu testte, kontrolü yapılacak sürücüden düz bir çizgi üzerinde yedi adım atması istenir. İlgili trafik polisi ise sürücünün yürüyüşünü gözlemleyerek aşağıdaki formu doldurur.

	Evet	Hayır
Sürücü açıklamalarının bitmesini beklemeden çizgi üzerinde yürümeye başladı.	[ ]	[ ]
Sürücü dengesini sağlayabilmek için kollarını yana açarak yürüdü.	[ ]	[ ]
Sürücü yedi adımdan daha az veya daha fazla adım attı.	[ ]	[ ]
Sürücü düz çizgiden saparak yürüdü.	[ ]	[ ]

38. Sürücünün gözlenmesi sonucunda, formdaki maddelerden iki veya daha fazlası "Evet" olarak işaretlenmiş ise sürücü, alkollü olarak değerlendirilip kendisine ceza kesilmektedir.

Sürücülerin alkol düzeylerinin ölçülmesinin ve değerlendirilmesinin türleri, aşağıdakilerden hangisinde doğru olarak verilmiştir? (2011)

	Ölçme	Değerlendirme
A)	Doğrudan	Mutlak
B)	Dolaylı	Bağıl
C)	Doğrudan	Bağıl
D)	Dolaylı	Mutlak
E)	Türetilmiş	Mutlak

39. "Yedi adım testi"nin uygulanması sonucunda sürücü için işaretlenen toplam "Evet" sayısı, kandaki alkol düzeyinin bir ölçümü olarak kabul edilmektedir.

Buna göre, "yedi adım testi"nden elde edilen alkol düzeyi ölçümleri hangi ölçek türündedir? (2011)

A) Sınıflama B) Sıralama C) Eşit dereceli D) Eşit aralıklı E) Eşit oranlı

40. (I) Eda Öğretmen, dersini işledikten sonra öğrencilerine bir sınav uygular. (II) Sınavı hazırlarken ilgili kazanımları ve kazanımların ağırlıklarını göz önüne alarak bir belirtke tablosu hazırlar. (III) Hazırladığı tabloya uygun olarak sınavında 10 tane açık uçlu maddeye yer verir. (IV) Her bir maddenin değerini 10 puan olarak belirler. (V) Uygulama sonunda sınavdan en az 50 puan alan öğrencileri başarılı sayacağını açıklar. (VI) Sınava giren Ayşe, 80 puan alarak sınavda başarılı olur.

Buna göre, ölçüt kaç numaralı cümlede belirtilmiştir? (2012)

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

41. VE 42. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

41. Bir Türkçe öğretmeni, 8A sınıfındaki öğrencilerden kendilerine verilen konuyla ilgili bir kompozisyon yazmalarını; 8B sınıfındaki öğrencilerden ise kendilerine verilen iki konudan birini seçip seçtikleri konuyla ilgili bir kompozisyon yazmalarını istemiştir. Öğretmen; öğrencilerin dönem boyunca okudukları kitap sayısı, kompozisyonu yazma süresi, konu seçme imkânı ve cinsiyet değişkenlerinin yazılı anlatım beceri düzeyini yordama gücünü incelemiştir.

Bu araştırmadaki bağımlı değişken aşağıdakilerden hangisidir? (2012)

A) Dönem boyunca okudukları kitap sayısı
B) Cinsiyet
C) Kompozisyonu yazma süresi
D) Konu seçme imkânı
E) Yazılı anlatım beceri düzeyi

42. **Bu araştırmadaki değişkenlerden hangisi dolaylı olarak ölçülür? (2012)**

A) Dönem boyunca okudukları kitap sayısı
B) Kayıtlı oldukları şube
C) Kompozisyonu yazma süresi
D) Konu seçme imkânı
E) Yazılı anlatım beceri düzeyi

egitim@hotmail.es

43. Bir Avrupa ülkesinde motosikletli polislerin hizmet içi eğitim programları dâhilinde ileri sürüş testleri ve eğitimleri yapılmaktadır. Motosikletli polisler, öncelikle yol sürüşü sınavına alınmaktadır. Bu sınavda, sınavı yapan gözlemci ayrı bir motosiklet ile adayı hemen arkasından takip eder. Gözlemci, sürüş bitiminde aşağıdaki forma 0 - 4 puan arasında uygun gördüğü puanları vererek gözlem sonuçlarını aktarır. 0 puan en düşük, 4 puan ise en yüksek beceri düzeyini göstermektedir.

Ölçütler	Puan
Motosikleti bakımları yapılmış ve yola çıkmaya uygun vaziyette hazır bulundurma	
Motosikletinin yönlendirmesini yumuşak ve çabuk hareketlerle yapabilme	
Trafikte uygun konumlanma noktaları seçme	
Debriyajı yumuşak, nazik ve vitesle uyumlu zamanlamayla kullanabilme	
Frenleri, tekerlerin kilitlenmesine izin vermeden etkili kullanabilme	
Virajlarda uygun viraj çizgilerini tayin ve takip edebilme	

Her bir maddeden adayın almış olduğu puanlar toplanarak bir toplam puan belirlenir. Gözlem sonucunda yeterli puanı alan motosikletli polisler görevine devam ederler. Yeterli puanı alamamış olan motosikletli polisler ise hizmet içi ileri sürüş eğitimine alınır.

Buna göre, motosikletli polislerin ileri sürüş becerilerinin değerlendirilmesinin amacına ve ölçütüne göre türleri aşağıdakilerden hangisinde birlikte verilmiştir? (2012)

	Amacına göre	Ölçütüne göre
A)	Tanıma ve yerleştirme	Mutlak
B)	Tanıma ve yerleştirme	Bağıl
C)	Değer biçme	Mutlak
D)	Değer biçme	Bağıl
E)	Süreç izleme	Bağıl

44. 2000 öğrencinin bulunduğu bir okuldan, matematik bilgi yarışması için 4 kişilik bir takımın oluşturulması hedeflenmektedir. Takıma alınacak başarılı öğrencileri seçmek amacıyla okuldaki tüm öğrenciler 50 maddelik çoktan seçmeli bir başarı testine alınacaktır.

Yapılan bu değerlendirme işleminde, mutlak ölçüt kullanılmasına bağlı olarak aşağıdaki sonuçlardan hangisi gözlenebilir? (2012)

A) Takım için yeterli sayıda öğrenci bulunamayabilir.
B) Öğrencilerin başarı düzeyleri belirlenemeyebilir.
C) Yeterli objektiflik sağlanamayabilir.
D) Ölçmenin standart hatası istenenden çok büyük olabilir.
E) Sınavın kullanışlılığı çok düşük olabilir.

ÖLÇMEDE HATA

2

ANAHTAR KAVRAMLAR

Ölçmede hata, sabit hata, sistematik hata, tesadüfi hata

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Ölçmede hatayı tanımlayabilecek,
- Hata türlerini açıklayabilecek,
- Hata türleri arasındaki farklılıkları tartışabilecek,
- Hata türleriyle ilgili örnekler verebilecek,
- Olası hata kaynaklarını açıklayabileceksiniz.

İÇİNDEKİLER

Ölçmede hata
Hata türleri
Hata kaynakları
Çalışma soruları

BAŞLAMADAN ÖNCE

Sizce sınavlardan beklediğiniz puanları alamadığınızda bunun sebepleri neler olabilir?

egitim@hotmail.es

ÖLÇMEDE HATA

Yapılan tüm ölçmelerde hatanın bulunuyor olması kaçınılmazdır. Ne kadar doğru ve dikkatli ölçme de yapsak yine de bir miktar hatanın bulunması söz konusudur. Hata kısaca aşağıdaki eşitlik ile tanımlanabilir:

Hata = Gözlenen ölçme sonucu – Gerçek ölçme sonucu

Bu eşitlikten de anlaşılacağı üzere, bizim gözlediğimiz ölçme sonuçlarının gerçek değerinden olan farkına ölçmede hata adı verilir. Her nesnenin ölçülmek istenilen özelliğine ilişkin gerçek bir değeri bulunmaktadır (Örneğin; masamızın boyunun gerçek (hiç hatasız) ölçümü 120cm olsun). Ancak biz bazen ölçme aracından, bazen ölçen bireyden ya da daha başka sebeplerden dolayı 121 ya da 118 gibi gerçek değerinden daha uzun ya da daha kısa olarak gözlemleyebiliriz. Bu durumda artı (121cm gözlemlemişsek +1 cm hata) ya da eksi (118 cm gözlemlemişsek -2 cm hata) yönde hata yapmak mümkün olmaktadır. Buradan da anlaşılacağı üzere hatanın hem miktarı hem de bir yönü (artı ya da pozitif, eksi ya da negatif) vardır.

Fen bilimlerinde ölçülen nesnelerin doğrudan gözlemlenebilmesi, bu nesnelere ilişkin yapılan ölçmelerin de hata miktarının çok az (bazen yok sayılabilecek kadar) olmasına imkân tanır. Hâlbuki sosyal bilimlerde ve eğitim bilimlerinde ölçmenin konusu daha çok gözlenemeyen kavramlar (başarı, bilgi, tutum, ilgi, kaygı vb.) olduğu için bu alanlarda yapılan ölçmeler çok daha fazla ve belirsiz hataların ortaya çıkmasına sebep olur.

Eğitim ortamında ölçme sonuçlarını puanlar olarak nitelendirirsek, yukarıda verilen eşitlik;

Hata= Gözlenen puan – Gerçek puan

olarak ifade edilir. Eğitimde, yapılan sınavlardan elde edilen puanlar, öğrencilerin bilgi, başarı gibi görünmeyen özelliklerinden elde edilen ölçme sonuçlarıdır ve maalesef öğrencilerin bu özelliklerine ilişkin gerçek puanlarını bulabilmek hemen hemen imkânsızdır. Eğitimde, öğrencilere ilişkin elde edilen puanlara karışan hatalar da değişik türlere ayrılmaktadır:

HATA TÜRLERİ

Sabit Hata	Sistematik Hata	Tesadüfi (Gelişi-güzel, Random, Rasgele) Hata
- Ölçmeden ölçmeye miktarı ve yönü değişmeyen, kaynağı belirlenebilen hatalardır. - Örneğin; ucu 1 cm kırık cetveli her seferinde bir kez kullanarak yapılan tüm ölçmelerde yapılan hata miktarı ve yönü aynıdır ve kaynağı bellidir: cetvel. - Eğitimde bu tür hataya örnek verirsek: Bir öğretmenin sınıftaki tüm öğrencilerine 5 puan fazla vermesi: Tüm öğrencilerin puanlarındaki hatanın miktarı ve yönü aynıdır ve hata kaynağı bellidir: öğretmen.	- Ölçmeden ölçmeye miktarı ve yönü değişen, kaynağı belirlenebilen hatalardır. - Örneğin; ucu 1 cm kırık cetveli her seferinde farklı kereler kullanarak yapılan tüm ölçmelerde yapılan hata miktarı değişecektir ancak kaynağı bellidir: cetvel. - Eğitimde bu tür hataya örnek verirsek: - Bir öğretmenin sınıftaki kız öğrencilerine 5 puan fazla vermesi: Tüm öğrencilerin puanlarındaki hatanın miktarı ve yönü aynı değildir. Kız öğrencilerinin puanları +5 hata içerirken erkek öğrencilerinin puanları bu hataya sahip değildir. Hata kaynağı bellidir: öğretmen. - Eğitimdeki bu tür hatalara "yanlılık" adı da verilmektedir.	- Ölçmeden ölçmeye miktarı ve yönü değişen, kaynağı belirlenemeyen hatalardır. - Eğitimde en çok karşılaşılan ve en büyük sorun olan hata türü tesadüfi hatalardır. Ne miktarı ya da yönü ne de kaynağı tam olarak belirlenemez. Pek çok etken elde edilen puanlarda hata bulunmasına sebep olabilmektedir. Örneğin; Bir öğretmenin istemeden ve farketmeden dikkatsizlik yaparak öğrencisinin kâğıdındaki puanları toplarken hata yapması, Öğrencinin çok iyi bildiği bir konudan sınav olurken rahatsızlanması sonucu alması gereken puanı alamaması, Sınav sorusunun açık ve net olmamasından dolayı öğrencilerin soruyu anlamayarak yanlış yapması vb...

Tesadüfi hatanın tanımından ve örneklerinden da anlaşılacağı üzere eğitim ortamında yapılan ölçmelerde birden fazla hata kaynağı bulunmaktadır.

Hata Kaynakları

Ölçmeyi yapan bireyden kaynaklanan hatalar (Örneğin; öğretmenin puanlama yaparkenki dikkatsizliği, psikolojisi, açlığı vb. etkenler)

Ölçmenin yapıldığı bireyden kaynaklanan hatalar (öğrencinin sınav sırasındaki kaygı düzeyi, rahatsızlanması, kopya çekmesi, şans başarı vb. etkenler)

Ölçmenin yapıldığı ortamdan kaynaklanan hatalar (gürültü, ışığın yetersiz olması, sınıfın sıcaklığının çok fazla ya da çok düşük olması vb. etkenler)

Ölçme aracından kaynaklanan hatalar (yazılı yoklamalarda puanlamanın subjektif olması, çoktan seçmeli testlerde şans başarısının olması vb. etkenler)

egitim@hotmail.es

Sabit ve sistematik hatalar ölçme sonuçlarının geçerliğini olumsuz yönde etkiler, güvenirliğini etkilemez. Tesadüfi hata ise ölçme sonuçlarının öncelikle güvenirliğini olumsuz yönde etkiler, dolaylı olarak da geçerliğini etkiler. Güvenirlik ve geçerlik kavramlarına dördüncü bölümde yer verilecektir:

DEĞERLENDİRME SORULARI

1. Aşağıdakilerden hangisi bir testte sabit hata bulunduğunun göstergesidir?

A) Öğrencilere sınavdan önce ipucu verilmesi
B) Öğretmenin puanlamada subjektif davranması
C) Öğretmenin, puanlarının %5'i kadar tüm öğrencilerine ek puan vermesi
D) Testteki bir soruyu anlamayan öğrencilerin doğru cevap verememesi
E) Testin uygulandığı tüm öğrencilerin puanlarına 10 puan eklenmesi

2. Aşağıdaki hata türlerinden hangisi <u>öncelikle</u> güvenirliği etkiler?

A) Tesadüfi hata
B) Mutlak hata
C) Sabit Hata
D) Sistematik hata
E) Bağıl hata

3. Bir öğretmen yapmış olduğu sınanvı puanlarken, sınavda en önde oturan öğrencilerine 5 puan fazla vermiştir.

Öğretmenin puanlamada yapmış olduğu hata türü aşağıdakilerden hangisidir?

A) Tesadüfi hata
B) Mutlak hata
C) Sabit Hata
D) Sistematik hata
E) Bağıl hata

KPSS'DE ÇIKMIŞ SORULAR

1. Bir sınavda her öğrenciye 5 puan fazla verilmesi ---- hataya örnektir. Öğrencilerin puanlarını, aldıkları puanların yüzde 10'u kadar artırmak ise ---- hataya örnek olabilir.

Bu parçadaki boşluklara, aşağıdakilerin hangisindeki sözcükler sırasıyla getirilmelidir? (2008)

A) rastgele - sabit
B) rastgele - sistematik
C) sabit - sistematik
D) sabit - rastgele
E) sistematik – sabit

2. Öğretmenler genellikle söz dinleyen ve kurallara uygun davranan öğrencilere sempati duyarlar. Yapılan araştırmalar, öğretmenlerin tepkilerinin, öğrencilerin derslerdeki başarısızlıklarından çok, kurallara uymamalarına yönelik olduğunu ortaya koymaktadır. Öğrencilere bu konuda bir açıklama yapılmasa da onlar, kurallara uygun davrandıklarında bunun başarı olasılığını etkileyebileceğini, sınavda doğru yanıtlar vermeseler de iyi not alabileceklerini düşünürler. Bu durumda, kurallara karşı çıktıklarında cezalandırılabileceklerini ya da öğretmenleri tarafından dışlanabileceklerini öğrenirler.

Kurallara uyan öğrencilere daha yüksek puan vererek gerçekleştirilen bu ölçmelerde aşağıdakilerden hangisi yapılmış olur? (2009)

A) Rastlantısal hata
B) Gözlem yetersizliği
C) Sabit hata
D) Mantık hatası
E) Sistematik hata

egitim@hotmail.es

KORELASYON

3

ANAHTAR KAVRAMLAR

Korelasyon, pozitif korelasyon, negatif korelasyon, sıfır korelasyon

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Korelasyonun ne olduğunu söyleyebilecek,
- Pozitif korelasyonu açıklayabilecek,
- Pozitif korelasyon gösteren değişkenlerle ilgili örnekler verebilecek,
- Negatif korelasyonu açıklayabilecek,
- Negatif korelasyon gösteren değişkenlerle ilgili örnekler verebilecek,
- Sıfır korelasyonu açıklayabilecek,
- Sıfır korelasyon gösteren değişkenlerle ilgili örnekler verebileceksiniz.

İÇİNDEKİLER

Korelasyon
Çalışma soruları

BAŞLAMADAN ÖNCE

Bir öğrencinin derse karşı tutumu – o dersteki başarısı arasındaki ilişki ile bir dersi çalışma süresi – dersi çalışırken gösterdiği konsantrasyon arasındaki ilişki arasındaki farklılık nedir?

egitim@hotmail.es

KORELASYON

- ✓ Korelasyonu genel olarak "iki değişken arasındaki ilişkinin matematiksel ifadesi" olarak tanımlayabiliriz. Kısacası korelasyon "ilişki" anlamı taşımaktadır.
- ✓ İki değişken arasındaki ilişki matematiksel olarak bir değere sahiptir ve bu değer korelasyon katsayısı (r_{xy}) ile gösterilir.
- ✓ Korelasyon katsayısı -1 ile +1 arasında değerler almaktadır. ($-1 \leq r_{xy} \leq +1$)
- ✓ İki değişken arasındaki ilişkiyi ifade eden korelasyon katsayısının bir sayısal değeri (büyüklüğü) bir de yönü (işareti) vardır.
- ✓ Korelasyon katsayısının işaretine bakılmaksızın 1'e yaklaşması ilişkinin mükemmele, 0'a yaklaşması ilişkinin yokluğa yaklaştığının göstergesidir.
- ✓ Korelasyon katsayısının işareti ise ilişkinin yönünü gösterir: Pozitif işaret "+", iki değişkenin birlikte arttığını ya da azaldığını yani aynı yönlü ilişkiye sahip olduklarını gösterirken; negatif işaret "-", iki değişkenden biri artarken diğerinin azaldığını ya da biri azalırken diğerinin arttığını yani ters yönlü ilişkiye sahip olduklarını gösterir.
- ✓ Eğer iki değişken arasındaki ilişki mükemmel ve aynı yönlüyse (bir değişkenin değeri artarken diğer değişkenin değeri de artıyorsa) korelasyon katsayısı +1 olacaktır.
- ✓ Tam tersi eğer iki değişken arasındaki ilişki mükemmel ve ters (zıt) yönlüyse (bir değişkenin değeri artarken diğer değişkenin değeri azalıyorsa) korelasyon katsayısı -1 olacaktır.
- ✓ Korelasyonun gösterimine ilişkin birkaç örnek durum aşağıdaki grafiklerde gösterilmeye çalışılmıştır.

Alıştırma 1: Bir Türkçe öğretmeni öğrencilerinin derse karşı olan tutumları ile dersdeki başarıları arasında ilişki olup olmadığını araştırmak için bir çalışma yapmıştır. Sınıfında bulunan 10 öğrencisinin Türkçe başarı puanları (X) ile Türkçe dersii tutum puanları (Y) aşağıdaki tabloda verilmiştir. Bu değişkenler arasındaki ilişki, korelasyon katsayısı kullanarak aşağıdaki şekilde hesaplanarak yorumlanır:

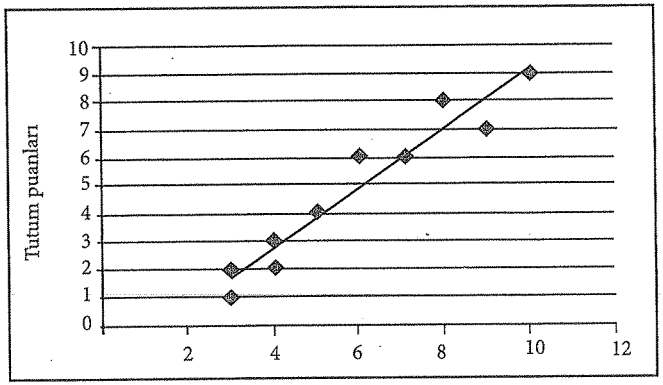
Korelasyon katsayısı: $$r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (X)^2\right]\left[n\sum Y^2 - (Y)^2\right]}}$$

Başarı puanları (X)	Tutum puanları (Y)	X^2	Y^2	XY
10	9	100	81	90
9	7	81	49	63
8	8	64	64	64
7	6	49	36	42
6	6	36	36	36
5	4	25	16	20
4	3	16	9	12
4	2	16	4	8
3	2	9	4	6
3	1	9	1	3
$\Sigma X = 59$	$\Sigma Y = 48$	$\Sigma X^2 = 405$	$\Sigma Y^2 = 300$	$\Sigma XY = 344$

$$r_{xy} = \frac{10 \cdot 344 - 59 \cdot 48}{\sqrt{\left[10 \cdot 405 - 59^2\right]\left[10 \cdot 300 - 48^2\right]}} = \frac{3440 - 2832}{\sqrt{[4050 - 3481][3000 - 2304]}}$$

$$= \frac{608}{\sqrt{569 \cdot 696}} = \frac{608}{629.3} = 0.97$$

olarak korelasyon değeri elde edilir. Bu değer başarı puanları ile tutum puanları arasında pozitif yönlü yüksek bir ilişki olduğunu göstermektedir. Aşağıdaki Şekil 1'de puanlar arasındaki pozitif yönlü bu ilişki grafiksel olarak da gösterilmeye çalışılmıştır.



Şekil 1. Pozitif yönlü ilişki

egitim@hotmail.es

Alıştırma 2: 10 öğrencinin matematik dersinde devamsızlık yaptıkları gün sayısıları (X) ile matematik testinden aldıkları puanlar (Y) aşağıdaki tabloda verilmiştir. Bu değişkenler arasındaki ilişki, korelasyon katsayısı kullanarak aşağıdaki şekilde hesaplanarak yorumlanır:

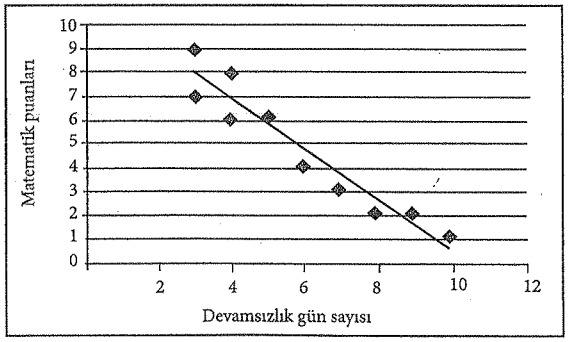
Korelasyon katsayısı: $r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (X)^2\right]\left[n\sum Y^2 - (Y)^2\right]}}$

Devamsızlık (X)	Matematik puanı (Y)	X^2	Y^2	XY
10	1	100	1	10
9	2	81	4	18
8	2	64	4	16
7	3	49	9	21
6	4	36	16	24
5	6	25	36	30
4	6	16	36	24
4	8	16	64	32
3	7	9	49	21
3	9	9	81	27
$\Sigma X = 59$	$\Sigma Y = 48$	$\Sigma X^2 = 405$	$\Sigma Y^2 = 300$	$\Sigma XY = 223$

$$r_{xy} = \frac{10 \cdot 223 - 59 \cdot 48}{\sqrt{\left[10 \cdot 405 - 59^2\right]\left[10 \cdot 300 - 48^2\right]}} = \frac{2230 - 2832}{\sqrt{[4050 - 3481][3000 - 2304]}}$$

$$= \frac{-602}{\sqrt{569 \cdot 696}} = \frac{-602}{629.3} = -0.96$$

olarak korelasyon değeri elde edilir. Bu değer soru çözme hızı ile soru çözme süresi arasında negatif yönlü yüksek bir ilişki olduğunu göstermektedir. Aşağıdaki Şekil 2'de soru çözme hızı ile soru çözme süresi arasındaki negatif yönlü bu ilişki grafiksel olarak verilmiştir.



Şekil 2. Negatif yönlü ilişki

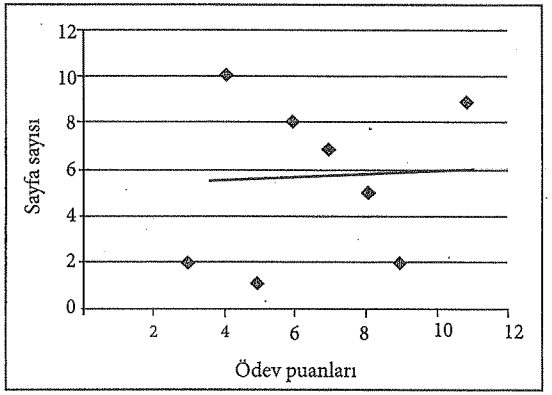
Alıştırma 3: 10 öğrencinin matematik dersi ödevinden aldıkları puanlar (X) ile teslim ettikleri ödevlerin sayfa sayıları (Y) aşağıdaki tabloda verilmiştir. Bu değişkenler arasındaki ilişki, korelasyon katsayısı kullanarak aşağıdaki şekilde hesaplanarak yorumlanır:

Korelasyon katsayısı: $r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (X)^2\right]\left[n\sum Y^2 - (Y)^2\right]}}$

Ödev puanları (X)	Ödev sayfa sayısı (Y)	X^2	Y^2	XY
10	9	100	81	90
9	2	81	4	18
8	5	64	25	40
7	7	49	49	49
6	8	36	64	48
5	1	25	1	5
4	10	16	100	40
4	6	16	36	24
3	7	9	49	21
3	2	9	4	6
$\Sigma X = 59$	$\Sigma Y = 57$	$\Sigma X^2 = 405$	$\Sigma Y^2 = 413$	$\Sigma XY = 341$

$$r_{xy} = \frac{10 \cdot 341 - 59 \cdot 57}{\sqrt{\left[10 \cdot 405 - 59^2\right]\left[10 \cdot 413 - 57^2\right]}} = \frac{3410 - 3363}{\sqrt{[4050 - 3481][4130 - 3249]}} = \frac{47}{\sqrt{569 \cdot 881}} = \frac{47}{708.02} = 0.07$$

olarak korelasyon değeri elde edilir. Bu değer matematik dersi ödevinden aldıkları puanlar ile teslim ettikleri ödevlerin sayfa sayıları arasında bir ilişki olmadığını göstermektedir. Aşağıdaki Şekil 3'te matematik dersi ödevinden aldıkları puanlar ile teslim ettikleri ödevlerin sayfa sayıları arasındaki ilişkisizlik grafiksel olarak verilmiştir.



Şekil 3. İlişkisiz değişkenler

egitim@hotmail.es

Sosyal bilimlerde ve eğitim bilimlerinde çalışılan pek çok özelliğe dair hesaplanan korelasyon değerlerinin ±1 ya da 0 değerleri alması imkansız denecek kadar azdır.

Çok sayıda değişken arasındaki ikili korelasyon değerleri genellikle korelasyon tabloları ile gösterilirler. Aşağıda korelasyon tablosuna ilişkin bir örnek verilerek, korelasyon tablolarının genel özellikleri açıklanmıştır.

Örnek: Aşağıdaki tabloda bir sınıfın beş farklı dersin sınavlarından aldıkları puanlar arasındaki korelasyon değerleri verilmiştir.

	Türkçe	Tarih	Fizik	Kimya	Matematik
Türkçe	**1.00**	0.67	-0.25	0.41	0.78
Tarih	0.67	**1.00**	-0.85	0.11	0.54
Fizik	-0.25	-0.85	**1.00**	0.62	0.80
Kimya	0.41	0.11	0.62	**1.00**	0.70
Matematik	0.78	0.54	0.80	0.70	**1.00**

Korelasyon tablolarının genel özellikleri:

1. Bir korelasyon tablosunda yer alan değerler daima [-1, +1] aralığındadır.
2. Korelasyon tablosunun esas köşegeni üzerinde yer alan değerler (bir değişkenin yine kendisiyle olan ilişkisi olduğu için) daima +1'dir. (pozitif değerler için "+" işareti kullanılmayabilir)
3. Korelasyon tablosundaki değerler daima esas köşegene göre simetriktir. Bu nedenle korelasyon tablolarının sadece alt köşesi ya da sadece üst köşesindeki değerleri verilmesi yeterlidir.

   Yukarıda verilen örnek tablo için;

   - Aralarında en yüksek ilişki bulunan dersler: fizik ile tarihtir.
   - Aralarında en düşük ilişki bulunan dersler: kimya ile tarihtir.

DİKKAT: Yukarıdaki örnekte verildiği gibi derslerdeki başarılar arasındaki ilişkileri gösteren bir korelasyon tablosu; dersler arasındaki başarıların ilişkisi dışında başka hiç bir bilgi **vermez.** Yani sınıfın hangi derste en başarılı olduğu, hangi dersteki başarının en düşük olduğu, hangi dersin sınavında öğrencilerin ortalama başarısının düşük olduğu vb... bilgileri **göstermez.**

NOT: *Bu bölümde açıklanan korelasyon "basit korelasyon"dur. Korelasyon konusuyla ilgili daha detaylı bilgi için kaynakçada yer alan istatistik kitaplarından yararlanabilirsiniz.*

DEĞERLENDİRME SORULARI

1. **Aşağıda verilen değişken çiftlerinden hangisi arasında negatif korelasyon olması beklenir?**

   A) ders karşı tutum- dersteki başarı
   B) derse ayrılan çalışma saati – dersteki başarı
   C) ders çalışma saati – yorgunluk
   D) yorgunluk – motivasyon
   E) motivasyon- dikkat

2. Öğretmen sınıftaki öğrencilerinin kardeş sayıları ile haftalık harçlık miktarları arasındaki ilişkiyi merak etmiştir.

   Öğretmenin buna yönelik yapacağı çalışma sonucunda bulacağı korelasyon değerinin aşağıdaki değerlerden hangisi olması beklenebilir?

   A) -0.85 B) -0.05 C) 0 D) 0.05 E) 0.85

3.

	Türkçe	Tarih	Müzik
Türkçe	1.00	0.88	-0.24
Tarih		1.00	0.04
Müzik			1.00

Yukarıdaki tabloda bir sınıfın üç farklı dersteki puanları arasındaki korelasyon değerleri verilmiştir.

Buna göre aşağıdakilerden hangisin söylemek doğrudur?

A) Tablo eksik verilmiştir, bir şey söylenemez.
B) Öğrencilerin üç dersteki puanlarına göre başarıları %100=1'dir.
C) Bu sınıfın en yüksek başarı gösterdiği ders tarihtir.
D) En düşük ilişkili iki ders tarih ile müziktir.
E) Sınıfın en başarılı olduğu iki ders tarih ile Türkçe'dir.

egitim@hotmail.es

KPSS'DE ÇIKMIŞ SORULAR

1. Negatif korelasyon, değişkenlerden birinin değeri artarken diğerinin değerinde azalma olduğunu gösterir.

Buna göre, aşağıdakilerin hangisinde verilen iki değişken arasında hesaplanacak korelasyonun negatif olması beklenir? (2004)

A) Düzenli çalışma - başarı düzeyi
B) Gelir düzeyi - satın alma gücü
C) İlgi - olumlu tutum
D) Zekâ - yaratıcılık
E) Yorulma - öğrenme hızı

2. Bir araştırmada X ve Y değişkenleri arasında sıfıra yakın bir ilişki bulunmuştur.

Bu değişkenler aşağıdakilerin hangisinde verilenler olabilir? (2008)

	X	Y
A)	Doğum yeri	Doğum yılı
B)	Boy	Kütle
C)	Yaş	Eğitim düzeyi
D)	Kaygı düzeyi	Başarı
E)	Spor yapma	Sağlık

ÖLÇME ARAÇLARINDA BULUNMASI GEREKEN NİTELİKLER

4

ANAHTAR KAVRAMLAR

Güvenirlik, geçerlik, kullanışlılık

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Ölçme araçlarında bulunması gereken özellikleri tanımlayabilecek,
- Güvenirliğin ne olduğunu açıklayabilecek,
- Güvenilir ölçme sonuçları elde edebilmek için neler yapılması gerektiğini tartışabilecek,
- Ölçme araçlarının güvenirliğini kestirebilme yöntemlerini açıklayabilecek,
- Standart hatayı tanımlayabilecek,
- Bir ölçme aracından elde edilen puanların belirli olasılıklarda güven aralığını hesaplayabilecek,
- Geçerliğin ne olduğunu açıklayabilecek,
- Geçerlik türlerini tartışabilecek,
- Güvenirlik ile geçerlik arasındaki ilişkiyi açıklayabilecek,
- Kullanışlığın ne olduğunu açıklayabileceksiniz.

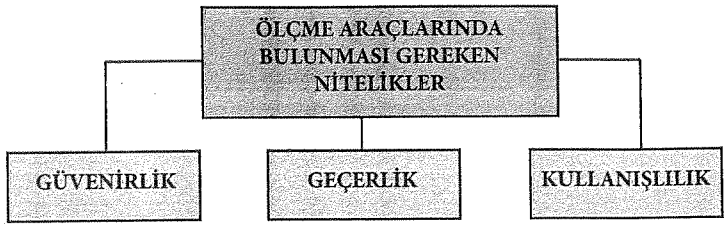
İÇİNDEKİLER

Ölçme araçlarında bulunması gereken özellikler
Güvenirlik
Güvenirlik türleri ve güvenirliği kestirme yöntemleri
Ölçme sonuçlarının standart hatası
Ölçme sonuçlarının güven aralığının hesaplanması
Geçerlik
Geçerlik türleri
Güvenirlik ve geçerlik arasındaki ilişki
Kullanışlılık
Çalışma soruları

BAŞLAMADAN ÖNCE

Güvenilir olmadığınız düşündüğünüz bir sınava girdiniz mi? Sizce neden güvenilir değildi?
Geçerli olmadığını düşündüğünüz bir sınava girdiniz mi? Sizce neden geçerli değildi?

egitim@hotmail.es



Ölçme sonuçlarının olabildiğince gerçeğe yakın ve objektif olabilmesi bu sonuçlara göre yapılacak değerlendirmelerin de o kadar yerinde ve isabetli olmasını sağlar. Bu sebeple ölçme araçlarından yararlanmak kaçınılmazdır. Ancak ölçme araçlarının da bir takım niteliklere sahip olması istenir: güvenirlik, geçerlik ve kullanışlılık.

Ölçme aracının güvenirlik ve geçerliği sahip olması mutlaka gerekli olan niteliklerdir. Kullanışlılık ise güvenirlik ve geçerliğin sağlanması koşuluyla istenen bir diğer niteliktir. Şimdi bu nitelikleri ayrıntılarıyla tanımaya çalışalım:

GÜVENİRLİK [0 , 1]

Güvenirlik genel tanımıyla "ölçme sonuçlarının tesadüfî hatalardan arınık olma derecesi" olarak ifade edilebilir (Baykul, Gelbal ve Kelecioğlu, 2003). Yukarıda bahsedilen hata kaynaklarından ortaya çıkabilecek tesadüfî hatalar ne kadar azaltılabilirse o kadar güvenilir ölçme sonuçları elde edilmiş olacaktır. Bu sebeple güvenirlikte "ya hep ya hiç" söz konusu değildir, güvenirlik bir derece ifadesidir.

Güvenirlik farklı kavramlarla da ifade edilebilir: tutarlılık, kararlılık ve duyarlılık.

Tutarlılık: Bir ölçme aracının tekrar tekrar kullanılması sonucu elde edilen sonuçların aynı ya da benzer olabilme derecesine tutarlılık anlamında güvenirlik denir.

Kararlılık: Bir ölçme aracının zaman aralıklarıyla tekrar tekrar kullanılması sonucu elde edilen sonuçların zamana rağmen aynı ya da benzer olabilme derecesine kararlılık anlamında güvenirlik denir.

Tutarlılık ile kararlılık arasındaki tek fark; kararlılığın elde edilen sonuçların zamana karşı direnç göstermesi ve benzerliğini koruyabilme derecesini de ifade ediyor olmasıdır.

Duyarlılık: Ölçme aracının biriminin olabildiğince hassas, küçük olmasının derecesidir. Ölçme aracında kullanılan birim ne kadar küçükse ölçmenin duyarlılığı o ölçüde artacaktır. Örneğin: uzunluk ölçerken kullanılacak cetvellerden mm bölmeli olanı cm bölmeli olana göre daha duyarlı sonuçlar verecektir. Eğitimde en çok kullandığımız ölçme araçlarımız da sınavlardır ve sınavları oluşturan her bir birim de sorularımızdır. Sınavların duyarlılığı da soru sayısının arttırılmasıyla sağlanır. Bir sınavda ne kadar çok soru sorulursa öğrencilere ilişkin o kadar duyarlı sonuçlar elde edilecektir. Düşününüz: Bir soru ile ölçülmek istenilen davranışa ilişkin elde edilen bilgi mi yoksa 10 soru ile elde edilen bilgi mi daha duyarlıdır?

Ölçme Aracının Güvenirliğini Arttırmanın Yolları:

1. *Soru sayısını arttırmak*: Yukarıda da belirtildiği gibi ne kadar çok soru sorarsak o kadar duyarlı yani güvenilir sonuçlar elde edebiliriz. Ancak soru sayısının da öğrencinin seviyesi, yaşı, sınavın ortamı, süresi gibi durumlara bağlı olarak mutlaka makul bir düzeyde tutulması gerekmektedir. Aksi taktirde güvenilir sonuç elde etmek için öğrencilerin cevaplayamayacakları kadar çok soru sormak hiç de doğru olmaz.
2. *Soruları açık,anlaşılır ve net ifadelerle hazırlamak*: Unutulmamalıdır ki sınavların amacı öğrenciye ilişkin ölçülecek özellik açısından olabildiğince gerçek verilere ulaşmaktır. Bu sebeple ölçülmek istenilen özellik için sorulan sorular net ve açık olmalıdır ki öğrencinin soruyu tam olarak anlayıp bilip-bilmediğini ortaya koyabilmesi mümkün olsun. Şayet öğrenci soruyu tam anlamazsa bizim öğrenmeye çalıştığımız bilgi düzeyini gözlemlememiz de mümkün olmaz. Sınavlar "ölçülmek istenilen özelliğe ilişkin bilgi toplamada ölçme aracı" olmanın dışında başka hiç bir amaçla kullanılmamalıdır.
3. *Öğrencilerin sınava olan motivasyonunu (güdülenmişliğini) arttırmak*: Bunun iki yolu vardır:
   - Öğretmenin ya da sınavı yapan kişi(ler)in tutumları: "Sınav" kavramı bile öğrencilerde (hatta herkeste) az ya da çok kaygı yaratır. Sınav sürecinde sınavı yapan kişilerin öğrencilerin bu kaygılarını azaltmaları, onların rahatlamasını sağlamaları çok önemlidir. Çünkü öğrencinin sınav kaygısı ne kadar azalırsa, gerçek durumunu ortaya koyma olasılığı o kadar artacak, yani daha az hata yapacak ve sınav sonuçları da böylece daha güvenilir olabilecektir.
   - Sınav sorularının kolaydan zora doğru sıralanması: Öğrencilerin sınavdaki kaygıları soruları çözebildikleri ölçüde azalmaya başlar. Bu sebeple özellikle ilk sorular kolay olacak şekilde düzenlenir ve soruların zorluğu giderek arttırılırsa öğrenciler de soruları yapabildikleri ölçüde kaygıları azalacak, sınava olan motivasyonları artacak ve öğrencilere ilişkin elde edilen sonuçlar böylece olabildiğince gerçeğe yakın, güvenilir olacaktır.
4. *Sınav süresini az ya da çok fazla olmayacak şekilde ayarlamak*: Sınav süresinin öğrencinin seviyesine, sınavın zorluk düzeyine göre olabilecek en makul düzeyde vermek de ölçme sonuçlarının güvenirliği açısından çok önemlidir. Yetersiz sınav süresi öğrencilerin soruları boş bırakmasına ya da atarak cevaplamalarına sebep olacaktır ki bu da sonuçların güvenirliğini düşürür. Benzer şekilde gerekenden çok daha fazla verilen sınav süresi de güvenirliği olumsuz yönde etkilemektedir. Sınav süresinin çok fazla verilmesi halinde öğrenci, ya yorgunluk vb. sebeplerle doğru verdiği cevaplarda çelişkiye düşmeye başlayarak yanlış cevaplara yönelmekte ya da kopya çekme eğilimi göstermektedir. Bu durumlar da yine ölçme sonuçlarının güvenirliğini düşürücü sebeplerdir.

egitim@hotmail.es

5. *Tesadüfi hata kaynaklarını olabildiğince azaltmak*: Hata konusunda bahsedilen hata kaynakları (öğretmen, öğrenci, sınav ortamı ve sınav araçları) ne kadar engellenirse, ölçme sonuçları da o kadar güvenilir olacaktır. Öğretmenin sınav sonuçlarını okurken göstereceği dikkat, öğrencinin hem fiziksel hem de psikolojik açıdan sınava hazır gelmesi, sınav ortamının ve sınav aracının sahip olması gereken özellikleri olabildiğince taşıması, sınavdan elde edilen sonuçların da o ölçüde güvenilir olmasını sağlayacaktır.

Güvenirliğin tanımı ve güvenirliğin arttırılmasına ilişkin yapılması gerekenleri anlatmaya çalıştık. Şimdi "Bir ölçme aracından (sınavlardan, testlerden) elde edilen sonuçların ne kadar güvenilir olduğunu nasıl bileceğiz, hesaplayacağız?" sorusunun cevabını bulacağız. Bunun için öncelikle "korelasyon" dediğimiz bir kavramı bilmemiz gerekecektir:

Güvenirlik Türleri ve Güvenirliği Kestirme Yöntemleri

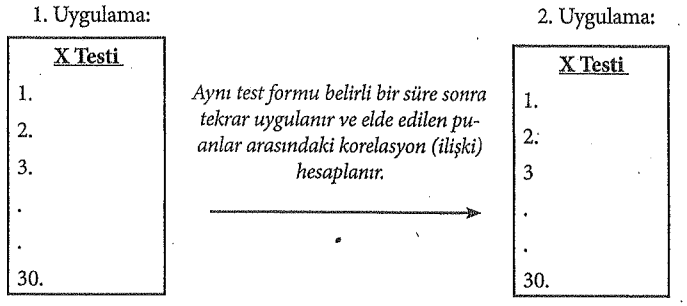
Aşağıda ölçme araçlarının (daha pratik olması açısından *test* olarak ifade edilmiştir) güvenirliğinin nasıl hesaplandığına ilişkin farklı yöntemler açıklanmıştır. Güvenirliğin kestirilmesinde korelasyon katsayısından yararlanılmaktadır. Yukarıda verilen korelasyon konusunda korelasyon değerinin [-1,+1] aralığında bulunduğunu söylemiştik. Ancak güvenirlik değeri [0, +1] arasında değer alır.

1. Test-TekrarTest Yöntemi: Bir test formunun farklı zamanlarda aynı gruba uygulanması sonucu elde edilen puanlar arasındaki korelasyon ile hesaplanır. Hesaplanan bu değer test-tekrar test güvenirliğini verir. Elde edilen değer 1'e ne kadar yakınsa testin güvenirliğinin o kadar yüksek, 0'a ne kadar yakınsa testin güvenirliğinin o kadar düşük olduğu sonucuna varılır.

Testin iki uygulaması sonucunda elde edilen puanlar arasındaki farklılığı (korelasyon değerinin 1'den uzaklaşması) ölçme aracının güvenirliğinin düşük olmasından kaynaklanabileceği gibi, test dışındaki fakörlerden de ortaya çıkmış olabilir:

1. Öğrencilerin iki uygulamadaki motivasyon, yorgunluk, sağlık gibi durumlarına ilişkin farklılıklardan meydana gelmiş olabilir.
2. Testin iki uygulamadaki ortamları arasındaki gürültü, ışık, ses, ısı gibi özellikleri arasındaki değişikliklerden kaynaklanmış olabilir.
3. İki uygulama arasında geçen zaman kısa tutulursa, hatırlama sonucu puanlarda değişim meydana gelecektir. İki uygulama arası zaman farkı uzun olursa bu kez de ölçülen özellikte değişimler meydana gelmiş, öğrenciler yeni bilgiler edinmiş, olgunlaşmış olabilir.

!!!!!Ölçülen değişkende zaman içinde değişmenin az olacağı düşünülüyorsa (duyuşsal alan davranışlarının ölçülmesi gibi) ve aracın kararlılığı hakkında bilgi edinilecekse test tekrar test yöntemiyle güvenirlik hesaplanır.



Test-Tekrar test yöntemi ile elde edilen güvenirlik katsayının hesaplanması, aşağıda verilen örnek ile pekiştirilmeye çalışılmıştır.

Örnek: 30 maddelik bir akademik tutum ölçeği iki hafta arayla aynı sınıfa uygulanmıştır. Bu ölçeğin test-tekrar test yöntemi güvenirliği için iki uygulamadan elde edilen puanlar arasındaki korelasyon katsayısı hesaplanır. Bu hesaplama aşağıda gösterildiği şekilde yapılır:

Korelasyon katsayısı: $$r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (X)^2\right]\left[n\sum Y^2 - (Y)^2\right]}} \qquad (4.1)$$

I. uygulama puanları (X)	II. uygulama puanları (Y)	X^2	Y^2	XY
15	15	225	225	225
12	10	144	100	120
9	11	81	121	99
21	21	441	441	441
3	5	9	25	15
27	27	729	729	729
18	20	324	400	360
15	15	225	225	225
27	27	729	729	729
21	21	441	441	441
$\Sigma X = 168$	$\Sigma Y = 172$	$\Sigma X^2 = 3348$	$\Sigma Y^2 = 3436$	$\Sigma XY = 3384$

$$r_{xy} = \frac{10\cdot 3384 - 168\cdot 172}{\sqrt{\left[10\cdot 3348 - 168^2\right]\left[10.3436 - 172^2\right]}} = \frac{33480 - 28896}{\sqrt{[33480 - 28224][34360 - 29584]}}$$

$$= \frac{5544}{\sqrt{5256\cdot 4776}} = \frac{4944}{5010\cdot 26} = 0.98$$

olarak korelasyon değeri elde edilir. Bu değer uygulanan akademik tutum ölçeğinin test-tekrar test güvenirliğini vermektedir.

egitim@hotmail.es

2. Paralel (Eşdeğer) Formlar Yöntemi: Paralel (eşdeğer) iki test formunun aynı gruba aynı zamanda uygulanması sonucu elde edilen puanlar arasındaki korelasyon hesaplanır. Testlerin paralel olabilmesi için testi oluşturan soruların güçlük düzeylerinin, ayırıcılık düzeylerinin vb. istatistiksel özelliklerin aynı olması gerekmektedir. Hesaplanan bu değer her iki formunda güvenirliğini verir. Elde edilen değer 1'e ne kadar yakınsa testlerin güvenirliğinin o kadar yüksek, 0'a ne kadar yakınsa testlerin güvenirliğinin o kadar düşük olduğu sonucuna varılır. Nadiren de olsa iki paralel form farklı zamanlarda da uygulanabilir. Böylece her iki formdan elde edilen ölçme sonuçlarının tutarlılığı anlamında güvenirliğiyle birlikte ölçme sonuçlarının zaman içindeki kararlılığı anlamına gelen güvenirliği de hesaplanmış olur.

Paralel formlar yöntemi, test tekrar test yönteminin bir dezavantajı olan soruların hatırlanması ya da öğrencilerin olgunlaşması, yeni bilgiler edinmeleri gibi faktörleri ortadan kaldırmaktadır. Ancak bu yöntemin de kendine özgü bazı dezavantajları bulunmaktadır:

1. İki test formunun paralelliğinin sağlanması oldukça güç olabilir. Testlerin güvenirliği de paralelliğin sağlanabildiği ölçüde doğru olacaktır.
2. İki paralel formun hazırlanması emek ve zaman açısından iki kat yük getireceğinden pratik bir yöntem değildir.

X Testi
1.
2.
3.
.
.
10.

Aynı kapsamı ölçen, aynı güçlük ve ayırıcılık düzeyinde fakat farklı sorulardan oluşan iki ayrı paralel test formu uygulanır. İki paralel formdan elde edilen puanlar arasındaki korelasyon (ilişki) hesaplanır.

Y Testi
1.
2.
3
.
.
10.

Paralel formlar yöntemi ile elde edilen güvenirlik katsayının hesaplanması, aşağıda verilen örnek ile pekiştirilmeye çalışılmıştır.

Örnek: Bir matematik öğretmeni hazırladığı 10'ar maddelik (soruluk) iki parelel test formunu sınıfına uygulanmıştır. Bu paralel formların güvenirliği için iki formdan öğrencilerinin aldığı puanlar arasındaki korelasyon katsayısını aşağıda gösterildiği şekilde hesaplamıştır.

Korelasyon katsayısı: $r_{xy} = \dfrac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (X)^2\right]\left[n\sum Y^2 - (Y)^2\right]}}$

X formu puanları	Y formu puanları	X^2	Y^2	XY
5	4	25	16	20
4	6	16	36	24
3	8	9	64	24
7	7	49	49	49
1	2	1	4	2
9	10	81	100	90
6	6	36	36	36
5	6	25	36	30
9	9	81	81	81
7	7	49	49	49
$\Sigma X = 56$	$\Sigma Y = 65$	$\Sigma X^2 = 372$	$\Sigma Y^2 = 471$	$\Sigma XY = 405$

$$r_{xy} = \frac{10 \cdot 405 - 56 \cdot 65}{\sqrt{\left[10 \cdot 372 - 56^2\right]\left[10.471 - 65^2\right]}} = \frac{4050 - 3640}{\sqrt{[3720 - 3136][4710 - 4225]}} =$$

$$= \frac{410}{\sqrt{584 \cdot 485}} = \frac{410}{532 \cdot 2} = 0.77$$

olarak korelasyon değeri elde edilir. Bu değer uygulanan matematik test formlarından elde edilen puanların güvenirliğini vermektedir

3. Eşdeğer Yarılar (İki yarıya ayırma) Yöntemi: Bu yöntemde hazırlanan tek bir test formu bir gruba uygulanır. Uygulanan bu testin puanlaması iki yarıya bölünerek yapılır. Öğrencilerin iki yarıdan aldıkları puanlar arasındaki korelasyon katsayısı hesaplanır. Elde edilen bu değer testin her bir yarısının güvenirlik değerini vermektedir. Aşağıda verilen formül ile testin tümünün güvenirliği hesaplanır:

$$r_{xx} = \frac{2r_{12}}{1 + r_{12}}$$

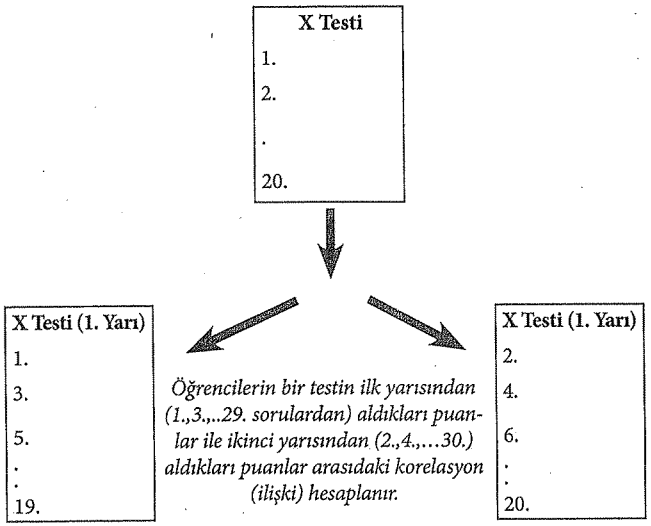
r_{xx} : tüm testin güvenirliği

r_{12} : eş değer yarıların güvenirliği

✓ Tek bir test formunun hazırlandığı ve tek bir uygulamanın yapıldığı bu yöntemde test-tekrar test ve paralel formlar yöntemindeki dezavantajlar ortadan kalkmaktadır.

egitim@hotmail.es

- ✓ Yalnız unutulmaması gereken çok önemli bir nokta vardır ki bu yöntem ancak testteki tüm sorular homojen (güçlük düzeyleri aynı, aynı kapsamı ölçen sorular) olduğu taktirde kullanılabilir. Aksi durumda bu yöntemle hesaplanan güvenirlik değeri doğru olmayacaktır.
- ✓ Puanlamada, testi iki yarıya ayırma işlemi değişik şekillerde yapılabilir. Ancak en yaygın olarak kullanılan yöntem, öğrencilerin testteki tek numaralı sorulara verdikleri cevaplar ile çift numaralı sorulara verdikleri cevapları ayrı ayrı puanlamaktır.



Öğrenci:	1.yarı puanı	2.yarı puanı
Sevgi	9	10
Özgür	10	8
.....	.....	.....
.....	.....	.....
.....	.....	.....
Ahmet	5	7

Eş değer yarılar yöntemi ile elde edilen güvenirlik katsayının hesaplanması, aşağıda verilen örnek ile pekiştirilmeye çalışılmıştır.

Örnek: Bir Türkçe öğretmeni, yazım kurallarına ilişkin hazırladığı homojen sorulardan oluşan 20 maddelik (soruluk) testi öğrencilerine uygulamıştır. Daha sonra bu testin güvenirliğini eş değer yarılar yöntemi aşağıda gösterildiği şekilde hesaplamıştır.

Korelasyon katsayısı: $r_{12} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (X)^2\right]\left[n\sum Y^2 - (Y)^2\right]}}$

1. yarı puanları (X)	2. yarı puanları (Y)	X^2	Y^2	XY
4	5	16	25	20
6	4	36	16	24
3	2	9	4	6
7	7	49	49	49
2	1	4	1	2
10	9	100	81	90
6	6	36	36	36
5	4	25	16	20
8	7	64	49	56
7	4	49	16	28
$\Sigma X = 58$	$\Sigma Y = 49$	$\Sigma X^2 = 388$	$\Sigma Y^2 = 293$	$\Sigma XY = 331$

$$r_{12} = \frac{10 \cdot 331 - 58 \cdot 49}{\sqrt{\left[10 \cdot 388 - 58^2\right]\left[10 \cdot 293 - 49^2\right]}} = \frac{3310 - 2842}{\sqrt{[3880 - 3364][2930 - 2401]}}$$

$$= \frac{468}{\sqrt{516 \cdot 529}} = \frac{468}{522.5} = 0.89$$

olarak korelasyon değeri elde edilir. Bu değer uygulanan Türkçe testinin her bir yarısının güvenirlik değeridir.

$$r_{xx} = \frac{2r_{12}}{1 + r_{12}} = \frac{2 \cdot 0 \cdot 89}{1 + 0.89} = \frac{1.78}{1.89} = 0.94$$

Elde edilen bu sonuçtan da görüleceği üzere Türkçe testinin güvenirliği yarılar güvenirliğinden daha yüksek bir değere sahiptir.

DÜŞÜNÜNÜZ!: Sizce eş değer yarılar güvenirliği hesaplanırken, tüm testin güvenirliği eş değer yarılardan elde edilen güvenirlik değerinden her zaman daha yüksek midir? Aksi durum mümkün olabilir mi? Eş değer yarıların güvenirliği ile testin tümünün güvenirliği aynı değere sahip olabilir mi?

egitim@hotmail.es

4. Kuder-Richardson KR-20 - KR-21 Formülleri: Bu formüller doğru cevabın "1" puan, yanlış veya boş bırakılmış soruya verilen puanın "0" olduğu tüm testlerin güvenirliğini hesaplamak için kullanılır.

- ✓ Diğer yöntemlerde ortaya çıkan dezavantajlar bu hesaplamalarda artık bulunmamaktadır.
- ✓ Bu formüllerle elde edilen değerler 1' e yakın olduğu ölçüde test güvenilir yani hatalardan o derecede arınık, 0'a yakın olduğu ölçüde de test hata içermekte yani o derecede güvenirliği düşük olmaktadır.
- ✓ ***KR-20:*** Her maddenin birbirine paralel olduğu varsayımına dayalı olarak hesaplanan bir korelasyon katsayısıdır. Maddelerin güçlük düzeyleri biliniyor ya da hesaplanabiliyorsa kullanılır.

$$KR-20=\frac{K}{K-1}\left[1-\frac{\sum s_x^2}{S_x^2}\right]$$

K: Testte yer alan madde (soru) sayısı

s_x^2: her bir maddenin puanları varyansı

S_x^2: test puanlarının varyansı

Not: Bu formül ile testin güvenirliğinin hesaplanması Bölüm 6'da ayrıntılı olarak açıklanmıştır.

- ✓ ***KR-21:*** Maddelerin güçlük düzeyleri bilinmiyor ya da hesaplanamıyorsa kullanılır. Tüm maddelerin güçlüklerinin birbirine eşit olduğu varsayılır.

$$KR-21=\frac{K}{K-1}\left[1-\frac{K\overline{X}-\overline{X}^2}{KS_x^2}\right]$$

K: Testte yer alan madde (soru) sayısı

$\overline{X}$: testin aritmetik ortalaması

S_x^2: testin varyansı

- ✓ *KR-21 ≤ KR-20 ilişkisi vardır. KR-21 değeri güvenirlik için bir alt sınır değerdir.*

Madde varyansı kavramı 6. bölümde, test varyansı ve aritmetik ortalama kavramları da 7. bölümde ayrıntılı olarak açıklanmıştır.

5. Cronbach alfa (α) Yöntemi: KR-20 yöntemine benzer. Tek farkı anket ve tutum ölçekleri gibi çoklu puanlanan (0-1 puanlamadan farklı puanlama yapılan) maddelerden (sorulardan) oluşan testlerde uygulanır.

KR-20, KR-21 ve Cronbach alfa (α) güvenirlikleri maddeler arası tutarlılığın bir ölçüsünü yani testin iç tutarlılığının ölçüsünü verir.

KR-20, KR-21 ve Cronbach alfa (α) güvenirlikleri yüksekse maddeler arası ve madde-test korelasyonlarının yüksek olduğu söylenebilir.

KR-20, KR-21 ve Cronbach alfa (α) güvenirlikleri, tek boyutlu özelliklerin ölçüldüğü testler için bir kapsam geçerliği ölçüsü olarak da işlev gösterebilir.

6. Puanlayıcı Güvenirliği: Bazı testlerin (örneğin: yazılı sınavlar, sözlü sunumlar, vb..) puanlanması subjektif (öznel) olabilmektedir. Bu tür testler birden fazla puanlayıcı tarafından puanlanır. Bu puanlayıcıların verdikleri puanlar arasındaki korelasyon (ilişki) hesaplanarak puanlayıcıların verdikleri puanların güvenirliği kestirilmiş olur.

Yukarıda açıklanan tüm güvenirlik türleri belirleme yöntemleri aşağıdaki tablo ile de özetlenmeye çalışılmıştır.

GÜVENİRLİK TÜRLERİ VE BELİRLEME YÖNTEMLERİ

(özet tablo)

Birden Çok Uygulama Gerektiren Yöntemler	
- Test Tekrar-Test Yöntemi:	1 Grup + 1 Test + 2 Uygulama (Aynı gruba aynı test farklı zamanlarda uygulanır)
TTT Güvenirlik Katsayısı;	*aynı bireylerin iki farklı zamandaki aynı testten aldıkları puanların dağılımları arasındaki korelasyon katsayısıdır.
	*ne kadar yüksekse, test puanları test dışındaki çevresel ve bireysel değişkenlerden o kadar az etkilenmektedir.
	* testten elde edilen sonuçların farklı zamanlarda ne derece kararlı olduğunun ölçüsünü verir.
	*1'e yaklaştıkça uygulama sonucu elde edilen puanlar birbirine yakın, 0'a yaklaştıkça uygulama sonucu elde edilen puanlar birbirinden farklı anlamına gelir.
Dezavantajları:	- İki uygulama sonucunda elde edilen puanlar arasındaki farklılıkların kaynağı ölçme aracının güvenirliğinin düşük olması olacağı gibi, test dışındaki kaynaklardan da ortaya çıkabilir:
	4. Bireylerde meydana gelen motivasyon, yorgunluk, sağlık gibi durumlarına ilişkin farklılıklar.
	5. Ortamdan kaynaklan an gürültü, ışık, ses, ısı gibi değişkenlerden kaynaklanabilir.
	6. Aradan geçen zaman etkisiyle ölçülen özellikte değişimler meydana gelmiş olabilir.
	- Zaman kısa tutulursa, hatırlama sonucu puanlarda değişim meydana gelebilir.
	- Uzun zaman so nra ikinci uygulama yapılırsa bu kez de bireylerin olgunlaşması sonucu puanlarda değişim meydana gelebilir.

egitim@hotmail.es

!!!!!Ölçülen değişkende zaman içinde değişmenin az olacağı düşünülüyorsa ve aracın kararlılığı hakkında bilgi edinilecekse kullanılır.	
- Eşdeğer (Paralel) Formlar Yöntemi:	1 Grup + 2 Test + 2 Uygulama (Aynı gruba paralel testler farklı/aynı zamanlarda uygulanır.
EF Güvenirlik Katsayısı (Eşdeğerlik Katsayısı);	* iki formdan elde edilen puanlar arasındaki korelasyon katsayısıdır. *1'e yakınsa eşdeğer form güvenirliği yüksek, 0'a yakınsa eşdeğer form güvenirliği düşük anlamına gelmektedir. *hem ölçme sonuçlarının zaman içindeki kararlılığını hem de formlara verilen cevapların tutarlılığının ölçüsünü verir.
!!!!! TTT Güvenirliğinin bir dezavantajı olan hatırlama ortadan kalkar. Zaman aralığı problemi de nispeten ortadan kalkar. İki formdaki sorular farklı olduğu için arka arkaya uygulanabilir.	
Dezavantajları:	- En önemli dezavantajlarından biri testlerin **paralelliğini sağlamak** oldukça güçtür. - Uygulamalar arası zaman çok kısa ya da arka arkaya uygulama yapılırsa, yorgunluk, bıkkınlık gibi etkenler ortaya çıkar ve test puanları etkilenir. - Uygulamalar arası zaman uzun tutulursa, ölçülen özellikte değişimler meydana gelmiş olabilir. - Formların veriliş sırası da, puanların değişimine sebep olan bir etken olabilir.

Tek Uygulamalı Yöntemler	
- Eşdeğer Yarılar (Testi Yarılama) Yöntemi:	1 Grup + 1 Test + 1 Uygulama (Aynı gruba aynı test farklı zamanlarda uygulanır)
!!!!!Ölçülen değişkende zaman içinde değişmenin az olacağı düşünülüyorsa ve aracın kararlılığı hakkında bilgi **edinilmeyecekse** kullanılır. Bu yöntem ile güvenirliğin belirlenebilmesi için mutlaka ölçme aracının **HOMOJEN** sorulardan (maddelerden) oluşması gerekir. Aksi taktirde elde edilen güvenirlik değeri doğru sonuç vermez.	
EY güvenirliği;	* kapsam tutarlığına ilişkni bir ölçü verir. * ölçme aracından elde edilen sonuçların (puanların) zaman içindeki kararlığına ilişkin **bilgi vermez**. * yarıların eşdeğerlik derecesini verir. * eşdeğer yarılardan birinin güvenirliğini verir, testin tümüne ilişkin güvenirlik de bu katsayılar yardımıyla hesaplanır.
!!!!! Her iki yarının birbirine paralel olduğu varsayılır. (= her iki yarının varyansları ve ortalamaları eşit kabul edilir)	
Dezavantajları:	- Testin nasıl ikiye bölüneceğine karar vermek bir problemdir.

EY Yönteminde;	- testteki ilk sorular ile son soruların cevaplandırılmasında yorgunluk, sıkılma, sorulara ve ortama alışma gibi faktörler aynı olamayacağından, soruların ayrımının tek-çift numaralı sorular şeklinde yapılması önerilir. - test tek boyutlu ve soruların güçlük düzeyleri birbirine yakınsa, iki yarı rasgele olarak belirlenebilir. - test maddeleri homojense, maddeler arası ilişki o derece yüksek ilişki göstereceği için maddeler arası tutarlık da o derece artacaktır. -maddeler heterojen hale geldikçe maddeler arası ilişki ve dolayısıyla maddeler arası tutarlık da düşecektir. Örn: 10 soruluk bir Türkçe sınavında 3 soru imla bilgisi, 4 soru cümlenin öğeleri, 3 soru da paragraf anlamına yönelikse, bu sınavdan 5 puan alan öğrencilerden biri; 1 soru imla bilgisi, 2 soru cümlenin öğeleri, 2 soru da paragraf anlamına doğru cevap verirken, diğer bir öğrenci 2 soru imla bilgisi, 2 soru cümlenin öğeleri, 1 soru da paragraf anlamına doğru cevap verdiyse Bu iki bireyin performansları eşit midir? Sorusunun cevabı bu tür heterojen testlerde belirsizdir.
Diğer İÇ TUTARLILIK katsayıları:	
0-1 Puanlama yapılan durumlarda:	
KR-20: Her maddenin birbirine paralel olduğu varsayımına dayalı olarak hesaplanan bir korelasyon katsayısıdır. Maddelerin güçlük düzeyleri biliniyor ya da hesaplanabiliyorsa kullanılır. ***KR-21:*** Maddelerin güçlük düzeyleri bilinmiyor ya da hesaplanamıyorsa kullanılır. Tüm maddelerin güçlüklerinin birbirine eşit olduğu varsayılır.	
!!!!!! KR-21 ≤ KR-20 ilişkisi vardır. KR-21 değeri güvenirlik için bir alt sınır değerdir	
0-1 Dışında Puanlama Yapılan Durumlarda:	
Cronbach Alfa (α) Güvenirliği: KR-20 ile aynı mantık üzerine kurulmuştur. ***Puanlayıcı Güvenirliği:*** Farklı puanlayıcılardan elde edilen puanlar arasındaki tutarlılığın derecesini verir.	

Ölçme Sonuçlarının Standart Hatası

2. bölümde yer alan "Ölçmede Hata" konusunda da açıklanmaya çalışıldığı gibi yapılan her ölçme işleminde az ya da çok hata bulunmaktadır. Söz konusu bir de, eğitimde olduğu gibi bilişsel ya da duyuşsal alanlarda ölçme yapmaksa hata kaçınılmaz olmaktadır. O halde diyebiliriz ki eğitimde öğrencilerimize uyguladığımız sınav, test vb. ölçme araçlarından elde ettiğimiz her ölçme sonuçlarında (puanlarda) bir miktar

egitim@hotmail.es

tesadüfi hata vardır. İşte bu hataların standart bir miktarını elde edebilmek için kullanılan eşitliğe *ölçme sonuçlarının* ya da kısaca *ölçmenin standart hatası* adı verilir. Standart hata aşağıdaki şekilde hesaplanır:

$$Sem = S_x\sqrt{1 - r_{xx}^2}$$

Sem: Ölçmenin standart hatası

S_x : Test puanlarının standart sapması

r_{xx} : Testin güvenirliği

Eşitlikten de görüleceği üzere bir sınavın ya da testin standart hatası o testin standart sapması ve güvenirliğine bağlıdır. Standart hata, standart sapma ile doğru orantılı güvenirlik ile ters orantılıdır. Daha açık söylemek gerekirse, testin ya da sınavın standart sapması arttıkça standart hatası artar, standart sapması azaldıkça standart hatası azalır ve testin ya da sınavın güvenirliği arttıkça standart hatası azalır, güvenirliği azaldıkça standart hatası artar.

Ölçme Sonuçlarının Güven Aralığının Hesaplanması

Yukarıdaki bölümde de açıklandığı gibi eğitimde kullandığımız ölçme araçlarından elde ettiğimiz sonuçlarda az ya da çok hata bulunmaktadır. Bu sebeple diyebiliriz ki eğitimde ne zaman bir sınav uygulamış ve bu sınavdan öğrencilerimizin almış olduğu puanları belirlemişsek bu puanların bir miktar hata barındırdığını da göz önünde bulunduracağız demektir. İşte bu noktada devreye standart hata girmektedir. Eğer öğrencilerimizin puanlarını çok daha gerçeğer yakın değerlerde ifade etmek istiyorsak, belirli olasılıklara gore standart hatayı kullanarak puanların (ölçme sonuçlarının) güven aralıklarını hesaplamak çok daha doğru bir yaklaşım olacaktır. Ölçme sonuçlarına karışan hata miktarın az ya da çok olduğu bilinmediğinden, bir başka deyişle hatanın artı ya da eksi yönlü olduğu bilinmediğinden standart hata miktarı elde edilen ölçme sonuçlarına bir eklenip bir de çıkarılarak bir aralık değeri elde edilir. Bu aralığa güven aralığı adı verilir. Ancak belirli olasılık değerlerine karşılık gelen farklı güven aralık değerleri hesaplanmaktadır. Bunu şu şekilde açıklayalım.

Eğer bir öğrencimiz girmiş olduğu sınavdan X puan almışsa, bu öğrencimizin;

- %68 olasılıkla puan güven aralığı: X ± 1Standart hata (Sem) → (X- Sem, X+ Sem)
- %95 olasılıkla puan güven aralığı: X ± 2Standart hata (Sem) → (X- Sem, X+ 2Sem)
- %99 olasılıkla puan güven aralığı: X ± 3Standart hata (Sem) → (X- 3Sem, X+ 3Sem)

Burada ifade edilen hesaplamaları aşağıda bir örnek ile açıklamaya çalışalım:

ÖRNEK: Ayşe matematik sınavından 80 puan almıştır. Matematik sınavının standart hatası 4 olduğuna göre, Ayşe'nin aldığı puanın %68, %95 ve %99 olasılıklarla güven aralıklarını hesaplayalım:

- %68 olasılıkla Ayşe'nin matematik sınavından aldığı puanın güven aralığı:

  80 ± 1Sem = 80 ± 4 ⇒ (76, 84) olacaktır.

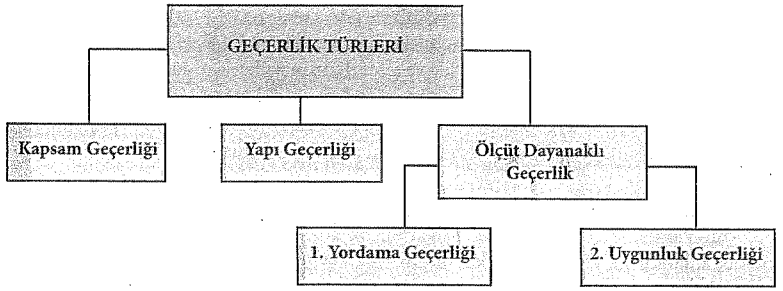
- Ayşe'nin puanı %68 olasılıkla (76, 84) puan aralığında yer alır.
  - **%95** olasılıkla Ayşe'nin matematik sınavından aldığı puanın güven aralığı:

    80 ± 2Sem = 80 ± 2.4= 80 ± 8 ⇒ (72, 88) olacaktır.
- Ayşe'nin puanı %95 olasılıkla (72, 88) puan aralığında yer alır.
  - **%99** olasılıkla Ayşe'nin matematik sınavından aldığı puanın güven aralığı:

    80 ± **3Sem** = 80 ± 3.4 = 80±12 ⇒ (68, 92) olacaktır.
- Ayşe'nin puanı %99 olasılıkla (68, 92) puan aralığında yer alır.

GEÇERLİK [-1 , 1]

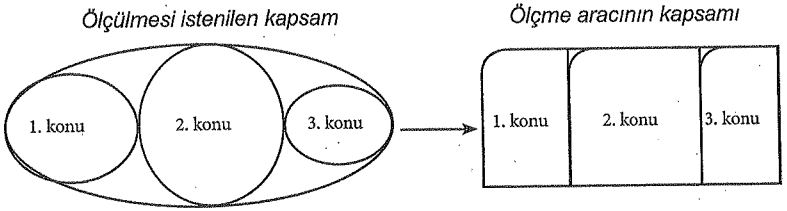
- Bir ölçme aracının ölçmesi istenilen özelliği ne derece ölçebildiğinin derecesine o ölçme aracının geçerliği adı verilir. Bir başak deyişle, ölçme aracının ölçtüğü değişkeni (özelliği, niteliği) başka hiç bir değişken karışmaksızın ölçebilmesinin derecesidir.
- Bir ölçme aracının (test, sınav, anket vb…) hatasız ölçebilmesi güvenirliğini, ölçülen özelliği başka özellik karışmaksızın ölçebilmesi de geçerliğini gösterir.
- Bir ölçme aracının *geçerliği*, o ölçme aracıyla ölçülmek istenen niteliği (örneğin; başarı) ölçmesi, başka bir nitelikle (öğrencinin şansı, tutumu, cinsiyeti vb…) karıştırmamasıdır.
- Ölçme aracının geçerli olabilmesi için güvenlir olması şarttır ancak yeterli değildir. Güvenilir olan her ölçme aracı geçerli olmayabilir. Bir ölçme aracının geçerliği varsa mutlaka güvenilirdir.



egitim@hotmail.es

1. Kapsam Geçerliği

Ölçme aracının ölçmesi istenilen kapsamı, dengeli bir şekilde, ne eksik ne de fazla olmaksızın ölçebilmesinin derecesidir. Örneğin;



Yukarıda şekille de açıklanmaya çalışıldığı gibi ölçülmesi istenilen kapsam farklı içerik ve yoğunluktaki üç konudan oluşuyorsa, bu kapsamı ölçecek ölçme aracı da kapsam dahilindeki üç konunun üçünü de ne eksik ne fazla olmaksızın ve yoğunluklarına uygun şekilde soru sayısına sahip olacak kapsamda oluşturulmalıdır.

- Kapsam geçerliğinin belirlenmesinde iki yol izlenebilir:

• ***Belirtke tablosu hazırlamak:*** Bu yola rasyonel (mantıksal) yaklaşım da denir.

  - Belirtke tablosu: Belirtke tablosu bir boyutunda (satır) ölçmenin yapıldığı konuların ana başlıklarının; diğer boyutunda (sütun) ise davranışların hangi bilişsel boyutta yer aldığını göstererek, hücrelerde satır ve sütunda kesişen davranışlardan kaç tane bulunduğunu belirten bir tablodur.
  - Ölçme aracının belirtke tablosundaki davranışları ne derecede ölçtüğü, konu alanında ve ölçme alanında uzman olan kişilerin görüşleri alınarak incelenir. Bunun sonucunda da kapsam geçerliğine ilişkin bir kanıya varılır. Önemli olan ölçme aracının tablodaki tüm davranışları yine tabloda belirtilen ağırlıkta yansıtıyor olmasıdır.

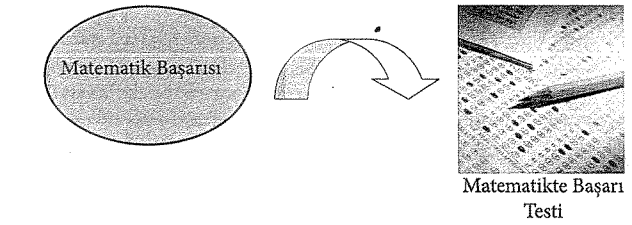
• ***İstatistiksel hesaplama:*** Ölçmek istenilen özellik ya da özelliklere ilişkin daha önceden hazırlanmış, güvenirlik ve geçerliğinin yüksek olduğu belirlenmiş bir ölçme aracı ile hazırlanan (kapsam geçerliği hesaplanmak istenen) yeni ölçme aracı aynı gruba uygulanarak aralarındaki korelasyon hesaplanır. Hesaplanan bu değer kapsam geçerliğinin bir göstergesi olarak yorumlanır.

2. Yapı Geçerliği

Bireylerde var olduğu kabul edilen ancak gözle gözlenemeyen özellikler "yapı" olarak adlandırılır. Örneğin; "zeka, tutum, ilgi, kaygı vb.." gibi bireysel özellikler birer yapıdır.

Bu yapılar gözlenemediği için bunlarla ilişkisi olduğu kabul edilen maddelere (sorulara) verilen cevaplarla, bireylerin bu yapılara sahip olup olmadıkları gözlenmeye çalışılır. Örneğin zekânın IQ testi (zeka testi) dediğimiz testler ile ölçülerek bir şekilde ne derece var olup olmadığının saptanmaya çalışılması gibi.

Yapı geçerliliği ölçme aracının teoride var olduğu düşünülen, gözle görülemeyen bu yapıları pratikte ölçebilme derecesine denir. Örneğin; matematik başarısını ölçmek amacıyla hazırlanan bir testin, matematik dışında (Türkçe, fizik, vb...) başka alanlardaki başarıyı ölçmeden, yalnızca matematik başarısını doğru bir şekilde ölçebilme derecesi o testin yapı geçerliği ile ilgilidir.



Bir başka deyişle, yapı geçerliği ölçme aracıyla ölçülmek istenilen özelliğin başka özellikler karışmadan ölçülebilme derecesidir. Örneğin; bir öğretmen 1. sınıf öğrencilerinin matematik dersine 10'a kadar sayılarla toplama işlemlerini doğru yapabilme başarılarını ölçmek istiyor olsun. Bu amaca yönelik hazırladığı 5 toplama işleminden oluşan sınav kağıdında bu işlemlerini yapmaları için 'aşağıdaki toplama işlemlerin yekününü bulunuz' ifadesini kulllanmış olsun. Uyguladığı bu sınav sonucunda, sınıfta toplama işlemlerini kolaylıkla çözebilen pek çok öğrencisi bu sınavda yer alan işlemleri yapmamış, boş bırakmış olabilir. Peki bu durumda, öğrenciler toplama işlemini yapmayı bilmedikleri için mi yoksa ifadede geçen 'yekün' kelimesinin anlamını bilmedikleri için mi bu işlemleri yapamamışlardır? Büyük olasılıkla kelimenin anlamını bilmediği için toplama işlemlerini yapabilecek öğrenciler de işlemleri yapmadan boş bırakmışlardır. Böylece öğretmen aslında amaçladığı toplama işlemlerini yapabilmedeki başarıyı değil bundan çok farklı bir yapıda yer alan 'kelime bilgisini' de ölçmüş olacaktır. Bu durumda öğretmini uyguladığı bu sınavın yapı geçerliği de düşmüş olur.

Benzer durum öğrencilerin fizik bilgisini ölçmek için hazırlanan bir soruda futbol terimlerinin kullanılması sonucu erkek öğrencilerin kız öğrencilere göre bu sorularda daha başarılı olması durumunda asıl ölçülen özelliğin fiizlk bilgisi değil futbol bilgisi olup sorunun cinsiyete göre yanlılık göstermesiyle ortaya çıkabilir. Sorulardaki bu tür yanlılık, bir başka deyişle, ölçme aracında yer alan maddelerin amacın dışında belirli bir grubun üstünlüğüne yol açması yapı geçerliğini tehdit eden bir sebeptir.

Yapı geçerliği her ölçme aracında olması gereken bir geçerlik türü olup çok önemlidir. Yapı geçerlğinin incelenmesinde alan uzmanlarının görüşünün alınmasıyla birlikte faktör analizi gibi istatistiksel yöntemler de kullanılmaktadır.

egitim@hotmail.es

3. Ölçüt Dayanaklı Geçerlik (Ölçüt Geçerliği)

✓ Ölçüt dayanaklı geçerlik, hazırlanan ölçme aracından elde edilen puanlar ile ölçüt (kriter) kabul edilen diğer bir ölçme aracı arasındaki korelasyon katsayısı hesaplanarak bulunur.

✓ Burada ölçüt olarak kabul edilen ölçme aracının güvenilir ve geçerli olması gerekmektedir. Ölçüt kabul edilen ölçme aracıyla elde edilen sonuçlar ile hazırlanan ölçme aracından elde edilen sonuçlar arasındaki korelasyon katsayısı ne kadar yüksek çıkarsa hazırlanan ölçme aracının da o derece geçerli ölçme sonuçları verdiği düşünülür. Kısacası hazırlanan ölçme aracının ölçüt dayanaklı geçerliğinin olduğu sonucuna varılır.

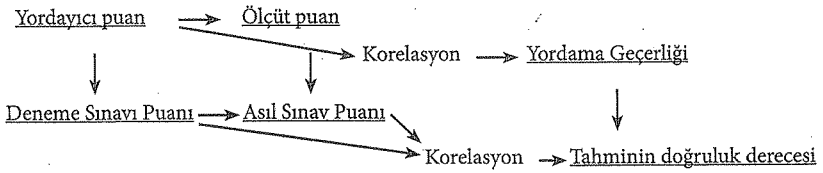
✓ Ölçüt dayanaklı geçerlik, ölçüt olarak seçilen ölçme aracının hazırlanan ölçme aracından önce ya da sonra uygulanmasına göre ikiye ayrılır:

a) Yordama (Tahmin) Geçerliği

- Ölçülen davranış ya da özelliğin gelecekteki durumu hakkında tahminde bulunmaya yordama adı verilir.
- Yordama geçerliğinde, hazırlanan ölçme aracıyla elde edilen puanlar gelecekteki bir durumu tahmin etmekte kullanılır. Bu sebeple ölçüt kabul edilecek puanlar henüz mevcut değil, gelecekte var olacak puanlardır. Hazırlanan ölçme aracından elde edilen puanların, gelecekte ortaya çıkacak ve asıl durumu gösterecek puanlar ile arasındaki korelasyon değeri ne kadar yüksek çıkarsa hazırlanan ölçme aracının yordama geçerliğinin o derece yüksek olduğu yorumu yapılır.

Örneğin; KPSS'ye hazırlık dershaneleri, asıl sınav öncesi öğrencilerine deneme sınavları (tahminde bulunmak üzere hazırlanan ölçme aracı) yaparlar. Bu sınavların amacı öğrencilerin asıl sınavdaki (gelecekte ortaya çıkacak, puanları ölçüt olarak kullanılanılacak ölçme aracı) durumları hakkında fikir edinebilmelerini sağlamaktır. İşte böyle bir deneme sınavından alınan puanlar ile öğrencilerin asıl sınavdan aldıkları puanlar arasındaki korelasyon hesaplanırsa, elde edilecek değer bize dershanenin uyguladığı sınavın yordama (tahmin) geçerliğinin ölçüsünü verecektir.

Bu örnekte deneme sınavından alınan puanlar yordayıcı (tahmin edici) puanlar, asıl sınavdan alınan puanlar da ölçüt puanlar olup, bu iki puan grubu arasındaki korelasyon katsayısı yordama geçerliğini verir.

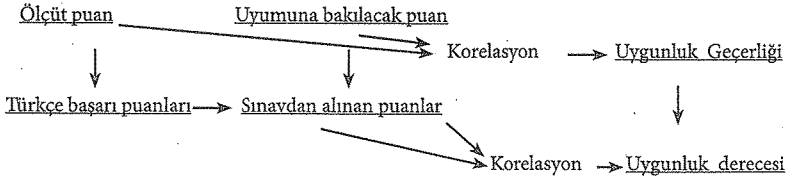


b) Uygunluk (Halihazır) Geçerliği

- Hazırlanan ölçme aracının, ölçülen davranış ya da özelliğin var olan durumu **uygun** bir şekilde ölçüp ölçmediğini belirlemek için uygunluk geçerliğinden yararlanılır.
- Uygunluk geçerliğinde, hazırlanan ölçme aracıyla elde edilen puanların var olan durumu doğru yansıtıp yansıtmadığı tespite çalışılır. Bu sebeple ölçüt kabul edilecek puanlar zaten mevcuttur ve bu puanların elde edildiği ölçme aracının güvenirlik ve geçerliğinin yüksek olması gerekmektedir. Hazırlanan ölçme aracından elde edilen puanların, var olan durumu gösteren puanlar ile arasındaki korelasyon değeri ne kadar yüksek çıkarsa hazırlanan ölçme aracının uygunluk geçerliğinin o derece yüksek olduğu yorumu yapılır.

Örneğin; Orta öğretimdeki bir sınıfın ders yılı ortalarında Türkçe öğretmeni değişmiş olsun. Bu sınıfta derse başlayan öğretmen bir süre sonra uyguladığı Türkçe sınavının uygunluk geçerliğini belirlemek istesin. Bu durumda Türkçe öğretmeni kendi uyguladığı sınavdan öğrencilerinin aldıkları puanlar ile öğrencilerinin önceki sınıftaki başarı puanları (bu puanların güvenilir ve geçerli olması gerekir) ile korelasyonuna bakarsa, bulacağı değer öğretmenin uygulamış olduğu sınavın uygunluk geçerliğinin ölçüsünü verecektir.

Bu örnekte önceki sınıftaki Türkçe başarı puanları ölçüt puanlar, öğretmenin hazırladığı sınavdan alınan puanlar da uyumuna bakılacak puanlar olup, bu iki puan grubu arasındaki korelasyon katsayısı uygunluk geçerliğini verir.



Geçerliği Arttırmanın Yolları

1. Güvenirliği arttırmak için yapılacaklar aynı zamanda geçerliği arttırmanın da bir yoludur. Ancak çok dikkat edilmesi gereken önemli bir koşul vardır: Soru sayısını arttırmak güvenirliği arttırır. Sınava eklenecek her soru güvenirliği arttırmasın karşın kapsam dışından eklenecek her soru kapsam geçerliğini tehdit eder. Bu sebeple her sorunun mutlaka kapsam dâhilinde olmasına dikkat edilmelidir.
2. Her bir soru tek bir davranışı ölçmelidir.
3. Soruların güçlük düzeyleri öğrencilerin düzeylerine ve sınavın amacına uygun olarak hazırlanmalıdır.
4. Her yıl aynı sorular kullanılmamalıdır.
5. Sınavdan önce sorular açıklanmamalı, sınavda ipucu verilmemelidir.
6. Sorular belirli bir kaynaktan aynen alınmamalıdır.

egitim@hotmail.es

Güvenirlik ile Geçerlik Arasındaki İlişki

- Bir sınav geçerli ise güvenilirdir. Ancak tersi her zaman doğru değildir. Yani bir sınav güvenilir ise geçerlidir ifadesi her zaman doğru değildir.
- Bir ölçme aracı güvenilir ise geçerlidir diyemeyiz. Güvenilir olup geçerli olmayabilir.
- Güvenirlik, bir ölçme aracının geçerliği için gerekli koşul olup, yeterli koşul değildir.
- Güvenirlik ve geçerliğin varlığı bir derece olarak ifade edilebilir. Güvenirlik ve geçerlik "var" ya da "yok" olarak ifade edilmez.

Güvenirlik ile geçerlik arasındaki ilişkiyi anlamak ve akılda tutmak bazı öğrenciler için zor olabilmektedir. Bu sebeple öğrencilerimin bu ilişkiyi kolayca anlamasını ve akılda kalıcılığı sağlayabilmek için sınıfta anlattığım bir benzetmeyi burada da sizlerle paylaşmak istedim:

Güvenilir Nişanlılıktan Geçerli Evliliğe

Genç kız ve delikanlı birbirlerinden hoşlanmaktadırlar ve güzel bir arkadaşlığa adım atarlar. Zamanla birlikte geçirdikleri güzel vakit onları birbirlerine daha da yakınlaştırır. Her şey çok daha güzel ve yaşanır olmuştur artık. Gelecek günleri de acısıyla tatlısıyla birlikte geçirebilmek adına birbirlerine olabildiğince ***güven*** duymaktadırlar ve bu *güvenin* sembolü olarak da *nişanlanırlar*. Ancak nişanlanmış olmaları onları resmiyette ve bulundukları herhangi bir ortamda geçerli bir çift yapmamıştır. Ne zaman ki *evlenirler* artık tüm ortam ve şartlar altında onlar *geçerli* bir çift olurlar.

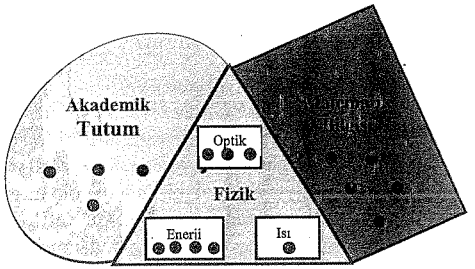
Dikkat: Bu hikâyenin **kuralı** çiftlerin evlenebilmesi için nişanlanmasının şart olmasıdır.

Buna göre hikayeden de anlaşılacağı üzere güvenirliğin sembolü nişanlılık geçerliğin sembolü evlilik gibi düşünüldüğünde güvenirlik ve geçerlik arasındaki ilişki kurulmuş olur. Özetle:

Nişanlı olan her çift evlenir diyemeyiz. Nişanlı olanlar evlenebilir de evlenmeyebilir de. Ancak evli olan her çift mutlaka nişanlanmıştır. (Kuralımızı unutmayalım!)

Güvenilir her ölçme aracı geçerlidir diyemeyiz. Güvenilir bir ölçme aracı geçerli olabilir de olmayabilir de. Ancak geçerli olan her ölçme aracı mutlaka güvenilirdir.

Güvenirlik ve geçerlik arasındaki farklılık ve ilişki aşağıdaki görsel açıklama ile pekiştirilmeye çalışılmıştır.



Varsayılım ki öğrencimizin ölçmek istediğimiz özelliğini (yapısını) görebiliyor, doğrudan ölçebiliyoruz. Öğrencimizin, gözlemleyemediğimiz özelliklerinin bir kesidinin yukarıda verilen sağda yer alan şekiller gibi olduğunu varsayalım. Bir başka deyişle, gözlenemeyen özelliklerinin bir kesidinde matematik bilgisi yapısı, fizik bilgisi yapısı ve akademik tutum yapısı yer alıyor olsun. Her bir şeklin içinde yer alan noktaların sayısı da öğrencimizin gerçek puanına karşılık gelsin. Bu durumda varsaydığımız gibi eğer ölçülmesi istenilen özellikleri doğrudan gözleyebiliyorsak öğrencimizin gerçek puanı: 17 puan olacaktır. Eğer ölçme aracımızla elde ettiğimiz sonuç da 17 puan ise gerçek puana ulaşmış olacak; tesadüfi hatamız hiç olmayacak güvenilir bir sonuç elde etmiş olacağız. (Yeniden hatırlayalım bu bir hayali durum). Ancak eğer ölçülmesi amaçlanan yapı fizik bilgisi yapısı ise elde ettiğimiz puan gerçek puan olmasına rağmen ölçülmesi istenilen yapı dışında bambaşka yapılar da (matematik bilgisi yapısı ve akademik tutum yapısı) ölçüldüğünden elde edilen ölçme sonucunun (17 puan) yapı geçerliği düşük olacaktır. Hatta fizik yapısı içinde sadece optik konusu ölçülmek isteniliyorsa bu durumda elde edilen puanın kapsam geçerliği de düşük olacaktır. Hem güvenilir hem de yapı ve kapsam geçerliğinin olduğu durumda elde edilmesi gereken puan 3 puan (optik parçası içindeki nokta sayısı) olmalıdır.

Eğer 17 yerine 15 puan elde etmişsek (sınav kaygısı, yorgunluk, ölçme ortamındaki gürültü ya da puanlama hatasından biri ya da birkaçı sonucu) ya da 17 yerine 18 puan elde etmişsek (kopya ya da şans başarısı gibi sebeplerden dolayı) gözlediğimiz puanda tesadüfi bir hata olacaktır. Bu durumda güvenilir bir sonuç bulunmadığından elde edilen sonuç zaten geçerli olmayacaktır.

egitim@hotmail.es

KULLANIŞLILIK

- ✓ Ölçme aracının kullanışlılığı, o ölçme aracının hazırlanmasının, uygulanmasının ve puanlanmasının ekonomik ve pratik olma derecesidir.
- ✓ Ölçme aracının kullanışlılığı, güvenirliği ve geçerliği artırıcı yönde rol oynamalıdır.
- ✓ Ölçme aracının kullanışlılığı, o ölçme aracının hazırlanmasının, uygulanmasının ve puanlanmasının ekonomik ve pratik olma derecesidir.

Ölçme aracının kullanışlılığı, güvenirliği ve geçerliği artırıcı yönde rol oynamalıdır. Kullanışlığı arttırmak amacıyla alınacak bir önlem, ölçme aracının güvenirliğini ya da geçerliğini düşürecekse uygulanmamalıdır. Güvenirliğe ve geçerliğie olumlu etkisi olabilecek kullanışlığı arttırıcı özelliklerin kullanılması yerinde olacaktır.

- ✓ Güvenirliği ve geçerliği olumsuz yönde etkilemediği sürece daha az masraf ve emekle oluşturulabilecek ölçme araçlarının kullanışlılığı da daha yüksek olacaktır.
- ✓ Ölçme aracının hazırlanması, uygulanması ve puanlaması için harcanan sürenin kısalığı da kullanışlığı arttırır. Buna göre yazılı sınavlar çoktan seçmeli sınavlara göre daha kısa sürede hazırlanabildiğinden hazırlanması açısından daha kullanışlıyken çoktan seçmeli sınavlar da yazılı sınavlara göre daha kısa sürede puanlanabildiğinden puanlanması açısından daha kullanışlıdır.
- ✓ Ölçme araçlarının cevaplanması öğrenci açısından ne kadar rahat ve kolaylıkla yapılabiliyorsa bu derecede kullanışlıdır. Ölçme aracında; sınavın süresi, kaç soru içerdiği, nasıl puanlanacağı, hangi amaç için kullanılacağı vb. açıklayıcı bilgilerin yer aldığı bir yönergenin bulunması kullanışlığı arttırıcı bir önlemdir.
- ✓ Ölçme aracının düzeni, yazı karakterinin büyüklüğü gibi özelliklerine bağlı olarak kolaylıkla okunabiliyor ve cevaplanabiliyor olması da kullanışlığı arttırıcı özellikler arasında yer almaktadır. Özellikle çoktan seçmeli sınavlarda her bir maddenin bir bütün olarak, bölünmeden bir sayfada yer alması; bir başka deyişle madde kökünün bir sayfada seçeneklerin diğer bir sayfada olmayacak şekilde düzenlenmesi cevaplayıcılar açısından oldukça önemlidir.
- ✓ Ölçme araçlarının puanlanması için bir ceva anahtarının oluşturulması da kullanışlığı arttıran özellikler arasında yer alır. Özellikle yazılı sınav, sözlü sınav gibi puanlaması subjektif olan ölçme araçları için bir cevap anahtarı ya da dereceli puanlama anahtarının hazırlanması hem puanlamanın kullanışlığını arttıracak hem de puanlamanın daha objektif yapılabilmesini sağlayarak ölçme aracının güvenirliğini ve geçerliğini arttıracaktır.

DEĞERLENDİRME SORULARI

1. Aşağıdaki seçeneklerde verilen standart sapma ve güvenirlik değerlerine göre hangisinin standart hatası <u>en yüksektir</u>?

	Standart S.	Güvenirlik
A)	3	0.70
B)	3	0.35
C)	5	0.50
D)	7	0.70
E)	7	0.35

2. Bir öğretmen yapacağı sınavın güvenirliğini arttırmak istiyorsa aşağıdakilerden hangisine <u>öncelikle</u> dikkat etmesi gerekir?

A) Sınav kapsamına giren tüm konulara ilişkin soru olmasına
B) Sınav sorularının cevaplanabileceği yeterli sürenin verilmesine
C) Sınavın puanlaması için bir cevap anahtarı hazırlamaya
D) Soruların açık, anlaşılır ve net olmasına
E) Sorularda geçen ifadelerin cevaplar için ipucu içermemiş olmasına

3. Bir testin iç tutarlılık katsayısı yüksek bulunmuşsa aşağıdakilerden hangisini söylemek <u>doğru olmaz</u>?

A) Test maddeleri arasındaki ilişki yüksektir.
B) Testin geçerliliği yüksektir.
C) Testin güvenirliği yüksektir.
D) Test puanlarındaki hata miktarı azdır.
E) Testin güçlük düzeyi yüksektir.

4. 'Okuduğu paragrafı anlayabilme becerisini' ölçmek için hazırlanan bir İngilizce sınavını inceleyen bir İngilizce öğretmeni bu sınavın anlama becerisinden çok kelime bilgisini ölçen bir sınav olduğu yorumunu yapıyorsa, bu sınavın hangi özelliğinin çok iyi <u>olmadığını</u> söyleyebiliriz?

A) Güvenirlik
B) Geçerlik
C) Doğruluk
D) Kullanışlık
E) Uygunluk

5. Bir sınavın kapsam geçerliği yüksek ise bu sınava ilişkin aşağıdaki yorumlardan hangisi <u>en doğru</u> olacaktır?

A) Sınav sorularının güçlük düzeyleri yüsektir.
B) Sınav sorularının ayırıcılık güçleri düşüktür.
C) Sınav soruları ölçülen konu kapsamını yeterli derecede örneklemektedir.
D) Sınavdaki soruların çoğu doğru cevaplandırılmıştır.
E) Tüm soruların cevaplanmasına yetecek sınav süresi verilmiştir.

egitim@hotmail.es

6. **Aşağıdakilerden hangisi bir testin güvenirliğini arttıracak sebepler arasında yer almaz?**

A) Soruların ipucu içermeyecek şekilde hazırlanması
B) Soruların kolay olması
C) Soruların net ve açık olması
D) Soru sayısının çok olması
E) Puanlamanın objektif yapılmış olması

7. Coğrafya deris için hazırlanmış ve rasgele hatalardan arındırılmış bir testin, yapılan inceleme sonunda matematik bilgisini de ölçtüğü gözlenmiştir.

Bu test aşağıdaki hangi iki özelliğe sahiptir?

I. Geçerli II. Güvenilir III. Kullanışlı
IV. Geçerli değil V.Güvenilir değil VI. Kullanışlı değil
A) I, V B) I, VI C) II, IV D) II, VI E) III, V

8. **Aşağıda soru türü ile soru sayısı verilen sınavlardan hangisinin tesadüfi hatasının en az olması beklenir?**

A) Dörder seçenekli 50 soru
B) Doğru-yanlış türü 50 soru
C) Üçer seçenekli 35 soru
D) Kısa cevaplı 25 soru
E) Boşluk tamamlamalı 25 soru

9. **Aşağıdakilerden hangisi bir sınavın geçerliğini düşüren sebeplerden biri değildir?**

A) Soru sayısının fazla olması
B) Sınav süresinin yetersiz olması
C) Soruların sınıfta çözülen örneklerin aynısının olması
D) Sınavda kopya çekilmiş olması
E) Şans başarısının olması

10. Bir şirketin eleman alımı için yapmış olduğu sınava giren bireylerin sınav sonrası aralarında geçen konuşmada, adaylardan biri: "Bu işe eleman seçebilmek için geçerli bir sınav yapıldığını düşünüyorum" demiştir.

Buna göre bu adayın yorumu yapılan sınavın hangi özelliğine vurguda bulunmaktadır?

A) Uygun bir ortamda yapıldığına
B) Soruların açık ve net olduğuna
C) Objektif olduğuna
D) Geçerli bir sınav uygulanmış olduğuna
E) Adayların yapabilecekleri güçlükte sorular olduğuna

KPSS'DE ÇIKMIŞ SORULAR

1. **Aşağıdakilerden hangisi, testin güvenirliğini düşürücü bir etken değildir? (2001)**

A) Cevaplama süresinin çok kısa olması
B) Sorularda yoruma açık ifadeler kullanılması
C) Puanların bir puanlayıcıdan diğerine değişebilmesi
D) Puanlama şeklinin öğrenciye önceden bildirilmesi
E) Soruların çok zor olması

2. Bir okulda aynı dersi veren iki öğretmenin, aynı konulardan yaptıkları sınavlardaki soru sayıları ve yanıtları puanlama biçimleri şöyledir:

1. öğretmen: 10 soru sormuş ve her soruya 2,5 puan vermiştir.

2. öğretmen: 25 soru sormuş ve her soruya 1 puan vermiştir.

"2. öğretmenin elde ettiği puanların güvenirliği daha yüksektir." **diyen bir kişi, bu iddiasını aşağıdakilerden hangisiyle destekleyebilir? (2001)**

A) Puanlamanın nesnelliği, ölçme sonuçlarına karışabilecek hatayı azaltır.
B) Bir sınavdaki yanıtların nasıl puanlanacağı kullanılan soru tipine göre değişebilir.
C) Bir ölçme aracının güvenirliği, araçta yer alan soruların güçlüğünden etkilenir.
D) Nesnellikten uzak puanlama yöntemlerinin kullanılması puanların güvenirliğini düşürür.
E) Testteki soru sayısı arttıkça test puanlarının güvenirliği artar.

3. İlköğretim 7. sınıf fen bilgisi öğretmeni, öğretim yılının başında öğrencilerinin bu derse hazır oluş düzeylerini belirlemek istemektedir.

Öğretmenin bu amaçla hazırlayacağı ölçme aracı, aşağıdakilerden hangisini kapsamalıdır? (2001)

A) 7. sınıf fen bilgisi dersinde öğrenilecek önemli davranışları
B) Uygulama düzeyinde veya daha yüksek düzeylerdeki hedeflere ilişkin davranışları
C) Derste, laboratuar deneyleriyle öğretilecek davranışları
D) Dersin güç olduğu düşünülen davranışlarını
E) 7. sınıf fen bilgisi dersinde gerçekleşecek öğrenmelerin önkoşullarını

egitim@hotmail.es

4. Bir matematik öğretmeni derste problem çözme becerilerine yönelik etkinliklere ağırlık verdiği halde, uyguladığı sınav daha çok, dört işlem yapabilme becerisini ölçen sorulardan oluşturmuştur.

Öğretmen bu yaklaşımıyla, aşağıdakilerden hangisini tehlikeye atmıştır? (2001)

A) Test-tekrar test güvenirliğini
B) Sınavın kullanışlılığını
C) Testin kapsam geçerliğini
D) Soruların ayırt etme gücünü
E) Puanlayıcı güvenirliğini

5. **Aşağıdakilerden hangisi, sınavın geçerliğini düşürücü etkenlerden biri değildir? (2001)**

A) Sınav sorularını yanıtlamada öğrencilerin birbirinden yararlanması
B) Soruların öğrencilere, sınavdan bir süre önce verilmesi
C) Öğrencilere soruların tümünü değil, seçtikleri bir bölümünü yanıtlama olanağı tanınması
D) Sınavda, her hedef davranışla ilgili soru sorulması
E) Yanıtların puanlanmasında yanlı davranılması

6. Bir öğretmen, yalnızca "okuduğunu anlama" becerisini ölçmede kullanılacak bir yazılı sınav hazırlamak istemektedir.

Aşağıdakilerden hangisi bu sınavdan elde edilecek puanların güvenirliğini ve geçerliğini artırmaz? (2002)

A) Objektif tipte olmayan sorular için ayrıntılı puanlama anahtarının hazırlanması
B) Bir öğrencinin bütün cevaplarının okunması ve sonunda öğrenci hakkındaki izlenimler de dikkate alınarak tümüne birden bir puan verilmesi
C) Sınav kâğıtlarının birden fazla puanlayıcı tarafından puanlanarak puanların ortalamasının alınması
D) Cevapların aynı puanlayıcı tarafından, sıraları değiştirilerek ikinci kez puanlanması ve puanlar arasında fark varsa bir daha puanlama yapılması
E) Çabuk cevaplanabilen türden, çok sayıda soru sorulması

7. "Güvenirliği yüksek bir test sonucunun geçerliğinin düşük olması" **riski, aşağıdaki durumların hangisinde en fazladır? (2002)**

A) Test puanlarının iç tutarlılık katsayısı düşük görünmekte; ancak test kapsamı yoklanmak istenen kritik davranışları içermektedir.
B) Test maddelerinin ayırt ediciliği düşük, test kapsamı dar tutulmuş ve uygulama denetim altında tutulamamıştır.
C) Test puanlarının standart hatası büyüktür; ancak testte çok sayıda ve çeşitli konulardan sorular vardır.
D) Test puanlarının kararlılığı yüksektir; ancak test kapsamı ölçülecek davranışları örneklemede yetersiz kalmaktadır.
E) Test puanları beklenene yakın biçimde dağılmakta, test kapsamı ölçülmek istenen davranışları örnekler görünmektedir.

8. "İlköğretim sekizinci sınıfta uygulanan bir başarı testinin sonuçları, öğrencilerin sonraki öğrenimlerinde farklı programlara yönlendirilmeleri amacıyla kullanılabilir mi?"

Yukarıdaki soruyu soran bir kişi yapacağı geçerlik çalışması için öncelikle testin hangi özelliğini incelemelidir? (2002)

A) İçerdiği sorularla ilgili puanların birbirleriyle korelasyonunu
B) Aynı sınıf düzeyindeki ders notlarıyla korelasyonunu
C) Testin yazım, çizim ve düzenleme kurallarına uygunluğunu
D) Ortaöğretim düzeyindeki başarıyı yordama gücünü
E) İlköğretim konularını kapsama derecesini

9. **Aşağıdakilerden hangisi, birden fazla ölçme sonucuna dayalı bir değerlendirmede verilecek kararın doğruluğunu artırıcı önlemlerden biri değildir? (2002)**

A) Niteliksel gözlem sonuçlarına daha büyük bir ağırlık verilmesi
B) Ağırlıklandırmada, araçların güvenirliğinin dikkate alınması
C) Ağırlıklandırmada, araçların geçerliğinin dikkate alınması
D) Karar vermede kullanılacak ölçütün eldeki ölçme sonuçları cinsinden ifade edilmesi
E) Amaca uygun nitelikli bir ölçütten yararlanılması

10. **Ölçme sonucunun hangi özelliği, ağırlıklı ortalama hesaplanırken bu ölçme sonucuna daha büyük ağırlık verilmesi için bir gerekçe olamaz? (2002)**

A) Kullanışlığı yüksek bir testten elde edilmiş olması
B) Standart hatasının küçük olması
C) Kapsamının daha geniş olması
D) Daha önemli hedeflerle ilgili olması
E) Geçerliğinin yüksek olması

11. **Aşağıdakilerden hangisi bir başarı testinden elde edilen puanların güvenirliğinin bir tanımı olabilir? (2003)**

A) Soruların konulara uygun şekilde dağıtılmış olması
B) Test puanlarının tesadüfi hatalardan arınık olma derecesi
C) Soruları, sadece yoklanan davranışı öğrenmiş olanların cevaplayabilmesi
D) Test puanlarıyla, öğrencinin gelecekteki başarısının yordanabilmesi
E) Soruların, testin amacına hizmet etme derecesi

12. **Bir başarı testinin** "kapsam geçerliğinin yüksek olduğu" **ifadesinden aşağıdaki sonuçların hangisi çıkarılabilir? (2003)**

A) Tutarlı ve kararlı sonuçlar veren bir testten yararlanıldığı
B) Testteki soruların pek çoğunu öğrencilerin doğru olarak yanıtlayabildikleri
C) Testteki soruların konu alanını iyi bir şekilde örneklediği
D) Testteki soruların güçlük düzeylerinin, gerçekleşmesi beklenen öğrenme düzeylerine uygun olduğu
E) Öğrencilerin, testteki soruları yanıtlamaya çalıştıkları

egitim@hotmail.es

13. **Aşağıdakilerin hangisinde, yordama geçerliğinin belirlenmesine çalışılmaktadır? (2003)**

A) Hazırlanan soruların, belirli bir alanda öğrenilenleri temsil edip etmediğinin incelenmesi

B) Ölçme aracının öğretimde bir üst düzeyde başarılı olacak öğrencileri seçme özelliğinin incelenmesi

C) Bir grup öğrencinin iki dersteki başarılarının karşılaştırılması

D) Bir ölçme aracının, birkaç hafta arayla aynı gruba tekrar uygulanması ve elde edilen puanlar arasındaki ilişkinin incelenmesi

E) Hazırlanan bir testin farklı düzeylerdeki öğrenci gruplarına uygulanması ve bu grupların başarılarının karşılaştırılması

14. VE 15. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Bir öğretmen, İngilizce kelime bilgisini yoklayan bir test hazırlamış, bunu birkaç hafta arayla öğrencilere iki kez uygulamış ve bu iki uygulamadan elde edilen puanlar arasındaki korelasyonu 0,90 olarak hesaplamıştır.

14. **Hesaplanan bu değer testin en çok hangi özelliği hakkında bilgi verir? (2004)**

A) KR-20 güvenirliği

B) Test-tekrar test güvenirliği

C) Kullanışlılığı

D) Yapı geçerliği

E) Kapsam geçerliği

15. **Bu korelasyon değerine dayanarak aşağıdaki yorumlardan hangisi yapılabilir? (2004)**

A) Sınıfın yüzde 10'u başarısızdır.

B) Sınıfın yüzde 10'u ortalamanın altında puan almıştır.

C) Test puanlarında yüzde 10 hata vardır.

D) Test puanlarının kararlılığı yüksektir.

E) Test, öğrenme düzeyi farklarının yüzde 90'ını açıklamaktadır.

16. **Aşağıdakilerden hangisi bir ölçme aracının yordama geçerliğini belirlemeye yönelik bir uygulamadır? (2004)**

A) Bir ölçme aracından elde edilen aritmetik ortalama ile aynı amaçla hazırlanmış farklı bir ölçme aracından elde edilen aritmetik ortalama arasında anlamlı fark olup olmadığını belirleme

B) Bir grubun belirli bir testten elde ettiği puanlarla, geçerliği daha önceden belirlenmiş başka bir testten elde ettiği puanlar arasındaki korelasyonu hesaplama

C) Öğrencilerin belirli bir öğretim programı sonundaki başarıları ile bu öğretim programına seçme amacıyla kullanılan ölçme aracından elde ettikleri puanlar arasındaki korelasyonu hesaplama

D) Aynı davranışları ölçtüğü düşünülen iki ölçme aracını aynı gruba aynı koşullarda uygulayıp bu uygulama sonunda elde edilen puanlar arasındaki korelasyonu hesaplama

E) Bir testi, ölçülmek istenen özelliğe sahip oluş dereceleri farklı iki gruba uygulama ve bu uygulama sonucunda elde edilen puanların gruplar arasındaki farkı yansıtıp yansıtmadığını belirleme

17. **Aşağıda standart sapması ve güvenirlik katsayısı verilen testlerden hangisiyle elde edilen puanların standart hatası en küçüktür? (2004)**

	Standart Sapma	Güvenirlik Katsayısı
A)	2	0,50
B)	4	0,50
C)	5	0,30
D)	5	0,60
E)	2	0,90

18. Güvenirlik, ölçme sonuçlarının hatadan arınıklık derecesidir.

Bu tanıma göre, bir ölçme aracı olarak testin güvenirliği aşağıdakilerden hangisinden etkilenmez? (2005)

A) Soruların konulara dengeli dağıtılması

B) Soru sayısının artırılması

C) Sorularda grubun düzeyine uygun bir dil kullanılması

D) Testin standart koşullarda uygulanması

E) Soruların açık ve anlaşılır bir biçimde yazılması

19. **Yabancı dil muafiyet sınavında kullanılan test için aşağıdakilerden hangisi bir geçerlik kanıtı olabilir? (2005)**

A) Testin hatasız puanlanabilmesi

B) Testin çoktan seçmeli maddelerden oluşması

C) Testten alınan puanların ranjının geniş olması

D) Yabancı dili iyi bilenlerin testten yüksek puan alması

E) Testte farklı yapılarda soruların bulunması

egitim@hotmail.es

20. **Bir firmaya eleman almak amacıyla yapılan bir sınav için, "Oldukça geçerli bir sınav yapıldı." diyen bir uzman sınavın hangi özelliğini ifade etmektedir? (2006)**

A) Zor olduğunu
B) Adayların yeteneklerine uygun olduğunu
C) Objektif olduğunu
D) Kolay olduğunu
E) İşteki başarıyı yansıtacak yapıda olduğunu

21. **Hazırladığı testin güvenirliğini artırmak isteyen bir öğretmen öncelikle aşağıdakilerden hangisini sağlamaya çalışmalıdır? (2006)**

A) Testteki soruların ve seçeneklerin dil bilgisi kurallarına uygun olmasını
B) Soruların açık, net ve anlaşılır olmasını
C) Seçenekler ile soru kökünün ifade bakımından birbirleriyle tutarlı olmasını
D) Test kapsamında ölçülmek istenilen her özelliğe ait soru bulunmasını
E) Seçeneklerin birbirine eşit ya da yakın uzunlukta olmasını

22. Pazarlama elemanları seçiminde kullanılmak üzere bir test geliştirilmiş ve işe alımlarda bu test uygulanmıştır. Daha sonra, işe alınan elemanların ilk aydaki başarılı satış sayıları ile işe girişte aldıkları test puanları arasındaki uyum incelenmiştir.

Yukarıda belirtilen inceleme, uygulanan testin hangi özelliği hakkında bilgi verir? (2007)

A) Yordama geçerliği
B) Yapı geçerliği
C) Kapsam geçerliği
D) İç tutarlığı
E) Puanlayıcılar arası güvenirliği

23. - 25. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Bir dersteki başarıyı ölçmek için geliştirilmiş beş testle ilgili olarak tablodaki bilgiler elde edilmiştir.

Test	Güvenirlik Katsayısı	Yordama Geçerliği	Kapsam Geçerliği	Ortalama Güçlük
1	0,72	0,45	Orta	0,70
2	0,65	0,29	Yüksek	0,35
3	0,83	0,64	Yüksek	0,58
4	0,92	0,18	Düşük	0,40
5	0,61	–0,55	Düşük	0,50

23. **Sınıfındaki öğrenciler arasında bu dersteki başarısı yüksek olanları seçmek isteyen bir öğretmenin hangi testi kullanması en uygun olur? (2007)**

A) 1 B) 2 C) 3 D) 4 E) 5

24. "Bir testin güvenirliğinin yüksek olması, geçerli bir test elde etmek için yeterli değildir." yargısını desteklemek için hangi test örnek gösterilebilir? (2007)

A) 1 B) 2 C) 3 D) 4 E) 5

25. **Diğerlerinden daha kolay olduğu hâlde kabul edilebilir düzeyde güvenilir ve geçerli olan test hangisidir? (2007)**

A) 1 B) 2 C) 3 D) 4 E) 5

26. Lisedeki bir rehber öğretmen öğrencilerin mesleki ilgilerini belirlemek için mesleki ilgi envanteri geliştirmiştir. Güvenirliğini belirlemek için bu envanteri aynı öğrencilere altı hafta arayla iki kez uygulamış ve iki uygulamadan elde edilen ölçüler arasındaki ilişkiyi (korelasyon) 0,45 bulmuştur.

Aşağıdakilerden hangisi, bu korelasyon değerinin düşüklüğü için gerekçe olarak gösterilemez? (2007)

A) Ölçülen özellik kolay değişme gösteren bir yapıdadır.
B) İki uygulamanın koşulları birbirinden oldukça farklıdır.
C) Öğrenciler maddelere yanıt verirken gerçekçi davranmamışlardır.
D) Sorular her okumada farklı anlamlara gelebilmektedir.
E) İki uygulama arasında geçen süre çok kısadır.

27. **Bir öğretmen aynı kapsamda ve düzeyde iki ayrı sınav hazırlamış ve aynı gruba uygulamıştır. Bu sınavların özellikleri şöyledir:**

- Birinci sınavda 25 soru vardır ve her soru 4 puan değerindedir.
- İkinci sınavda 10 soru vardır ve her soru 10 puan değerindedir.

"Birinci sınavdan elde edilen puanların güvenirliği daha yüksektir." **diyen bir kişi bu iddiasını aşağıdaki açıklamalardan hangisiyle destekleyebilir? (2008)**

A) Bir sınavda yanıtların nasıl puanlanacağı kullanılan soru tipine göre değişir.
B) Puanlamanın nesnelliği ölçmeden kaynaklanabilecek hatayı azaltır.
C) Soruların ayırt edicilik dereceleri test puanlarının güvenirliğini etkiler.
D) Soruların güçlük dereceleri test puanlarının güvenirliğini etkiler.
E) Testteki soru sayısı arttıkça ölçme sonuçlarına karışan hata miktarı azalır.

28. Bir araştırmacı çoklu zekâ alanlarını ölçmek amacıyla geliştirdiği çok boyutlu testi, dil bilimi, matematik, resim, heykel, müzik, beden eğitimi, iletişim, psikoloji bölümlerindeki öğrencilere uygulamıştır. Uygulama sonucunda farklı bölümlerdeki öğrencilerin kendi alanıyla ilgili sorulardan yüksek, diğer alanlardaki sorulardan ise daha düşük puan aldıklarını saptamıştır.

Buna göre, ölçme aracı için aşağıdakilerden hangisi söylenebilir? (2009)

A) Amaca hizmet etme derecesi düşüktür.
B) Hata düzeyi düşüktür.
C) Geçerliği yüksektir.
D) Kullanışlılığı yüksektir.
E) Güvenirliği düşüktür.

egitim@hotmail.es

29. Bir okulda, fen bilgisi proje ödevi verilmiştir. Ürünlerin puanlanmasının sınıfta derse giren öğretmen yerine, diğer sınıfın öğretmeni tarafından yapılması sağlanmıştır. Projelere kod numaraları verilmiş ve öğrenci isimleri, projeleri değerlendiren öğretmenden gizlenmiştir. Ayrıca değerlendirme için (kapsamlı, detaylı ve açık) bir puanlama rehberi oluşturulmuş, öğretmenin bu rehbere göre puan vermesi istenmiştir.

Yukarıda anlatılan işlemler, kullanılan ölçme aracının hangi özelliğini iyileştirmek amacıyla gerçekleştirilmiştir? (2009)

A) Objektiflik
B) İç tutarlık
C) Kullanışlılık
D) Görünüş geçerliği
E) Yapı geçerliği

30.

- Portfolyolardan elde edilen puanlarla aynı konulara ilişkin hazırlanmış objektif testlerden elde edilen puanlar arasındaki ilişkinin incelenmesi
- Portfolyolardan elde edilen puanlarla bu portfolyonun kapsamına ilişkin daha sonradan ortaya çıkacak performans ölçüleri arasındaki ilişkinin incelenmesi
- Farklı öğretmenlerin aynı portfolyolara verdiği puanlar arasındaki ilişkinin incelenmesi

Portfolyo değerlendirmesine ilişkin yukarıda verilen yöntemler aşağıdaki amaçlardan daha çok hangisini gerçekleştirmek için kullanılmaktadır? (2009)

A) Portfolyo değerlendirmesinde kullanılacak puanlama yönergelerini hazırlamak
B) Portfolyolardan elde edilen verilerin güvenirliğini ve geçerliğini sağlamak
C) Öğrencilerin işlenen konulara ilişkin öğrenme düzeylerini saptamak
D) Öğrencilerin sergileyebileceği en iyi çalışmaları belirlemek
E) Öğrencilerin önceki eğitim hayatında gösterdikleri performansı ortaya çıkarmak

31. Genel yetenek testinde matematiksel ilişkilerden yararlanma gücünü ölçmeyi amaçlayan bir araştırmacı, ilk olarak 80 soruluk bir sınav hazırlamış ve ön deneme uygulaması yapmıştır. Ön deneme sonunda iç tutarlık katsayısını 0,93 olarak hesaplamıştır. Sınavın cevaplanmasının çok uzun zaman aldığını gören araştırmacı, en düşük ayırt ediciliğe sahip 30 maddeyi sınavdan çıkarmıştır. Kalan 50 maddelik form için 0,85 olarak hesapladığı iç tutarlık katsayısını yeterli gören araştırmacı, testini bu şekliyle kullanmaya karar vermiştir.

Bu araştırmacı ön deneme uygulaması ve devamındaki işlemlerde sınavının hangi özelliğini artırmıştır? (2010)

A) Güvenirliğini
B) Kapsam geçerliğini
C) Kullanışlılığını
D) Objektifliğini
E) Yapı geçerliğini

32. **Anekdot kayıtlarının güvenirliğini ve geçerliğini etkileyen aşağıdaki faktörlerden hangisi anekdotu yazan kimsenin kişisel yanlılığıyla ilgili değildir? (2010)**

A) Olayın karışık ve anlaşılması zor bir ifadeyle yazılmış olması
B) Olayın yazımı sırasında gerçeğin bazı yönlerinin atlanması veya değiştirilmesi
C) Olay kahramanından söz ederken övücü ya da yerici sözcükler kullanılması
D) Olay sırasında önem verilen davranışların daha dikkatle gözlenip ayrıntılı olarak yazılması
E) Sadece olumsuz davranışların gözlenip kaydedilmesi

33. Osman Öğretmen'in elinde her bir konu için güçlük ve ayırt edicilikleri belirlenmiş maddeler yer almaktadır. Osman Öğretmen uygulayacağı sınavda dört konudan beşer madde kullanmayı planlamaktadır ancak ikinci konuyla ilgili elinde yeterli ayırıcılıkta madde olmadığını görür. Bunun üzerine ikinci konuya ilişkin sorular yerine, üçüncü konu kazanımlarını ölçmek üzere yazılmış ve yüksek ayırt edicilik sağlayan maddeleri kullanır.

Osman Öğretmen'in bu kararı sınavın güvenirliğini ve geçerliğini nasıl etkilemiştir? (2010)

A) Güvenirliği etkilememiş, geçerliği artırmıştır.
B) Güvenirliği artırmış, geçerliği etkilememiştir.
C) Güvenirliği artırmış, geçerliği azaltmıştır.
D) Güvenirliği ve geçerliği azaltmıştır.
E) Güvenirliği azaltmış, geçerliği etkilememiştir.

34. Bir öğretmen yıl sonunda, sözlü sınav puanlarını da dikkate alarak öğrencileri değerlendirmektedir. Öğrencilerin sözlü sınav puanlarını ise her ders sonunda sorduğu sorulara verilen cevaplara birer puan vererek elde etmektedir.

Bu öğretmenin, öğrencilere her ders sonundaki değerlendirmelerin toplamıyla sözlü sınav notu vermesinin en önemli nedeni aşağıdakilerden hangisidir? (2010)

A) Öğrencileri ders çalışmaya teşvik etmek
B) Öğrenciler arasında rekabet yaratmak
C) Sınavın kapsam geçerliğini artırmaya çalışmak
D) Sınava ilişkin kaygı düzeyini azaltmak
E) Bireyler arası farkları ortaya koymak

egitim@hotmail.es

35. Bir öğretmen sınavında yer alan beş maddenin niteliğini incelemek amacıyla öğrencilerin her bir maddeden aldıkları puanlar ile testin tümünden aldıkları puanlar arasındaki korelasyon katsayısını hesaplamıştır. Bilenle bilmeyeni en iyi ayırt eden maddeyi bu şekilde görebileceğini iddia eden öğretmen aşağıdaki katsayıları elde etmiştir. Maddeler Korelasyon katsayısı

Maddeler	Korelasyon katsayısı
1	0,2
2	-0,1
3	0,5
4	0,6
5	-0,7

Bu öğretmenin en iyi ayırt edici olduğunu iddia edeceği madde hangisidir? (2010)

A) 1 B) 2 C) 3 D) 4 E) 5

36. Tüm öğrencileri bilgisayar okuryazarı olan ve yeterli düzeyde bilgisayar laboratuvarı bulunan bir okulun uzmanı aşağıdaki açıklamayı yapmaktadır: "Bazı sınavlar bilgisayar ortamında uygulanmaktadır. Bu uygulamalarda, öğrencilerin her cevabından sonra yetenek kestirimleri ve bu kestirimlerin hassasiyetleri yeniden hesaplanır. Öğrenci, testin maddelerini yanıtladıkça, bir sonraki madde, öğrencinin yetenek düzeyine yakın olarak seçilir. Böylece öğrenciler kendi yetenek düzeylerinin çok altında veya çok üstünde kalan maddeleri cevaplamak zorunda kalmazlar. Yeterli düzeyde hassasiyetle bir yetenek kestirimine ulaşıldığında test sona erdirilir. Bu algoritma sayesinde, bilgisayar tabanlı testler, geleneksel kâğıt-kalem testleri ile aynı ölçme hassasiyetini daha az sayıda madde ile sağlar."

Bu uzman, bilgisayar tabanlı testleri hangi özelliği bakımından geleneksel kâğıt-kalem testlerinden üstün bulmaktadır? (2010-2)

A) Kapsam geçerliği
B) Kullanışlılık
C) Objektiflik
D) Ölçüm yansızlığı
E) Yordama geçerliği

37. Kamyonların üretildiği bir fabrikada çalışan Semih, kamyonların tartıldığı kantarla kendi kütlesini ölçmek ister. Kantarın üzerine çıkar ve göstergenin sıfır civarında olduğunu görür. Kantarın bozulduğunu düşünür. Bakımını yaptırır. Tekrar kantarın üzerine çıkar ve göstergenin yine sıfıra yakın olduğunu görür.

Bu durum, ölçme araçlarının hangi özelliği açıklanırken kullanılabilir? (2010-2)

A) Yeterliği
B) Büyüklüğü
C) Duyarlığı
D) Yansızlığı
E) Kullanışlılığı

38. Serap Öğretmen, hazırladığı ve ilk defa uyguladığı bir sınavda geçme puanını 60 olarak belirlemiştir. Ancak sınav sonuçlarını değerlendirirken 59 puan alanları da başarılı saymıştır.

Serap Öğretmen'in 59 puan alan öğrencileri de başarılı saymasındaki temel gerekçe aşağıdakilerden hangisi olursa ölçme ve değerlendirme açısından geçerli bir uygulama yapmış olur? (2011)

A) Başarıyı teşvik etmek istemiştir.
B) Derse karşı motivasyonu artırmaya önem vermiştir.
C) Ölçmenin standart hatasını göz önüne almıştır.
D) Kendi başarısını artırmak istemiştir.
E) Derse katılımı ödüllendirmek istemiştir.

39. Başarı testi geliştiren bir uzman, pilot uygulama için on öğrenciyi tek tek sınava almıştır. Bu uygulamada öğrencilerden maddeleri sesli olarak çözmelerini isteyen uzman, öğrencilerin maddeleri çözerken kullandıkları bilişsel süreçleri açığa çıkarmayı amaçlamıştır.

Uzmanın bu çalışması, geliştirilen testin en çok hangi özelliği hakkında bilgi sağlar? (2011)

A) İç tutarlılığı B) Objektifliği C) Kullanışlılığı
D) Yapı geçerliği E) Yordama geçerliği

40. **Bir öğretmen, bir testin iki paralel formunu 2 hafta arayla öğrencilerine uygulamıştır. Formların uygulanması veya puanlanması esnasında karşılaşılabilecek olası durumlar aşağıda verilmiştir.**

I. Öğrencilerin ikinci uygulamada yorgun olmaları
II. Öğretmenin her iki uygulamada da toplam puanları hesaplarken işlem hataları yapması
III. Toplam puanları düşük bulan öğretmenin, her iki uygulamada da puanlar ortalaması 65, standart sapması 10 olacak şekilde artırması

Bu durumlardan hangileri, paralel formlar yöntemi ile hesaplanan güvenirlik katsayısını düşürür? (2012)

A) Yalnız I B) Yalnız II C) I ve II
D) I ve III E) II ve III

egitim@hotmail.es

41. Aşağıda belirtilen değişkenler arasındaki ilişkilerden hangisi, Kamu Personel Seçme Sınavı Eğitim Bilimleri Testinin yordama geçerliği hakkında bilgi verir? (2012)

	1. Değişken	2. Değişken
A)	Testten alınan toplam puanlar	Lisans mezuniyet not ortalamaları
B)	Testten alınan toplam puanlar	Yeni atanan öğretmenlerin ilk yıl sonundaki yeterlik puanları
C)	2010 yılında uygulanan testin madde ayırt edicilik indeksleri	2011 yılında uygulanan testin madde ayırt edicilik indeksleri
D)	Testin 2010 yılı uygulamasından alınan toplam puanlar	Testin 2011 yılı uygulamasından alınan toplam puanlar
E)	2010 yılında uygulanan testin madde güçlük indeksleri	2011 yılında uygulanan testin madde güçlük indeksleri

42. Bir şiddet eğilimi ölçeği geliştirme çalışmasında şu işlemler yapılmıştır:

- Maddelerin, saldırganlık durumlarını uygun biçimde örneklediği, uzman kanılarına başvurularak irdelenmiştir.
- Öğrencilerin ölçekte yer alan maddeleri cevaplamak için kaç dakikaya ihtiyaçlarının olacağı, deneme uygulamalarında gözlenmiş ve ölçekteki madde sayısı bu doğrultuda azaltılmıştır.
- Ölçek uygulanmasından elde edilen puanlara bakıldığında erkek öğrencilerin daha yüksek puanlara sahip oldukları belirlenmiştir. Bu durumun, erkek öğrencilerin kız öğrencilere göre şiddete daha eğilimli oldukları bilgisiyle uyum içinde olduğu belirlenmiştir.
- Ölçekten elde edilen puanların merkezi eğilim ve değişkenlik istatistikleriyle Cronbach'ın α katsayısı hesaplanmıştır.

Yapılan bu işlemler, ölçme aracının hangi özelliği hakkında bilgi vermez? (2012)

A) Kullanışlılık
B) Kapsam geçerliği
C) Yapı geçerliği
D) Ölçmenin standart hatası
E) Objektiflik

EĞİTİMDE KULLANILAN ÖLÇME ARAÇ VE YÖNTEMLERİ

5

ANAHTAR KAVRAMLAR

Sözlü sınavlar, yazılı sınavlar, kısa cevaplı testler, doğru yanlış testleri, çoktan seçmeli testler, performans değerlendirme, öğrenci ürün dosyası (portfolyo), dereceli puanlama anahtarı (rubrik)

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Sözlü sınavların nerede ve nasıl kullanılacağını, özelliklerini, sözlü sınavların avantajlarını ve dezavantajlarını açıklayabilecek,
- Yazılı sınavların, nerede ve nasıl kullanılacağını, özelliklerini, yazılı sınavların avantajlarını ve dezavantajlarını açıklayabilecek,
- Kısa cevaplı testlerin nerede ve nasıl kullanılacağını, özelliklerini, kısa cevaplı testlerin avantajlarını ve dezavantajlarını açıklayabilecek,
- Doğru yanlış testlerinin nerede ve nasıl kullanılacağını, özelliklerini, doğru yanlış testlerinin avantajlarını ve dezavantajlarını açıklayabilecek,
- Çoktan seçmeli testlerin nerede ve nasıl kullanılacağını, özelliklerini, doğru yanlış testlerinin avantajlarını ve dezavantajlarını açıklayabilecek,
- Performans değerlendirmenin ne olduğunu söyleyebilecek,
- Performans değerlendirmenin klasik ölçme araç ve yöntemlerinden farklılıklarını açıklayabilecek,
- Öğrenci ürün dosyasının ne olduğunu söyleyebilecek,
- Öğrenci ürün dosyasının klasik ölçme araç ve yöntemlerinden farklılıklarını açıklayabilecek,
- Dereceli puanlama anahtarının (rubrik) ne olduğunu söyleyebilecek
- Dereceli puanlama anahtarının (rubrik) türlerini açıklayabileceksiniz.

İÇİNDEKİLER

Sözlü sınavlar
Yazılı sınavlar
Kısa cevaplı testler
Doğru-yanlış testleri
Çoktan seçmeli testler
Performans değerlendirme
Öğrenci ürün dosyası (portfolyo)
Dereceli puanlama anahtarı (rubrik)
Çalışma soruları

BAŞLAMADAN ÖNCE

Yukarıda "içindekiler" başlığı altında verilen ölçme araçlarından hangisini ilk kez duyuyorsunuz? Eğitim yaşantınızda hangi ölçme araçları size uygulandı? Hangi tür sınav size en çok kaygı yaşatmakta? Neden?

egitim@hotmail.es

EĞİTİMDE KULLANILAN ÖLÇME ARAÇ VE YÖNTEMLERİ

Eğitim sürecinin her aşamasında (girdi, süreç, ürün) ölçme ve değerlendirmenin önemi tartışılmazdır. Değerlendirmenin doğru kararlarla sonuçlanabilmesi için öncelikle ölçmenin olabildiğince hatasız ve objektif yapılması gerekmektedir. Ölçmenin doğru yapılabilmesi için de ölçme araç ve yöntemlerini iyi tanımak, özelliklerini bilmek, avantaj ve dezavantajlarının farkında olmak çok önemlidir. Aşağıda eğitimde kullanılan, sizlerin de oldukça aşina olduğunuz ölçme araçları tanıtılmaya çalışılmıştır.

Unutulmamalıdır ki her bir ölçme aracını diğerlerinden avantajlı ya da dezavantajlı kılan özellikleri bulunmakta, her birinin kullanılmasının daha uygun olduğu farklı durumlar bulunmaktadır. Hiç bir ölçme aracı her açıdan en güvenilir ve en geçerli olarak tanımlanamaz. Bununla birlikte, genel anlamda en güvenilir ve geçerli ölçme aracının çoktan seçmeli testler, en az güvenilir ve geçerli ölçme aracının da sözlü sınavlar olduğu söylenebilir.

SÖZLÜ SINAVLAR

Sözlü Sınav Nedir?

Belki de bilinen en eski sınav türüdür. Sözlü sınavların, milattan önce Çin'de ve Yunanistan'da kullanıldığı bilinmektedir. Soruların sözlü ya da yazılı olarak verildiği öğrencilerin cevapları kendisinin düşünerek sözlü olarak verdiği sınav türüdür. Öğrencilerin, derse katılımının sağlanmasında, aktif kılınmasında ve iletişim becerisi kazandırılmasında önemi büyüktür.

Sözlü Sınavlar Ne Zaman Tercih Edilebilir?

- Özellikle öğrencilerin sözel iletişim becerilerinin, kelimeleri telaffuz yeteneklerinin ya da okuma becerilerinin ölçülmesinin gerektiği durumlarda sözlü sınavlar kaçınılmazdır.
- Derse başlarken ve ders aralarında öğrencilerin dikkatini çekebilmek için sorular sorarak katılımlarını sağlayarak öğrencilerin aktif olmasına yardımcı olmak amacıyla sözlü sınavlar tercih edilebilir.
- Öğrencilerin derse hazırbulunuşluğunu sınamak, ders aralarında anlaşılmayan ya da yanlış anlaşılan noktaları tespit etmek, konu ya da ünite sonlarında genel bir tekrarın yapılması için sözlü sınavların kullanılması aynı zamanda öğrencileri de ders çalışmaya teşvik eden, öğrendiklerini pekiştirmeye yardım eden bir araç olmaktadır.

Sözlü Sınavlarda Nelere Dikkat Edilmelidir?

- Sözlü sınavların öğrencilerin tahtaya kaldırılarak yapılması pek tercih edilmemelidir. Öğrencilerin sınav kaygısını en fazla yaşadıkları sınav türü sözlü sınavlardır. Özellikle çekingen, içine kapanık öğrencilerin tahtaya kaldırılarak soruları cevaplandırmalarının istenmesi, sınav kaygısının çok yoğun şekilde yaşanmasına sebep olmakta ve bu özellikteki öğrencilerin cevabı çok iyi bilseler bile cevap verememelerine sebep olmaktadır.
- Yukarıdaki sebeplerden dolayı, sözlü sınavların sorulan soruya öğrencilerin kendi istekleriyle ve yerlerinde cevap vermelerini sağlayacak şekilde yapılması önerilebilir. Puanlamanın daha çok öğrencinin cevap vermeye motivasyonunu sağlamak amacıyla yapılmalı, öğrencinin puan kaygısını arttırmayacak şekilde uygulanmalıdır. Böylece öğrencilerin eleştirel düşünme, yaratıcı düşünme ve iletişim becerisi kazanmasına, öğrencinin toplum içinde konuşabilme öz güveni oluşturmasına yardımcı olur.
- Sözlü sınavlarda öğrencilere verilen puanların ders notuna etkisi olabildiğince azaltılmalıdır. Sözlü sınav puanıyla öğrenciler hakkında önemli kararlar verilmemelidir. Sözlü sınavlar, genel anlamda güvenilirliği ve geçerliği en düşük ölçme aracı olarak bilinir.
- Sınavdan önce sınav planı yapılmalıdır. Soru sayısı tüm konuları kapsamalıdır. (Belirtke tablosuna göre soru hazırlanmalı). Puanlamanın olabildiğince objektif yapılabilmesini sağlamak için cevap anahtarı hazırlanarak puanlama yapılmalıdır. Sınavda uyulması gereken kurallar varsa mutlaka önceden bildirilmelidir.
- Dili uygun, etkili kullanma becerisi ölçülmeyecekse kısa ve çok sayıda soru sorulmalıdır.

Sözlü Sınavların Avantajları Nelerdir?

- İletişim becerisini, anadilini veya yabancı dili etkili olarak kullanma gibi becerilerin ölçülmesi gerektiğinde tercih edilecek en geçerli ve güvenilir ölçme yöntemidir. Ayrıca bir iş ya da görev için en uygun özelliklere sahip bireylerin seçilmesi gerekli durumlarda, bireyleri derinlemesine tanımayı sağlayabilecek en elverişli ölçme aracıdır.
- Diğer sınav türlerinde puanlayıcının, puan vermede kararsız kaldığı, yarım bırakılmış, açıklanmamış, gerekçe gösterilmemiş cevaplar için yeni sorular sorma şansı yoktur. Sözlü sınavlarda ise sorular sorularak, derinlemesine yanıtlar alınabilir.

egitim@hotmail.es

- Soruların hazırlanması kolaydır ve az zaman alır.
- Şans başarısı ve kopya çekme olasılığı olmadığı için güvenirlik artar.
- Öğrencinin soruyu yanlış anlayarak farklı cevap vermesi, soruyu istediği yöne çekme olasılığı olmadığı için güvenirlik ve geçerlik artar.

Sözlü Sınavların Dezavantajları Nelerdir?

- Sınavı yapan ile sınav olan kişi arasındaki etkileşim en fazladır. Ölçülmek istenen değişkenlerden başka değişkenlerin (öğrencinin duruşu, ses tonu, mimikleri, ifade gücü, iletişim becerisi, hatta kılık-kıyafeti vb.) öğrencilerin alması gerekenden farklı puanlar almasına sebep olabilir.
- Uygulanması açısından kullanışlı değildir. Çok zaman alır. Her öğrenciye aynı sorunun sorulması adaletsizliğe (ilk öğrenciye soru sorulduktan sonra artık diğer öğrencileri soruyu bilmektedir), farklı soruların sorulması da öğrencilerin doğru bir şekilde karşılaştırılmasına engel olduğu için her iki durumda da güvenirlik ve geçerlik düşmektedir.
- Soru sorulduktan hemen sonra cevap vermek durumunda oldukları için öğrencilerin düşünme ve yanıtı gözden geçirme şansı çoğunlukla yoktur.
- Sözlü sınavların uygulanması çok zaman aldığından az soru sorulduğu için kapsam geçerliği ve güvenirlik düşer.
- Yanıtların doğruluk derecesi puanlayıcı tarafından belirlendiği için yanlılık vardır. Bir başka deyişle puanlaması subjektif (öznel)dir.
- Sözlü sınavlarda cevaplar kaydedilmediği için cevapların ve puanlamanın tekrar gözden geçirilme imkanı yoktur.

YAZILI SINAVLAR

Yazılı Sınav Nedir?

Soruların genellikle yazılı olarak verildiği, cevapları öğrencilerin kendilerinin düşünerek yazılı olarak verdikleri sınav türüdür. Öğrencilerin cevapları kendilerinin organize ederek vermelerinin gerekli olduğu durumlarda, kompozisyon yazma becerilerinin ölçülmesine imkan veren ölçme araçlarıdır.

Yazılı sınavların farklı türleri bulunmaktadır:

1) *Klasik Yazılı Sınavlar*: En çok bilinen yazılı sınav türüdür. Öğrencinin kendisine sorulan her bir soruyu yazılı olarak cevaplaması gerekmektedir. Sorulara verdiği cevapların doğruluğuna göre puan alır.
2) *Tercihli Sınavlar:* Sınavda yer alan soruların tümünü cevaplandırmanın zorunlu olmadığı yazılı sınav türüdür. Belirli sayıdaki sorulardan bir kısmının seçilerek cevaplandırılması istenir.

En büyük avantajı; öğrencilere cevaplamada bir miktar özgürlük sağladığı için sınav kaygısını azaltarak, motivasyonu arttırır bu da güvenirliğin artmasına yardımcı olur.

Ancak tercihli sınavlara ilişkin bazı eleştiriler de bulunmaktadır:

- Yazılı sınavların az soru içermesinden dolayı kapsam geçerliğinin düşük olması özelliği bu tür sınavlarda daha da düşmektedir
- Her öğrencinin sınavda ölçülmek istenilen tüm davranışları kazanıp-kazanmadıkları ölçülemez.
- Her öğrencinin farklı soruları cevaplandırmasının mümkün olması elde edilen puanların karşılaştırılmasına olanak tanımaz, puanlar karşılaştırılamadığı için de güvenirlik düşer.

> ***DİKKAT !!!*** *Tercihli sınavların avantaj ve dezavantajları birlikte değerlendirildiğinde, hem zorunlu hem de seçmeli (tercihli) maddeler (sorular) dan oluşan bir yazılı sınavların hazırlanması güvenirlik ve geçerliğin arttması açısından tavsiye edilebilir.*
>
> *Zorunlu sorularla öğrencilerin puanlarının karşılaştırılması mümkün kılınırken, öğrencinin istediği soruya cevap verme özgürlüğünün de bulunması sınav kaygısını azaltacaktır.*

3) *Açık Kitap Sınavları: Sınav sırasında defter, ders kitabı, vb. kaynaklara bakılmasının serbest olduğu yazılı sınav türleridir.* İstatistik, matematik, fizik vb… derslerde kullanılması daha yaygındır. Denklem, formül vb… ezberlenmesi yerine kaynak kullanma özgürlüğü verilir. Bu tür sınavlarda yer alan sorular kaynaklardan aynen alınmamalı, çözümleri ya da çözümlerine ilişkin örnekler ya da ipuçları bulunmamalıdır.

En büyük avantajı; sınav kaygısının bir miktar azalmasını sağladığı için motivasyonu böylece güvenirliği arttırır.

Ancak tercihli sınavlara ilişkin bazı eleştiriler de bulunmaktadır:

Sınavın asıl amacı olan sorulara verilen cevapların doğruluğu dışında kaynaklara aşinalığı ölçer hale gelebilir ki bu da yapı geçerliğini düşürür. Okuma hızı ve dikkat gücü fazla olan öğrencilerin daha yüksek puan almasına sebep olabilir ki bu da yine yapı geçerliğini düşürür.

> ***DİKKAT !!!!*** Açık kitap sınavlarını tercih etmek yerine sınavın kapsamında yer alan formül, denklem vb… ezbere bilgilerin sınavda öğrencilere verilmesi daha uygun bir yöntem olarak önerilebilir.

egitim@hotmail.es

Yazılı Sınavlar Ne Zaman Tercih Edilebilir?

Sınav sorusu hazırlamak için fazla zaman yok ancak puanlama yapmak için yeterli süre olduğu durumlarda tercih edilebilir. Özellikle öğrencilerin, analiz, sentez ve değerlendirme gibi üst düzey bilişsel davranışlarının ölçülmesi gereken durumlarda kullanılacak en uygun ölçme aracıdır. Çok fazla sayıda öğrencinin bulunduğu durumlarda yazılı sınavlar tercih edilmez.

Öğrencilerin kompozisyon becerisinin, anlatım ve yazım gücünün ölçülmesinin gerektiği, öğrencilerin cevaplarıyla orjinal bir fikir ortaya koymalarının, yaratıcılıklarını ve kendi düşüncelerini sergilemelerinin istendiği durumlarda kullanılır.

Yazılı Sınavlarda Nelere Dikkat Edilmelidir?

1. Sınavda yer alan sorular mutlaka açık-anlaşılır olmalı ve olabildiğince az kelimeyle ifade edilmelidir. Sorularda yazım ve imla hatası yapılmamalıdır.
2. Mümkün olduğu durumlarda, az sayıda uzun cevap gerektirecek sorular sormak yerine, çok sayıda kısa cevap gerektirecek sorulara yer verilmesi sınavın güvenirliğini ve kapsam geçerliğini arttırır.
3. Sorulara verilecek cevaplar birbirinden bağımsız olmalıdır. Aksi taktirde soruların puanlanması güçleşir.
4. Her soru tek bir davranışı ölçüyor olmalıdır. Bir soruyla birden fazla davranışı ölçmek puanlamayı güçleştirir.
5. Sınav soruları ders kitabından ya da kaynağından aynen alınmamalıdır ki öğrenciler ezbere teşvik edilmiş olmasın.
6. Sınav soruları çok önceden hazırlanmalı, mümkünse ölçülecek davranışlarla ile ilgili çokça soru yazılıp, bunların içinden en iyileri seçilerek kullanılmalıdır.
7. Sınav sorularının çoğaltılma işlemi de önceden yapılmalıdır.
8. Yazılı sorularına verilen cevapların puanlanmasında mutlaka önceden hazırlanmış bir cevap anahtarı ya da dereceli puanlama anahtarı kullanılmalıdır.
9. Yazılı sorularına verilen cevaplar puanlanırken, yazılı kâğıdının hangi öğrenciye ait olduğu bilinmemelidir. Aksi taktirde, subjektif olan puanlama daha da subjektif hale gelir ki bu da puanların güvenirliğini düşürür.
10. Yazılı sınavın amacının yazı güzelliği, kompozisyon becerisini ölçmek olmadığı durumlarda; yazı güzelliği, kompozisyon becerisi vb. özelliklerin puanlamayı etkilemesine izin verilmemelidir. Aksi taktirde ölçme aracının yapı geçerliği düşer.
11. Yazılı sınavların puanlaması yapılırken, çok iyi ve çok kötü kâğıtların okunmasından sonra okunacak kâğıtlar üzerindeki etkiyi kaldırabilmek için kâğıtlar önce sınıflandırılıp sonra puanlama yapılabilir.
12. Mümkünse birden fazla puanlayıcıya puanlatılmalıdır.
13. Puanlama yaparken puanlayıcının yorgunluk, açlık gibi olumsuz etkiler altında olmamasına dikkat etmesi puanların güvenirliği açısından önemlidir.
14. Eğer öğrencilerin istenilen cevabın yanı sıra fazladan da bir şeyler yazma olasılığı varsa, yazılan bu ekstra cevaplarda doğru cevaba ters bir ifade bulunduğunda eksi puanlama uygulanacağına dair öğrenciler de uyarılarak puanlama bu şekilde yapılmalıdır.

Yazılı Sınavların Avantajları Nelerdir?

1. Yazılı sınav sorularının hazırlanması kolay ve az zaman alır. Bu sebeple hazırlanması açısından kullanışlıdır.
2. Üst düzey bilişsel davranışların ölçülmesinde (analiz, sentez ve değerlendirme) kullanılabilecek geçerliği en yüksek olan ölçme aracıdır.
3. Şansla doğru cevap verme olasılığı neredeyse hiç yoktur, bu da güvenirliği arttırıcı bir özelliktir.
4. Öğrencileri daha çok çalışmaya teşvik eder. Öğrenme ve öğretme sürecinde etkili bir ölçme aracıdır.

Yazılı Sınavların Dezavantajları Nelerdir?

1. Belirli bir süre içinde öğrencilerin soruları okuyarak, düşünüp cevapları yazılı olarak vermeleri sebebiyle yazılı sınavların az sayıda soru içermesi güvenirliğinin ve kapsam geçerliğinin düşmesine sebep olur.
2. Özellikle cevaplamada belirli bir sınırlamanın (bir cümleyle açıklayınız, üç maddesin ifade ediniz vb.) olmadığı durumlarda, cevaplarda kesin bir doğru-yanlış ayrımı yapmak zordur. Bu da puanlamanın objektif yapılmasını zorlaştırır.
3. Şişirme yolla cevaplamaya elverişlidir.
4. Özellikle soruların anlaşılırlığının azaldığı ölçüde öğrencilerin soruları istediği yöne çekerek cevap verme olasılıkları artar. Bu durum sınavın güvenirlik ve geçerliğini düşürür.
5. Soruların güçlüğü, güvenirliği ve geçerliği önceden kestirilemez.

KISA CEVAPLI TESTLER

Kısa Cevaplı Test Nedir?

Cevapların öğrencilerin kendilerinin bir kelime, rakam ya da en fazla bir cümle ile verdiği sınav türüdür. Kısa cevaplı testler farklı şekillerden olabilir:

1) Soru cümlesi şeklinde olan maddeler:

- Tanımayı gerektiren maddeler: Bir tanım, açıklama, durum vb.'nin verilerek, verilenin ne olduğunun bilinmesinin istendiği maddelerdir.

egitim@hotmail.es

Örneğin:

Soru: Ölçme aracının tesadüfi hatalardan arınık olma derecesine ne denir?

Cevap: Güvenirlik

- Tanımlamayı gerektiren maddeler: Bir terim, olgu, kavram vb'nin açıklanmasının, tanımlanmasının istendiği maddelerdir.

Örneğin:

Soru: Bir ölçme aracının güvenirliği ne demektir?

Cevap: Ölçme aracının tesadüfi hatalardan arınık olma derecesine güvenirlik denir.

2) *Eksik cümle şeklinde olan maddeler (Boşluk doldurma maddeleri):* Verilen bir cümle, paragraf ya da bir bütün içinde eksik (boş) bırakılan kısmın doğru bir şekilde tamamalanmasının istendiği kısa cevaplı madde türüdür.

Örneğin:

Ölçme aracının tesadüfi hatalardan arınık olma derecesine .....denir.

Cevap: güvenirlik

Kısa Cevaplı Testler Ne Zaman Tercih Edilebilir?

Bilişsel davranış basamaklarının alt düzeylerindeki (bilgi, kavrama gibi) davranışlar çok sayıda soruyla ölçülmek isteniyorsa tercih edilebilir.

Kısa Cevaplı Testlerin Hazırlanmasında Nelere Dikkat Edilmelidir?

1. Ölçülmek istenilen davranış önemli olmalıdır. Gereksiz bir ayrıntıyı ölçen maddeler (sorular) kullanılmamalıdır.
2. Her maddede tek bir davranış (kazanım) ölçülmelidir. Bir maddede birden fazla soru ya da boşluk olmamaldır.
3. Soruların yalnızca bir doğru cevabı olmalıdır. Birden fazla doğru cevabı olabilecek sorular tercih edilmemelidir. Yine de böyle bir durum söz konusuysa, alternatif bir cevabın olması durumunda cevap anahtarı buna göre hazırlanmalıdır.
4. Kısa cevaplı test maddelerinde ipucu olabilecek ifadeler bulunmamalı, gereksiz bilgi verilmemelidir. Maddeler açık, anlaşılır ve olabildiğince az kelime ile ifade edilmelidir.
5. Boşluk doldurmalı maddelerde var olan kelimeler, boşluğa ne cevap verilmesini açıkça ifade edecek şekilde düzenlenmelidir.

Örneğin:

"İstanbul'un fethi.....dır/dir." Şeklindeki bir maddede yer alan boşlukta cevap olarak tarih mi, fethi gerçekleştiren padişahın adı mı, yoksa sonucu mu istendiği açıkça anlaşılmamaktadır. Böyle bir madde kısa cevaplı test maddesi olarak kullanılmamalıdır.

6. Kısa cevaplı test maddeleri ders kaynaklarından aynen alınmamalıdır. Ezbere bilgi ölçen maddelerden kaçınılmalıdır.
7. Maddeler birbirinden bağımsız olarak cevaplanabilmelidir. Bir maddenin cevabı diğer bir maddenin içinde yer almamalı, bir maddenin cevabına dayalı olarak diğer bir maddenin cevabı verilebilir olmamalıdır.
8. Boşluk doldurmalı maddelerden oluşan kısa cevaplı testlerde boşluk uzunlukları her maddede eşit şekilde verilmelidir. Aksi taktirde daha uzun bırakılan boşluk cevabın daha uzun, kısa bırakılan boşluk cevabın daha kısa olması gerektiğine ilişkin öğrenciyi yönlendirebilir.
9. Kısa cevaplı testlerde cevap yeri için verilen boşlular cevaplamayı ve puanlamayı kolaylaştıracak şekilde düzenlenmelidir. Görünüş geçerliğine dikkat edilmelidir.

Kısa Cevaplı Testlerin Avantajları Nelerdir?

1. Cevaplamak az zaman aldığı için çok sayıda soru sorulabilir. Bu durum da güvenirliği ve kapsam geçerliğini arttırır.
2. Cevaplar en fazla bir cümle gerektirdiği için yazılı sınavlarda olduğu kadar yazı güzelliği, ifade becerisi gibi özellikler puanlamada etki etmez. Kısacası puanlaması sözlü ve yazılı sınavlara göre daha objektif ve kolaydır. Bu da kısa cevaplı testleri daha güvenilir ve kullanışlı kılar.
3. Kısa cevaplı maddelerde cevabı öğrencilerin kendisi vermesine rağmen yazılı sınavlarda olduğu gibi şişirme cevap verme olasılığı yoktur, bu da güvenirliği ve geçerliği arttırır.
4. Her eğitim düzeyinde kolaylıkla uygulanabilecek bir sınav türüdür.
5. Cevapların kısa, belli ve genellikle tek olması güvenirliği arttırır.
6. Bilgi ve kavrama gibi alt düzey bilişsel davranışların ölçülmesinde kullanıldığı için hazırlanması oldukça kolaydır. Uygulanması ve puanlanması da kolaydır. Tüm bu durumlar açısından kullanışlığı yüksektir.

Kısa Cevaplı Testlerin Dezavantajları Nelerdir?

1. Üst düzey bilişsel davranışların (analiz, sentez ve değerlendirme) ölçülmesinde kullanılması uygun değildir.
2. Çoktan seçmeli ya da doğru-yanlış testlere göre puanlaması daha subjektiftir.

egitim@hotmail.es

DOĞRU-YANLIŞ TESTLERİ
(Sınıflama Gerektiren Testlerin En Çok Bilineni)

Doğru-Yanlış Test Nedir?

Cevapların sadece doğru ya da yanlış olarak verildiği sınav türüdür. Doğru-yanlış test maddeleri soru cümlelerinden oluşmaz. Cümleler önermeler şeklinde, doğruluğu ve yanlışlığı net olan ifadelerden oluşur.

Doğru-Yanlış Testleri Ne Zaman Tercih Edilebilir?

Bilgi ya da en fazla kavrama düzeyindeki bilişsel davranışların çok soruyla objektif bir şekilde ölçülmesinin istendiği durumlarda tercih edilebilir.

Doğru-Yanlış Testlerin Hazırlanmasında Nelere Dikkat Edilmelidir?

1. Her maddede tek bir davranış (kazanım) ölçülmeldir. Bir maddede birden fazla doğru ya da yanlış olabilecek önerme bulunmamaldır.
2. Her madde kesin doğru ya da yanlış olacak önermelerden oluşmalıdır. Bazen, çoğu zaman vb. gibi belirsiz ifadeler yer almamalıdır.
4. Bir maddenin yanlışlığı önemsiz bir ayrıntıda olmamalıdır.
5. Maddeler açık, net ve olabildiğince kısa cümlelerden oluşmalıdır.
6. Olumsuz ifade, özellikle de çift olumsuz ifade kesinlikle bulunmamalıdır.
7. Bir kanıya dayalı ifadeler bir kaynağa dayalı olarak, açık bir şekilde ifade edilmelidir.
8. Doğru-yanlış sayıları yaklaşık eşit olmalıdır.
9. Doğru-yanlış belli bir örüntüde yerleştirilmemelidir.
10. Test maddesinin doğru ya da yanlış olduğunu açığa çıkaracak şekilde ipuçları içermesine izin verilmemelidir.
11. Maddeler aynen bir kaynaktan alınmamalı, ezbere bilgi ölçülmemelidir.
12. Düzeltme formülü ("iki yanlış bir doğruyu götürecektir" gibi) kullanılacaksa, yönergede mutlaka ifade edilmelidir.

Doğru-Yanlış Testlerin Avantajları Nelerdir?

1. Testin hazırlanması, uygulanması ve puanlanması kolaydır. Bu nedenle her açıdan kullanışlı bir sınav türüdür.
2. Çok sayıda soru sorulabilir. Bu da güvenirliği ve kapsam geçerliğini arttırır.
3. Puanlaması objektif ve kolaydır.
4. Yazı güzelliği, ifade becerisi gibi durumlar puanlamayı etkilemez, bu da yapı geçerliğini arttırır.

5. Şans başarısı en yüksek sınav türüdür, güvenirliği düşüktür. Yanlış cevabın doğrusu istenirse şans başarısı azaltılabilir.
6. Puanlaması kolaydır.

Doğru-Yanlış Testlerin Dezavantajları Nelerdir?

1. Doğru cevabın tahminle bulunma olasılığının en yüksek olduğu sınav türüdür. Doğru cevabı bilmeyen bile %50 şansla doğruyu bulabilir. Bu da güvenirliği düşürür.
2. Test maddeleri dikkatli ve özenli hazırlanmazsa sadece ezbere dayalı, önemsiz ya da çok basit bir bilgi ölçmekten öteye gidilemez.
3. Maddeler dikkatli hazırlanmadığı taktirde ipucu veren noktalar içerebilir, bu da testin geçerliğini düşürür.
4. Öğrenme eksikliklerinin belirlenmesi ve öğretimin değerlendirilmesi amacı ile kullanılan izleme testleri için uygun değildir.

!!!!Yanlış öğretmek amaç olmadığı için ilköğretimde özellikle de ilk sınıflarında kullanılmaması tavsiye edilir.

ÇOKTAN SEÇMELİ TESTLER
(Seçme Gerektiren Testler)

Çoktan Seçmeli Test Nedir?

Cevapların cevaplayıcılar tarafından verilmediği, doğru cevabın testi hazırlayan/lar tarafından seçenekler arasında verilerek cevaplayıcılardan doğru cevabın bulunmasının istendiği test türüdür.

Çoktan seçmeli testlerde yer alan madde formu 4 bölümden oluşmaktadır:

1. Madde kökü: Cevaplandırılması istenilen soru yer alır.
2. Seçenekler (şıklar): Madde kökünde yer alan soruya uygun cevap olma olasılığı olan ifadelerden oluşur.
3. Anahtarlanmış (doğru) cevap: Sorunun doğru cevabıdır.
4. Çeldiriciler: Doğru cevap dışındaki, bilmeyeni ve az bileni yanıltmak üzere verilen seçeneklerdir.

egitim@hotmail.es

Örneğin;

MADDE KÖKÜ: Aşağıdakilerden hangisi "bir testin tesadüfi hatalardan arınık olma derecesi" ni ifade eder?

SEÇENEKLER
- A) güvenirlik → *Doğru cevap*
- B) geçerlik → *Çeldiriciler*
- C) kullanışlılık → *Çeldiriciler*
- D) uygunluk → *Çeldiriciler*
- E) ölçülebilirlik → *Çeldiriciler*

Çoktan Seçmeli Testler Ne Zaman Tercih Edilebilir?

Çoktan seçmeli testler; z- Sınavın uygulanacağı kişi sayısı çok olduğunda,

- Seçme ve yerleştirme gerektiren durumlarda,
- Soru sayısının çok olması gerektiğinde,
- Sınav hazırlamak için yeterince vakit var ve puanların çabuk bildirilmesinin gerektiği durumlarda,
- Çoktan seçmeli test maddesi hazırlamak için gerekli uzmanlık ya da bilgiye sahip olunduğunda,
- Testin güvenirlik ve geçerliğinin yüksek olmasının gerekli olduğu durumlarda

tercih edilebilir.

Çoktan Seçmeli Testlerin Hazırlanmasında Nelere Dikkat Edilmelidir?

1. Çoktan seçmeli testlerin madde kökü ve tüm seçenekleri açık, anlaşılır ve olabildiğince kısa olmalıdır.
2. Çoktan seçmeli test maddelerinin seçeneklerinde "çoğunlukla, bazen, nadiren" gibi kesinlik ifade etmeyen kelimeler kullanılmamalıdır.
3. Çoktan seçmeli testlerin cevaplandırılmasında, cevaplayıcı zamanın çoğunluğunu test maddelerini okuyarak ve cevabı düşünerek geçirecektir. Bu sebeple test maddeleri olabildiğince az kelimeyle açık ve net olarak ifade edilmelidir.
4. Test maddelerinin yazımında sınava girecek bireylerin düzeylerine uygun bir dil kullanılmasına dikkat edilmelidir.
5. Test maddelerinin gerek kökü gerekse seçeneklerindeki ifadeler belirli bir kaynaktan aynen alınmamalıdır. Aksi taktirde, kaynaktan olduğu gibi alınarak oluşturulan maddeler ezbere bilgiyi ölçecektir. Eğitimin amacı, öğrenciyi ezbere yönlendirmek değildir.
6. Çoktan seçmeli testte yer alan her bir test maddesinin cevaplandırılması birbirinden bağımsız olmalıdır. Aksi taktirde, test maddelerinin puanlandırılmasında zorluk yaşanacaktır.
7. Çoktan seçmeli testteki her hangi bir maddedenin cevabı diğer bir maddenin kökünde ya da seçeneklerinde yer almamalı, ipucu verilmemelidir.
8. Çoktan seçmeli testteki her bir maddenin tüm seçeneklerindeki ifadeler, tarz, uzunluk vb. yönlerden benzer ve uyumlu olmaldır. (Örn; tüm seçeneklerdeki ifadelerin geniş zaman kipiyle yazılması gibi)
9. Test maddelerinin çeldiricileri sadece *bilmeyeni ve yanlış bileni* yanıltmalıdır. İfade bozukluğu, gramer hatası vb. yapılarak, *bileni* yanıltacak test maddeleri kesinlikle oluşturulmamalıdır.
10. Mümkünse "Yukarıdakilerin hepsi/ yukarıdakilerin hiçbiri" seçenekleri tercih edilmemelidir, tercih edildiğinde ise dikkatli kullanılmalıdır. Kesin doğru gerektiren maddelerde ise asla kullanılmamalıdır.
11. Çoktan seçmeli test maddelerinin seçenekleri belli bir düzene göre yerleştirilmelidir. (Örneğin: alfabetik sıraya göre).
12. Çoktan seçmeli bir testte yer alan tüm maddelerin seçenek sayısı birbirine eşit sayıda olmalıdır.
13. Çoktan seçmeli test maddelerinin seçenek sayısı cevaplayıcıların seviyesine uygun olmalıdır. (İlköğretim dördüncü sınıf seviyesine kadar üç seçenek, lise seviyesine kadar dört seçenek, lise ve üzeri beş seçenek tercih edilmelidir.) Ayrıca testte yer alan bütün maddelerin seçenek sayıları aynı olmalıdır.
14. Doğru cevap belli bir örüntüde yerleştirilmemelidir.
15. Çoktan seçmeli testteki herbir doğru seçenek sayısı (A'ların, B'lerin... toplam sayısı) yaklaşık aynı oranda yerleştirilmelidir.
16. Çoktan seçmeli test maddelerinin madde kökünde ve seçeneklerinde olabildiğince olumsuz ifade kullanılmamalıdır. Şayet kullanılmış ise olumsuz ifade dikkat çekici şekilde (altı çizili, kalın vb.) belirgin yazılmalıdır.
17. Çoktan seçmeli testte yer alan tüm maddeler belirli bir düzende verilmelidir. (Aralardaki boşluklar vb...)
18. Yazım yanlışlığı yapılarak çeldirici oluşturulmamalıdır.
19. Birbirinin zıddı seçeneklerin kullanılmamasına dikkat edilmelidir.
20. Çoktan seçmeli testte yer alan bir maddenin çeldiricileri diğer maddelerin doğru cevapları arasından seçilebilir.

egitim@hotmail.es

Çoktan seçmeli testlerde yer alan soruların kolaydan zora doğru sıralanması önerilir. Bu durum cevaplayıcıların sınava olan motivasyonlarının artmasına yardımcı olacaktır.

Çoktan Seçmeli Testlerin Avantajları Nelerdir?

1. Puanlaması objektiftir. (*güvenirlik* açısından olumlu bir özellik)
2. Uygulaması kolaydır. (*kullanışlık* açısından olumlu bir özellik)
3. Puanlaması kolaydır (bazen puanlamayı makinalar yapar) ve puanlama süresi kısadır. (*kullanışlık* açısından olumlu bir özellik)
4. Çok sayıda soru sorulabilir. (*güvenirlik* ve *kapsam geçerliği* açısından olumlu bir özellik)
5. Zamanın çoğu maddeyi okuma ve cevabı bulmaya harcanır. Sınav süresini öğrenci cevabı yazmak için harcamak zorunda kalmaz.

Çoktan Seçmeli Testlerin Dezavantajları Nelerdir?

1. Doğru cevap çoğu zaman madde içinde verilir. Bu da bilginin hatırlatılmasına sebep olduğu için üst düzey davranışların ölçülmesinde yetersizdir. (*Yapı geçerliği* açısından olumsuz bir özellik)
2. Şans başarısı vardır. (*güvenirlik* açısından olumsuz bir özellik)
3. İleri düzey davranışların ölçülmesi bu test türü için çok zordur. (*Yapı geçerliği* açısından olumsuz bir özellik)
4. Hazırlama süresi uzun ve zahmetlidir, uzmanlık gerektirir. (*kullanışlık* açısından olumsuz bir özellik)

Çoktan Seçmeli Madde Tipleri →

A) Doğru Cevaba Göre Sınıflama:

1. Doğru cevabı kesin ve tek olan maddeler.
2. Anahtarlanmış cevabı en doğru olan maddeler.
3. Anahtarlanmış cevabı birden fazla olan maddeler.(Puanlaması ve istatistiksel analiz yapılması zor olduğu için pek tercih edilmemeli.)
4. Birleşik cevap gerektiren maddeler.
5. Doğru cevabı gizli maddeler. (Hatırlamayı en aza indirgeyen madde türüdür).

B) Madde Köküne Göre Sınıflama:

1. Kökü soru tipinde olan maddeler.
2. Kökü eksik cümle yapısında olan maddeler.
3. Kökü olumsuz olan maddeler.

C) Maddelerin Gruplanışına Göre Sınıflama:

1. Ortak köklü maddeler. (Soru çok olursa, soruların bağımsızlığı zorlaşır.)
2. Ortak seçenekli maddeler. (Soru sayısı ile seçenek sayısı eşit ise şans başarısı artar.)

PERFORMANS DEĞERLENDİRME

Performas değerlendirmede, öğrencilerin akademik bilgi ve becerilerini gerçek yaşamda karşılaşılan problemlere uygulayabilme düzeylerini belirleyebilmek amaçlanır. Performans değerlendirme; öğrencinin bir resim yaparken izlediği süreci, bir kimya deneyinde deney düzeneğini hazırlama ve deneyi gerçekleştirebilme becerilerini gözlemlemeyi, kompozisyon yazma becerisinin ölçülmesini ya da bir matematik problemini çözerken izlediği yolları adım adım takip etmeyi mümkün kılar. Tüm bu durumların ortak özelliği verilen görevlerin gerçek yaşamdaki durumlarla ilgili olmasıdır. Baker, O'Neil ve Linn (1993) performansın ölçülmesinin altı genel özelliğini sıralamıştır. Performansın ölçülmesi,

1. Çoktan seçmeli maddelerden oluşmaz, açık-uçlu sorular biçimindedir.
2. Üst düzey bilişsel becerilere (analiz, sentez, değerlendirme) odaklanma eğilimindedir.
3. Öğrencilerin maddeleri ya da görevleri doğru cevaplandırmaları için gerekli stratejileri doğrudan uygulamalarını gerektirir.
4. Çoğunlukla tamamlamak için daha fazla zamanın gerektiği karmaşık görevlerden oluşmaktadır.
5. Bireysel olarak ya da geniş gruplara uygulanabilir.
6. Genellikle verilen soruların (görevlerin) cevaplandırılmasında öğrencilere daha fazla özgürlük tanımaktadır.

Böylece performans değerlendirme, öğrencilerin öğrendiklerine ilişkin çok daha geniş açıdan bilgi sahibi olabilme olanağı sağlar. Bireysel ayrılıkları göz önünde bulundurarak, öğrencilerin zayıf ve güçlü yönlerinin belirlenmesini mümkün kılar. Performans değerlendirmenin avantajlarından biri de çoktan seçmeli testlerle ölçmenin zor olduğu üst düzey bilişsel beceriler için kolaylıkla kullanılabilmesidir. Performans değerlendirmede, çoktan seçmeli testlere göre ölçülmek istenilen özellikler daha açık olarak gözlemlendiği için, çoktan seçmeli testlerden daha geçerli sonuçlara ulaşmak mümkündür (Alharby, 2006).

Performans değerlendirmenin klasik ölçme araç ve yöntemlerinden farklı özellikleri aşağıdaki tabloda kısaca özetlenmeye çalışılmıştır:

egitim@hotmail.es

Performans Değerlendirme	Klasik Ölçme Araç ve Yöntemleri
Sonuçla birlikte süreç de gözlenerek puanlama yapılır.	Sadece sonuca göre puanlama yapılır.
Üst düzey bilişsel davranışların ölçülmesini mümkün kılar.	Üst düzey bilişsel davranışların ölçülmesi bazen çok zor bazen de imkansız olabilir.
Tek bir doğru cevap ya da en iyi cevap yoktur.	Genellikle tek bir doğru cevap vardır.
Öğrenci istenilen görevi yerine getiriken ya da bir soruyu cevaplandırırken her aşamasında dönüt verme imkanı vardır.	Öğrencinin cevabına yönelik dönüt ya sonuçta verilir ya da hiç verilmez.
Öğrenciden yerine getirmesi istenilen görev ya da cevaplandırması istenilen soru gerçek yaşamda karşılaşacağı durumlarla ilişkili, gerçek yaşamdaki problemleri çözmesine yardımcıdır.	Öğrenci, cevaplaması gereken sorular ile gerçek yaşamdaki problemleri ilişkilendiremeyebilir.
Puanlamada dereceli puanlama anahtarı (rubrik, bknz sf:) gibi tek bir doğru cevaba göre puanlamanın yapılmadığı, belirli kriterlerin yer aldığı araçlar kullanılır. Öğrenciler, puanlamanın yapılacağı bu kriterlerden haberdardır.	Puanlamada cevap anahtarı gibi genelde tek bir cevabın (bazen de alternatif başka bir cevabın) yer aldığı araçlar kullanılır. Öğrenciler, cevap anahtarını ya hiç görmezler ya da puanlama süreci tamalandıktan sonra görebilirler.
Puanlaması daha subjektif ve zaman alıcıdır.	Puanlaması daha objektiftir ve daha kolaydır.
Performans değerlendirmede ortaya çıkan, gözlenebilir bir performans ya da somut bir üründür.	Ortaya somut bir ürün çıkmayıp, genelde kağıt üzerinde verilen cevapla sonuçlanır.
Öğrencilerin, orjinal fikirler oluşturabilmedeki başarısının, yaratıcılığının da ölçülebilmesine imkân tanır.	Öğrencilerin, yaratıcılığını her zaman ölçmek mümkün olmayabilir. Zaman zaman ölçülen ezbere bilginin ötesine geçemeyebilir.

ÖĞRENCİ ÜRÜN DOSYASI (PORTFOLYO)

Portfolyo olarak adlandırılan öğrenci ürün dosyaları, öğrencilerin belirli bir süreçte belirli bir amaca yönelik gösterdikleri performansı, gelişimi, çalışma ve çabalarının bir arada bulunduğu gelişim dosyalarıdır. Bir başka deyişle, öğrencilerin belirli bir zaman periyodunda sergiledikleri çalışma ve ilerlemenin açıkça gözlenebildiği öğrenci ürünlerinin bir koleksiyonudur.

Bir portfolyoda, öğrencilerin bitmiş ödevlerinin yanısıra, bu ödevi oluştururken ortaya koydukları taslak çalışmalar, karalamalar yer alabilir. Öğrencilerin tamaladıkları bir projenin ses ya da görüntü kayıtları, taslak çalışmaları, projenin tamamlanmış hali bulunabilir. Öğrencilerin yazdıkları kompozisyonlar, şiirler, yaptıkları resimler, mektuplar, öğrenci günlükleri, bir problem çözmede gösterdikleri aşamalar, bu problem çözerken gösterdiği bilişsel performansın yanısıra duyuşsal olarak da neler hissettiği, düşünceleri vb. yer alabilir. Ayrıca portfolyolarda öğretmenin değerlendirmelerinin yanı sıra öğrencilerin kendi değerlendirmeleri ve arkadaşlarının değerledirmeleri de bulunabilir.

Örneklerden de anlaşılacağı üzere portfolyo, öğrencilerin bilişsel, duyuşsal ve hatta psiko-motor alanlarda gösterdiği gelişim sürecini izlemeye imkân tanıyan çok boyutlu ve kapsamlı değerlendirmeyi sağlayan bir ürün dosyasıdır. Değerlendirme sürecinde öğretmenin yanısıra öğrencinin kendisinin de yer alması, öğrenciye hem sorumluluk alma duygusunu kazandırır hem de değerlendirme sürecinde öğrenmenin gerçekleşmesini sağlar.

Portfolyolar, farklı amaçlar için hazırlanmış farklı içeriklere sahip olabilir:

- Öğrencinin çalışmalarının tüm evrelerinin (karalamalar, taslak çalışmalar, vb.) yer aldığı, öğrencinin gelişim sürecinin açıkça izlenmesinin mümkün olduğu, not verme amaçlı olmayan çalışma dosyaları olabilir.
- Öğrencinin "en iyi" olarak nitelendirdiği çalışmaların yer aldığı dosyalardır. Genellikle öğrencinin kendisinin seçtiği ve gösterdiği en iyi performansın yer aldığı çalışmalardan oluşur. Oluşturulan bu tür portfolyolar öğrencinin bir vitrini gibidir. Bu tür portfolyolarda, yapılan değerlendirmeler daha çok sonuca ya da ürüne dayalıdır.
- Öğrencinin en iyi ortaya koyduğu çalışmasının dışında diğer çalışmalarının da yer aldığı, en iyiyi ortaya koyana dek gerçekleştirdiği aşamaların izlenebildiği, öğrenim amaçlarına yönelik öğrencinin gösterdiği gelişimi gösteren dosyalardır. Ürünün yanısıra sürecin değerlendirilmesine de imkan tanır. Bu tür dosyalarda her çalışma puanlanarak öğrencinin gelişim aşamaları açıkça gözlenebilir.

Portfolyolar, matematik ve bilim alanlarında, resim, edebiyat, müzik gibi sanatsal alanlarda oluşturulabileceği gibi iki veya daha fazla alanı kapsayan projeler için de oluşturulabilir. Genellikle öğretmenin yanısıra öğrenci de dosyasında yer alacak çalışmaların neler olabileceğine karar vermeye dâhil edilir. Portfolyolar, öğrencilerin kendilerini açıkça yansıtabilmelerine, yaratıcılıklarını ortaya koyabilmelerine yardımcı olduğu gibi kendi kendilerini analiz etme özelliklerini geliştirmelerine de katkı sağlar. Portfolyolarda yer alacak çalışmalar, aşağıdaki soruların yanıtlanmasında yardımcı olabilir:

egitim@hotmail.es

Öğrenci;

1. bu çalışma ile ne öğrendi?
2. öğrenirken nasıl bir yol izledi?
3. ortaya koyduğu çalışma sürecinde neler düşündü?
4. ortaya koyduğu çalışma sürecinde kendine ne gibi sorular sordu?
5. çalışma sürecinde kendini nasıl analiz etti?
6. bu çalışmayı gerçekleştirirken bilgisini nasıl yapılandırdı?
7. diğer kişilerle nasıl iletişim kurdu?
8. öğrenirken ne gibi güçlüklerle karşılaştı?

Portfolyolar, öğrencinin öğrenim sürecindeki gelişimini tüm boyutlarıyla izleyebilmeyi, öğrencilerin çalışmalarını olabildiğince en doğru şekilde değerlendirebilmeyi sağlar. Bir teste benzemez; çünkü portfolyoda, öğrenci başarısını tek bir açıdan ve tek bir zamanda ölçmek amaç değildir. Ancak, portfolyolar sadece öğrenci çalışmalarının dosyalanması olarak da algılanmamalıdır. Öğrencilerin çalışmalarının yer aldığı bir koleksiyon olmanın çok daha ötesinde amaçları vardır. Bir öğrenci ürün dosyası aynı zamanda;

1. Öğrencilerin kendileri hakkındaki görüşlerini, kendi kendilerini değerlendirmelerini içerir.
2. Öğrencilerin, ebeveynleriyle yaptıkları çalışmalarda daha iyi bir iletişim kurmalarını sağlar.
3. Ebeveynlerin öğretmenler ile iletişimini kolaylaştırır.
4. Değerlendirme aşamasına dâhil edilerek, öğrencilerin sorumluluk duygusunun gelişmesine yardımcı olur ve değerlendirme sürecinde aynı zamanda öğrenmenin de gerçekleşmesine katkı sağlar.
5. Eğitim programı ve öğretimin değerlendirilmesi, aynı zamanda öğretmenin de kendi kendisini değerlendirmesi için bir araç sağlar.

Portfolyo ile değerlendirmenin klasik ölçme araç ve yöntemlerinden farklı özellikleri aşağıdaki tabloda kısaca özetlenmeye çalışılmıştır:

Portfolyo	Klasik Ölçme Araç ve Yöntemleri
Öğrencinin gelişimi, öğrenim sürecinde sürekli olarak ölçülebilir.	Öğrencinin bilgisi, yeteneği, vb. belirli ve tek bir zamanda ölçülebilir.
Değerlendirmede, öğretmenle birlikte öğrenci de bulunmaktadır. Öğrenci değerlendirmedeki ölçütlerden haberdardır.	Değerlendirmeyi sadece öğretmen yapar. Öğrenci değerlendirme sürecinde yer almaz ve ölçütlerden habersizdir.
Öğrenciye sorumluluk alma duygusu kazandırır.	Öğrenciye sorumluluk verilmez.
Öğrencinin bilişsel davranış düzeyinin gelişimiyle birlikte duyuşsal ve hatta psiko-motor davranış düzeylerindeki gelişim de izlenebilir.	Öğrencinin genellikle sadece bilişsel davranış düzeyine ilişkin gözlemler yapılabilir.
Bireysel ayrılıklar göz önünde bulundurulur. Her öğrenci kendi gelişim sürecinde değerlendirilir, başka öğrencilerle kıyaslama yapılmaz.	Her bir öğrencinin diğer öğrencilerle kıyaslanması mümkündür.
Öğrenciler diğer arkadaşlarının çalışmalarını değerlendirebilirler (akran değerlendirmesi). Böylelikle birbirlerine yardımcı olma, ekip çalışması yapabilme alışkanlıklarını da kazanılmaya başlar. (AKRAN DEĞERLENDİRME FORMU ÖRNEĞİ KONULABİLİR)	Öğrencileri birbirlerini değerlendirmez. Değerlendirmeyi sadece öğretmen yapar.
Ürünle birlikte süreç de değerlendirilir.	Sadece ürün değerlendirilir.
Değerlendirme sürecinde öğrenme de gerçekleşir.	Ayrıntılı dönüt verilmezse değerlendirme sürecinde öğrenme gerçekleşmez.



Bireysel gelişim dosyasında, geleneksel yöntemde olmasına karşın, hiçbir zaman "Hangi öğrenci daha çok biliyor?" sorusuna yanıt aranmaz. Bu öğrenci 'Ne biliyor?'un yanıtı aranır.

Portfolyo ile değerlendirmede öğretmen, öğrenci ve velinin değişik rolleri vardır;

egitim@hotmail.es

Öğretmenin rolü 	- Öğrencilerine rehberlik eder, öğrencilerinin çalışmalarında yardımcı olur. - Oluşturulacak portfolyonun kapsamını belirler. - Portfolyonun değerlendirilmesi öğretmenin sorumluluğundadır. - Değerlendirme ölçütlerinin belirlenmesinde öncülük eder. Bu ölçütleri öğrencilerine ve velilere açık ve anlaşılır bir şekilde bildirir.
Öğrencinin rolü 	- Portfolyoda yer alacak çalışmaların neler olacağına karar vermede öğretmeniyle birlikte rol alır. - Gelişimini gözlemleyerek eksiklerini saptar ve bunların giderilmesi için neler yapması gerektiğine ilişkin sorumluluk alır. - Öğrencilerinin birbirlerine değerlendirmesiyle ekip çalışması yapma alışkanlığı ve sorumluluğu kazanır.
Velinin rolü 	- Öğrenim sürecinin bir parçası olarak, portfolyonun nasıl kullanıldığını ve öğrenciler için önemini çok iyi kavraması istenir.

Kısaca özetlersek, portfolyonun genel amaçları şunlardır:

- Öğrencinin gelişimini izleyebilmek,
- Öğrencideki öz disiplin ve sorumluluk bilincini geliştirmek ve kendi kendini değerlendirme becerisine ulaşmasını sağlamak,
- Öğrencinin gerçekte ve öğrendiğinin açık resmini çizebilmek,
- Müfredata bağlı olarak yapılan yazılı ve sözlü değerlendirmelerin dışına çıkıp, alternatif ölçme ve değerlendirme yöntemi geliştirmek,
- Gelecekteki öğretmenlerine bilgi sunmak,
- Öğrencinin gelişimini daha sağlıklı izleyebilmek,
- Öğrencilerin yeteneklerini sergilemek ve ilgi alanlarını çoğaltmak,
- Öğrencinin gerçekte ne öğrendiğini ortaya koymak,
- Öğrencilerin diğer arkadaşlarının yaptığı portfolyoları izleyip, yardımcı olarak gelecekteki takım çalışmasına zemin hazırlamaktır.

Portfolyo ile Değerlendirmenin Öğrenci Açısından Yararları

1. Bir sınav kâğıdı üzerindeki öğrenci başarısını değerlendirmek yerine bir öğretim süreci içerisinde öğrencinin gelişimini takip etmeyi mümkün kılar.
2. Bir öğrenme süreci içerisinde öğrencinin harcadığı zamanı, çalışmalarını, performansını, karalamalarını, eksikliklerini ve düzeltmelerini ayrıntılı bir şekilde gösterir.
3. Öğrencinin karalamalarından düzeltmelerine kadar çalışmasının her bölümünün önemli olduğunu ve birbiriyle ilişkili olduklarını fark etmesini sağlar.
4. Bir öğrenme ünitesi içerisindeki birçok adımdan sonra öğrencideki gelişmeleri gösterir.
5. Öğrenci en iyi yaptığı çalışmaları bireysel gelişim dosyasına koymayı ve çalışmaları ile iftihar etmeyi, çalışma için daha fazla sorumluluk almayı öğrenir.
6. Öğrencinin kendi çalışmalarını değerlendirmesine rehberlik eder.
7. Öğrencinin yaratıcılığını geliştirmede daha bilinçli olmasını sağlar.

Portfolyo ile ile Değerlendirmenin Öğretmen Açısından Yararları

1. Öğretmenin öğrenme-öğretme süreci içerisinde her bir öğrencinin, gelişimleri ile ilgili müsvedde, karalama, düzeltme ve sonuca giden performansını içeren geniş bir kaydının elinde olmasını sağlar.
2. Öğretmenler arasında öğretim materyallerinin ve tekniklerinin paylaşımına katkıda bulunur.
3. Başka sınıflardaki öğrencilerin çalışmalarını incelerken, öğretimle ilgili fikirlerini geliştirme olanağı bulur.
4. Bir öğrencinin bir bütün olarak gelişim aşamalarını izleme fırsatını elde eder.

DERECELİ PUANLAMA ANAHTARI (RUBRİK)

Dereceli puanlama anahtarları, öğrencilerin gösterdikleri performansların sürecini ve sonucunu analiz edebilmek için öğretmenler ya da ölçme uzmanları tarafından geliştirilen açıklayıcı puanlama şemalarıdır. Puanlama öncesi bu tür bir şemanın hazırlanmış olması, puanlamanın subjektif olduğu performansın ölçülmesinde yer alan görev ya da maddelerin daha objektif puanlanabilmesini sağlar. Dereceli puanlama anahtarlarında her bir puanın hangi özellikleri taşıyan cevaplara karşılık geldiği açıkça tanımlanır. Örneğin;

- deney düzeneğinin eksiksik, ve gerekli sırada hazırlanmış olması: 5 puan,
- deney düzeneğinin eksiksiz ancak gerekli sırasının gerektiğinden farklı düzenlenmiş olması: 4 puan vb.

Böylelikle farklı puanlayıcıların bir soruya verilen cevaba aynı puanı verme olasılığı artar. Performansın ölçülmesinde yer alan görev ya da maddelerin dereceli puanlama anahtarı kullanılmadan puanlanmasında (örneğin; bir öğrencinin bir matema-

egitim@hotmail.es

tik sorusuna verdiği cevap için on puan üzerinden yedi puan alması gibi) öğrencinin performansının hangi aşamasında nasıl bir başarı ya da başarısızlık gösterdiğine ve bir sonraki performansında nasıl bir gelişme göstermesi gerektiğine ilişkin açıklayıcı bir bilgi bulunmaz. Halbuki dereceli puanlama anahtarlarıyla yapılan puanlamalarda, öğrencinin her bir görev ya da maddedeki performansının hangi düzeyde olduğu, performansını geliştirmek için hangi kriterleri sağlaması gerektiği görülebilmektedir.

Bir performansın ölçülmesinde dereceli puanlama anahtarının uygun bir puanlama yöntemi olup olmadığına karar vermek, hangi ders ya da sınıf düzeyi için kullanılacağından çok yapılan ölçmenin amacına bağlıdır. Dereceli puanlama anahtarları, grup aktivitelerinin, projelerin ya da sözel sunumların puanlanmasının yanı sıra Türkçe, yabancı dil, matematik ve fen derslerinde, lise ve öncesi sınıf düzeylerinde kullanılabilirler. Öğrencinin gelişmesi için geri bildirimde bulunmak ya da öğrencinin belirli ölçütlere göre nereye ulaştığı açıkça saptanmak isteniyorsa dereceli puanlama anahtarlarının kullanılması uygun olacaktır.

Dereceli puanlama anahtarları bütünsel (holistik) ve analitik olmak üzere iki sınıfa ayrılmaktadır.

Bütünsel (Holistik) Dereceli Puanlama Anahtarları

Bütünsel dereceli puanlama anahtarları ilk kez 1960'larda ortaya çıkmış ve ilk olarak genel izlenimle puanlama yöntemi olarak tanımlanmıştır. Bütünsel dereceli puanlama anahtarları öğrencilerin her bir görev ya da maddede göstermiş oldukları performansın bir bütün olarak, alt bileşenlerine ayrılmaksızın puanlanmasında kullanılır. Bütünsel dereceli puanlama anahtarlarıyla yapılan puanlamada, her bir görev ya da maddenin bütününde gösterilen performans yüksek ise bazı bileşenlerindeki hatalar görmezden gelinebilir. Bütünsel dereceli puanlama anahtarlarıyla puanlamada, puanlayıcı her bir görev ya da cevabı bir bütün olarak inceleyerek, bu bütüne ilişkin bir puan vermektedir. Pek çok araştırmacı bu yöntemin, puanlayıcının daha subjektif olmasına sebep olabilecek daha gereksiz ayrıntılara takılmaksızın puanlama yapmasına imkan verdiğini savunmaktadır. Bu puanlama yaklaşımının bütünün parçalarından daha büyük önem taşıdığı durumlarda yani cevapların organizasyon ve ikna etme gücü gibi genel özelliklerinin kalitesinin değerlendirildiği durumlarda daha uygun olduğunu vurgulamışlardır.

Tablo 1. Bütünsel Dereceli Puanlama Anahtarları İçin Bir Örnek

Puanlar	Tanımlar
5 puan	Problemin *tümüyle* anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin *tüm gereklilik* cevapta yer almıştır.
4 puan	Problemin *oldukça* anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin *tüm gereklilik* cevapta yer almıştır.
3 puan	Problemin *kısmen* anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin *çoğu gereklilik* cevapta yer almıştır.
2 puan	Problemin *çok az bir kısmının* anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin *çoğu gereklilik* cevapta yer almamıştır.
1 puan	Problemin *anlaşılmamış* olduğu gösterilmiştir.
0 puan	Cevap verilmemiştir /görev hiçbir şekilde gerçekleştirilmemiştir.

Analitik Dereceli Puanlama Anahtarları

Analitik dereceli puanlama anahtarlarında, bütünsel olanların aksine, öğrencinin her bir görev ya da soruda gösterdiği performans öncelikle alt bileşenlerine ayrılarak puanlanır. Görev ya da soruya ilişkin puan bu bileşenlerden alınan puanların toplamıdır. Bir görev ya da cevapta gerçekleştirilen performansın her bir alt boyutunda öğrencinin güçlü ve zayıf yönleri belirlenebilir. Pek çok araştırmacıya göre, bir alanın alt alanlarına ayrılarak yapılan puanlamada, puanlayıcıların subjektifliğinden daha az etkilenme olmakta ve puanlayıcılar arası farklılık azalmaktadır.

Tablo 2. Analitik Dereceli Puanlama Anahtarları İçin Bir Örnek

	1 Puan	2 Puan	3 Puan	4 Puan	Toplam
Ölçüt 1	Performansın başlangıç seviyesi gösterilmiştir.	Performansta üst düzeye doğru bir gelişme gösterilmiştir.	Üst düzey bir performans gösterilmiştir.	En üst düzeyde performans gösterilmiştir.	Puan
Ölçüt 2	Performansın başlangıç seviyesi gösterilmiştir.	Performansta üst düzeye doğru bir gelişme gösterilmiştir.	Üst düzey bir performans gösterilmiştir.	En üst düzeyde performans gösterilmiştir.	Puan
Ölçüt 3	Performansın başlangıç seviyesi gösterilmiştir.	Performansta üst düzeye doğru bir gelişme gösterilmiştir.	Üst düzey bir performans gösterilmiştir.	En üst düzeyde performans gösterilmiştir.	Puan
Ölçüt 4	Performansın başlangıç seviyesi gösterilmiştir.	Performansta üst düzeye doğru bir gelişme gösterilmiştir.	Üst düzey bir performans gösterilmiştir.	En üst düzeyde performans gösterilmiştir.	Puan

egitim@hotmail.es

Performansı ölçmek isteyen kişi pek çok faktörü düşünerek bütünsel ya da analitik dereceli puanlama anahtarlarından hangisini kullanmanın daha uygun olacağına karar verir. Göz önünde bulundurulması gereken en önemli faktörlerden biri, ölçmede yer alan görevlerin (soruların) türüdür. Örneğin, fen bilgisi dersinde öğrencinin yapacağı bir deneyde göstereceği performansın, deney düzeneğini oluşturabilme, analizleri doğru yapabilme, doğru yorum ve sonuçlara ulaşabilme gibi alt bileşenlerinin ölçüleceği durumlarda analitik dereceli puanlama anahtarı kullanmak daha uygun olabilir. Diğer taraftan performansın alt bileşenlerinin ölçülmesine gerek olmadığı performansın ölçülmesinde ise holistik dereceli puanlama anahtarı tercih edilebilir.

Dereceli puanlama anahtarı seçimindeki bir diğer önemli faktör, ölçmenin yapılacağı kişi sayısıdır. Bu açıdan bakıldığında çok sayıda bireyin katıldığı performans ölçmelerinde, zaman açısından tasarruf sağlayacağı için bütünsel dereceli puanlama anahtarı kullanmak kolaylık sağlayacaktır.

Dereceli puanlama anahtarlarından hangisinin daha uygun olacağının belirlenmesindeki bir başka faktör ise performansın ölçülmesiyle elde edilen sonuçların hangi amaç için kullanılacağıdır. Özellikle sınıf içi ölçmelerde olduğu gibi detaylı geri bildirimlere ihtiyaç duyulan durumlarda analitik dereceli puanlama anahtarı tercih edilebilir. Analitik dereceli puanlama anahtarıyla sınıf çalışmalarının güçlü ve zayıf yönleri kolaylıkla saptanabilmekte ve sonraki öğretim sürecinin belirlenmesinde yol gösterici olabilmektedir. Değer biçmeye dönük bir değerlendirme yapılacaksa, performansa ilişkin bütüncül bir puan elde edilmek isteniyorsa bütünsel dereceli puanlama anahtarı kullanmak uygun olabilir. Puanlayıcıların dereceli puanlama anahtarı kullanmasına yönelik bilgilendirilmeleri ve uzmanlaştırılmaları hangi tür için daha kolay olacaksa o dereceli puanlama anahtarı türü tercih edilebilir. Ancak şunu belirtmekte yarar vardır ki analitik dereceli puanlama anahtarı için eğitim vermek bütünsel dereceli puanlama anahtarından çok daha zaman alıcı ve zahmetli olmaktadır. Son olarak, eğer performansın alt bileşenlerinin teoride taşıdıkları önem, farklılık gösteriyorsa analitik dereceli puanlama anahtarı bu farklılıklara göre ağırlıklandırmayla puanlama yapılmasına imkân vermektedir. Bu ise bir ya da daha fazla alt bileşenin diğerlerinden daha kritik olduğu durumlarda oldukça önem taşımaktadır.

NOT: *Bu bölümde yer alan konulara ilişkin daha ayrıntılı bilgiye kaynakçada yer alan (Atılgan, Kan ve Doğan, 2006) ile (Linn ve Gronlund,1995)'ten ulaşabilirsiniz.*

DEĞERLENDİRME SORULARI

1. Bir öğretmen yapacağı sözlü sınavın puanlamasını olabildiğince objektif yapabilmek istiyor. Bu öğretmene aşağıdakilerden hangisini yapmasını önerirdiniz?

A) Puanlamayı cevap anahtarı ile yapmasını.
B) Sözlü sırasında sınıfta sessizlik ve düzeni sağlamasını.
C) Genel izlenimle puanlama yapmasını.
D) Üst düzey bilişsel davranışları ölçecek sorular sormasını.
E) Çok sayıda kısa soru sormasını.

2. Aşağıdakilerden hangisi yazılı sınavların çoktan seçmeli sınavlara göre bir avantajıdır?

A) Puanlaması subjektiftir.
B) Az sayıda, uzun cevap gerektiren sorular içerir.
C) Öğrencilerin sınava olan motivasyonu yüksektir.
D) Şans başarısı çok azdır.
E) Bilgi ve kavrama düzeyindeSki bilişsel davranışların ölçülmesi için idealdir.

3. Çoktan seçmeli sınavlarda seçenek sayısının belirlenmesinde aşağıdakilerden hangisi öncelikle göz önünde bulundurulmalıdır?

A) Sınava girecek öğrenci sayısı
B) Sınavda sorulacak soru sayısı
C) Sınava girecek öğencilerin sınıf düzeyi
D) Sınavın süresi
E) Sınav sorularının güçlük düzeyi

4. "Öğrencilerin öğrenmelerindeki tüm aşamaları kapsamlı bir şekilde gözleyebilmek, öğrenmeleri ile ilgili değerlendirmelere öğrencileri de dahil etmek, öğrencilerin yaratıcı özelliklerini ve üst düzey bilişsel davranışlarını geliştirmek" isteyen bir öğretmenin kullanması gereken <u>en etkili</u> değerlendirme yöntemi aşağıdakilerden hangisi olabilir?

A) Gözlem formları B) Standart testler C) Anketler
D) Portfolyo E) Performans testleri

5. Eğitimde kullanılan aşağıdaki ölçme araçlarından hangisinin kapsam geçerliğinin <u>en düşük</u> olması beklenir?

A) Yazılı yoklamalar
B) Kısa cevaplı testler
C) Boşluk doldurmalı testler
D) Doğru-yanlış testleri
E) Çoktan seçmeli testler

egitim@hotmail.es

6. **Aşağıdakilerden hangisi yazılı yoklamalardan elde edilen puanların daha güvenilir ve geçerli olmasını sağlayacak yollardan biri değildir?**

A) Cevap anahtarıyla puanlama yapılması.
B) Cevap kağıtlarının kime ait olduğunun bilinmemesi.
C) Cevap kağıtlarının farklı zamanlarda birden fazla kez puanlanması.
D) Cevap kağıtlarının farklı puanlayıcılar tarafından birden fazla kez puanlanması.
E) Cevap kağıdı düzgün ve okunaklı olanlara fazla puan verilmesi.

7. **Çoktan seçmeli bir sınavda yer alan bir sorunun doğru cevabı bir başka sorunun kökünde verilmiştir. Buna göre aşağıdakilerden hangisi bu sınav için söylenebilecek en doğru açıklama olacaktır?**

A) Soru kolaylaştığı için sınavın güvenirliği düşer.
B) Soru dikkat ölçer hale geldiği için sınavın geçerliği düşer.
C) Şans başarısı azaldığı için sınavın güvenirliği artar.
D) Sınavın güvenirlik ve geçerliğinde hiç bir değişiklik olmaz.
E) Sınavın hem güvenirliği hem de geçerliği artar.

8. **Aşağıdakilerden hangisi çoktan seçmeli madde yazımında bulunması gereken özelliklerden biri değildir?**

A) Madde kökü yoruma açık olmalıdır.
B) Madde kökünde gereksiz açıklamalardan kaçınılmalıdır.
C) Doğru cevabı bileni şaşırtıcı ifadelerden kaçınılmalıdır.
D) Seçeneklerin uzunlukları birbirine yakın olmalıdır.
E) Testin ilk soruları daha kolay sorulardan seçilmelidir.

9. **Aşağıdakilerden hangisi portfolyoların eğitimde kullanılma amaçlarından birisi değildir?**

A) Öğrencinin eğitim sürecindeki gelişimini izlemek.
B) Öğrencilerin birbirlerine göre durumlarını kıyaslamak.
C) Öğrencinin de değerlendirme sürecinde sorumluluk almasını sağlamak.
D) Öğrencilerin eğitim sürecinde öğrendiklerini ayrıntılı olarak tespit etmek.
E) Öğrencilerin velilerine ve gelecekteki öğretmenlerine yardımcı olacak bilgiler sunmak.

10. **Aşağıdakilerden hangisinde öğrenci de değerlendirme sürecinde aktif rol almaktadır?**

A) Performans testleri B) Yazılı yoklamalar C) Ödev ve projeler
D) Portfolyolar E) Sözlü yoklamalar

KPSS'DE ÇIKMIŞ SORULAR

1. Bir öğretmen, sınav sorularına verilen yanıtları dört, beş gün arayla, puanları ayrı yerlere kaydederek iki kez puanlamış; bu iki puanlamada elde ettiği sonuçları karşılaştırarak aralarındaki tutarlılığa bakmıştır.

Öğretmen bu çabasıyla, aşağıdakilerden hangisini belirlemeye çalışmaktadır? (2001)

A) Soruların bilimsel doğruluğunu
B) Soruların güçlük derecesini
C) Puanlama anahtarının doğruluğunu
D) Sınavın kapsam geçerliğini
E) Puanlamadaki nesnellik derecesini

2. Bir Türkçe öğretmeni öğrencilerin, yazı yazarken yazım kurallarına ne ölçüde uyduklarını saptamak istemektedir.

Öğretmenin, bu amaçla yapacağı sınavda aşağıdaki soru tiplerinin hangisinden yararlanması en uygun olur? (2001)

A) Doğru-yanlış
B) Uzun yanıtlı (klasik)
C) Çoktan seçmeli
D) Eşleştirme
E) Tamamlama (boşluk doldurma)

3. Okul eğitiminde başarının ölçüsünü "not" olarak görenler, karnesinde yüksek notlar bulunan öğrencileri daha iyi öğrenmiş kabul ederler. Ancak eğitim uygulamaları dikkate alındığında, yüksek not alanların iyi öğrenenlerden çok, iyi ezberleyenler olduğu görülmektedir.

Bunları söyleyen bir kişinin aşağıdakilerden hangisini kanıtlaması, bu görüş için güçlü bir destek oluşturur? (2001)

A) Sınavlarda daha çok hatırlama düzeyinde sorular sorulduğunu
B) Eğitimde başarıya ilişkin kararların çoktan seçmeli testlere dayandırıldığını
C) Öğrenci notlarının öğretmenin kanılarından etkilendiğini
D) Ders konularının öğrenci düzeyinin üzerinde olduğunu
E) Öğretimin öğrenci merkezli değil, öğretmen merkezli olduğunu

egitim@hotmail.es

4. Bir öğretmen tamamlama (boşluk doldurma) tipinde sorulardan oluşan bir sınav hazırlamıştır. Bu soruları yazarken ders kitabından cümleler seçmiş ve her cümleden bir ya da birkaç sözcüğü çıkararak soruları oluşturmuştur.

Öğretmenin kullandığı bu yöntemin en önemli sakıncası aşağıdakilerden hangisidir? (2001)

A) Öğrencilerin öğrenme güdülerini zayıflatması
B) Puanlamadan kaynaklanabilecek hataları artırması
C) Sınavın kapsam geçerliğini düşürmesi
D) Sınavdaki soruların önceden tahmin edilmesini kolaylaştırması
E) Öğrencileri hatırlama düzeyinde öğrenmelerle yetinmeye yöneltmesi

5. Bir öğretmen, çoktan seçmeli bir test için soru yazarken, öğrencilerin sıklıkla düştükleri hataları çeldirici olarak kullanmaktadır.

Bu durum, öğretmene en çok hangi konuda yardımcı olur? (2001)

A) Test için verilecek süreyi kısaltmak
B) Sorunun anlaşılmasını kolaylaştırmak
C) Puanlamayı kolaylaştırmak
D) Yanlış öğrenmeleri belirlemek
E) Testin kullanışlığını artırmak

6. Eğitimde kullanılacak sınavların bir plana göre hazırlanması gerekir.

"Sınav planı" **adı verilen bu hazırlıkta ilk olarak aşağıdakilerden hangisi belirlenir? (2001)**

A) Sınavda kapsanacak konular
B) Sınavda yer alacak soruların güçlük düzeyi
C) Sınavın yapılış amacı
D) Sınavdaki yanıtları puanlama yöntemi
E) Sınavda kullanılacak soru tipi

7. **Bir ölçme aracı veya yönteminin objektif olması ne demektir? (2002)**

A) Bağımsız puanlayıcılar kullanıldığında bunların aynı sonuca ulaşması
B) Maddelerin (soru) uygun güçlük düzeylerinden seçilmesi
C) Uygulama koşullarının değişmemesi (standart olması)
D) Puanlamanın yalın ve kolay olması
E) Cevaplama sırasında kopya çekilmesine olanak vermemesi

8. Son yıllarda, klasik ölçme ve değerlendirme yöntemlerinin yanı sıra, araştırma yapma becerileri, problem çözme becerileri, sosyal beceriler, duyuşsal alana yönelik gelişmeler, yaratıcılık gibi özelliklerde her ders ve konu alanının içerisinde ele alınmakta; öğrencilerin öğrenme sürecinde gösterdikleri performans ve ortaya koydukları ürün, öğrenci ile öğretmen tarafından birlikte değerlendirilmektedir.

Bu yaklaşımın amacı aşağıdakilerden hangisi olamaz? (2002)

A) Üst düzey becerilerin ölçülmesine önem verme
B) Kişilerde toplumsal becerilerin gelişmesini destekleme
C) Öğretim etkinliklerini daha çok öğretmen merkezli hale getirme
D) Öğretmenleri çeşitli yöntemlerden yararlanmaya yöneltme
E) Günlük hayattaki problemlerle başa çıkma becerilerini geliştirme

9. "Ürün örneklerine ve bireysel gelişime bakarak (portfolyo tipi) değerlendirme" **yaklaşımının öğretmene sağladığı en önemli yarar aşağıdakilerden hangisidir? (2002)**

A) Sınıflar arasındaki öğrenci başarısını karşılaştırması için veri sağlaması
B) Öğrencinin herhangi bir zamanda, neyi ne kadar öğrendiğinin kolayca görülebilmesini sağlaması
C) Öğrenci başarısını diğer öğrencilerin başarılarıyla karşılaştırma kolaylığı sağlaması
D) Kullanılan öğretim materyal ve tekniklerinin etkililik derecesine ışık tutması
E) Öğrencinin gelişimindeki aşamaları bir bütün olarak izlemeye olanak tanıması

10. Okul öğrenmelerinin ölçülmesinde çoğunlukla kâğıt-kalem testleri kullanılır.

Aşağıdakilerden hangisi bu testlerin sonuçlarına karışabilen hata kaynaklarından biri değildir? (2003)

A) Puanlama işlemini yapan kişinin yansızlık derecesi
B) Ölçme aracının ölçülecek davranışları kapsama derecesi
C) Ölçme yönteminin, ölçülen davranışlara uygunluğu
D) Öğrencilerin testteki soruları yanıtlamaya isteklilik dereceleri
E) Öğrencilerin yoklanacak davranışları öğrenme düzeyleri

11. Bir yüksek lisans giriş sınavına ait cevap kâğıtlarını, ilgili anabilim dalında bulunan dört öğretim elemanı okumuştur. Daha sonra, bu öğretim elemanlarının verdikleri puanların tutarlılığına bakılmıştır.

Bu uygulama, aşağıdakilerin hangisi için uygun bir örnek oluşturur? (2003)

A) Anahtara göre puanlama yapma
B) Puanlama anahtarı hazırlama
C) Düzeltme puanı kullanma
D) Puanlama güvenirliğini saptama
E) Değerlendirme için ölçüt belirleme

egitim@hotmail.es

12. Okul öğrenmelerinin ölçülmesinde öğretmenler, "tamamlama" veya "boşluk doldurma" tipi sorular hazırlarken genellikle, kitaptaki bir ifadeyi aynen almakta, bu ifadenin içinden bir sözcüğü ya da sözcük grubunu çıkarmakta ve öğrenciden, çıkarılan bu sözcüğü ya da sözcük grubunu bulmasını istemektedir.

Bu tür soruları kullanmanın en önemli sakıncası aşağıdakilerden hangisidir? (2003)

A) Puanlamanın çok zaman alması
B) Kesin bir yanıt bulmanın zor olması
C) Kopya çekmeyi kolaylaştırması
D) Öğrencileri ezbere yöneltmesi
E) Öğrenme farklarını gösterme gücünün az olması

13. Bir lise öğretmeni, daha önce kullandığı dört seçenekli sorularını, seçenek sayısını artırarak beş seçenekli duruma getirmiştir.

Öğretmenin, bu değişikliği aşağıdakilerden hangisini gerçekleştirmek amacıyla yaptığı söylenebilir? (2003)

A) Şans başarısını azaltmak
B) Testin kapsamını zenginleştirmek
C) Puanlamayı kolaylaştırmak
D) Çeldiricilerin aynı derecede çalışmasını sağlamak
E) Soruları güçleştirmek

14. Portfolyo, öğrencinin geçirdiği aşamaları gösterecek biçimde seçilmiş örneklerden oluşan ve eğitim süreci içerisinde öğrencinin gelişimini kendisinin, velisinin ve öğretmenlerinin izleyebilmesine olanak sağlayan bir ürünler bütünüdür.

Buna göre, aşağıdakilerden hangisi eğitim sürecinde portfolyoların kullanım amaçlarından biri değildir? (2003)

A) Öğrencinin gelişimini kanıtlarıyla ve daha sağlıklı olarak izlemek
B) Öğrencinin yetenekleri ve ilgi alanları hakkında bilgi edinmek
C) Öğrencinin öğretim programındaki kritik davranışlarla ilgili eksik öğrenmelerini belirlemek
D) Derslerin öğretim programındaki sınırlarına bağlı kalmadan, öğrencinin çok yönlü gelişimi hakkında bilgi vermek
E) Öğrencinin gelecekteki öğretmenlerine onun hakkında bilgi sunmak

15. **Test planında yapılması gereken bazı işlemler, öncelik sırasına göre aşağıda verilmiştir.**

1. ——
2. Yoklanacak davranışların belirlenmesi
3. Belirtke tablosunun hazırlanması
4. ——
5. Test maddelerinin (soruların) yazılması
6. Maddelerin (soruların) test düzenine konması

Bu sıralamada, 1 ve 4 numaralı satırlara aşağıdakilerin hangisinde verilen işlemler getirilmelidir? (2003)

A) 1. Sınav amacının belirlenmesi
4. Kullanılacak madde (soru) türünün belirlenmesi

B) 1. Sınav amacının belirlenmesi
4. Cevapları puanlama şeklinin belirlenmesi

C) 1. Test yönergesinin yazılması
4. Cevaplama süresinin belirlenmesi

D) 1. Sınavda kapsanacak konuların belirlenmesi
4. Cevapların nereye, nasıl yazılacağının belirlenmesi

E) 1. Test yönergesinin yazılması
4. Sınavda kapsanacak konuların belirlenmesi

16. **Aşağıdaki davranışlardan hangisi, yazılı yoklama türü sınavlarla daha geçerli bir biçimde ölçülebilir? (2004)**

A) Görüşünü, nedenleriyle birlikte açıklama
B) Bir kavramı veya ilkeyi tanımlama
C) Bir bütünü oluşturan öğeler arasındaki ilişkiyi belirtme
D) Bilgilerinden yararlanarak yeni bir problemi çözme
E) Olgu ve olayları verilen bir ölçüte göre gruplama

17. - 18. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

17.

I. Sentez basamağındaki davranışların yoklanmasına elverişli olması
II. Puanlamada objektifliğin daha fazla olması
III. Kapsam geçerliğinin daha yüksek olması
IV. Öğrencilerin cevaplarını istediği biçimde yazma olanağına sahip olması
V. İlköğretimin ilk sınıflarına daha uygun olması

Yukarıdakilerden hangileri çoktan seçmeli sorulardan oluşan testlerin yazılı yoklamalara olan üstünlükleri arasındadır? (2004)

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

egitim@hotmail.es

18. **Yukarıdakilerden hangileri kısa cevaplı yazılı yoklamaların uzun cevaplı yazılı yoklamalara olan üstünlükleri arasındadır? (2004)**

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

19. Öğrencilerin aktif olduğu öğretme-öğrenme yöntemlerinde, öğretmen, öğrencileri değerlendirirken, sınav sonuçlarıyla birlikte öğrencilerin sınıf içi etkinliklere katılma düzeylerini; bilimsel tutum ve davranışlarını; gözlem, araştırma ve inceleme yapabilmelerini; sahip oldukları ve sergiledikleri fikir zenginliklerini; sorumluluk alma düzeylerini; ekiple çalışabilmelerini; edindikleri bilgi ve bulguları paylaşabilme vb. özelliklerini de dikkate alarak başarıları hakkında karar verir.

Aşağıdakilerden hangisi, öğretmenin bu bilgileri elde etmesinde daha kapsamlı katkı sağlar? (2004)

A) Yazılı kâğıt kalem testleri
B) Performans testleri
C) Gözlem formları
D) Günlükler
E) Portfolyolar

20. **Aşağıda soru türü ve soru sayısı verilen ölçme araçlarından hangisinin diğerlerine göre daha güvenilir ölçme sonuçları vermesi beklenir? (2005)**

A) Kısa cevaplı 10 soru
B) Boşluk tamamlama türü 25 soru
C) Doğru-yanlış türü 50 soru
D) Üçer seçenekli 25 soru
E) Beşer seçenekli 50 soru

21. **Belli bir öğrenme süreci içerisinde öğrencinin geçirdiği aşamaları tüm yönleriyle izlemek ve sürece katılan herkesten (öğretmen, öğrencinin kendisi, arkadaşları ve velisi) bu gelişimsel süreç hakkında bilgi almak isteyen bir öğretmen, hangi değerlendirme aracını tercih etmelidir? (2005)**

A) Standart test
B) Rubrik
C) Anket
D) Portfolyo
E) Mülakat

22. **Aşağıdakilerden hangisi yazılı yoklamalarda puanlama güvenirliğini artırmaya yönelik önlemlerden biri değildir? (2006)**

A) Puanlamayı birden fazla puanlayıcının yapması ve verilen puanların ortalamasının alınması
B) Bir soruya öğrencilerin tümünün verdiği cevaplar puanlandıktan sonra diğer soruya geçilmesi
C) Cevabın değişik bölümlerine verilecek puanların ayrı ayrı belirlenmesi
D) Öğrencinin sınıf-içi etkileşimiyle ilgili puanının sınav puanlarına katılması
E) Puanlama için ayrıntılı yönergelerin bulunduğu cevap anahtarının kullanılması

23. **Çoktan seçmeli bir test maddesini cevaplayan bir kişinin çeldiricileri eleyerek doğru cevaba ulaşmasında, aşağıdakilerden hangisi belirleyici olmalıdır? (2006)**

A) İfadelerin, kök ifadesiyle tutarlılığı
B) Seçeneklerin soyutluk-somutluk derecesi
C) Yoklanan davranışın öğrenilmiş olması
D) Seçeneklerin birbirini içerip içermemesi
E) İfadelerin kısalık ve uzunluğu

24. **Test maddesinde, ders kitabından aynen alınan ifadelerin kullanılması eğitimsel doğurguları açısından nasıl bir etki yaratır? (2006)**

A) Maddeyle yoklanan davranış hatırlama düzeyinde kalır.
B) Kitabı okumamış öğrenci maddeyi kesinlikle cevaplayamaz.
C) Maddenin anahtarlanması kolaylaşır.
D) Maddenin okunup anlaşılması zorlaşır.
E) Maddede gereksiz ipuçları bulunmaz.

25. Rubrikler, süreç ya da ürün değerlendirmede, çeşitli becerileri tanımlayan dereceli puanlama anahtarlarıdır.

Aşağıdakilerden hangisi rubrikler oluşturulurken dikkat edilmesi gereken özelliklerden biri değildir? (2006)

A) Öğrencilerin bilişsel, duyuşsal ve psiko-motor gelişim düzeylerini belirlemek
B) Değerlendirme sırasında öğrencilerin ilgi ve ihtiyaçlarını dikkate almak
C) Her konu ve alan için standart bir form geliştirmek
D) Geliştirirken alanda birden fazla uzman kişinin görüşünü almak
E) Öğrencilerde kazandırılması istenilen özellikleri belirlemek

egitim@hotmail.es

26. Öğrencinin belirli bir süreçteki gelişimini incelemek amacıyla geliştirilen bireysel gelişim dosyalarının değerlendirilmesine "portfolyo değerlendirme" adı verilir.

Buna göre, portfolyo değerlendirme sürecinde öğrencilerden aşağıdakilerin hangisini yapmaları beklenmez? (2006)

A) Süreç içerisindeki gelişimlerini yansıtacak günlükler oluşturmaları
B) Dersle ilgili beklentilerini ortaya koydukları bir değerlendirme raporu hazırlamaları
C) Kendi kendilerini değerlendirdikleri öz değerlendirme formunu doldurmaları
D) Arkadaşlarıyla ilgili duygu ve düşüncelerini yansıttıkları akran değerlendirme formunu doldurmaları
E) Derste işlenen konularla ilgili sorular yazmaları

27. **Çoktan seçmeli bir testteki sorularda, çeldiricilerin doğru cevaba anlamca yaklaştırılması bu testin hangi özelliğini doğrudan etkiler? (2007)**

A) Güvenirliğini
B) Geçerliğini
C) Güçlüğünü
D) Kullanışlılığını
E) Objektifliğini

28. Yazılı yoklamalardaki sorularda, öğrencilerin cevaplarını düşünerek bulmaları ve buldukları bilgileri organize ederek yazmaları gerekir. Bu tür sorularla üst düzeydeki zihinsel beceriler ölçülebilir.

Aşağıdaki sorulardan hangisi, yazılı yoklamaların bu avantajlarını ortaya koyan uygun bir örnek değildir? (2007)

A) Enflasyonu düşürmek için ne tür önlemler alınabilir? Örnek vererek irdeleyiniz.
B) Avrupa Birliği'ne katılma konusundaki görüşleriniz nelerdir? Bu şekilde düşünmenizin nedenlerini açıklayınız.
C) İzlediğiniz en iyi sinema filmi hangisidir? Bu filmi neden beğendiğinizi gerekçeleriyle açıklayınız.
D) Başarı ile yetenek arasındaki ilişkiyi bir örnekle açıklayınız. Sizce insan yaşamında hangisi daha ön plandadır?
E) Ölçme nasıl tanımlanır? Eğitimde yapılan ölçmeler kaç gruba ayrılır?

29. Sosyal bilgiler sınavında sorulan bir soru şöyledir:

> ***Türkiyenin pek çok sınır komşusu vardır. En uzun sınırı paylaştığımız sınır komşumuz aşağıdakilerden hangisidir?***
>
> *A) Bulgaristan'la sınırımız en uzundur.*
> *B) Irak'la sınırımız en uzundur.*
> *C) İran'la sınırımız en uzundur.*
> *D) Suriye'yle sınırımız en uzundur.*
> *E) Yunanistan'la sınırımız en uzundur.*

Yukarıdaki sorunun kusurları arasında aşağıdakilerden hangisi yoktur? (2007)

A) Kökte yazım hatasının olması
B) Kökte gereksiz sözcük kullanılması
C) Kök ile seçenekler arasında ifade uyumsuzluğu olması
D) Seçeneklerde aynı sözcüklerin tekrarlanması
E) Seçeneklerin uygun olmayan bir sırada verilmesi

30. **Öğrencilerinin öğretim süreci boyunca gelişimlerini görmek ve değerlendirmeye katılımlarını sağlamak isteyen bir öğretmen için en uygun ölçme ve değerlendirme yaklaşımı aşağıdakilerden hangisidir? (2007)**

A) Performans değerlendirme
B) Akran değerlendirme
C) Mutlak değerlendirme
D) Tümel (portfolyo) değerlendirme
E) Bağıl değerlendirme

31. Bir ölçme aracının özelliklerinden bazıları şunlardır:

1. Bir işin yapılması sırasında nelerin, hangi sırada ve nasıl yapılacağını göstermesi nedeniyle öğretim amaçlı da kullanılabilir.
2. Gözlenecek davranışların varlığını veya yokluğunu gösterme olanağı tanır.
3. Performansı oluşturan davranışlardan ne kadarına sahip olunduğunu belirlemeyi kolaylaştırır.
4. Performansın en önemli ve gözlenebilir yanlarını kapsar.
5. Performansı oluşturan davranışların her birine ne düzeyde sahip olunduğu hakkında bilgi veremez.

Bu ölçme aracı aşağıdakilerden hangisidir? (2007)

A) Dereceleme ölçeği
B) Kontrol listesi
C) Tutum ölçeği
D) Anket
E) Puanlama yönergesi (rubrik)

egitim@hotmail.es

32. Bir ölçme uzmanı Fen ve Teknoloji dersi için hazırlanan başarı testini incelemiş ve ortak köke dayalı soruların 5-6 civarında olduğunu saptamıştır. Bu durumu eleştiren uzman, bir köke dayalı olarak sorulabilecek soru sayısının üçü geçmemesini önermiştir.

Ölçme uzmanının bu eleştirisinin <u>en önemli</u> gerekçesi aşağıdakilerden hangisi olabilir? (2008)

A) Öğrencilere aynı kökle ilgili çok sayıda soru sormak motivasyonu düşürür.
B) Sorular tek bir test yaprağına sığmayacağı için test düzeni bozulabilir.
C) Benzer davranışı ölçen soru sayısının artması güvenirliği düşürür.
D) Soruların birbirinden bağımsız cevaplanabilme kuralı bozulabilir.
E) Okuduğunu anlama değişkeninin etkisi artarak geçerliği düşürür.

33. Bir sınav türünün özellikleri şunlardır:

1. Kısa sürede çok soru sorulmasına olanak tanır.
2. Puanlaması kolay ve nispeten nesneldir.
3. Şans başarısı yoktur.

Bu özellikleri taşıyan sınav türü aşağıdakilerden hangisidir? (2008)

A) Doğru – yanlış
B) Boşluk doldurma
C) Çoktan seçmeli
D) Açık uçlu
E) Sözlü

34. **Aşağıdakilerden hangisi performans değerlendirmesini gerektiren bir durumdur? (2008)**

A) Bir elementin yerini periyodik tabloda gösterme
B) Mevsim normallerini gösteren bir tablodan yararlanarak belli bir tarihteki hava sıcaklığını tahmin etme
C) Verilenler arasından bir romanın ana fikrini yansıtan tümceyi seçme
D) Verilen molekül formüllerinin açılımlarını yazma
E) Işığın çukur ve tümsek aynada yansımasını gösterecek düzenek geliştirme

35. Öğrencilerin, belli bir konuda kendileriyle ilgili bilgi, beceri, tutum ve davranışlarını yorumlamalarına ---- ; araç - gereç kullanma, deney yapma ya da proje hazırlama vb. becerilerinin, puanlama kriterleri kullanılarak değerlendirmesi sürecine ---- denir.

Bu parçadaki boşluklara, aşağıdakilerin hangisindeki sözcükler sırasıyla getirilmelidir? (2008)

A) öz değerlendirme - rubrik (puanlama yönergesi)
B) performans değerlendirme - grid
C) öz değerlendirme - performans değerlendirme
D) portfolyo (öğrenci ürün dosyası) - dallanmış ağaç
E) performans değerlendirme - akran değerlendirme

36. **Aşağıdakilerden hangisi performans görevlerine dayalı ölçmelerin geleneksel kâğıt - kalem testlerine göre üstün özelliklerinden biridir? (2008)**

A) Öğrencilerin gerçek hayata yakın işlemler gerçekleştirmesi
B) Puanlamanın daha objektif olması
C) Uygulamanın daha az zaman alması
D) Velilerin öğrenci çalışmalarına yardım etmesi
E) Öğrencilerin hazırbulunuşluk düzeyine uygun olması

37. **Dönem boyunca az sayıda sınav yapıp, öğrencilerin öz değerlendirme ve akran değerlendirme formlarını, performanslarını ve hazırladıkları ürün dosyalarını değerlendirmeye daha çok zaman ayıran bir öğretmen bu uygulamasıyla aşağıdakilerden hangisine <u>en az</u> hizmet eder? (2008)**

A) Değerlendirmenin öğrenmeye katkısını artırmaya
B) Özdenetim becerilerini geliştirmeye
C) Değerlendirme becerilerini kazandırmaya
D) Bireysel gelişim düzeylerini belirlemeye
E) Öğrencilerin başarılarını karşılaştırmaya

38. Fen ve teknoloji öğretmeni, öğrencilerin sınıf içi başarılarını yazılı ve sözlü yoklamalarla değerlendirmekte, öğrencilerin, hazırladıkları ürün dosyalarıyla değerlendirilmesine ise daha az yer vermektedir.

Bu öğretmen aşağıdakilerden hangisine <u>daha çok</u> önem vermektedir? (2009)

A) Öğrencinin kendi öğrenmeleriyle ilgili güçlü ve zayıf yönlerini analiz etmesini sağlamaya
B) Öğrencilerin bireysel gelişim aşamalarını izlemeye
C) Öğrencileri başarılarına göre sıralamaya
D) Öğrencide sorumluluk bilincini geliştirmeye
E) Öğrencilere günlük hayata dair beceriler kazandırmaya

39. Sınıf içi testlerin kullanım amacı, temelde eğitim öğretim etkinliklerinin niteliğini artırmaktır. Test sonuçları öğretmene, sınıf içerisinde yapmış olduğu etkinliklerin ne ölçüde başarılı olduğuna dair geri bildirimde bulunmakta, aynı şekilde öğrencilere de sınıf içi etkinlikler sonucunda neyi ne kadar öğrendikleri konusunda bilgi vermektedir.

Test sonuçlarına göre böyle bir geri bildirim alan öğretmenden beklenen aşağıdakilerden hangisi <u>olamaz</u>? (2009)

A) Eksik öğrenmeler varsa bunları tamamlaması
B) Kullandığı öğretim yöntem, teknik ve materyallerinin etkililiğini gözden geçirmesi
C) Öğretim etkinlikleri boyunca öğrencilerin gelişimlerini izlemesi
D) Öğrenciler arasındaki öğrenme güçlüklerini belirlemesi
E) Öğrencileri yetenek düzeylerine göre gruplaması

egitim@hotmail.es

40. Mehmet Öğretmen hazırladığı bir Türkçe sorusunda bir sözcüğün cümlede hangi anlamda kullanıldığını sormuştur.

Mehmet Öğretmen, bu sorusuyla aşağıdaki süreçlerden hangisini ölçmek istemektedir? (2009)

A) Metin içerisinde açıkça verilmiş bilgiden doğrudan çıkarım yapma
B) Metinde kullanılan cümle yapısını irdeleme
C) Bilgi ve fikirleri yorumlama ve birleştirme
D) Metnin kapsamını ve dilini gözden geçirme ve değerlendirme
E) İlişki kurma ve karar verip çözümleyebilme

41. Fen ve teknoloji dersine ait bir kazanım şöyledir: "Farklı yoğunluğa sahip sıvıların cisimlere uyguladığı kaldırma kuvvetini karşılaştırır ve sonuçları yorumlar."

Bu kazanımın ne ölçüde edinildiğini belirlemek isteyen bir öğretmen aşağıdaki uygulamalardan hangisini gerçekleştirmelidir? (2009)

A) Çoktan seçmeli test
B) Yapılandırılmış grid
C) Kavram haritası
D) Performans değerlendirme
E) Sözlü sınav

42. Öğrencilerin küresel ısınmanın getirdiği tehlikelerle ilgili ödevlerini inceleyen bir öğretmenin, ölçme sonuçlarına karışacak ölçmeciden kaynaklı hata miktarıyla ilgili endişeleri vardır.

Puanlamasına karışacak olan öznelliği azaltmak isteyen bu öğretmene aşağıdakilerden hangisini yapması önerilir? (2010)

A) Dereceli puanlama anahtarı (rubrik) kullanması
B) Ödevin kapsamını genişletmesi
C) Ödev konularını seçmeli olarak tasarlaması
D) Ödev teslim tarihini öğrencilerle birlikte kararlaştırması
E) Öğrencilerin düzeyine uygun ödev konuları vermesi

43. İlköğretim 3. sınıf hayat bilgisi dersine ilişkin bazı kazanımlar şunlardır:

I. Cumhuriyetimizi korumak, güçlendirmek ve geliştirmek için yapılması gerekenler hakkında fikirler üretir.
II. Geçmişten günümüze iletişim teknolojilerinde meydana gelen değişimi araştırır.
III. Çeşitli ortamlarda liderlerin gruplar ve toplumlar üzerindeki etkilerini araştırır.

Aşağıdakilerden hangisi bu kazanımlara ulaşılma derecesini belirlemede kullanılabilecek en uygun yöntemdir? (2010)

A) Sözlü sunum
B) Yazılı yoklama
C) Çoktan seçmeli test
D) Öz değerlendirme
E) Performans görevi

44. Bilişsel davranışların ölçülmesinde, ölçmenin geçerliğini artırmak için aynı sürede uygulanan uzun cevaplı yazılı sınavlara göre, kısa cevaplı maddelerden oluşan testler tercih edilir.

Bu durumun başlıca nedeni aşağıdakilerden hangisidir? (2010)

A) Kısa cevaplı maddelerin kolayca cevaplanabilmesi
B) Kısa cevaplı maddelere verilecek cevabın sınırlandırılmış olması
C) Kısa cevaplı maddelerin çoktan seçmeli test maddesi köküne benzemesi
D) Daha çok kısa cevaplı maddeyle kapsamın daha iyi örneklenebilmesi
E) Kısa cevaplı maddelerle her düzeydeki davranışların yoklanabilmesi

45. **Öğretim sürecinde öz değerlendirme ve akran değerlendirmeye ağırlık veren bir öğretmenin bu uygulamasının öncelikli nedeni aşağıdakilerden hangisi olabilir? (2010)**

A) Değerlendirme sürecinde sorumluluğu öğrencilere vererek iş yükünü hafifletmek istemesi
B) Öğrencilerin kendi öğrenmelerinin sorumluluklarını üstlenmesini planlaması
C) Değerlendirme sonunda verilen kararların daha geçerli olmasına çalışması
D) Verdiği notların öğrenciler tarafından beğenilmemesinden rahatsızlık duyması
E) Değerlendirme sürecinin daha kısa sürede tamamlanması

46. Bir ilköğretim öğretmeni öğrencilerinden yaptıkları ürünlerini, araştırma raporlarını, ünite testlerini ve etkinlik raporlarını bir dosyada toplamalarını ve sene sonunda kendisine teslim etmelerini istemiştir.

Aşağıdakilerden hangisi öğretmenin bu uygulamasının gerekçelerinden biri olamaz? (2010)

A) Öğrencilerin kendi gelişimlerini takip etmelerini sağlamak
B) Velileri öğrencilerde meydana gelen değişim ve gelişimler açısından bilgilendirmek
C) En iyi dosyayı seçerek öğrenciler arasında rekabet ortamı oluşturmak
D) Öğretim sürecini de değerlendirme sürecine katmak
E) Öğrencilerin farklı özelliklerini ortaya çıkarabilmek

egitim@hotmail.es

47. Dilek Öğretmen, öğrencilerine, yapacağı yazılı sınavdaki 15 sorudan istedikleri 10 soruyu seçerek cevaplandırabileceklerini söyler.

Dilek Öğretmen'in bu uygulaması,

I. öğrenci motivasyonu
II. puanların kıyaslanabilirliği
III. kapsam geçerliği

özelliklerinden hangilerini olumsuz etkilemiştir? (2010-2)

A) Yalnız I B) Yalnız II C) I ve II
D) I ve III E) II ve III

48. Bir ölçme değerlendirme uzmanı, soru bankasına kaydedilmek üzere aşağıdaki dört madde çiftinden birer tanesini tercih edecektir.

1A	Hava, maddenin üç hâlinden biri olan ….. hâlinde bulunur. *(İstenen cevap: gaz)*
1B	Hava bir ….. . *(İstenen cevap: gazdır)*
2A	Yağ suda çözünmez. Doğru ( ) Yanlış ( ) *(İstenen cevap: Doğru)*
2B	Yağ suda çözünür. Doğru ( ) Yanlış ( ) *(İstenen cevap: Yanlış)*
3A	Ağaçların yeşil yaprakları vasıtasıyla güneş enerjisini, kendi ihtiyaçları olan yaşam enerjisine dönüştürmelerine ne ad verilir? a. Fotosentez b. Fototropizm c. Kemotropizm *(İstenen cevap: a)*
3B	Ağaçların yeşil yaprakları vasıtasıyla güneş enerjisini sentezleyerek, kendi ihtiyaçları olan yaşam enerjisini elde etmelerine ne ad verilir? a. Fotosentez b. Fototropizm c. Kemotropizm *(İstenen cevap: a)*
4A	İstanbul ile Ankara'yı kıyaslayınız.
4B	İstanbul ile Ankara'yı nüfus, iklim, ekonomik ve tarihi özellikleri bakımından kıyaslayınız.

Ölçme değerlendirme uzmanının, madde yazım ilkeleri doğrultusunda tercih etmesi beklenen maddeler aşağıdakilerin hangisinde doğru olarak yerilmiştir? (2010-2)

A) 1A, 2B, 3A, 4B
B) 1A, 2A, 3B, 4B
C) 1B, 2B, 3A, 4A
D) 1B, 2A, 3B, 4B
E) 1A, 2B, 3A, 4A

49.

I. Objektif sınav türleriyle uygun biçimde ölçülemeyecek bilgi ve becerilerin ölçülmesi için kullanılmalıdır.
II. Puanlama anahtarı hazırlanmalıdır.
III. Puanlama aşamasında bir öğrencinin bütün cevapları puanlandıktan sonra diğer öğrencinin kâğıdının puanlanmasına geçilmelidir.

Uzun cevaplı yazılı yoklamalarla ilgili olarak yukarıdakilerden hangileri doğrudur? (2010-2)

A) Yalnız I B) Yalnız II C) Yalnız III
D) I ve II E) I, II ve III

50. **Aşağıdaki öğrenme çıktılarından hangisi, bilişsel alan sınıflamasında diğerlerine göre daha üstte yer alır? (2010-2)**

A) Akıcı ve bütünlüğü olan bir konuşma metni hazırlama
B) Bilimsel bir önermeyi test edecek bir deney tasarlama
C) Bir problemin çözümü için gerekli bilgileri ayırt etme
D) Bir teknik raporun güçlü ve zayıf yönlerini belirleme
E) Verileri grafik üzerinde ifade etme

51. Eğitimde öğrenme çıktılarının değerlendirildiği geleneksel uygulamalarla, süreç becerilerinin değerlendirildiği tamamlayıcı ölçme ve değerlendirme uygulamaları birlikte kullanılmalıdır.

Tamamlayıcı ölçme ve değerlendirme uygulamalarından olan portfolyo değerlendirme sürecinde aşağıdaki sorulardan hangisinin cevabı en az öneme sahiptir? (2010-2)

A) Portfolyo hazırlanırken hangi beceriler gelişti?
B) Portfolyoya konan sınav sonuçlarına göre öğrencinin başarı sırası nedir?
C) Portfolyoya konan bir çalışma nasıl yapıldı?
D) Portfolyoda bulunan çalışmalarla ilgili değerlendirmeler nasıldır?
E) Portfolyoda bulunan çalışmalar neden seçildi ve portfolyoya konuldu?

52. İlköğretim düzeyinde okutulan medya okuryazarlığı dersi öğretim programında yer alan bir kazanım şöyledir: *"İnternette bilgiye erişim, haber okuma, sohbet, e-posta, uzaktan eğitim gibi etkinlikleri uygulamalı olarak gerçekleştirir."*

Bu kazanımın gerçekleştirilme derecesini belirlemek amacıyla kullanılabilecek en uygun değerlendirme yöntemi aşağıdakilerden hangisidir? (2010-2)

A) Kavram haritası
B) Performans görevi
C) Sözlü sınav
D) Ven diyagramı
E) Yapılandırılmış grid

egitim@hotmail.es

53. Özlem Öğretmen, bir ölçme ve değerlendirme uzmanından uygulayacağı sınavda kullanacağı madde türleriyle ilgili görüş istemiştir. Uzman, sınavda hem çoktan seçmeli, doğru-yanlış veya eşleştirme gibi seçimlik maddeler hem de kısa cevaplı ve açık uçlu maddeler kullanmasını önermiştir.

Ölçme ve değerlendirme uzmanının birden çok madde türünün kullanılmasını önermesinin gerekçesi aşağıdakilerden hangisi olamaz? (2011)

A) Farklı becerilerin ölçülmesinde farklı madde türlerinin daha elverişli olması
B) Farklı öğrencilerin, farklı madde türlerini daha rahat cevaplandırması
C) Sınavın güçlük düzeyinin öğrenci seviyesinde tutulabilmesi
D) Öğrencilerin birden çok beceri türünde yetkinlik kazanmalarının özendirilmesi
E) Ölçülmesi hedeflenen kazanımlar ile maddeler arasındaki ilişkinin güçlendirilmesi

54. İlköğretim okullarında not verme amacıyla sözlü sınav uygulamasına son verilmiştir. Ancak öğretmenler, öğrencilerin dikkatini uyanık tutmak, sözlü anlatım becerilerini geliştirmek, öğrenme düzeylerini hızlıca kontrol etmek, öğrencilere önemli noktaların hangileri olduğunu işaret etmek ve sınıf içi etkinliklere yön vermek amacıyla sınıf içinde sözlü olarak soru sormaktadır.

Bu türden sözlü soru-cevap etkinliklerinin uygulanmasında aşağıdakilerden hangisi önerilmez? (2011)

A) "Bu sonuca nasıl vardın?", "Neden?" gibi takip soruları kullanmak
B) Bazen cevap vermeye istekli olmayan öğrencilere soru sormak
C) Cevabı sadece "Evet" ya da "Hayır" olan sorulardan kaçınmak
D) Sınıfın hep birlikte cevaplayacağı sorular sormak
E) Soru sorduktan sonra öğrenciye düşünmesi için süre tanımak

55. Standart testleri eleştiren bir öğretmen aşağıdaki açıklamayı yapmıştır:

"Çoktan seçmeli maddelerden oluşan standart testlerde çoğu zaman gerçek hayatta karşılıkları olmayan, sınav için üretilmiş problemler kullanılmaktadır. Ancak biz öğrencilerimizin gerçek yaşamda karşılaşacakları türden problemler çözmelerini isteriz. Üstelik bu standart sınavlarda öğrenciler problemleri çözerken hiçbir araç, kaynak ve materyal kullanmamaktadır. Oysa gerçek yaşam problemlerinin çözümünde farklı araçlar ve kaynaklar kullanılır. Ayrıca, çoktan seçmeli madde türü genellikle alt düzey becerileri öne çıkarmaktadır. Gerçek yaşam problemlerinin çözümünde ise temel becerilerin yanı sıra yaratıcılık ve üst düzey düşünme becerileri de kullanılmaktadır. Bu nedenlerle, okullarımızda farklı bir değerlendirme yaklaşımının kullanılmasını destekliyorum."

Bu açıklamayı yapan öğretmenin, okullarda hangi değerlendirme yaklaşımını desteklemesi beklenir? (2011)

A) Biçimlendirici B) Doğrudan C) Geleneksel D) Mutlak E) Tamamlayıcı

56. VE 57. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Bir Avrupa ülkesinde motosikletli polislerin hizmet içi eğitim programları dâhilinde ileri sürüş testleri ve eğitimleri yapılmaktadır. Motosikletli polisler, öncelikle yol sürüşü sınavına alınmaktadır. Bu sınavda, sınavı yapan gözlemci ayrı bir motosiklet ile adayı hemen arkasından takip eder. Gözlemci, sürüş bitiminde aşağıdaki forma 0 - 4 puan arasında uygun gördüğü puanları vererek gözlem sonuçlarını aktarır. 0 puan en düşük, 4 puan ise en yüksek beceri düzeyini göstermektedir.

Ölçütler	Puan
Motosikleti bakımları yapılmış ve yola çıkmaya uygun vaziyette hazır bulundurma	
Motosikletinin yönlendirmesini yumuşak ve çabuk hareketlerle yapabilme	
Trafikte uygun konumlanma noktaları seçme	
Debriyajı yumuşak, nazik ve vitesle uyumlu zamanlamayla kullanabilme	
Frenleri, tekerlerin kilitlenmesine izin vermeden etkili kullanabilme	
Virajlarda uygun viraj çizgilerini tayin ve takip edebilme	

56. Her bir maddeden adayın almış olduğu puanlar toplanarak bir toplam puan belirlenir. Gözlem sonucunda yeterli puanı alan motosikletli polisler görevine devam ederler. Yeterli puanı alamamış olan motosikletli polisler ise hizmet içi ileri sürüş eğitimine alınır.

Motosikletli polislerin ileri sürüş becerilerinin değerlendirilmesinde kullanılan bu yaklaşımın, aşağıdakilerden hangisiyle adlandırılması uygundur? (2012)

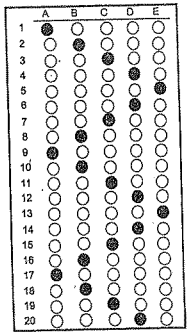
A) Performans değerlendirme
B) Akılcı değerlendirme
C) Modern değerlendirme
D) Ürün değerlendirme
E) İleri değerlendirme

57. Gözlemcinin kullanmış olduğu ölçme aracı aşağıdakilerden hangisidir? (2012)

A) Kısa yanıtlı sınav
B) Bütüncül dereceli puanlama anahtarı
C) Kontrol listesi
D) Derecelendirme ölçeği
E) Ürün dosyası

egitim@hotmail.es

58. Bir öğretmen, geliştirdiği 5 seçenekli 20 maddelik, çoktan seçmeli testin cevap anahtarını aşağıdaki örüntüyü oluşturacak şekilde hazırlamıştır.



Cevap anahtarındaki ilk 10 maddedeki örüntüyü fark eden bazı öğrenciler, geride kalan maddeleri okumadan doğru cevaplamışlardır.

Bu öğrenciler için elde edilen ölçme sonuçlarından hangi niteliği düşüktür? (2012)

A) Puan ortalaması
B) Kullanışlılık
C) Yapı geçerliği
D) Puanlama objektifliği
E) Kapsam geçerliği

59. Bir öğretim üyesi, üniversite 2. sınıf öğrencilerine ölçme ve değerlendirme ara sınavında sormak üzere aşağıdaki çoktan seçmeli maddeyi hazırlamıştır.

Korelasyon, iki değişken arasındaki ilişkinin derecesini verir. İki değişken arasındaki ilişkinin derecesini veren katsayıya korelasyon adı verilir ve rx,y ile gösterilir. Aşağıdaki korelasyon değerlerinden hangisi daha zayıf bir ilişkiyi gösterir?

a. 0.05
b. 0.11
c. 0.21
d. 0.57
e. 0.68

Öğretim üyesinin hazırlamış olduğu maddede, çoktan seçmeli maddelerin yazılmasında dikkat edilmesi gereken esaslardan hangisi gözetilmemiştir? (2012)

A) Maddenin yalnız bir doğru yanıtı olmalıdır.
B) Seçeneklerin uzunlukları birbirine yakın olmalıdır.
C) Seçenek sayısı, testin hitap edeceği öğrenci düzeyine uygun olmalıdır.
D) Madde kökünde gereksiz açıklamalardan kaçınılmalıdır.
E) Noktalama işaretlerine ve imla kurallarına uyulmalıdır.

60. Aşağıda 8. sınıf öğrencileri için hazırlanmış, sözlü ve yazılı sınav sorusu örnekleri ile ev ödevi örnekleri verilmiştir.

Sözlü Sınav Sorusu

A1. 18. yüzyıl öncesinde, Osmanlı İmparatorluğu"nda merkezi hazinenin gelir kaynaklarından ikisini açıklayınız.

A2. Osmanlı İmparatorluğu'nda maliye sistemini anlatınız.

Yazılı Sınav Sorusu

B1. Ülkemizin 81 ilinden biri olan ve güneyde yer alan Antalya'nın, Akdenize kıyısı olması nedeniyle sahip olduğu coğrafi özellikleri yazınız.

B2. Antalya'nın bitki örtüsü ve iklim özellikleri arasındaki ilişkiyi açıklayınız.

Ev Ödevi

C1. 30 x 60 cm boyutlarında, üzerinde en büyük beş gölümüzün, sıra dağlarımızın ve en büyük üç platomuzun gösterildiği bir kabartma Türkiye haritası hazırlayınız.

C2. Beyaz renkte hazırlanmış ve size verilmiş olan kabartma Türkiye haritasında en büyük beş gölümüzün, sıra dağlarımızın ve en büyük üç platomuzun adlarını üzerine yazınız ve bunları farklı renklerde boyayınız.

Soru yazım ilkeleri göz önüne alındığında bir öğretmenin sırasıyla sözlü sınav sorusu, yazılı sınav sorusu ve ev ödevi örneklerinden hangilerini seçmesi beklenir? (2012)

A) A1, B1, C1
B) A1, B2, C1
C) A1, B2, C2
D) A2, B1, C1
E) A2, B1, C2

egitim@hotmail.es

TEST GELİŞTİRME SÜRECİ VE MADDE ANALİZİ

6

ANAHTAR KAVRAMLAR

Test geliştirme süreci basamakları, madde analizi, basit yöntem (alt-üst grup yöntemi), 2. yöntem, madde güçlüğü, madde ayırıcılığı, madde varyansı, madde standart sapması, madde güvenirliği, test parametrelerinin hesaplanması.

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Test geliştirme sürecinin basamaklarını açıklayabilecek,
- Çoktan seçmeli test maddelerinin analizini klasik test kuramında yer alan basit yöntemle (alt- üst grup yöntemi) yapabilecek,
- Çoktan seçmeli test maddelerinin analizini klasik test kuramında yer alan 2. yöntem ile yapabilecek,
- Test parametrelerini hesaplayabileceksiniz.

İÇİNDEKİLER

Test Geliştirme Sürecinin Basamakları
Madde Analizi
Basit Yöntem (alt – üst grup yöntemi)
2. Yöntem
Test Parametrelerinin Hesaplanması
Çalışma Soruları

BAŞLAMADAN ÖNCE

"Madde analizi" kavramını ilk kez duyduğunuzda size neyi çağrıştırdı? "Madde analizi" bölümünün içindekiler kısmına göz gezdirdikten sonra bu kavram hakkında düşünceniz değişti mi? Peki, şimdi ne olduğunu düşünüyorsunuz?

egitim@hotmail.es

TEST GELİŞTİRME SÜRECİ VE MADDE ANALİZİ

TEST GELİŞTİRME SÜRECİNİN BASAMAKLARI

Amacın Belirlenmesi

Kapsamın Belirlenmesi

Maddelerin yazılması ve düzeltilmesi

Ön uygulamanın yapılması

Madde analizi ve seçimi

Son test formunun oluşturulması

Son testin uygulanması

Puanlama

1. Amacın Belirlenmesi: Sadece testlerden değil sınav türü ne olursa olsun öncelikle uygulanacak sınavın amacı belirlenmelidir. Uygulanacak testle, öğrencilerin öğrenme eksiklerinin belirlenmesi, belirli bir eğitim süreci sonundaki başarı düzeylerinin tespit edilmesi amaçlanmış olabilir. İşte test hazırlama sürecinin başında bu amacın ne olduğuna karar verilmesi çok önemlidir.

2. Kapsamın belirlenmesi: Testin uygulanmasının amacı belirlendikten sonra bu amaca hizmet edecek konu kapsamının belirlenmesi gerekir. Testin amacı öğrencilerin belirli periyotlar sonunda öğrenmelerinin izlenmesi, eksiklerinin belirlenerek geri dönüt verilmesi ise testte her hedef-davranışı ölçen en az üç madde bulunmalıdır. Eğer testle amaçlanan, öğrenim süreci sonunda öğrencilerin düzeylerini belirlemekse bu durumda her bir kritik davranışı ölçen en az iki maddeden oluşan bir testin hazırlanması uygun olacaktır. Böylece testin kapsamını geçerli bir şekilde ölçecek madde türü (doğru-yanlış, kısa cevaplı, çoktan seçmeli vb.) ve sayısı da belirlenmiş olacaktır.

3. Maddelerin yazılması ve düzeltilmesi: Testin kapsamı belirlenerek bu kapsamı ölçecek uygun madde türü ve yeterli madde sayısı belirlendikten sonra dersin uzmanları tarafından madde yazımına geçilir. Maddeler yazıldıktan sonra bir kez daha gözden geçirilir, yazım hatalarının, cümle düşüklüklerinin vb. olup olmadığına bakılır. Testin görünüş geçerliğinin olup olmadığı kontrol edilir. Ayrıca testi hazırlayan kişi dışında başka bir uzman tarafından maddelerin incelenmesi tavsiye edilir. Böylece hazırlayan kişinin gözden kaçırdığı hatalar ya da eksiklikler daha iyi belirlenmiş olacaktır. Tüm bu düzeltme işlemleri yapıldıktan sonra ön uygulamanın yapılabilmesi için test çoğaltılarak uygulamaya hazır hale getirilir.

4. Ön uygulama: Testte bulunması amaçlananın en az iki katı maddenin bulunduğu ön test asıl (nihai) uygulamanın yapılacağı gruba benzer bir gruba madde analizinin yapılabilmesi için uygulanır.

5. Madde analizi ve madde seçimi: Ön uygulamanın yapılarak öğrencilerden alınan cevaplardan yararlanılarak, asıl teste alınabilecek en güvenilir ve uygun maddelerin seçimine geçilir. Madde analizinin nasıl yapılması gerektiği ve madde seçiminin nasıl olacağı bir sonraki "madde analizi" başlığı altında ayrıntılı olarak açıklanmıştır.

6. Son testin oluşturulması, uygulanması ve puanlanması: Testin amacına uygun hazırlanan en az iki katı maddeden oluşan testin ön uygulaması sonucu elde edilen cevapların kullanılarak madde analizi yapıldıktan sonra, son (nihai) testi oluşturacak en güvenilir sorular belirlenir. Bu sorulardan oluşan son test, ölçmenin yapılması istenilen asıl gruba uygulanır. Uygulama sonucu elde edilen cevaplar puanlanır ve ölçme süreci tamamlanmış olur. Bundan sonraki süreçte elde edilen puanlarla öğrencilerin başarı düzeyleri vb. durumlar (test analizi başlığı altında verilmiştir) hakkında bilgi edinmek mümkün olacaktır.

NOT: *Test geliştirme süreciyle ilgili daha detaylı bilgiye kaynakçada yer alan (Özçelik, 1989)'ten ulaşabilirsiniz.*

MADDE ANALİZİ

Eğitimde kullanılan ölçme araçlarının (kısaca sınavların) biriminin sorular olduğuna güvenirlik konusunda değinmiştik. Sınavlarda yer alan sorular her zaman "soru cümlelerinden" oluşmadığı için (örneğin; doğru-yanlış testlerinde) sınav sorularına daha genel anlamda "madde" adı verilmektedir. Kısacası madde analizi demek soru analizi demek anlamına gelmektedir.

Ölçme araçlarının olabildiğince güvenilir ve geçerli olması gerekmektedir. Bu özelliklerin sağlanabilmesi için testin ön uygulama sonuçlarından yararlanılarak, testteki her bir maddenin analizi yapılır ve maddelerin istenilen özelliklere sahip olup olmadığı araştırılır.

Bu bölümde çoktan seçmeli test maddelerinin analizini Klasik Test Kuramı'nda yer alan iki yöntemle açıklayacağız. Bu yöntemler:

1. **Basit yöntem (Alt-üst grup yöntemi)**: Basit yöntemde; en yüksek puan alan öğrenciler ile en düşük puan alan öğrencilerin oluşturduğu gruplar analize alınır. Öğrenci sayısının 300-400 ya da daha çok olduğu durumlarda bu yöntemin kullanılması tavsiye edilebilir. Bu yöntemin kullanılması için test puanları dağılımının simetrik olması istenir. Bir başka deyişle, puanların ortalama etrafında yoğunlaştığı, en yüksek ve en düşük puan alan öğrenci sayılarının olabildiğince birbirine eşit sayıda olduğu durumun gerçekleşmiş olması beklenir. (Not: Simetrik dağılım bir sonraki bölümde ayrıntılı şekilde açıklanmıştır.)

egitim@hotmail.es

2. **Yöntem:** Bu yöntemde testin uygulandığı tüm öğrenciler analize dâhil edilir. Öğrenci sayısının çok daha az olduğu (60-70 gibi) durumlarda tercih edilebilir.

Şimdi bu yöntemlerle madde analizinin nasıl yapıldığı örnekler üzerinde anlatılmaya çalışılacaktır.

Öncelikle kullanılan yöntem ne olursa olsun öğrencilerin her bir maddeye verdiği cevapların yer aldığı madde cevaplarının gösterildiği matris oluşturulmalıdır. Bu matrisin nasıl oluşturulacağını bir örnek ile anlatmaya çalışalım. 10 öğrencinin cevapladığı 10 sorunun yer aldığı beş seçenekli çoktan seçmeli bir testteki öğrenci cevapları Tablo 1'de gösterildiği gibi olsun. Bu testin cevap anahtarına uygun olarak, doğru cevaplar 1 puan; yanlış, boş ve geçersiz cevaplar da 0 puan olacak şekilde puanlandığında, öğrencilerin her bir maddeden aldığı puanlar Tablo 2'de gösterildiği şekilde olacaktır.

Tablo 1: Madde Cevapları Matrisi

	Test Maddeleri									
	1	2	3	4	5	6	7	8	9	10
Doğru cevap / ÖĞRENCİLER	**A**	**B**	**E**	**E**	**D**	**C**	**A**	**D**	**B**	**C**
1. Aslı	A	B	E	E	C	A	E	B	B	-
2. Cüneyt	A	B	C	A	D	E	D	B	B	A
3. Ebru	A	A	C	C	D	B	B	C	E	C
4. Can	A	B	E	E	D	C	B	A	E	C
5. Ali	A	E	B	C	B	E	B	A	C	B
6. Eda	A	B	E	E	D	C	E	D	B	C
7. Özge	A	B	E	E	C	D	-	D	B	A
8. Emre	A	C	D	C	D	C	D	D	E	C
9. Sevda	A	B	E	E	D	C	E	D	B	C
10. Burak	A	B	E	C	D	C	E	A	B	C

Tablo 2: Madde Puanları Matrisi

	Test Maddeleri										
	1	2	3	4	5	6	7	8	9	10	Öğrenci Toplam puanı (X_i)
Doğru cevap / ÖĞRENCİLER	A	B	E	E	D	C	A	D	B	C	
1. Aslı	1	1	1	1	0	0	0	0	1	0	**5**
2. Cüneyt	1	1	0	0	1	0	0	0	1	0	4
3. Ebru	1	0	0	0	1	0	0	0	0	1	**3**
4. Can	1	1	1	1	1	1	0	0	0	1	**7**
5. Ali	1	0	0	0	0	0	0	0	0	0	**1**
6. Eda	1	1	1	1	1	1	0	1	1	1	**9**
7. Özge	1	1	1	1	0	0	0	1	1	0	**6**
8. Emre	1	0	0	0	1	1	0	1	0	1	**5**
9. Sevda	1	1	1	1	1	1	0	1	1	1	**9**
10. Burak	1	1	1	0	1	1	0	0	1	1	**7**
Madde Toplam puanı (I_i)	10	7	6	5	7	5	0	4	6	6	56

Madde analiz süreci yukarıda gösterilen, öğrencilerin maddelere verdikleri cevaplara göre oluşturulan madde puanları matrisinden yararlanılarak gerçekleştirilir. Öncelikle basit yönteme göre madde analizleri verilecek ardından Henrysson yöntemi açıklanacaktır.

Basit Yöntem (alt – üst grup yöntemi)

Madde analizinde basit yöntem uygulanırken sırasıyla aşağıdaki basamaklar takip edilir:

1. Öğrenci cevapları, doğru cevaplar 1 puan; yanlış, boş ve geçersiz cevaplar da 0 puan olacak şekilde puanlandıktan sonra her bir öğrencinin toplam test puanı hesaplanır. Bu durum, Tablo 2'nin son sütununda görülmektedir. Örneğin; Aslı'nın toplam test puanı 5, Cüneyt'in toplam test puanı 4 vb.
2. Öğrencilerin toplam test puanlarına göre, öğrenci cevap kâğıtları en yüksek puandan en düşük puana doğru sıralanır. Bizim yukarıdaki örneğimizde öğrenci sayısı çok az olduğu için bu yöntemin uygulanması doğru olmayacaktır.
3. Test puanı en yüksek olan grubun *%27'si ile en düşük puanı alan grubun %27'si alınarak, sırasıyla üst grup ve alt grup adları verilir. Ortada kalan %46'lık grup ise analiz dışında tutulur.

egitim@hotmail.es

Üstteki %27'lik grup testteki soruları doğru bilenleri temsil eden "Bilenler" den oluşan üst grup, alttaki %27'lik grup ise testteki soruları doğru bilemeyenleri temsil eden "Bilmeyenler" den oluşan alt grup olarak ifade edilir.

4. Bundan sonraki adımlarda, her bir maddenin analizini yapabilmek için aşağıda bir örnekle gösterilecek olan madde analiz tabloları oluşturularak, her bir maddenin madde güçlüğü, madde ayırıcılığı vb. değerleri bu tablolar yardımıyla hesaplanır.

Madde analiz tabloları sadece üst ve alt grubun yer aldığı tablolardır. Oluşturulan üst ve alt gruplar toplam öğrenci sayısını oluşturur. Bir başka deyişle, madde analizinde, toplam öğrenci sayısı üst grup ve alt grup öğrnci sayısının toplamından oluşmaktadır ve daima üst gruptaki öğrenci sayısı alt gruptaki öğrenci sayısına eşittir.

*NOT: Belirtilen bu "%27" değer istatistiksel bir değerdir ve istatistik uzmanları tarafından belirlenmiştir. Nedenin bilinmesi istatistik konusunda uzmanlığı gerektirdiği için burada açıklaması verilmeyecektir.

Örnek Madde Analiz Tablosu:

1. madde	A	B	C	*D	E	TOPLAM
Üst Grup	10	13	10	15	2	**50**
Alt Grup	22	10	10	5	3	**50**
TOPLAM	32	23	20	20	5	100

*Maddenin doğru cevabı D seçeneğidir.

Madde Güçlüğü (p_j)

Bir maddeyi doğru cevaplayan öğrenci sayısının testin uygulandığı tüm öğrencilerin sayısına oranı olarak tanımlanabilir.

$$p_j = \frac{\text{doğru cevaplayan öğrenci sayısı}}{\text{toplam öğrenci sayısı}} \qquad (6.1)$$

Bu eştlik daha açık olarak aşağıdaki şekilde de ifade edilebilir:

$$p_j = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı + alt grupta doğru cevaplayan öğrenci sayısı}}{\text{toplam öğrenci sayısı}} \qquad (6.2)$$

Buna göre örneğimizde yer alan maddenin güçlüğü;

$p_j = \frac{15 + 5}{100} = 0{,}2$ olarak bulunur.

Peki, bu değer nasıl yorumlanacaktır? Bu soruyu cevaplayabilmek için madde güçlüğünün alt ve üst sınırını, bir başka deyişle alabileceği en yüksek ve en düşük değerlerin ne olduğunu açıklayalım:

Düşününüz ki testi alan tüm öğrenciler bu maddeyi bilseydi, bu durumda madde güçlüğü (100/100=1) "1" çıkacaktır. Burada elde edilen değer madde güçlüğünün alabileceği maksimum değerdir. (100 kişinin katıldığı bir testte 100'den daha fazla sayıda öğrenci doğru cevaplayamayacağı için...) Diğer taraftan da bir maddenin hiçbir öğrenci tarafından doğru cevaplanmadığını düşünürsek bu kez de madde güçlüğü (0/100=0) "0" değerini alarak, madde güçlüğünün alabileceği minimum değer ortaya çıkacaktır.

Böylelikle artık çok rahat söylenebilir ki madde güçlüğü en düşük 0 ile en yüksek 1 arasında değerler alabilen bir istatistiktir.

$0 \le p_j \le 1$

0 — 0.50 — 1

Madde güçlüğünün orta değeri 0.50'dir. Bu değer maddenin orta güçlükte (zorlukta) olduğunun göstergesidir. Madde güçlüğü azaldıkça yani 0'a yaklaştıkça (bilemeyen sayısı arttıkça) madde zorlaşır, madde güçlüğü arttıkça yani 1'e yaklaştıkça (bilen sayısı arttıkça) madde kolaylaşır.

!! DİKKAT: Bu ifadedeki güçlük ve zorluk arasındaki zıtlık, hesaplanan değere "madde güçlüğü" isminin verilmesinden kaynaklıdır. Bu nedenle lütfen yukarıdaki açıklamaları dikkatle okuyunuz ve anlamaya çalışınız.

Bu açıklamalara göre, örneğimizde yer alan maddenin güçlük düzeyinin düşük yani zor bir madde olduğunu söyleyebilriz ($p_j = 0.20$).

Madde Ayırıcılığı (r_{jx})

Bir maddenin bilenleri bilmeyenlerden ayırabilmesinin derecesini verir. Madde ayırıcılığı, bir maddenin en önemli istatistiğidir. Maddenin güvenirliğinin bir ölçüsünü verir. Bu sebeple, bir maddenin teste alınıp alınmamasına karar verirken öncelikle madde ayırıcılığının yüksek olup olmamasına bakılır. Madde ayırıcılığı;

egitim@hotmail.es

$$r_{jx} = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı - alt grupta doğru cevaplayan öğrenci sayısı}}{\text{toplam öğrenci sayısı / 2}} \quad (6.3)$$

ya da

$$r_{jx} = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı - alt grupta doğru cevaplayan öğrenci sayısı}}{\text{üst gruptaki öğrenci sayısı}} \quad (6.4)$$

ya da

$$r_{jx} = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı - alt grupta doğru cevaplayan öğrenci sayısı}}{\text{alt gruptaki öğrenci sayısı}} \quad (6.5)$$

eşitliklerinden biri kullanılarak hesaplanabilir.

Buna göre örneğimizde yer alan maddenin ayırıcılığı

$r_{jx} = \frac{15 - 5}{50} = 0{,}2$ olarak bulunur.

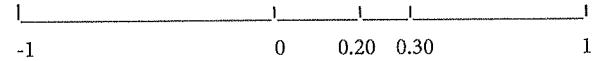
!! DİKKAT: Yukarıdaki örnek için hesaplanan madde güçlüğü ile madde ayırıcılığının her ikisinin de 0.20 olarak aynı değer bulunması TAMAMEN BİR TESADÜFTÜR. Bu nedenle madde güçlüğü ile madde ayırıcılığı arasında bir eşitlik, biri yüksek ise diğeri de yüksek olacaktır ya da biri yüksek oldukça diğeri azalacaktır gibi bir bağıntı olduğu kesinlikle düşünülmesin.

Şimdi de madde ayırıcılığının nasıl yorumlanacağına bir bakalım:

- Aynen madde güçlüğünde olduğu gibi madde ayırıcılığının da alt ve üst sınırını, bir başka deyişle alabileceği en yüksek ve en düşük değerlerin ne olduğunu bulmaya çalışalım:
  - Düşününüz ki üst ve alt gruptaki tüm öğrenciler bu maddeyi doğru bilseydi, bu durumda madde ayırıcılığı [(50-50)/50 = 0] "0" çıkacaktı. Benzer şekilde bu kez de üst ve alt gruptaki hiçbir öğrenci bu maddeyi doğru cevaplayamasaydı, bu durumda da madde ayırıcılığı [(0-0)/50=0] "0" olarak bulunacaktı. Dikkat edilirse her iki durumda da yani bir maddenin tüm öğrenciler tarafından bilinmesi ya da hiçbir öğrenci tarafından bilinememesi durumunda, madde ayırıcılığı 0 çıkmaktadır. Yani tüm öğrencilerin bildiği ya da hiç bir öğrenci tarafından bilinmeyen bir madde, bileni bilmeyenden ayıramamaktadır. Bu tür bir madde hiçbir şekilde teste alınmak istenmez.

  - Bu kez de düşünelim ki üst gruptaki tüm öğrenciler maddeyi doğru cevaplarken, alt gruptaki hiçbir öğrenci doğru cevap verememiş olsun (50-0/50=1) bu durumda madde ayırıcılığı 1 çıkacaktır. Dikkat edilirse madde ayırıcılığının 1'den daha büyük bir değer alması mümkün değildir. Bu nedenle madde ayırıcılığının maksimum değerinin 1 olduğunu söyleyebiliriz. Madde ayırıcılığının 1 olması, maddenin bilenler ile bilmeyenleri mükemmel bir şekilde ayırıabildiğinin göstergesidir.
  - Şimdi de düşünelim ki üst gruptaki öğrencilerin hiç biri maddeyi doğru cevaplayamazken, alt gruptaki tüm öğrenciler maddeye doğru cevap vermiş olsun (0-50/50=-1) bu durumda madde ayırıcılığı -1 çıkacaktır. Burada dikkat edilirse madde ayırıcılığının -1'den daha küçük bir değer alması mümkün değildir. Bu nedenle madde ayırıcılığının minimum değerinin -1 olduğunu söyleyebiliriz.

$-1 \le r_{jx} \le 1$



(a) $-1 \le r_{jx} \le 0$**:** Madde ayırıcılığının negatif değer alması istenmeyen bir durumdur, maddede bir sorun olduğunun işaretidir. Cevap anahtarında bir hata yapılmış olabileceğinin de bir göstergesi olabilir. Bu ayırıcılık düzeyine sahip olan bir madnenin teste alınması istenmez.

(b) $0 \le r_{jx} \le 0.20$**:** Madde ayırıcılığının 0.20'nin altında bir değer alması, maddenin iyi düzeyde ayırıcılığının olmadığını; bilenle bilmeyeni birbirinden ayıramadığını işaret eder. Bu tür bir madde teste alınmamalıdır.

(c) $0.20 \le r_{jx} \le 0.30$**:** Madde ayırıcılığının 0.20 ile 0.30 arasında bir değer alması, maddenin orta düzeyde ayırıcılığının olduğunu, bilenle bilmeyeni birbirinden iyi düzeyde ayıramadığını işaret eder. Bu tür maddeler zorunlu olmadıkça teste alınmamalıdır. Eğer mutlaka bu tür maddelerin teste alınması gerekiyorsa, mutlaka maddede gerekli düzeltmeler yapıldıktan sonra teste alınmalıdır.

(d) $r_{jx} \ge 0.30$**:** Madde ayırıcılığının 0.30 ve üstünde bir değer alması, maddenin iyi düzeyde ayırıcılığının olduğunu; bilenle bilmeyeni birbirinden ayırabildiğini işaret eder. Bu tür maddelerin teste alınması istenir.

egitim@hotmail.es

- ***Bir maddenin teste alınması için:***
  1. Madde ayırıcılığının olabildiğince yüksek (1'e yakın) olması,
  2. Madde güçlüğünün orta düzeyde (0.50 ve civarı) olması,
  3. Çeldiricileri işaretleyen üst gruptaki öğrenci sayısının alt gruptaki öğrenci sayısından az olması,
  4. Çeldiricileri işaretleyen öğrenci sayısının olabildiğince dengeli bir dağılım göstermiş olması istenir.

Tüm bu bilgilere dayanarak yukarıda madde analiz tablosu için verilen örneğimizi değerlendirirsek;

- Madde ayırıcılık düzeyi: r_{jx}=0.20 olup, madde ayırıcılığı düşüktür, madde ya teste alınmamalıdır ya da gerekli düzeltmeler yapılarak teste alınabilir.
- Madde güçlük düzeyi p_j=0.20 olup, madde güçlük düzeyi düşük, oldukça zor bir maddedir.
- Maddenin çeldiricileri incelendiğinde; A çeldiricisinde, bilmeyenleri (alt grubu) bilenlere göre daha fazla çeldirdiği için bir sorun olmadığı görülmektedir. Hâlbuki B çeldiricisi tam tersine bilenleri (üst grup) bilmeyenlere göre daha fazla çeldirmiştir. Bu seçeneğin düzeltilmesi gerekmektedir. C çeldiricisini bilenlerle bilmeyenler aynı sayıda işaretlemiştir. E çeldiricisini bilmeyenler daha fazla işaretlemiş olsa da işaretleyen toplam öğrenci sayısı çok azdır. Nerdeyse bu seçenek hiç işlememiştir. Bir bakıma bu madde 4 seçenekliymiş gibi işlev göstermiştir. Bu nedenledir ki C ve E çelidiricilerinin de gözden geçirilmesi önerilebilir.

Dikkat edilirse madde ayırıcılığının düşük olması bize maddede sorun olduğunun sinyalini vermişti. Bu sebeple bu madde teste alınmayacak ya da mutlaka alınacaksa çeldiriciler için belirtilen gerekli düzeltmeler uygulandıktan sonra teste alınması uygun olacaktır.

Madde Varyansı (s_x^2) ve Standart Sapması (s_x)

Madde varyansı ve standart sapması, bir maddeye verilen cevapların (0-1 puanlarının) değişiminin derecesini gösteren değerlerdir. Kısacası madde puanlarının değişim ölçüleridir. Madde varyansı ve standart sapması aşağıda verilen eşitlikler yardımıyla hesaplanır:

$$s_x^2 = pj\,(1 - pj) \qquad (6.6)$$

ve

$$s_x = \sqrt{pj(1-pj)} \qquad (6.7)$$

s_x^2 : madde varyansı

s_x : madde standart sapması

p_j: madde güçlüğü

Eşitliklerden de görüleceği gibi, standart sapma, varyansın kareköküdür. Bu durum ifadelerin sembollerinden de anlaşılmaktadır. Şimdi, bu değerleri daha iyi anlayabilmek ve ne ifade ettiklerini daha iyi görebilmek için şu şekilde örneklemeye çalışalım:

Tablo 2'de yer alan 1., 6.ve 7. maddelerin puanlarını inceleyelim:

1.maddeyi, öğrencilerin tümü doğru cevapladığı için maddeye ilişkin sütunudaki tüm puanlar 1'lerden oluşmaktadır. Kısacası maddeye ilişkin puanlarda hiç bir değişim olmamaktadır.

1. maddenin güçlük düzeyini hesaplarsak p_j= 10/10= 1 ve buradan

$s_x^2 = 1\,(1-1) = 0$ ve $s_x = 0$ bulunur.

Değişim 0'dır ki biz de yukarıda 1. maddedeki puanların değişiminin olmadığını söylemiştik.

6. maddeyi, toplam 10 öğrencinin 5'i doğru, 5'i yanlış cevapladığı için puanlarının %50'si 1 ve %50'si 0'lardan oluşmaktadır. Kısacası bu maddeye ilişkin puanlar %50-50 bir değişim göstermektedir.

6. maddenin güçlük düzeyini hesaplarsak p_j= 5/10= 0.5 ve buradan

$s_x^2 = 0.50\,(1-0.50)=0.25$ ve $s_x = 0.5$ bulunur.

Değişim 0.5'tir ki biz de yukarıda 1. maddedeki puanların değişiminin 0.5 olduğunu söylemiştik.

maddeyi, öğrencilerin hiçbiri doğru cevaplayamadığı için tüm puanlar 0'lardan oluşmakta, kısacası maddeye ilişkin puanlarda hiç bir değişim olmamaktadır.

8. maddenin güçlük düzeyini hesaplarsak p_j= 0/10= 0 ve buradan

$s_x^2 = 0\,(1-0) = 0$ ve $s_x^2 = 0$ bulunur.

Değişim 0'dır ki biz de yukarıda 8. maddedeki puanların değişiminin olmadığını söylemiştik.

Böylelikle çok daha fazla sayıda öğrenciye uygulanan sınavlarda her bir maddenin puanlarındaki değişim miktarını hesaplayabilmek için varyans ve standart sapma eşitlikleri bize yardımcı olmaktadır.

Varyansın minimum değeri 0 ve maximum değeri 0.25'tir: $0 \le s_x^2 \le 0.25$

Standart sapmanın minimum değeri 0 ve maximum değeri 0.5'tir: $0 \le s_x \le 0.5$

Madde Güvenirliği (r_x)

$$r_x = r_{jx} \cdot s_x \qquad (6.8)$$

Her bir maddenin güvenirliği, maddenin ayırıcılığı ve standart sapmasıyla doğru orantılı olarak artar. Buradadan da anlaşılacağı üzere, maddenin ayırıcılığı doğrudan güvenirliğini arttıran bir özelliktir. Bu sebepledir ki madde ayırıcılığının olabildiğince

egitim@hotmail.es

yüksek olması istenilen bir özelliktir. Madde standart sapması en yüksek değerine (yukarıdaki standart sapma eşitliğinden görülmektedir) madde güçlüğü 0.50 olduğunda ulaşmaktadır. Bu sebepledir ki madde güçlüğünün 0.50 veya civarında olması maddenin güvenirliğini dolaylı olarak arttıran bir diğer faktördür.

2. Yöntem

- Bu yöntemde, testin uygulandığı tüm öğrenciler analize dahil edilir.
- Basit yöntem ile bu yöntemin sonuçları birbirinin aynı olmayabilir. Ancak testin uygulandığı öğrenci sayısı arttıkça her iki yöntemle elde edilen sonuçlar arasındaki fark azalır.
- Bu yöntem basit yönteme göre biraz daha zor hesaplamalar içerdiğinden elle hesaplama yapmak daha güçtür. Hesap makinası, bilgisayar gibi elektronik araçlarla hesaplama yapma olanağı olduğunda tüm öğrencilerin cevap kağıtları analize dahil edildiği için bu yöntem tercih edilmelidir.

2. Yöntemin Basamakları:

1. Öğrencilerin cevapları, öğrenciler satırlarda, test maddeleri sütunlarda yer alacak biçimde Tablo 2'de gösterildiği üzere, doğru cevaplar 1 puan, yanlış, boş ve geçersiz cevaplar 0 puan olacak şekilde madde puanları matrisi olarak düzenlenir.
2. Her sütundaki doğru cevaplar (1 puanlar) toplanarak her bir maddenin kaç öğrenci tarafından doğru cevaplandırıldığı hesaplanarak, her bir maddenin (Ij) madde puanı bulunur.
3. Benzer şekilde, her satırdaki doğru cevaplar (1 puanlar) toplanarak her bir öğrencinin tüm soruların kaçını doğru cevaplandırdığı hesaplanarak, her bir öğrencinin (Xi) toplam test puanı bulunur.
4. Test puanlarının aritmetik ortalaması ($\overline{X}$) ve standart sapması (Sx) aşağıda ifade edilen eşitlikler kullanılarak hesaplanır.

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (6.9)$$

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n}}{n-1}} \qquad (6.10)$$

Eşitliklerde verilen, sigma işaretleri "Σ" önünde bulunan terimlerin toplamlarını göstermek üzere,

X_i: test puanlarını,

n: öğrenci sayısını göstermektedir.

Buna göre Tablo 2'de yer alan test puanlarının standart sapması;

$$S_x = \sqrt{\frac{372 - \frac{56^2}{10}}{10-1}} = \sqrt{\frac{372\text{-}313 \cdot 6}{9}} = \sqrt{\frac{58.4}{9}} = 2.55$$

1. **Madde güçlüğü (p_j)** ve **ayırıcılık gücü (r_{jx})** aşağıdaki eşitliklerle hesaplanır.

$$P_j = \frac{n(d)}{N} \qquad (6.11)$$

$$r_{jx} = \frac{\overline{X}(d) - \overline{X}}{S_x} x \sqrt{\frac{p_j}{q_j}} \qquad (6.12)$$

Eşitlikte;

n(d): Maddeye doğru cevap veren öğrencilerin sayısı,

N: Öğrenci sayısı,

X(d): maddeye doğru cevap verenlerin test puanları aritmetik ortalaması,

($\overline{X}$) : test puanların aritmetik ortalaması,

S_x : test puanlarının standart sapması,

p_j : madde güçlüğü ve

q_j : 1-p_j'dir.

2. Maddelerin ayırıcılıklarına ve güçlüklerine, çeldiricilerin çalışıp çalışmadığına bakılarak son test için madde seçimi yapılır. Son test için madde seçerken özellikle *ayırıcılık gücü indeksi 0.30'dan büyük* maddeleri dikkate alınmalıdır.

egitim@hotmail.es

Test Parametrelerinin Hesaplanması

Son (nihai) testin uygulanmasıyla öğrencilerden elde edilen puanlarla test puanlarının aritmetik ortalaması hesaplanabilir. Test puanlarının aritmetik ortalamasının (kısaca testin aritmetik ortalaması) aşağıdaki eşitlikle hesaplanabilir:

$$\bar{X} = \sum_{j=1}^{K} p_j \qquad (6.13)$$

$\bar{X}$: test puanlarının aritmetik ortalaması

p_j : maddelerin güçlük düzeyleri

Testin aritmetik ortalamasının yüksek olması isteniliyorsa, madde güçlük düzeyleri yüksek olan (kolay) maddelerin teste seçilmesi, ortalamasının düşük olması isteniyorsa, madde güçlük düzeyleri düşük (zor) maddelerin seçilmesi uygun olacaktır.

Testin ortalama güçlüğü; testin aritmetik ortalamasının, testte yer alan madde sayısına bölünmesiyle hesaplanır. Yukarıdaki eşitlikle hesaplanan aritmetik ortalama kullanılarak aşağıdaki eşitlik ile testin ortalama güçlüğü kestirilebilir.

$$\bar{P} = \frac{\bar{X}}{K} \qquad (6.14)$$

$\bar{P}$: testin ortalama güçlüğü

K: testte yer alan madde sayısı

Son (nihai) testin standart sapması, yukarıda yer alan 6.10 eşitliği ile de hesaplanabileceği gibi testte seçilen her bir maddenin güvenirliği ile o maddenin standart sapmasının çarpımlarının toplamı olarak ifade edilen aşağıdaki 6.15 eşitliği ile de hesaplanabilir. Maddelerin standart sapmasının ve güvenirliğinin ne demek olduğu ve nasıl hesaplanacağı verilmişti. Böylece testin standart sapması aşağıda verilen eşitlik yardımıyla da kestirilebilir:

$$S_x = \sum_{j=1}^{K} (r_{jx} \times s_j) \qquad (6.15)$$

S_x : testin standart sapması

r_{jx} : maddelerin güvenirlik indeksi

s_j : maddelerin standart sapması

Bir testin güvenirliği, testin standart sapmasından etkilenir. Standart sapma arttıkça testin güvenirliği de artma eğilimindedir. Bu nedenle güvenirliği yüksek maddeleri seçmek, testin standart sapmasını ve böylece testin güvenirliğini artırır.

Son testin güvenirliği "güvenirlik" konusunda bahsedilen KR-20 yöntemi ile hesaplanabilir. Nihai testin güvenirliğinin belirlenmesinde aşağıdaki eşitlikten yararlanılır.

$$KR-20 = \frac{K}{K-1}\left(1 - \frac{\sum_{j=1}^{K}(pj \cdot qj)}{S_x^2}\right) \qquad (6.16)$$

K: testteki madde sayısı

$(p_j \cdot q_j) = s_j^2$: testte yer alan maddelerin standart sapmalarının kareleri (varyansları)

S_x^2 : testin standart sapmasının karesi (testin varyansı)

Bu bölümün başında verilen örnek tabloda yer alan 10 öğrencinin 10 maddeden aldıkları puanların güvenirliğini hesaplayalım. Test puanlarının standar sapması yukarıda yer alan 6.10 eşitliği yardımı ile 2.55 olarak hesaplanmıştır. Buradan $Sx^2 = 2.55^2 = 6.50$ olarak bulunur. Tablo 2'de yer alan madde puanlarının standart sapmalarının kareleri (varyansları) ve bunların toplamı aşağıdaki Tablo 3'te yer almaktadır.

Tablo 3: Tablo 2'de yer alan madde puanları varyansları

maddeler	1	2	3	4	5	6	7	8	9	10	□
pj	1	0.7	0.6	0.5	0.7	0.5	0	0.4	0.6	0.6	
qj	0	0.3	0.4	0.5	0.3	0.5	1	0.6	0.4	0.4	
pj . qj	0	0.21	0.24	0.25	0.21	0.25	0	0.24	0.24	0.24	1.88

Tüm bu değerler, 6.16 eşitliğinde yerine koyulduğunda;

$$KR-20 = \frac{10}{10-1} \cdot \left[1 - \frac{1.88}{6.50}\right] = \frac{10}{9}[1-0.29] = 1.11 \cdot 0.71 = 0.79$$

olarak test puanlarının güvenirliği elde edilmiş olur.

Öğrencilerin başarılarını yorumlamasının doğruluğu ya da öğrenciler hakkında alınacak kararların isabetli olmasının derecesi, öğrencilerden elde edilen ölçme sonuçlarının doğruluğuna, bu sonuçların doğruluğu da uygulanan ölçme araçlarının

egitim@hotmail.es

güvenilir ve geçerli olabilme derecesine bağlıdır. (Ölçme sonuçlarının öğrenci başarılarını ne derecede yansıttığına bağlıdır.) Uygulanan testlerle (ölçme araçları) elde edilen puanlardaki tesadüfî hata miktarı azaldığı ölçüde öğrencilere ilişkin elde edilecek bilgiler daha doğru olacaktır. Ölçme sonuçlarına karışabilecek tesadüfî hatanın azalması ölçme ortamın uygunluğunun, ölçmenin yapıldığı öğrencinin ve ölçmeyi yapan öğretmenin yanısıra ölçme araçlarının olabildiğince dikkatli ve hatasız şekilde hazırlanmasına bağlıdır. Olabildiğince az hata içeren bir ölçme aracı hazırlamak, madde yazma tekniklerini çok iyi bilmeyi, deneyimli ve konu alanına hâkim olmayı gerektirir. Bununla birlikte, hazırlanan bir testin (test hazırlama sürecinde de belirtildiği gibi) uygun bir öğrenci grubuna ön uygulaması yapılarak ve madde analiziyle madde seçimi gerçekleştirilerek, geliştirilmesi olabildiğince iyi hale getirilmesi mümkündür. Ancak her zaman ön uygulama yapmak mümkün olmayabilir. Özellikle okullarda derse giren öğretmenlerin ne zamanı ne de olanakları böyle bir uygulamaya elverişli olmayacaktır. Bu durumda dersin öğretmeni her dönem uyguladığı testlerdeki maddelerin analizlerini yapabilirler. Belirli bir süre sonra (bir kaç eğitim yılı sonunda) öğretmenlerin elinde maddelerin güçlük ve ayırıcılık düzeylerini bildikleri pek çok maddeden oluşan bir madde havuzu ortaya çıkacaktır. Bundan sonra uygulayacakları testlerde amaçlarına uygun bu havuzdan en güvenilir soruları seçebilirler. Maddeler üzerinde gerekli düzeltme ve iyileştirmeleri işlemlerini yapabilirler. Böylece ileriki uygulamalarında çok daha güvenilir ve geçerli testler uygulayarak öğrencileri hakkında çok daha doğru ve yerinde kararlara varabilirler.

DEĞERLENDİRME SORULARI

Aşağıdaki tabloda numaralanmış beş sorunun soru cevaplanma yüzdeleri (p_j) ile ayırt edicilik değerleri (r_{jx}) verilmiştir.

1.- 3. Soruları aşağıdaki tabloya göre cevaplayınız.

Soru	p_j	r_{jx}
1	0,27	0,58
2	0,51	-0,18
3	0,85	0,07
4	0,28	0,15
5	0,81	0,62

1. Hazırlanacak bir test için yukarıdaki sorulardan hangisi kesinlikle tercih edilmemelidir?

A) 1 B) 2 C) 3 D) 4 E) 5

2. Alınacak öğrenci sayısının başvuran öğrenci sayısından çok daha az olduğu durumlarda kullanılacak soruların çoğu, güçlük ve ayırt edicilik bakımından yukarıdaki sorulardan hangisine benzer nitelikte olmalıdır?

A) 1 B) 2 C) 3 D) 4 E) 5

3. Yukarıdaki sorulardan hangisi kolay ve ayırt edici nitelikte bir sorudur?

A) 1 B) 2 C) 3 D) 4 E) 5

4. Aşağıdakilerden hangisi madde analizi ile elde edilecek bilgiler arasında yer almaz?

A) Testin güçlük düzeyi
B) Testteki maddelerin tutarlılık derecesi
C) Testteki maddelerin çeldiricilerinin işlevselliği
D) Testi cevaplayan öğrencilerin aldıkları puanların değişkenlik düzeyi
E) Testin başarılı ve başarısız öğrencileri birbirinden ayırabilme derecesi

Aşağıdaki tabloda 200 öğrencinin katıldığı bir sınavın 1. maddesinin analizi sonuçları verilmiştir. Buna göre, 5.- 7. soruları bu tabloya göre cevaplayınız.

	A	B	C	D	E	Boş	N
Üst Grup	16	30	12	12	8	22	100
Alt Grup	4	26	13	16	8	33	100

egitim@hotmail.es

5. **Bu maddenin doğru cevabı hangi seçenek olursa ayırt ediciliği en yüksek olacaktır?**

A) A B) B C) C D) D E) E

6. **Maddenin doğru cevabı B ise, madde ayırıcılık gücü indeksi kaç olur?**

A) 0,00 B) 0,04 C) 0,26 D) 0,30 E) 1,00

7. **Tabloda verilen bilgilere dayanarak, bu madde hakkında aşağıdaki yargılardan hangisine ulaşılabilir?**

A) A seçeneği çeldirici ise, istenen yönde işlev göstermiştir.
B) Maddeyi üst ve alt gruplardan eşit sayıda öğrenci cevaplamıştır.
C) B seçeneği çeldirici ise, çok az kişiyi çeldirmiştir.
D) Doğru cevap C ise, soru negatif ayırıcılığa sahiptir.
E) Seçeneklerin tümü dengeli bir şekilde işlev göstermiştir.

Aşağıdaki tabloda bir sınıftaki öğrencilere uygulanmış bir testte yer alan 3. maddeye ilişkin seçenekleri işaretleyen öğrencilerin alt ve üst gruba ait dağılımları verilmiştir.

8.-10. soruları aşağıda verilen tabloya göre cevaplandırınız.

3. Madde	A	B*	C	D	E	Toplam
Üst Grup	10	66	4	2	18	100
Alt Grup	28	22	24	14	12	100
Toplam	38	88	28	16	30	200

8. **Tabloya göre maddenin güçlük indeksi kaçtır?**

A) 0.22 B) 0.44 C) 0.49 D) 0.66 E) 0.98

9. **Tabloya göre maddenin ayırıcılık indeksi kaçtır?**

A) 0.22 B) 0.44 C) 0.49 D) 0.66 E) 0.98

10. **Bu maddeye ait seçenekler hakkında aşağıdakilerden hangisi söylenebilir?**

A) Çeldiriciler eşit oranda işlev göstermiştir.
B) En iyi E çeldiricisi işlev göstermiştir.
C) Doğru seçenek yeterli işlev göstermiştir.
D) C ve D çeldiricileri gözden geçirilmelidir.
E) Çeldiricilerin hepsi çok iyi işlev göstermiştir.

KPSS'DE ÇIKMIŞ SORULAR

1. - 3. SORULARI AŞAĞIDA VERİLEN BİLGİLERE GÖRE CEVAPLAYINIZ.

Aşağıdaki tabloda numaralanmış beş sorunun doğru cevaplanma yüzdeleri (P_j) ile ayırt edicilik değerleri (R_{jx}) verilmiştir.

Soru	1	2	3	4	5
P_j	0,35	0,40	0,80	0,25	0,85
R_{jx}	0,60	-0,35	0,10	0,15	0,50

1. **Yukarıdaki sorulardan hangisi, düzeltilemiyorsa testten mutlaka çıkarılmalıdır? (2003)**

A) 1 B) 2 C) 3 D) 4 E) 5

2. **Başvuran sayısının çok, seçilecek kişi sayısının az olduğu bir sınavda kullanılacak soruların çoğu, güçlük ve ayırt edicilik bakımından yukarıdaki sorulardan hangisine benzer nitelikte olmalıdır? (2003)**

A) 1 B) 2 C) 3 D) 4 E) 5

3. **Yukarıdaki sorulardan hangisi kolay ve ayırt edici bir sorudur? (2003)**

A) 1 B) 2 C) 3 D) 4 E) 5

4. **Aşağıdaki sorulardan hangisi madde analizi ile elde edilecek bilgilerden yararlanılarak cevaplanamaz?(2004)**

A) Testin zorluk derecesi amaca uygun mudur?
B) Testteki sorularla 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?

egitim@hotmail.es

5. - 7. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

DİKKAT! HER SORUYU BİRBİRİNDEN BAĞIMSIZ OLARAK CEVAPLAYINIZ.

	A	B	C	D	E	Boş	N
Üst Grup	18	30	10	10	10	22	100
Alt Grup	2	26	15	14	10	33	100

5. Yukarıda analizi verilen maddede doğru cevap hangi seçenekte olursa ayırt edicilik en yüksek olur? (2006)

A) A B) B C) C D) D E) E

6. Yukarıda analizi verilen maddenin doğru cevabı E ise, madde ayırıcılık gücü indeksi kaç olur? (2006)

A) 0,00 B) 0,10 C) 0,16 D) 0,20 E) 1,00

7. Yukarıdaki madde analizi tablosunda verilen bilgilere dayanarak, bu madde hakkında aşağıdaki yargılardan hangisine ulaşılabilir? (2006)

A) A seçeneği çeldirici ise, beklenen yönde çalışmıştır.
B) Maddeyi üst ve alt gruplardan eşit sayıda aday cevaplamıştır.
C) B seçeneği çeldirici ise, çok az kişiyi çeldirmiştir.
D) Doğru cevap C ise, soru negatif ayırıcıdır.
E) Seçeneklerin tümü dengeli bir şekilde çalışmıştır.

8. - 9. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Bir testteki maddelere ilişkin aşağıdaki değerler elde edilmiştir.

Maddeler	I	II	III	IV	V
Madde güçlülüğü	0,1	0,15	0,5	0,6	0,9
Madde ayırıcılık gücü	0,08	– 0,15	0,5	0,28	0,43

8. Tabloda verilen bilgilere göre hangi maddenin güvenirliğe katkısı en yüksektir? (2008)

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

9. Yapılandırılacak bir testte bu maddeleri kullanma zorunluluğu olsaydı, en köklü değişikliğin hangi maddede yapılması gerekirdi? (2008)

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

10. – 12. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Aşağıdaki tabloda 10 soruya ait madde güçlüğü (p) ve ayırtedicilik (r) değerleri verilmiştir.

Soru	p	r
1	0,35	0,25
2	0,05	0,65
3	0,15	0,15
4	0,90	0,05
5	0,70	-0,35
6	0,40	0,40
7	0,15	0,10
8	0,75	0,60
9	0,15	-0,40
10	0,95	0,25

10. Uygulama sonucunda yukarıdaki değerleri elde eden bir öğretmen, ayırtediciliği düşük olan maddeleri düzeltilemediği için testten çıkarmak istemektedir.

Bu maddelerin çıkarılması aşağıdakilerden hangisine neden olur? (2009)

A) Testin ölçüt geçerliğinin düşmesine
B) Kalan maddelerin güçlük indeksinin düşmesine
C) Çeldiricilerin iyi çalışmamasına
D) Testin iç tutarlığının düşmesine
E) Testin kapsam geçerliğinin düşmesine

11. Tablodaki maddelerden hangisi hem kolay hem de ayırtedici bir sorudur? (2009)

A) 3 B) 4 C) 8 D) 9 E) 10

12. Tablodaki maddelerden hangisi, testle ölçülen özellikler açısından oldukça üst düzey yeterliğe sahip olan öğrencileri seçmek için kullanılabilir? (2009)

A) 1 B) 2 C) 5 D) 7 E) 10

egitim@hotmail.es

13. – 15. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Aşağıdaki tabloda aynı kapsamı ölçmek üzere tasarlanmış beş farklı test yer almaktadır. Bu testlerin her biri beşer adet çoktan seçmeli maddeden oluşmaktadır. Tabloda bu maddelerin güçlük indeksleri verilmiştir.

Testler	Madde güçlük indeksleri				
	Madde 1	Madde 2	Madde 3	Madde 4	Madde 5
I.	0,4	0,5	0,6	0,5	0,6
II.	0,4	0,8	0,9	0,7	0,3
III.	0,3	0,4	0,5	0,2	0,9
IV.	0,8	0,7	0,6	0,5	0,9
V.	0,2	0,4	0,3	0,1	0,5

13. III. testte yer alan <u>en zor</u> madde hangisidir? (2010)

A) 1 B) 2 C) 3 D) 4 E) 5

14. V. testte madde ayırt edicilik indeksi <u>en yüksek</u> olması beklenen madde hangisidir? (2010)

A) 1 B) 2 C) 3 D) 4 E) 5

15. Yukarıdaki testler içinde <u>en kolay</u> olanı hangisidir? (2010)

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

16. Aşağıdaki tabloda beş seçenekli çoktan seçmeli bir maddenin analizi verilmiştir.

	A	B	C	D	E	Toplam
Üst grup	1	6	8	12	3	30
Alt grup	13	6	7	2	2	30

Tablodaki bilgilere göre,

I. A seçeneği doğru cevap olursa madde ayırt ediciliği en yüksek olur.
II. B seçeneği doğru cevap olursa kolay bir madde olur.
III. E seçeneği doğru cevap olursa zor bir madde olur.

ifadelerinden hangileri doğrudur? (2010-2)

A) Yalnız I B) Yalnız II C) Yalnız III
D) I ve II E) II ve III

17. Bir matematik öğretmeni aşağıdaki çoktan seçmeli maddeyi sınavında kullanmıştır.

Madde:

Evren'in ablası, Evren'den 10 yaş büyüktür. 2 yıl sonra Evren'in yaşı ablasının yaşının yarısı kadar olacaktır. Evren'in ablası kaç yaşındadır?

A) 8 B) 10 C) 18 D) 20 E) 28

Çözüm:

Evren'in yaşı x ise ablasının yaşı x + 10'dur. $\frac{x+2}{x+10+2}=\frac{1}{2}$ *ise* $2x + 4 = x + 12$ *olur.*
Dolayısıyla x = 8'dir.
Evren'in yaşı 8 ise ablasının yaşı 18'dir.
Cevap: C

Problemde sözel olarak ifade edilen durumu cebirsel denkleme dönüştüren ve bu denklemi çözen ancak yeterince dikkatli olmayan öğrenciler, Evren'in yaşının değil, ablasının yaşının sorulduğunu unutarak A seçeneğini işaretlemişlerdir.

Yukarıda açıklanan durumun, maddenin

I. doğru cevaplanma oranı,
II. ayırt edicilik düzeyi,
III. geçerlik

özelliklerinden hangilerini düşürmesi beklenir?

A) Yalnız I B) Yalnız II C) I ve II
D) II ve III E) I, II ve III

egitim@hotmail.es

18. **Aşağıda, 8. sınıf öğrencilerine uygulanan uluslararası bir sınavın fen bilimleri alt testinde yer alan çoktan seçmeli bir madde, bu maddenin ölçtüğü bilişsel düzey ve madde istatistikleri verilmiştir.**

Canlılarda büyük ve karmaşık yapıdaki moleküller parçalanarak daha basit ve küçük moleküller hâline dönüşür.

Bu olaya ne denir?

a. Boşaltım
b. Emilim
c. Sindirim
d. Dolaşım

Bilişsel düzey	Hatırlama
Madde güçlük indeksi (p)	0,21
Madde ayırt edicilik	0,61

Bu maddeyle ilgili olarak aşağıdakilerden hangisi söylenebilir? (2012)

A) Başarı düzeyi yüksek öğrencilerle düşük olanları ayırt etmekte yetersizdir.
B) Madde testten çıkarılırsa iç tutarlık katsayısı artar.
C) Testi alan öğrencilerin çoğunluğu tarafından doğru cevaplanmıştır.
D) Öğrencilerin test puanları arttıkça maddeyi doğru cevaplama olasılıkları artmaktadır.
E) Üst düzeyden bir düşünme becerisini ölçmektedir.

19. **Bir okuldaki 7. sınıf öğrencilerine 5 seçenekli bir İngilizce testi uygulanmıştır. Testteki ilk üç maddenin alt grup-üst grup yöntemiyle madde analizini gösteren tablo aşağıda verilmiştir. Altı çizili seçenekler anahtar cevaptır.**

Madde No	Grup	Öğrenci Sayısı	Seçenekler				
			A	B	C	D	E
1	Üst	25	0	15	6	4	0
	Alt	25	4	12	3	6	0
2	Üst	25	2	2	20	1	0
	Alt	25	5	8	6	2	4
3	Üst	25	4	3	9	4	5
	Alt	25	6	4	5	2	8

Bu maddelerle ilgili olarak aşağıdakilerden hangisi söylenebilir?(2012)

A) Madde 1 ve madde 3'ün ayırt edicilikleri birbirine eşittir.
B) Madde 2, madde 1'den daha ayırt edici bir maddedir.
C) Madde 1, diğer ikisine göre daha zordur.
D) İlk üç maddeden alınan puanların ortalaması 1,25'tir.
E) Madde 3'ün C çeldiricisi, beklenen yönde işlemiştir.

TEST PUANLARI ÜZERİNDE İSTATİSTİKSEL İŞLEMLER

7

ANAHTAR KAVRAMLAR

Veri, grup sayısı, grup aralık sayısı, gerçek puan aralığı, grup orta noktası, grupların grafikle gösterimi, betimleyici istatistikler, merkezi yığılma ölçüleri, mod, medyan, aritmetik ortalama, merkezi dağılım ölçüleri, ranj, çeyrek sapma, standart sapma, bağıl değişkenlik katsayısı.

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Topladığınız verileri düzenleyebilecek,
- Düzenlenen verilerin grafik gösterimini yapabilecek,
- Merkezi yığılma ölçülerini (mod, medyan, aritmetik ortalama) hesaplayabilecek,
- Merkezi yığılma ölçüleri arasındaki ilişkiyi açıklayabilecek,
- Merkezi dağılım ölçülerini (ranj, çeyrek sapma, standart sapma) hesaplayabileceksiniz.

İÇİNDEKİLER

Verilerin düzenlenmesi
Merkezi yığılma ölçüleri
Merkezi yığılma ölçüleri arasındaki ilişki
Merkezi dağılım ölçüleri
Çalışma soruları

BAŞLAMADAN ÖNCE

Bir sınavdan alınan puanlar, sınavı uygulayanlara ne tür bilgiler sunar hiç düşündünüz mü?

egitim@hotmail.es

TEST PUANLARI ÜZERİNDE İSTATİSTİKSEL İŞLEMLER

Verilerin Düzenlenmesi

Yukarıda açıklanan madde analizi sürecinin uygulanmasıyla olabildiğince güvenilir ve geçerli maddelerden oluştuğu düşünülen testin uygulanması sonucu öğrencilerden elde edilen puanlar, üzerlerinde hiç bir düzenleme yapılmadan, bazı işlemler uygulanmadan bir puan yığını olmanın dışında bize hiç bir bilgi sunmaz. Bunu bir örnek vererek açıklamaya çalışalım.

Örneğin; aşağıdaki tabloda 50 öğrencinin tarih dersine ait ölçme sonuçları görülmektedir:

Tablo 1: Tarih Dersine Ait Ölçme Sonuçları

(50 öğrencinin 100 üzerinden aldığı puanlar)

Öğrenci	Puanı	Öğrenci	Puanı	Öğrenci	Puanı	Öğrenci	Puanı	Öğrenci	Puanı
1	45	11	79	21	45	31	43	41	65
2	69	12	53	22	65	32	53	42	69
3	69	13	49	23	69	33	59	43	65
4	55	14	59	24	53	34	53	44	69
5	43	15	55	25	77	35	49	45	53
6	77	16	53	26	87	36	31	46	77
7	60	17	49	27	16	37	45	47	45
8	39	18	37	28	25	38	49	48	60
9	23	19	35	29	35	39	53	49	59
10	15	20	27	30	39	40	36	50	39

Yukarıdaki tabloda öğrencilerin tarih sınavından aldıkları puanlar karışık bir şekilde verilmiştir. Bu tabloya bakarak öğrencilere ilişkin bilgide bulunmak, sınıfın başarı düzeyi vb. hakkında yorum yapmak oldukça zordur. Hâlbuki bu tablo aşağıda Tablo 2'de görüldüğü şekilde en düşük puandan en yüksek puana doğru düzenlenebilir. Hatta puanların sıraya dizilmesiyle birlikte her puanın kaç öğrenci tarafından alındığını gösteren *frekans* sütunu da oluşturulabilir.

Tablo 2'de puan ve frekans sütununun yanında bir de "tf" ile gösterilmiş bir sütun daha verilmiştir. Bu sütun "toplam frekans" bir başka deyişle "yığılmalı frekans" değerlerini gösteren sütundur. Bu sütundaki değerler her bir puanın sahip olduğu frekans ile önceki frekansların toplanmasıyla bulunur. Örneğin; 16 puanının frekansı 1, bir önceki puanların toplam frekansı (ki sadece tek bir puan vardır) 1 olduğu için ikisinin toplamından 16 puanının toplamı 2 olarak hesaplanır. 49 puanının frekansı 4, önceki puanların toplam frekansı 19 olduğu için 19 + 4 = 23 olarak 49 puanının toplam frekansı hesaplanmış olur.

Aynı zamanda toplam frekans, her bir puandan önce kaç tane puan olduğunu ve her bir puanın kaçıncı puan olduğunu da görebilmemizi sağlayan bir değerdir. Örneğin; 23 puanı 3. sırada yer alır ve 23'ten önce iki puan bulunmaktadır; 11., 12., ve 13. puanlar 39'tur ve 39'dan önce 10 puan daha bulunmaktadır vb.

Tablo 2: Tarih sınavı puanlarının en düşükten en yükseğe doğru sıralanarak frekanslarıyla verilmesi

puan	frekansı	tf	puan	frekansı	tf
15	1	1	45	4	19
16	1	2	49	4	23
23	1	3	53	7	30
25	1	4	55	2	32
27	1	5	59	3	35
31	1	6	60	2	37
35	2	8	65	3	40
36	1	9	69	5	45
37	1	10	77	3	48
39	3	13	79	1	49
43	2	15	87	1	50

Tablo 2'de en düşük puan ve en yüksek puanla birlikte aradaki tüm puanların kaç öğrenci tarafından alındığı açıkça görülmekle birlikte, öğrencilerin puanlarının hangi puanlarda yoğunlaştığı, bir başka deyişle puanların ağırlık merkezi de daha kolay bir şekilde görülebilmektedir.

Her ne kadar bu şekilde düzenlenmiş bir tabloyla sınıfın genel durumu hakkında bilgi sahibi olmak mümkün olsa da zaman zaman puanların grafiklerle görsel olarak ifade edilmesi çok daha bilgi verici olabilir. Bu nedenle puanların grafiklerle gösteriminin nasıl yapılabileceği konusunda da bilgi vermeye çalışalım.

Tablo 2'de görüldüğü şekilde puanlar düzenlendikten sonra, puanların grafiklerle gösterilmesi için aşağıdaki adımlar izlenir:

1. **Grup sayısının belirlenmesi:** Yukarıdaki örnekte öğrenci sayısı 50'dir ve 22 farklı puan bulunmaktadır. Puan sayısı çok fazla olduğunda puanların grafikte tek tek gösterilmesi zor olabilir bunun yerine puanlar gruplanarak gruplar halinde grafikte göstermek çok daha pratik olacaktır. Puanların kaç gruba ayrılacağı uygulayıcının kararına bağlı olmakla birlikte genellikle sonraki işlemlerde kolaylık sağlaması için grup sayısı 'tek sayı' olacak şekilde tercih edilir. Biz de burada grup sayısını 9 olarak seçeceğiz.
2. **Grup aralık katsayısının belirlenmesi:** Puanların gruplandırılacağı aralığın genişliğini, grup aralık katsayısı belirler. Aralık genişliğinin hesaplanması için en yüksek puandan en düşük puanı çıkarılarak tercih edilen grup sayısına bölünür: (87-15)/9=72/9=8. Böylelikle grup sayısı 9 olarak belirlenen puanlar en küçük

egitim@hotmail.es

puandan itibaren 8 arttırılarak gruplandırılacaktır. Bu durumda ilk grup en düşük puan olan 15 ile başlarken ikinci grup 15+8=23 ile başlayacak, üçücü grubun başlangıç değeri ise 23+8=31 olacaktır.

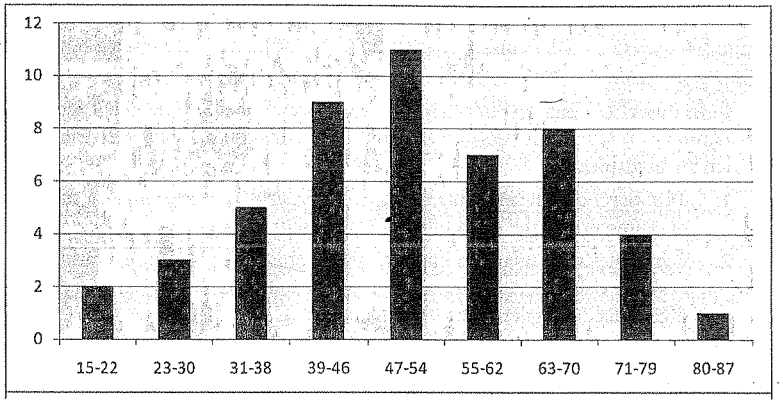
3. **Puanların gerçek puan aralıklarının belirlenmesi:** Başarı sürekli değişken olduğundan histogram, çizgi, frekans poligonu (çokgeni) gibi grafik türlerinden biri kullanılabilir. Ancak yukarıdaki grup aralık katsayılarına göre örneğin histogram çizilmek istense şekil 1'de görülen kesikli bir grafik ortaya çıkar. Hâlbuki amacımız sürekli bir grafik elde etmektir. Bu sebeple grup aralık katsayılarından yararlanılarak gerçek grup aralıkları belirlenir ki puanların sürekliliği sağlanabilsin. İkinci adımda elde edilen grup aralıklarının alt sınırından 0,5 puan çıkarılarak ve üst sınırına 0,5 puan eklenerek puanlar sürekli hale getirilir. Bu şekilde elde edilen yeni değerlerle gerçek puan aralıkları belirlenmiş olur. Örneğimizin gerçek grup aralıkları Tablo 3'te 3. sütunda görülmektedir.
4. **Grup orta noktasının belirlenmesi:** Her bir grup aralığının alt değeri ile üst değeri toplanarak yarısı alınarak grup orta noktaları bulunur. Grup orta noktaları Tablo 3'te 4. sütunda görülmektedir.
5. **Grup aralıklarının frekanslarının belirlenmesi:** Tablo 2'de gösterilen frekans gibi grup aralıklarının frekansları da benzer şekilde hesaplanır. Her bir grup aralığında bulunan puanların kaç öğrenci tarafından alındığı hesaplanarak her bir grubun frekansı belirlenir. Örneğin; ilk grup olan 15 ile 22 puanları arasında puan alan öğrenci sayısı Tablo 2'den de görüleceği üzere 2'dir.

Böylece Tablo 2'deki puanlara yukarıdaki basamaklardaki işlemler uygulanarak aşağıdaki Tablo 3 elde edilir.

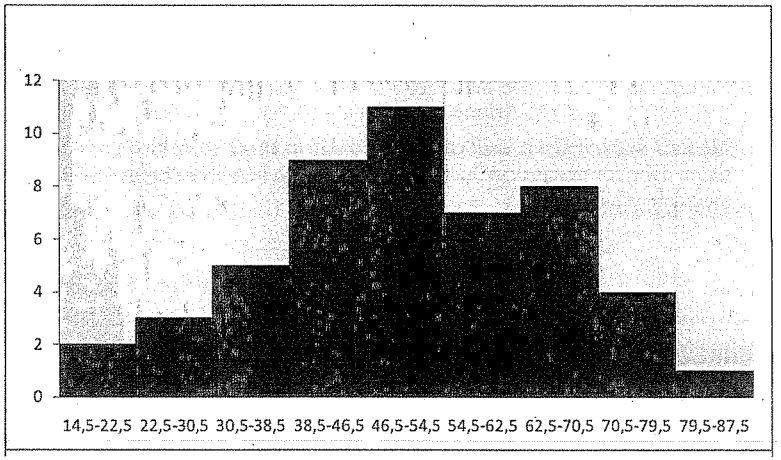
Tablo 3: Sıralı Verilerin Gruplandırılması

Gruplar	Grup Aralığı	Gerçek Grup Aralığı	Grup Orta Noktası	Frekans	tf
1. Grup	15-22	14,5-22,5	18,5	2	2
2. Grup	23-30	22,5-30,5	26,5	3	5
3. Grup	31-38	30,5-38,5	34,5	5	10
4. Grup	39-46	38,5-46,5	42,5	9	19
5. Grup	47-54	46,5-54,5	50,5	11	30
6. grup	55-62	54,5-62,5	58,5	7	37
7. Grup	63-70	62,5-70,5	66,5	8	45
8. Grup	71-78	70,5-78,5	74,5	4	49
9. Grup	79-87	78,5-87,5	82,5	1	50

6. **Puanların grafikle gösterimi:** Tablo 3'teki puanlar kullanılarak aşağıda gösterilen histogram grafiği elde edilir. Histogram çizilirken puanların gerçek grup aralıkları yatay eksende frekans değerleri düşey eksende gösterilir.



Şekil 1: Grup Aralıklarına Göre Bar Grafiği (4)



Şekil 2: Gerçek Grup Aralıklarına Göre Puanların Histogramı (5)

Şekil 2'de yer alan histograma bakıldığında, puanların yığıldığı puan aralıklarını, öğrenciler tarafından en fazla ve en az alınan puan aralıklarını, puanların hangi aralıkta dağıldığını vb. görmek mümkündür. Bu tür görsel olarak bize bilgi sunan grafiklerle öğrencilerin başarı durumu hakkında genel bir izlenime sahip olmak mümkündür.

egitim@hotmail.es

Sınıftaki öğrenmeler hakkında bilgi sahibi olmak, öğrencilerin durumunu takip etmek vb. hakkında bilgi sahibi olabilmek için puanların yukarıdaki şekilde düzenlenmesiyle birlikte puanlara ilişkin istatistiksel işlemler yapmak da oldukça açıklayıcı bilgiler vermektedir.

7. **Betimleyici (Açıklayıcı) istatistiklerin hesaplanması:** Betimleyici (özetleyici, tanımlayıcı) istatistikler, puan dağılımının yapısı hakkında bilgi verir. Elde edilen bu istatistikler:
   1. Tüm öğrencilerin genel olarak başarı-başarısızlık durumunu ya da sınıftaki öğrenmenin yüksek ya da düşük düzeyde olduğunu,
   2. Öğrencilerin puanlarının birbirine yakın (benzer) ya da uzak (farklı), bir başka deyişle puanların homojen mi yoksa heterojen bir dağılım mı gösterdiğine ilişkin bilgiler verir.

Başarı- başarısızlık hakkında bilgi veren istatistikler "Merkezi yığılma ölçüleri" başlığı altında toplanırken, homojen-heterojen dağılım hakkında bilgi veren istatistikler ise 'Merkezi dağılım ölçüleri" olarak adlandırılırlar.

- ***Merkezi yığılma ölçülerinin*** en basit hesaplanan ancak en az bilgi veren istatistikten daha zor hesaplanan ancak en iyi bilgi veren istatistiğe doğru sıralanışı aşağıdaki şekildedir:

  *Mod (tepe değer)

  **Medyan (ortanca)

  ***Aritmetik ortalama

- ***Merkezi dağılım ölçülerinin*** en basit hesaplanan ancak en az bilgi veren istatistikten daha zor hesaplanan ancak en iyi bilgi veren istatistiğe doğru sıralanışı ise aşağıdaki şekildedir:

  *Ranj (dizi genişliği)

  **Çeyrek sapma

  ***Standart sapma - Varyans

MERKEZİ YIĞILMA ÖLÇÜLERİ

Puanların merkezi yığılma ölçüleri, öğrencilerin başarı-başarısızlıkları hakkında bilgi veren değerler olup, puanların hangi merkez çok bilgi verene doğru sıralanarak verilecektir. Bu cümlede yer alan "en az bilgi – en çok bilgi" ifadeleri aşağıda daha net olarak anlaşılacaktır.

1) **Mod (Tepe Değer):** Puanlar içinde en çok tekrarlanan puandır. Bir başka deyişle, frekansı en yüksek puandır. Puan dağılımında bu şekilde "en çok tek alınan" puan bulunuyorsa puanların dağılımının yoğunluğu bu puana doğru yaklaşır.

Örneğin; Tablo 2'de yer alan düzenlenmiş puanlara bakarsakı en fazla tekrarlanan puanın 53 olduğunu görebiliriz (yedi kez tekrarlanmıştır). O halde örneğimizde verilen puanların modu (tepe değeri) 53'tür.

- Tablo 2'deki puanlara ilişkin genel bir yorum yapıldığında elli öğrencinin puanlarının 53 etrafında yığıldığı söylenebilir. Buradan da öğrencilerin başarılarının orta düzeyde olduğu sonucuna varılabilir.
- Eğer bir puan dağılımında en fazla alınan bir puan (mod) varsa, puan dağılımı mod etrafında yoğunlaşacaktır. Ancak bazen en çok tekrarlanan puan birden fazla olur ki bu durumda tek bir mod bulmak mümkün değildir. Tek bir modun bulunmadığı puan dağılımında birden çok tepe değer olacaktır, bu dağılımın da ağırlık merkezinin hangi puan olduğunu belirlemek de zorlaşacaktır. Böylelikle anlaşılacağı üzere, bazı puan dağılımlarının modunu bulmak mümkün olmamaktadır. Bu nedenl puanların "mod"unu bulmak çok basit bir işlem gerektiriyor olmakla birlikte bazen merkezi yığılım hakkında bilgi verebilecek özellikte olmamaktadır. Bu nedenle "mod" merkezi yığılma ölçüleri içinde en basit bulunabilen ancak en az bilgi veren özelliktedir.

!!! Puanların frekanslarına bakılarak, en yüksek frekansa karşılık gelen puanın mod olarak belirlenmesi yoluyla modun çok kolay bulunabilmesi "mod"un olumlu yanı iken, puanların ağırlık merkezinin yani öğrencilerin başarısının hangi puanda yoğunlaştığının tek bir puana bakılarak belirlenmesi, modun önemli eksiklikleri arasında yer alır.

Moda göre daha iyi bilgi veren merkezi yığılma ölçüsü: MEDYAN'dır.

2) **Medyan (Ortanca):** Sıralanmış verilerin tam ortasında yer alan puandır. Bir başka deyişle, sıralanan puanları %50 iki yarıya ayıran puandır. Sıralanmış puanları iki eşit yarıya ayıran puan (medyan), dağılımın ağırlık merkezini "mod"a göre daha iyi göstermektedir.

(Dikkat! Bu tanıma göre, acaba puanları küçükten büyüğe mi büyükten küçüğe mi sıralamalayız? **Dikkatli düşününüz!)**

Ortancanın hesaplanması, puan sayısının tek ya da çift sayı olmasına göre değişiklik gösterir:

Tek sayıda puan varsa:

Puan dağılımındaki puan sayısının tek olması durumunda medyan kolayca hesaplanabilir. Sıralanmış puan dağılımını iki eşit parçaya bölen tek bir puan bulunur ve bu puan da medyandır. Örneğin; dokuz öğrencinin yer aldığı bir sınavdan elde edilen pualar sıraya dizildiğinde tam ortada yer alan tek bir puan vardır ve bu puan beşinci puan olacaktır. Beşinci puanın altında ve üstünde yer alan eşit sayıda (dörder puan) puan vardır. Bu durumu daha net açıklamaya çalışalım:

Örneğin; dokuz öğrencinin bir sınavdan aldıkları puanlar 11, 18, 21, 19, 12, 21, 13, 11,15 olsun. Bu puanlar sıralandıktan sonra tam ortada yeralan puan 15 olacaktır. Böylece bu puan dağılımının medyanı (ortancası) 15 olarak bulunur.

Çift sayıda puan varsa:

Puan dağılımındaki puan sayısının çift olması durumunda medyanı hesaplamak biraz farklılaşır. Puan sayısı çift olduğunda, sıralanmış puan dağılımını iki eşit parça-

egitim@hotmail.es

ya bölen tek bir puan doğrudan bulunmamaktadır. Bu durumda tam ortada yer alan iki puan vardır. Medyan da bu iki puanın ortalaması olacaktır. Örneğin: on öğrencinin yer aldığı bir sınavdan elde edilen pualar sıraya dizildiğinde tam ortada yer alan tek bir puan yoktur ve tam ortada beşinci ve altıncı puanlar yer alır. Böylece beşinci puan ile altıncı puan toplanır ve ikiye bölünerek medyan (ortanca) basitçe hesaplanmış olur. Bu durumu daha net açıklamaya çalışalım:

Örneğin; on öğrencinin bir sınavdan aldıkları puanlar 11, 18, 21, 19, 12, 21, 13, 11,15, 14 olsun. Bu puanlar sıralandıktan sonra tam ortada yeralan puanlar 14 ve 15 olacaktır. Böylece bu puan dağılımının medyanı (ortancası) 14+15/2= 14,5 olarak bulunur.

Eğer sıralanmış puanlar (Tablo 2) frekansları ile verilmişse, puan dağılımındaki puan sayısının tek ya da çift olması da dikkate alınarak medyan (ortanca) aşağıdaki şekilde hesaplanır.

$$X_{ort} = A_{ort} + \left(\frac{\frac{n}{2} - tf_A}{tf_b} \right) . a \qquad (1)$$

Formülde kullanılan terimlerin açıklaması:

X_{ort}: Ortanca değeri

A_{ort}: Ortancanın bulunduğu aralığın gerçek alt sınır değeri

n: Puan (veri) sayısı

tf_A: Ortancanın bulunduğu aralığa kadar olan toplam frekans sayısı

tf_b: Ortancanın bulunduğu aralığın frekans sayısı

a: Aralık katsayısı

(1) eşitliği incelendiğinde, Tablo 2'teki değerlerden yararlanarak medyan şu şekilde hesaplanacaktır:

Tablo 4 : Sıralı Verilerin Gruplandırılması

Gruplar	Grup Aralığı	Gerçek Grup Aralığı	Grup Orta Noktası	Frekans	Tf
1. Grup	15-22	14,5-22,5	18,5	2	2
2. Grup	23-30	22,5-30,5	26,5	3	5
3. Grup	31-38	30,5-38,5	34,5	5	10
4. Grup	39-46	38,5-46,5	42,5	9	19
5. Grup	47-54	46,5-54,5	50,5	11	30
6. grup	55-62	54,5-62,5	58,5	7	37
7. Grup	63-70	62,5-70,5	66,5	8	45
8. Grup	71-78	70,5-78,5	74,5	4	49
9. Grup	79-87	78,5-87,5	82,5	1	50

1. Medyanın hesaplanması için öncelikle 50 puanın tam ortasında yer alan puanın kaçıncı puan olduğunu bulalım: 50 puanı tam iki yarıya ayıran puanlar 25. ve 26. puanlarıdır. Böylelikle ortancanın 5. gurpta 47-54 puan aralığında yer aldığı (toplam frekans sütununa bakınız) görülecektir.
2. 47 - 54 puan aralığının gerçek sınır değerleri 46,5-54,5 olup alt sınır değeri (A_{ort})=46,5'tir.
3. Ortancanın bulunduğu aralığa kadar olan toplam frekans sayısı: (tf_A)=19.
4. Ortancanın bulunduğu puan aralığının frekansı (tf_b)= 11 ve
5. Grup aralık katsayısı (a)= 8 (Veriler gruplandırılmamış, tek tek verilmişse bu durumda grup aralık katsayısı 1 olarak alınır.)
6. Puan sayımız da n=50'dir.

Puan (veri) sayı tek olsaydı formüldeki "n" değeri 1 ile toplanarak 2'ye bölünecekti.

Böylelikle tüm bu değerleri yerine koyarak Tablo 4'te yer alan puanların medyanını hesaplarsak;

$$X_{ort} = 46,5 + \left[\frac{\frac{50}{2} - 19}{11} \right] \cdot 8 = 50,86$$

buluruz.

Eğer gruplandırılmış puanların (verilerin) (tablo 4'teki puanlar) medyanını bu formülle uğraşmadan kısaca bulmak istersek gruplandırılmış puan aralıkları içinde en yüksek frekansa sahip puan aralığı bulunur (5. grup: 47-54) bu puan aralığını temsil eden grup orta noktası 50,5 medyan olarak kabul edilebilir.

!!! Puanların medyanı bulunurken puanlar arasındaki farklılık göz önünde bulundurulmaz ve puanlar sadece büyüklüklerine göre sıraya konulur. Bu yönüyle yetersiz kalmaktadır. Ancak 'mod' a göre daha iyi bilgi veren bir merkezi eğilim ölçüsüdür.

!!! Tüm puanların dikkate alınarak hesaplamanın yapıldığı bir değer olması nedeniyle aritmetik ortalama; mod ve medyana göre daha duyarlı, daha üst düzey bilgi veren bir değerdir. Ancak aşırı uç değerleri bulunduğunda (dağılım sağa ya da sola çarpık olduğunda) aritmetik ortalama yerine tercih edilen merkezi eğilim ölçüsüdür.

3) Aritmetik ortalama: Puanların toplamının puan sayısına (öğrenci sayısına)bölünmesiyle elde edilen değerdir. Kısaca "ortalama" olarak da bilinir. *Ancak "ortanca" ile karıştırılmamalıdır.*

egitim@hotmail.es

Aritmetik ortalama aşağıdaki eşitlikle hesaplanır:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$\overline{X}$: Aritmetik Ortalama

$\Sigma(X)$: Puanların toplamı

n: Puan (öğrenci) sayısı

Örnek verilerimiz için ortalama hesaplandığında, puanların toplamı 2574'tür. Toplam 50 puan olduğuna göre aritmetik ortalama=2574/50 =51,48'dir.

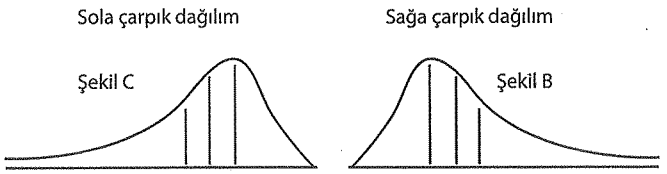
- 51,48 puanların ağırlık merkezi konumundadır. Dağılımda yer alan puanlar bu aritmetik ortalama değeri etrafında yoğunlaşmaktadır.
- Hesaplanan ortalama değeri, öğrencilerin başarı düzeyi hakkında fikir sahibi olmamızı sağlar. Verilen örnekte en yüksek puanın 87 olduğu göz önüne alındığında, Türkçe dersinden sınav olan 50 öğrencinin başarı düzeyinin yüksek olduğu yorumu yapılabilir.

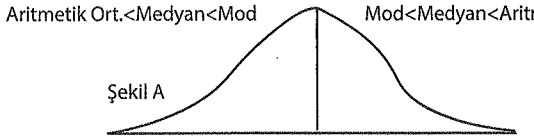
MERKEZİ YIĞILMA ÖLÇÜLERİ ARASINDAKİ İLİŞKİ

Puan dağılımında en iyi merkezi yığılma ölçüsü şudur gibi bir yorum yapmak pek doğru bir yaklaşım değildir. Dağılımın şeklinde göre kullanılması tercih edilecek yığılma ölçüsü değişmektedir. Örneğin; eğer şekil A'daki gibi dağılım simetrikse; mod, medyan ve aritmetik ortalama değerleri birbirine eşit olacaktır (Simetrik dağılım). Bu durumda bu ölçülerden birini tercih etmek söz konusu olmayıp aynı anda üç değer birlikte yorumlanmış olacaktır. Dikkat edilirse simetrik dağılımda öğrencilerin yarısı (aritmetik) ortalamanın altında puan alırken diğer yarısı da ortalamanın üstünde puan almışlardır (Doğal olarak "medyan"ın da tanımı gereği). Eğer puanlar simetriğe yakın bir dağılıma sahipse, ortalama ya da ortancayı tercih etmek daha doğru olacaktır ki genelde de ortalama daha çok tercih edilir.

Ancak bazı durumlarda medyan ya da mod değeri ortalamaya tercih edilebilir. Örneğin; şekil B'de görüleceği gibi, yüksek puan alan öğrenci sayısı oldukça az, düşük puan alan öğrenci sayısı oldukça fazla ise mod, medyan ve ortalama birbirinden farklı değerlere sahip olacaktır (Sağa çarpık dağılım ya da pozitif kayışlı dağılım). Puanların çoğunluğuna göre daha düşük olan uçta yer alan puanlar mod ve meyana göre daha çok ortalamayı etkiler. Puanların ortalaması puanların ağırlık merkezi konumunda olduğu için ortalama uçtaki değerlere doğru kayma eğilimi gösterir. Bu nedenle, ortalama uçta yer alan yüksek puanlara yakın bir değer olup, mod ve medyandan daha yüksek değere sahip olacaktır. Mod en düşük değerken; medyan, mod ve ortalamanın arasında bir değere sahip olur.

Benzer şekilde, şekil C'de görüleceği gibi, düşük puan alan öğrenci sayısı oldukça az, yüksek puan alan öğrenci sayısı oldukça fazla ise mod, medyan ve ortalama yine birbirinden farklı değerlere sahip olacaktır (Sola çarpık dağılım ya da negatif kayışlı dağılım). Puanların çoğunluğuna göre daha düşük olan uçta yer alan puanlar mod ve meyana göre daha çok ortalamayı etkiler. Puanların ortalaması puanların ağırlık merkezi konumunda olduğu için ortalama uçtaki değerlere doğru kayma eğilimi gösterir. Bu nedenle, ortalama uçta yer alan düşük puanlara yakın bir değer olup, mod ve medyandan daha düşük değere sahip olacaktır. Mod en yüksek değerken; medyan, mod ve ortalamanın arasında bir değere sahip olur.





Simetrik dağılım

Mod=Medyan=Aritmetik Ort.

Yukarıdaki tablolarda verilen 50 öğrencinin Türkçe dersi sınavından aldığı puanların mod, medyan ve aritmetik ortalama değerlerine bakarsak, mod: 53, medyan: 53 ve aritmetik ortalam: 51,5 olup (Artimetik Ort.<Medyan<Mod) puan dağılımı sola çarpık bir dağılım gösterecektir. Bu durumda öğrencilerin yarısından fazlasının ortalamanın üzerinde puan aldığı, sınıfın başarılı olduğu yorumu yapılabilir.

MERKEZİ DAĞILIM ÖLÇÜLERİ

Merkezi dağılım ölçüleri, puanların aritmetik ortalamay göre hangi aralıkta dağıldığını gösteren değerlerdir. Bir başka deyişle puanların birbirine göre değişimlerinin, puan farklılıklarının bir ölçüsünü verirler. Puanların birbirinden farkı büyüdüğü ölçüde puan dağılımının genişliği artacaktır. Puanlar arasındaki farklılık arttıkça puanların heterojenliği artarken, puanlar arasındaki farklılık azaldıkça (puanlar birbirine yakın değerlerse) puaların homojenliği artar.

egitim@hotmail.es

1) **Dizi genişliği (Ranj):** Puanların hangi aralıkta değiştiğini gösteren en basit merkezi dağılım ölçüsüdür. Bir başka deyişle, en yüksek puan ile en düşük puan arasındaki farktır.

"Ranj= En yüksek puan – En düşük puan" eşitliği ile hesaplanır.

Örneğin, yukarıda verilen tablolarda yer alan puanların ranjını hesaplarsak; görüleceği gibi en yüksek puan 87 ve en düşük puan 15 olup, puanların ranjı= 87 -15= 72 olarak bulunacaktır. Hesaplanan bu değer, Türkçe sınavına giren öğrencilerin puanların birbirinden farklı olduğunu gösterir. Bu sınavda düşük puan alan öğrencilerin yanı sıra yüksek puan alan öğrenciler de bulunmaktadır.

Bir merkezi dağılım ölçüsü olan ranj, kolay hesaplanmasıyla bir avantaj sağlarken sahip olduğu dezavantajları nedeniyle de pek sık kullanılan bir dağılım ölçüsü değildir: Ranj, tamamen uçta yer alan puanlara dayalı hesaplandığı için hatalı yorum yapmaya açık bir dağılım ölçüsüdür. Bu durumu aşağıdaki örnekle açıklamaya çalışalım:

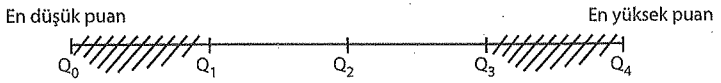
Örnek1: 10 öğrencinin aldığı puanlar: 5, 20, 35, 45, 50, 55, 55, 70, 95, 100 ise puanların ranjı= 100-5= 95'tir.

Örnek 2: 10 öğrencinin aldığı puanlar: 5, 60, 65, 65, 65, 70, 70, 70, 75, 100 ise puanların ranjı= 100-5= 95'tir.

Şimdi örnek 1 ve 2'deki 10'ar öğrencinin aldığı puanların ranjına bakıldığında her iki durumda da öğrencilerin puanları arasındaki farklılığın (değişimin) aynı olduğu yorumu yapılacaktır. Halbuki puanların kendisine bakıldığında durumun hiç de böyle olmadığı, örnek 1'de puanların farklılığının çok daha fazla olup heterojen bir durum gösterdiği, örnek2'deki puanların ise (uçtaki değerler olmasa) çok daha birbirine benzer, homojen bir dağılım gösterdiği fark edilecektir. Bu örneklerden de anlaşılacağı üzere ranj, uç değerlerden oldukça etkilenen bir ölçü olup bazı durumlarda yanıltıcı yorumlara sebep olabilmektedir.

2) **Çeyrek sapma:** Puanların değişkenliğinin bir ölçüsü olan ranjın uç değerlerden etkilenmesi kısıtlılığını gideren bir ölçüdür. Çeyrek sapma, %75. puan ile %25. puan arasındaki farkın yarısı hesaplanarak elde edilir. Daha açık olarak şu şekilde açıklamaya çalışalım:

Öncelikle puanlar küçükten büyüğe sıralanır. Sıralanan pualar dört parçaya ayrılır (çeyreklere). İlk çeyrek ve son çeyrek atıldıktan sonra geriye kalan puanların en yükseğinden en düşüğü çıkarılarak ikiye bölünür, elde edilen değer çeyrek sapmayı verir.



Q sembolü ile gösterilen çeyrek sapma aşağıdaki formül ile hesaplanabilir.

$Q=(Q_3-Q_1)/2$

Q: Çeyrek sapma

Q_3: Üçüncü çeyrek (küçükten büyüğe doğru sıralı verilerde %75'inci puan)

Q_1: Birinci çeyrek (küçükten büyüğe doğru sıralı verilerde %25'inci puan)

Ortanca %50'inci (Q_2 ya da ikinci çeyrek) puan olarak alındığında, Q_1 ve Q_3'ün ortancanın hesaplanması ile tamamen aynı yöntem kullanılarak hesaplanabilir. Tablo 4'teki veriler kullanılarak Q_1 ve Q_3 aşağıdaki gibi hesaplanmıştır. Tablo 4'te birinci çeyrek (12.5'inci puan 13. puan alınabilir) tek bir puana karşılık geldiğinden (13'üncü sıradaki 39 puanı) Q_1=39'dur. Üçüncü çeyrek olan Q_3 de (37.5'inci puan 38. puan alınabilir) 38'inci sıradaki 65'e karşılık gelmektedir. Buna göre çeyrek sapma aşağıda verilen şekilde hesaplanmaktadır.

$$Q_3 = L + \left(\frac{\%75.n - t_{fa}}{f_b}\right).a \quad \text{ve} \quad Q_1 = L + \left(\frac{\%25.n - t_{fa}}{f_b}\right).a$$

L: istenilen %'lik dilime karşılık gelen aralığın gerçek alt sınırı

t_{fa}: karşılık gelen aralığa kadar olan toplam frekans

f_b: karşılık gelen aralığın frekansı

a: grup aralık katsayısı

Hesaplandıktan sonra $Q=(Q_3-Q_1)/2$ eşitliği ile çeyrek sapma değeri bulunmuş olur.

Buna göre örneğimizin çeyrek sapmasını hesaplarsak:

$$Q_3 = 62.5 + \left(\frac{\%75.50 - 37}{8}\right).8 = 63 \quad \text{ve} \quad Q_1 = 38.5 + \left(\frac{\%25.50 - 10}{9}\right).8 = 36.3$$

Buradan Q= 63 – 36.3 / 2 =13.4 bulunur.

Çeyrek sapma hesaplanırken ilk çeyrek ve son çeyrek hesaplamaya katılmadığı için çeyrek sapmanın ranjda olduğu gibi puanların uç değerlerinden etkilenmesi durumu ortadan kalkar. Ancak bu kez de puanaların yarısı göz önüne alınmamaktadır.

egitim@hotmail.es

egitim@hotmail.es

Ancak puan dağılımının sağa ya da sola çarpık olduğu durumlarda yığılma ölçüsü olarak medyanın tercih edilmesi gibi dağılım ölçüsü olarak da çeyrek sapma tercih edilebilir.

Çeyrek sapma da ranj gibi standart bir değer olmayıp yorumlanması güçtür. Bu nedenle her ikisine göre çok daha yaygn kullanılan merkezi dağılım ölçüsü standart sapma ve bunun karesi olan varyans değeridir.

3) **Standart sapma:** En çok tercih edilen ve en yaygın kullanılan merkezi dağılım ölçüsüdür. Puanların aritmetik ortalamadan farklılıklarının (uzaklıklarının) standart değerini verir. Puanların farklılığı arttıkça standart sapma değeri artar, puaların farklılığı azaldıkça (benzerliği arttıkça) standart sapma değeri azalır.

Standart sapma aşağıda verilen eşitlik yardımıyla hesaplanır:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

Eşitlikten de görüleceği üzere standart sapma (S_x), her bir puanın (X_i) aritmetik ortalamadan ($\bar{X}$) farklarının karelerinin toplamının puan sayısının (n) bir eksiğine bölünerek, karekökünün alınmasıyla hesaplanır. Bu eşitliğin karesi (karekökün olmadığı eşitlik) puanların varyansını verir. Standart sapma puanların orijinal birimi cinsinden bir değer olduğu için dağılım ölçüsü olarak daha çok tercih edilen bir değerdir.

Yukarıdaki eşitlik, aritmetik ortalama değerinin bilinmesini gerektirmektedir. Eğer ortalama bilinmiyor ise aşağıdaki eşitlik yardımıyla da standart sapma değeri hesaplanabilir:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} X_i^2 - \frac{\left[\sum_{i=1}^{n} X_i\right]^2}{n}}{n-1}}$$

Tablo 2'deki sıralı verilere ait standart kaymanın hesaplanabilmesinde işlem kolaylığı için, tabloya puanların karelerinin ve bunların frekansları ile çarpımlarının bulunduğu iki sütun eklenmiş ve Tablo 5'te verilmiştir. Tablo 5'te son sütunun altında ise puanların karelerinin frekansları ile çarpımlarının toplamı verilmiştir. Buradaki ve aritmetik ortalamanın hesaplanmasındaki toplam puanlar (2574) kullanılarak puanların standart kaymaları aşağıda hesaplanmıştır.

Tablo 5: Tarih Dersi Puanlarının Standart Sapmasının Hesaplanması

Puan	Frekans	Kareleri (X^2)	f. X^2	Puan	Frekans	Kareleri (X^2)	f.X^2
15	1	225	225	**45**	4	2025	8100
16	1	256	256	49	4	2401	9604
23	1	529	529	53	7	2809	19663
25	1	625	625	55	2	3025	6050
27	1	729	729	59	3	3481	10443
31	1	961	961	**60**	**2**	3600	7200
35	2	1225	2450	**65**	**3**	4225	12675
36	1	1296	1296	69	5	4761	23805
37	1	1369	1369	77	3	5929	17787
39	3	1521	4563	**79**	**1**	6241	6241
43	2	1849	3698	**87**	**1**	7569	7569
Karelerin Toplamı = 145838				Puanların toplamı=2574			

Toplam veri sayısı (n) = 50 olduğundan standart sapma,

$$S_x = \sqrt{\frac{145838 - \frac{(2574)^2}{50}}{50-1}} = 14{,}82$$

olarak bulunur.

Test puanlarının standart sapmasının doğru yorumlanabilmesi için merkezi yığılma ölçülerinden aritmetik ortalamanın bilinmesine ihtiyaç vardır. Aritmetik ortalama ve standart sapma değeri kullanılarak hesapalan *"Bağıl değişkenlik katsayısı (V)"* puan dağılımının farklılığını daha doğru yorumlamamızı sağlayacak bir başka dağılım ölçüsüdür.

* **Bağıl değişkenlik katsayısı (V):** Standart sapmanın aritmetik ortalamaya oranının yüzdesini veren bir dağılım ölçüsüdür. Aşağıdaki eşitlik yardımıyla hesaplanır:

$$V = \frac{S_x}{\bar{X}} \cdot 100$$

egitim@hotmail.es

Bağıl değişkenlik katsayısının yorumlanması aşağıdaki kriterlere göre yapılır:

- **20 < V <25 ise (bağıl değişkenlik katsayısı 20 ile 25 arasında bir değer alıyorsa);** puanlar normal dağılım göstermektedir.
- **V<20 ise (bağıl değişkenlik katsayısı 20'den küçük bir değer almışsa);** puanların homojen, birbirine benzer olduğunu, puanlar arasındaki farklılığın az olduğunu gösterir ki bu durumda puanlar normale göre daha sivri bir dağılıma sahip olacaktır.
- **V <25 ise (bağıl değişkenlik katsayısı 25'ten büyük bir değer almışsa);** puanların heterojen, birbirininden farklı olduğunu, puanlar arasındaki farklılığın fazla olduğunu gösterir ki bu durumda puanlar normale göre daha basık bir dağılıma sahip olacaktır.

Örneğin; puanlarımız için hesaplanan bağıl değişkenlik katsayısına bakıldığında,

$$V = \frac{14,82}{51,48} 100 = 28,79$$

normal dağılım sınırlarının üzerinde olduğu görülmektedir. Bu ise örnekte verilen puanların heterojen bir dağılım gösterdiğini; normalden daha basık, alınan puanların birbirinden farklı olduğunu gösterir. Puanlara bakıldığında birbirinden farklı ve geniş bir boyutta puanların dağıldığı görülmektedir.

NOT: *Test istatistikleri ile ilgili daha ayrıntılı bilgiye kaynakçada yer alan (Baykul, 1999) ve (Büyüköztürk,Çokluk ve Köklü, 2010)'den ulaşabilirsiniz.*

DEĞERLENDİRME SORULARI

Tabloda beş farklı derse ilişkin öğrencilerin aldıkları puanlara ait bazı istatistiksel değerler yer almaktadır.
1.- 4. soruları bu tabloya göre cevaplandırınız.

Ders	Arit.Ort.	Medyan	Mod	St. Sapma
Türkçe	65	60	57	8
Tarih	70	67	65	10
Matematik	60	60	60	5
Kimya	45	51	57	14
Fizik	63	54	46	7

1. Puanların simetrik dağılım gösterdiği ders hangisidir?
A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

2. Puanlar arasındaki farklılaşmanın en fazla olduğu ders hangisidir?
A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

3. Hangi derse ilişkin puanların dağılımı sola çarpıktır?
A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

4. Hangi dersteki başarı en yüksektir?
A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

Tabloda beş farklı derse ilişkin öğrencilerin aldıkları puanlara ait bazı istatistiksel değerler yer almaktadır.
5.- 8. soruları bu tabloya göre cevaplandırınız.

Ders	Arit. Ort.	Medyan	Mod	St. Sapma
Felsefe	63	67	72	5
Tarih	80	85	91	10
Psikoloji	75	75	75	7
Coğrafya	63	59	54	4
Sosyoloji	80	70	65	3

egitim@hotmail.es

5. **Hangi derse ait puanların dağılımı normal dağılım gösterir?**

A) Felsefe B) Tarih C) Psikoloji D) Coğrafya E) Sosyoloji

6. **Hangi derste öğrencilerin başarı düzeyleri birbirine en yakındır?**

A) Felsefe B) Tarih C) Psikoloji D) Coğrafya E)Sosyoloji

7. **Hangi derse ait başarı düzeyi en yüksektir?**

A) Felsefe B) Tarih C) Psikoloji D) Coğrafya E) Sosyoloji

8. **Hangi derste öğrenciler tarafından en çok alınan puan en yüksektir?**

A) Felsefe B) Tarih C) Psikoloji D) Coğrafya E) Sosyoloji

Aşağıdaki tabloda 5 sınıfa ait en düşük ve en yüksek puanlar ile en yüksek frekansı olan puanlar verilmiştir.

9 - 10. soruları aşağıda verilen tabloya göre cevaplandırınız.

Sınıflar	En düşük – En yüksek puanlar	En yüksek frekansa sahip puan
A	15-75	60
B	5-90	65
C	20-80	55
D	10-100	60
E	25-100	45

9. **Tabloya göre hangi sınıfın ranjı en büyüktür?**

A) A B) B C) C D) D E) E

10. **Tabloya göre hangi sınıfın modu en büyüktür?**

A) A B) B C) C D) D E) E

KPSS'DE ÇIKMIŞ SORULAR

1. - 3. SORULARI AŞAĞIDA VERİLEN BİLGİLERE GÖRE CEVAPLAYINIZ.

Bir sınıftaki öğrencilerin beş farklı testteki yanıtları 100 üzerinden puanlanmış ve bu puanlardan, aşağıdaki tabloda gösterilen istatistikler hesaplanmıştır.

	Aritmetik Ortalama	Tepe Değer (Mod)	Ortanca (Medyan)	Standart Sapma
1)	70	80	75	8
2)	40	50	45	6
3)	50	50	50	10
4)	50	30	35	10
5)	60	40	50	15

1. **Bu bilgilere göre, hangi testten elde edilen puanların dağılımı, "tek modlu ve simetrik"tir? (2001)**

A) 1 B) 2 C) 3 D) 4 E) 5

2. **Sınıfın ortalama öğrenme düzeyi, hangi testle yoklanan öğrenmeler için en yüksek, hangi testle yoklanan öğrenmeler için en düşüktür? (2001)**

	En yüksek	En düşük
A)	1. test	4. test
B)	1. test	2. test
C)	2. test	5. test
D)	3. test	4. test
E)	5. test	2. test

3. **Hangi testle yoklanan öğrenmeler açısından, sınıftaki öğrenciler arasındaki farklılaşma (değişkenlik, varyans) en büyüktür? (2001)**

A) 1. B) 2. C) 3. D) 4. E) 5.

egitim@hotmail.es

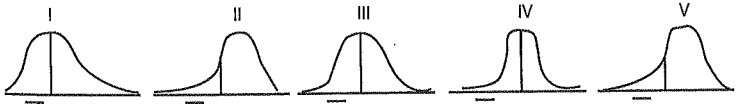
4. Aşağıda, aynı değişken üzerindeki bir ölçümler dizisi ile ilgili dört tanım verilmiştir.

I. En yüksek frekansa sahip olan değer
II. Büyüklük sırasına konmuş ölçümler dizisinin tam ortasındaki değer
III. Ölçümlerin toplamının ölçüm sayısına bölünmesiyle elde edilen değer
IV. En büyük ölçüm ile en küçük ölçüm arasındaki fark

Bu tanımlara karşılık gelen kavramlar, aşağıdakilerin hangisinde doğru olarak verilmiştir? (2002)

	I	II	III	IV
A)	Medyan	Mod	Standart sapma	Ranj
B)	Medyan	Ranj	Aritmetik ortalama	Mod
C)	Ranj	Standart sapma	Aritmetik ortalama	Medyan
D)	Mod	Medyan	Standart sapma	Aritmetik ortalama
E)	Mod	Medyan	Aritmetik ortalama	Ranj

5. Her biri 0 veya 1 şeklinde puanlanan 100 soruluk genel başarı sınavında, eşit sayıda öğrencisi olan beş sınıfın puan dağılımlarının genel şekilleri ile aritmetik ortalama ve standart sapma değerleri aşağıdaki grafiklerde gösterilmiştir.



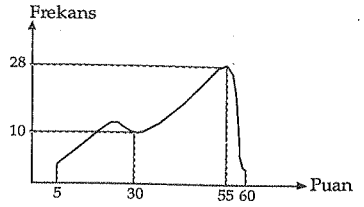
Buna göre, testte grup olarak en başarılı olan sınıf hangisidir? (2002)

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

6. **Bir testten alınan puanlarla ilgili olarak yalnızca standart sapmanın küçük olduğunu, diğer bir deyişle bağıl değişkenliğin az olduğunu bilen bir kişi aşağıdakilerden hangisini söyleyebilir? (2002)**

A) Test puanlarının ranjının (en büyük ve en küçük puan arasındaki farkın) büyük olduğunu
B) Test için verilen cevaplama süresinin yeterli olduğunu
C) Test puanlarının "normal dağılım"a benzer bir dağılım gösterdiğini
D) Öğrencilerin testte başarılı olduklarını
E) Öğrencilerin testte ölçülen özellik bakımından birbirlerine benzer olduğunu

7.



Yukarıdaki dağılımda mod kaçtır? (2004)

A) 10 B) 28 C) 30 D) 55 E) 60

8. Farklı dersler için hazırlanan ve soru sayıları aşağıdaki tabloda parantez içinde verilen beş test bir gruba uygulanmış ve ham puanlar üzerinden elde edilen aritmetik ortalamalar ($\bar{X}$) tabloda gösterilmiştir.

Test (soru sayısı)	$\bar{X}$
Biyoloji (10)	5,2
Matematik (45)	38,1
Fizik (19)	13,4
Kimya (12)	6,9
Tarih (14)	11,7

Sınıfın hangi testteki mutlak başarı düzeyi en yüksektir? (2004)

A) Biyoloji
B) Matematik
C) Fizik
D) Kimya
E) Tarih

9. **Bu beş testteki puanlar üzerinden, bu testlerin hangisi için hesaplanacak ortalama madde güçlüğü en küçüktür? (2004)**

A) Biyoloji
B) Matematik
C) Fizik
D) Kimya
E) Tarih

egitim@hotmail.es

10. VE 11. SORULARI AŞAĞIDA VERİLEN BİLGİLERE GÖRE CEVAPLAYINIZ.

196 öğrencinin katıldığı bir testten alınan puanların frekansları ve yığmalı frekansları aşağıda verilmiştir.

Ham Puan (X)	Frekans (f)	Yığılmalı Frekans (Σf)
107	3	3
106	2	5
105	3	8
104	1	9
103	3	12
102	8	20
101	14	34
100	39	73
99	21	94
98	40	134
97	62	196

10. Bu öğrencilerin aldıkları puanların toplamı (ΣfX) 19 404 olduğuna göre aritmetik ortalaması kaçtır? (2005)

A) 98 B) 98,5 C) 99 D) 99,5 E) 100

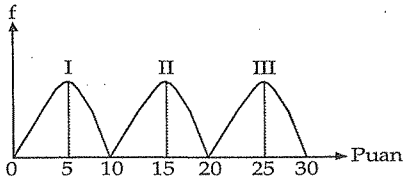
11. Yukarıdaki frekans dağılımında, medyan değeri hangi puanlar arasındadır? (2005)

A) 98-99
B) 99-100
C) 100-101
D) 101-102
E) 102-103

12. Bir ölçme işleminde, ölçümlerin frekansları birbirine eşit ise bu ölçümler için aşağıdakilerden hangisi hesaplanamaz? (2005)

A) Aritmetik ortalama
B) Mod
C) Standart sapma
D) Medyan
E) Ranj

13.



Grafikte I, II ve III numarayla gösterilen üç dağılım için aşağıdaki ifadelerden hangisi doğrudur? (2005)

A) Ranj değerleri aynı, mod değerleri farklıdır.
B) Ranj değerleri farklı, mod değerleri aynıdır.
C) Ranj değerleri farklı, medyan değerleri aynıdır.
D) Ranj ve mod değerleri aynıdır.
E) Ranj ve medyan değerleri aynıdır.

14. Bir grup öğrencinin tarih, coğrafya, felsefe ve yabancı dil testlerinden aldıkları puanlar arasındaki korelasyonlar aşağıdaki tabloda gösterilmiştir.

DERS	Tarih	Coğrafya	Felsefe	Yabancı Dil
Tarih	1,00	0,60	0,85	-0,35
Coğrafya	0,60	1,00	0,40	-0,25
Felsefe	0,85	0,40	1,00	0,30
Yabancı Dil	-0,35	-0,25	0,30	1,00

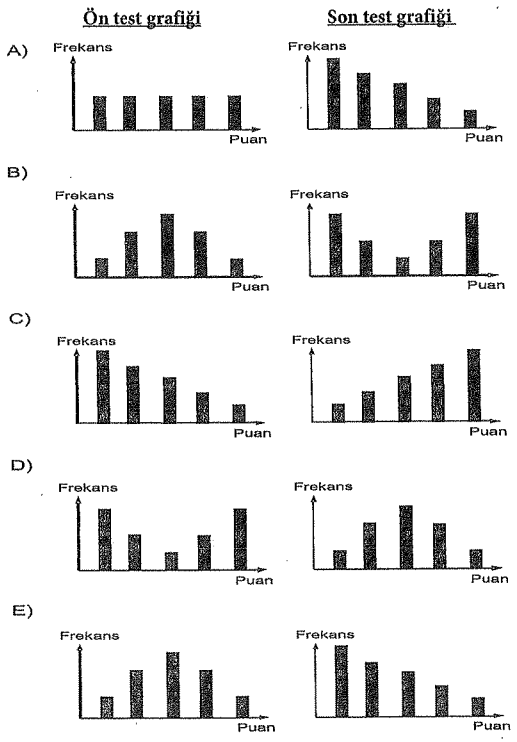
Bu tablodaki bilgilere dayanarak aşağıdaki sonuçlardan hangisine varılabilir? (2006)

A) Yabancı dil testi en düşük ortalamaya sahiptir.
B) Öğrenciler tarih ve coğrafya derslerinde çok başarısızdırlar.
C) Felsefe ve yabancı dil testleri çok farklı özellikleri ölçmektedir.
D) En yüksek ilişki felsefe ile tarih puanları arasındadır.
E) En düşük ilişki tarih ile yabancı dil puanları arasındadır.

egitim@hotmail.es

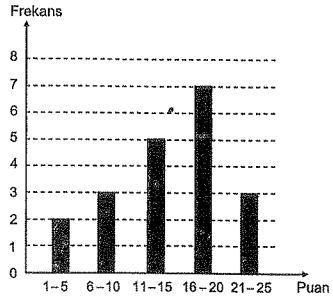
15. Bir öğretmen, geliştirmiş olduğu bir başarı testini derse başlamadan önce (ön test) ve dersin bitiminde (son test) sınıftaki öğrencilere uygulamıştır. Uygulama sonucunda öğrencilerin ön test ve son test puanlarına ait sütun grafiklerini inceleyerek kullandığı öğretim yöntem ve tekniklerinin öğrenme üzerinde etkili olduğu sonucuna varmıştır.

Bu öğretmenin elde ettiği grafikler aşağıdakilerden hangisi olabilir? (2007)



16. - 18. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

5A sınıfındaki öğrencilerin 25 soruluk matematik testinden aldıkları puanlar gruplanarak aşağıdaki frekans grafiğinde gösterilmiştir.



16. **5A sınıfının mevcudu kaçtır? (2007)**

A) 5 B) 7 C) 8 D) 20 E) 25

17. **5A sınıfının matematik testi puanlarının aritmetik ortalaması kaçtır? (2007)**

A) 5,7 B) 7,4 C) 14,5 D) 18,6 E) 20

18. **5A sınıfının matematik testi puanlarının tepe değeri (mod) kaçtır? (2007)**

A) 5 B) 7 C) 8 D) 18 E) 25

egitim@hotmail.es

19. Bir grup öğrenciye uygulanan dönem sonu sınavlarına ilişkin istatistikler tablodaki gibidir.

Dersler	Soru sayısı	Aritmetik ortalama	Medyan	Mod	Standart sapma
Kimya	20	8	6	4	4
Tarih	40	30	35	39	7
Türkçe	30	15	18	23	6
Fizik	50	30	30	30	11
Felsefe	40	14	10	7	9

Öğrencilerin en başarılı olduğu ders hangisidir? (2008)

A) Kimya
B) Tarih
C) Türkçe
D) Fizik
E) Felsefe

20. Mehmet Öğretmen, öğrencilerine uyguladığı bir test sonucunda elde ettiği puanlardan bazı istatistikleri hesaplamıştır. Daha sonra, öğrencilerden birinin puanını yanlış yazdığını fark etmiş ve gerekli düzeltmeyi yaparak istatistikleri tekrar hesaplamıştır.

Buna göre, Mehmet Öğretmen'in düzeltmesi sonucunda aşağıdaki istatistiklerden hangisi kesin olarak değişmiştir? (2008)

A) Ortalama
B) Ortanca
C) Mod
D) Ranj
E) Yığmalı frekans

21. Bir coğrafya öğretmeni 100 soruluk bir başarı testi uygulamış ve test puanlarının ortalamasını 75, ortancasını 80, tepe değerini (mod) 90, standart sapmasını 8 ve ranjını 90 bulmuştur. Öğrencilerin dersten geçebilmeleri için en az 65 puan almaları gerekmektedir.

Test puanlarına ilişkin istatistiklere ve değerlendirme ölçütüne dayanarak bu dağılım için aşağıdaki yorumlardan hangisi yapılamaz? (2009)

A) Ölçümlerin yüzde elliden fazlası ortalamadan büyüktür.
B) Merkezî eğilim ölçüsü olarak ortancadan yararlanılmalıdır.
C) Sınıfın büyük çoğunluğu geçer not alır.
D) Tepe değeri, gözlenen en büyük değerdir.
E) Ortalamayı çarpıtan uç değer(ler) bulunmaktadır.

22. – 24. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

10 soruluk bir sınava ait öğrenci yanıtları aşağıdaki tabloda verilmiştir. Her soruda doğru yanıta "1", yanlış yanıta "0" puan verilmiştir. Yanıtlanmayan soru bulunmamaktadır.

		SORULAR									
		1	2	3	4	5	6	7	8	9	10
ÖĞRENCİLER	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

22. **Öğrencilerin bu sınavdan aldıkları puanların ranjı kaçtır? (2009)**

A) 3 B) 5 C) 6 D) 9 E) 10

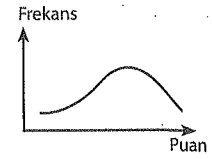
23. **Hangi sorular ölçülen özellikler açısından öğrencilerin bireysel farklılıklarını ortaya çıkarmada hiç etkili olmamıştır? (2009)**

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

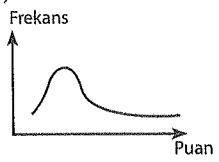
egitim@hotmail.es

24. Bu sınavdaki test puanlarının frekans dağılımı yaklaşık olarak aşağıdakilerden hangisinde verilmiştir? (2009)

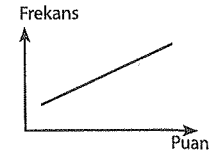
A)



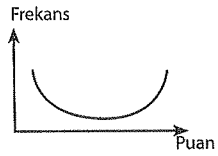
B)



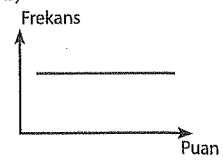
C)



D)



E)



25. Bir testten elde edilen puanların standart sapması küçük bulunmuşsa bu testi alan bireyler hakkında aşağıdaki sonuçlardan hangisine ulaşılabilir? (2010)

A) Ölçülen özellik bakımından birbirlerine benzemektedirler.
B) Bu tür testleri almaya alışmışlardır.
C) Testi, verilenden daha kısa sürede cevaplamışlardır.
D) Testte yer alan sorular kendi düzeylerine uygundur.
E) Soruları cevaplarken gerçek performanslarını ortaya koymuşlardır.

26. VE 27. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Altı maddelik bir sınavdan alınan puanlara ait frekans tablosu aşağıda verilmiştir.

Puan	Frekans
1	2
2	10
3	5
4	4
5	3
6	2

26. Bu sınav kaç öğrenciye uygulanmıştır? (2010)

A) 6 B) 10 C) 21 D) 26 E) 80

27. Bu puan dağılımıyla ilgili olarak aşağıdakilerden hangisi doğrudur? (2010)

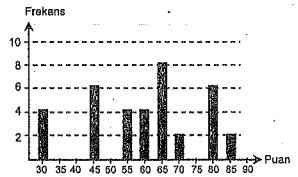
A) Tek modlu simetriktir.
B) Tek modlu sağa çarpıktır.
C) Modu yoktur.
D) Çift modlu simetriktir.
E) Çift modlu sağa çarpıktır.

28. Bir yükseköğretim programında öğrencilerin genel notları ara sınavın % 40'ı ile genel sınavın % 60'ı toplanarak oluşturulmaktadır.

Bu yükseköğretim programında geçme notu 60 olduğuna göre, ara sınavdan 45 puan alan bir öğrencinin dersi geçmesi için genel sınavdan en az kaç puan alması gerekir? (2010)

A) 65 B) 70 C) 75 D) 80 E) 85

29. Duygu Öğretmen, birinci sınavda modun (tepe değerin) üzerinde puan alan öğrencilerine işbirliğine dayalı grup çalışmasında, istedikleri arkadaşlarıyla çalışma şansı vereceğini açıklar. Bu sınava ilişkin puan dağılımı aşağıdaki grafikte gösterilmiştir.



Bu sınıfta kaç öğrenci grup arkadaşlarını seçebilecektir? (2010-2)

A) 4 B) 6 C) 8 D) 10 E) 12

egitim@hotmail.es

30. Bir okulda bulunan sekizinci sınıf öğrencileri beş şubeden oluşmaktadır. Bu öğrenciler eğitim öğretim yılı başında, altıncı ve yedinci sınıf matematik dersinin önemli kazanımlarının ölçüldüğü 100 maddelik bir sınava girmiştir. Şubelere göre, öğrencilerin puan ortalamaları ve standart sapmaları aşağıdaki tabloda verilmiştir.

Şubeler	Aritmetik ortalama	Standart sapma
I	50,0	5
II	50,5	15
III	51,0	20
IV	51,5	10
V	52,0	30

Buna göre, diğer değişkenler sabit tutulduğunda, bu şubelere derse girecek bir matematik öğretmeninin hangi şubede daha çok zorlanması beklenir? (2010-2)

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

31. – 33. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

Aşağıdaki tabloda beş ayrı teste ait bilgiler verilmiştir.

Test	Güvenirlik Katsayısı	Testin standart sapması	Testin aritmetik ortalaması	Madde sayısı
Felsefe	0,96	5	15	50
Fizik	0,91	15	60	100
Kimya	0,84	10	25	100
Biyoloji	0,75	4	20	50
Tarih	0,64	15	40	100

31. Tabloya göre, adaylara en zor gelen test hangisidir? (2010-2)

A) Felsefe B) Fizik C) Kimya D) Biyoloji E) Tarih

32. Tabloya göre, ölçmenin standart hatası en büyük olan test hangisidir? (2010-2)

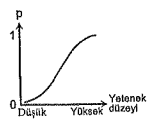
A) Felsefe B) Fizik C) Kimya D) Biyoloji E) Tarih

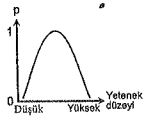
33. Bir aday, şans başarısı düzeltme formülü kullanılmayan ve bağıl değerlendirme yapılan bu testlerin her birinde 40 soruyu doğru cevaplamışsa, adayın en başarısız olduğu test hangisidir? (2010-2)

A) Felsefe B) Fizik C) Kimya D) Biyoloji E) Tarih

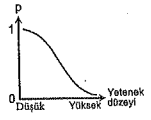
34. Bir ölçme kuramına göre, bireyin bir maddeye doğru yanıt verme olasılığının bireyin yeteneğiyle bağlantılı olduğu varsayılmakta ve bu ilişki Madde Karakteristik Fonksiyonu (MKF) ile ifade edilmektedir. MKF, bireyin ilgili yetenek düzeyi ile maddeye doğru yanıt verme olasılığı (p) arasındaki ilişkinin grafik gösterimidir. İyi işleyen bir madde için, yetenek düzeyi arttıkça bireyin maddeye doğru yanıt verme olasılığı artmakta, yetenek düzeyi azaldıkça da doğru yanıt verme olasılığı düşmektedir.

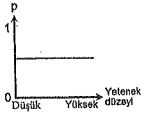
Buna göre, ayırt ediciliği uygun olan bir madde için Madde Karakretistik Fonksiyonu (MKF) aşağıdakilerden hangisi olabilir? (2010-2)

A)


B)


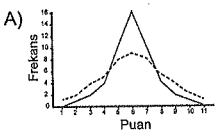
C)

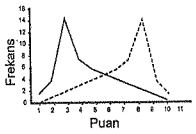

D)


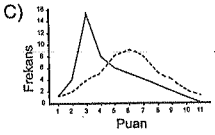
E)


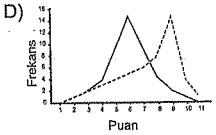
35. Durmuş Ali Öğretmen, iki farklı lise 1. sınıf şubesine matematik sınavı uygulamıştır. İki uygulama sonucunda her iki sınıfta da puan dağılımı tek modlu simetrik bir yapıda olmuştur. İki şubenin puan ortalaması aynı, standart sapmaları ise farklı hesaplanmıştır.

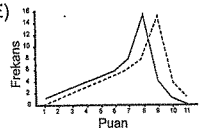
Bu iki şubenin puan dağılımları grafiği aşağıdakilerden hangisi gibi olabilir?

A)


B)


C)


D)


E)


egitim@hotmail.es

36. Bir okulda öğrencilerin gelişimini izlemek amacıyla, bir hafta arayla, 100'er maddelik paralel iki test uygulanmıştır. Bu iki uygulamadan elde edilen doğru cevap ortalamaları, standart sapmalar ve Zeynep adlı öğrencinin doğru cevap sayıları aşağıdaki tabloda verilmiştir.

	Birinci Uygulama	İkinci
Uygulama		
Testin Ortalaması	50	60
Testin Standart Sapması	10	10
Zeynep'in Doğru Cevap Sayısı	60	65

Zeynep'in durumuyla ilgili olarak

I. Mutlak başarı düzeyi artmıştır.

II. T standart puanı artmıştır.

III. Sınıf içindeki sıralamada yükselmiştir.

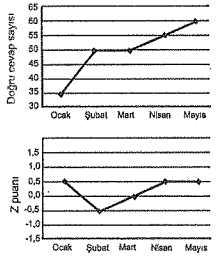
ifadelerinden hangileri kesinlikle doğrudur?

A) Yalnız I B) Yalnız III C) I ve II D) II ve III E) I, II ve III

37. VE 38. SORULARI AŞAĞIDAKİ BİLGİLERE GÖRE CEVAPLAYINIZ.

İlköğretim 6. sınıf öğrencilerinin okuma becerilerinin zaman içindeki gelişimini izlemek amacıyla, 80'er maddelik çoktan seçmeli beş test geliştirilmiştir. Bu testler ocak ayından itibaren her ay okuldaki tüm 6. sınıf öğrencilerine uygulanmıştır. Puanlamada doğru cevaplanmış maddelere 1, yanlış cevaplara ve cevaplanmamış maddelere ise 0 puan verilmiştir.

Aşağıda, Burcu'nun bu testlerdeki doğru cevap sayılarına ve z standart puanlarına ait grafikler verilmiştir.



37. **Burcu'nun durumuyla ilgili olarak aşağıdakilerden hangisi kesinlikle doğrudur?**

A) Burcu'nun mutlak başarı düzeyinin aynı kaldığı dönemde okul sıralamasındaki yeri de aynı kalmıştır.

B) Burcu'nun mutlak başarı düzeyinin % 50'nin altında olduğu iki test vardır.

C) Mart ayında Burcu'nun doğru cevap sayısı testin ortalamasına eşittir.

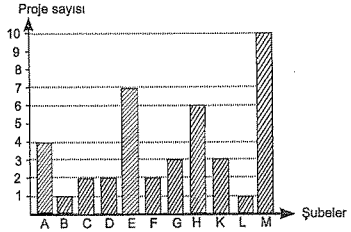
D) Nisan ayında uygulanan testte Burcu'nun doğru cevap sayısı okul ortalamasının altında kalmıştır.

E) Ocak-şubat döneminde Burcu'nun okul içi sıralamadaki yeri yükselmiştir.

38. **Nisan ayında uygulanan testin standart sapması 6 olduğuna göre, bu testten 64 soruyu doğru cevaplayan bir öğrencinin T standart puanı kaç olur?**

A) 40 B) 50 C) 60 D) 70 E) 80

39. **Bir okuldaki bilim şenliğinde 11 şubenin yaptığı proje sayıları aşağıdaki grafikte verilmiştir.**



Buna göre, proje sayıları dağılımının ranjı kaçtır?

A) B) 3 C) 9 D) 10 E) 41

40. Bir matematik öğretmeni dönem sonunda 24 kişilik sınıfında birçok konuyu kapsayan çoktan seçmeli bir test uygulamıştır. Bu testte ondalık sayılar konusundan beş madde bulunmaktadır. Bu beş maddelik alt testin her bir maddesindeki seçenekleri işaretleyen öğrenci sayılarını aşağıdaki tabloda özetlemiştir. Doğru cevap olan seçenekleri yıldız işareti (*) ile göstermiştir.

Maddeler	Seçenekler				Toplam
	A	B	C	D	
M1	3	3*	15	3	24
M2	4	2*	14	4	24
M3	0	9	8	7*	24
M4	6*	6	5	7	24
M5	4	4	12*	4	24

Bu beş maddelik alt testin ortalaması kaçtır?

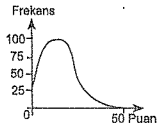
A) 0,75 B) 1,25 C) 1,75 D) 2,25 E) 2,75

egitim@hotmail.es

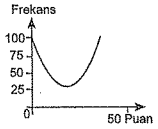
41. 2000 öğrencinin bulunduğu bir okuldan, matematik bilgi yarışması için 4 kişilik bir takımın oluşturulması hedeflenmektedir. Takıma alınacak başarılı öğrencileri seçmek amacıyla okuldaki tüm öğrenciler 50 maddelik çoktan seçmeli bir başarı testine alınacaktır.

Belirtilen amaç doğrultusunda hazırlanan bir sınavdan alınan puanların frekans grafiğinin aşağıdakilerden hangisine en yakın olması beklenir? (2012)

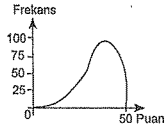
A)



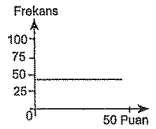
B)



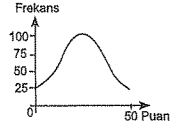
C)



D)



E)



42. Bir tarih öğretmeni, 6A ve 6B şubelerindeki öğrencilere aynı testi uygulamıştır. Şubelerdeki puan dağılımına ait betimsel istatistikleri aşağıdaki tabloda vermiştir.

	6A	6B
Öğrenci sayısı	36	34
En Düşük Puan	8	4
En Yüksek Puan	40	36
Aritmetik Ortalama	25,50	18,74
Standart Sapma	5,13	9,11

Bu tabloya göre, aşağıdakilerden hangisi kesinlikle söylenebilir? (2012)

A) 6A şubesinin puan aralığı, 6B şubesinin puan aralığından daha geniştir.

B) 6A şubesindeki puanlar, 6B şubesindeki puanlardan daha homojen bir dağılım göstermiştir.

C) 6B şubesi, tarih testinde 6A şubesinden daha başarılı olmuştur.

D) 6A şubesinde, testteki tüm maddeleri doğru cevaplayan en az bir öğrenci vardır.

E) 6B şubesindeki puan dağılımı sola çarpıktır.

STANDART PUANLAR

8

ANAHTAR KAVRAMLAR

Z puanı, T puanı

ÖĞRENME ÇIKTILARI

Bu bölümü tamamladıktan sonra;

- Standart puanların neden gerekli olduğunu açıklayabilecek,
- En yaygın kullanılan ve en çok bilinen standart puanlardan standart normal dağılım istatistiği olan Z puanı'nı hesaplayabilecek,
- Z puanından yararlanarak t puanını hesaplayabileceksiniz.

İÇİNDEKİLER

Standart puanlar
Z puanı
T puanı

BAŞLAMADAN ÖNCE

Standart puan sizce ne olabilir?

egitim@hotmail.es

STANDART PUANLAR

Öğrencilerin sınavlardan doğru ve yanlış cevaplarına göre hesaplanarak aldıkları pualar ham puan olarak adlandırılır. Örneğin; 20 soruluk bir testte her doğru cevap 1 puan ve yanlış ya da boş cevap 0 puan ise 17 doğrusu olan bir öğrencinin bu sınavdan alacağı ham puan 17 olacaktır. Eğer bu sınava her doğru cevap 5 puan ve yanlış ya da boş cevap 0 puan ise bu kez da öğrencinin alacağı ham puan 85 olacaktır. Elde edilen bu ham puanlar sıralama ölçeğinde ölçme sonuçlarıdır. Ham puanlar birbirine göre daha yüsek – daha düşük olarak değerlendirilmenin ötesinde bir işleve sahip değildirler. Bu nedenl ham puanlar üzerinden yorum yapmak her zaman doğru olmamaktadır. Bu durumu aşağıda verilen örnekle açıklamaya çalışalım:

Örnek: Farklı sınıflarda bulunan dört öğrencinin matematik dersinden oldukları başarı sınavı sonuçları aşağıda verilen tabloda gösterildiği şekilde olsun:

Sınıf	Öğrenciler	Öğrencilerin hampuanları	Sınavdan alınabilecek En yüksek puan	Sınıf puanlarının aritmetik ort.	Sınıf puanlarının stand.sapm.
A	Ayşe	72	100	64	8
B	Emre	41	50	35	5
C	Selin	8	10	4	2
D	Mete	32	40	34	4

- Yukarıdaki dört öğrencinin başarı durumunu karşılaştırmak istediğimizde, öğrencilerin ham puanlarına göre değerlendirme yapmak elbette ki hatalı olacaktır. Bir başka deyişle A sınıfındaki Ayşe'nin Emre'den daha başarılı olduğunu söylemek ya da bu öğrenciler içinde en başarısız öğrenci Selin en başarılı Ayşe vb. yorumlar yapmak hiç de doğru olmayacaktır.

İşte bu durumda en doğru yorumlar standart puanlaın hesaplanması ile yapılabilmektedir. Öğrencilerin başarı durumlarını doğru bir şekilde yorumlamak ancak ham puanların standart puanlara çevrilmesinden sonra yapılabilmektedir. Ham puanların aritmetik ortalama ve standart sapma kullanılarak yeni bir puana dönüştürülmesine puanların *standartlaştırılması* denir. En yaygın kullanılan ve en çok bilinenen standart puanlar standart normal dağılım istatistiği olan "*Z puanı*" ve Z puanından yararlanılarak elde edilen "*t puanı*"dır.

Z PUANI: HAM puanların, aritmetik ortalaması 0 (sıfır) ve standart sapması 1 olan puanlara dönüştürülmesiyle elde edilen standart puanlara "Z puanı" adı verilir ve Z puanı aşağıda verilen eşitlik yardımıyla hesaplanır:

$$Z_{puanı} = \frac{X - \bar{X}}{S_x}$$

X: öğrencinin puanı

$\bar{X}$: aritmetik ortalama

S_x: standart sapma

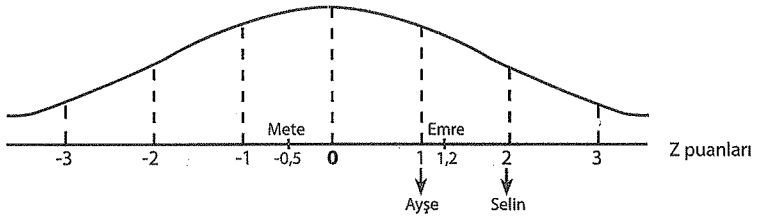
Bu eşitliğe göre şimdi her bir öğrencinin Z puanını hesaplayalım:

Ayşe'nin Z puanı: 72 – 64 / 8 = 1

Emre'nin Z puanı: 41 – 35 / 5 = 1.2

Selin'in Z puanı: 8 – 4 / 2 = 2

Mete'nin Z puanı: 32 – 34 / 4 = - 0.5



Böylelikle artık her bir öğrencinin puanının eşit aralıklı bir ölçek üzerinde nerede olduğunu görmek, birbirlerine göre durumunu doğru bir şekle yorumlamak mümkündür. "Bu dört öğrenci içinde en başarılı öğrenci Selin, en başarısız öğrenci Mete'dir" yorumunu yapmak artık doğru olacaktır.

- Z puanı teorik olarak -∞ ile +∞ arasında değişir.
- Ancak uygulamada puanların %99'u (-3, +3), yaklaşık %95'i (-2, +2) ve yaklaşık %68'i (-1, +1) arasındadır.
- Z puanının **negatif değer** olması öğrencinin ham puanının **sınıf ortalamasının altında**, Z puanının **pozitif değer** olmasıysa **sınıf ortalamasının üstünde** olduğunu gösterir.

egitim@hotmail.es

- Örnekte de görüleceği üzere Z puanları negatif ya da ondalık değerler alabilmektedir. Bu da puanların yorumlanmasında, anlaşılmasında sıkıntılara yol açabilmektedir. Örneğin; bir öğrenci velisine kızın – 2 puan aldı ya da sizin oğlunuz 0,6 puan aldı demek, veli için çok da anlaşılır bir açıklama olmayacaktır. Bu nedenle anlaşılmasının ve yorumlanmasının daha kolay olması nedeniyle genellikle Z puanları diğer bir standart puan çeşidi olan ve en yaygın kullanılan T puanlarına dönüştürülür.

T PUANI: Ham puanların, ortalaması 50 ve standart sapması 10 olan standart puanlara dönüştürülmesiyle T puanı elde edilir. T puanının hesaplanmasında, ham puanlar öncelikle Z puanına çevrilmelir ve daha sonra aşağıdaki eşitlik kullanılır:

$T_{puanı} = 50 + (Z_{puanı} \cdot 10)$

Yukarıda verilen tablodaki öğrencilerin şimdi de T puanlarını hesaplarsak:

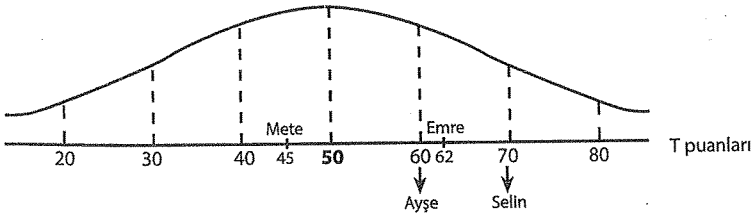
Ayşe'nin T puanı: 50 + (1.10) = 60

Emre'nin Tpuanı: 50 + (1.2.10) = 62

Selin'in T puanı: 50 + (2.10) = 70

Mete'nin T puanı: 50 + (- 0.5 . 10) = 45

elde ederiz. Görüleceği üzere öğrencilerin birbirlerine göre durumu değişmemiş, sadece puanların birimi yönünden bir farklılık oluşmuştur.



DEĞERLENDİRME SORULARI

1.- 3. SORULARI AŞAĞIDA VERİLEN TABLOYA GÖRE CEVAPLANDIRINIZ.

Ders	Ezgi'nin Puanı	Aritmetik Ortalama	Standart Sapma
Türkçe	65	57	8
Tarih	70	65	10
Matematik	70	55	5
Kimya	80	50	10
Fizik	60	40	10

1. **Ezgi'nin <u>en başarısız</u> olduğu ders hangisidir?**
   A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

2. **Hangi derste öğrencilerin puanları <u>en homojendir ?</u>**
   A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

3. **Hangi derste Ezgi'nin başarı düzeyi sınıfın başarı düzeyine <u>en yakındır</u>?**
   A) Türkçe B) Tarih C) Matematik D) Kimya E) Fizik

4. Bir coğrafya sınavı beş farklı sınıfta uygulanmış ve aşağıdaki tabloda görülen aritmetik ortalama ve standart sapma değerleri elde edilmiştir.
   Bu sınavdan 60 puan alan Sedat hangi sınıfta en yüksek ve en düşük başarıya sahiptir?

Sınıflar	A	B	C	D	E
Arit. Ort.	58	48	52	58	56
St. Sapma	10	6	8	6	4

	<u>En Yüksek</u>	<u>En Düşük</u>
A)	B	A
B)	C	B
C)	E	A
D)	D	A
E)	B	E

egitim@hotmail.es

5. VE 6. SORULARI AŞAĞIDA VERİLEN PARAGRAFA GÖRE CEVAPLANDIRINIZ.

*** Didem'in biyoloji sınavı ham puanı 75 ve bu sınav sonucu sınıf ortalaması 50 ve standart sapması 10 olarak hesaplanmıştır. Kimya sınavındaki Z puanı ise biyoloji sınavı Z puanından 1.5 puan daha fazladır ve bu sınavın aritmetik ortalaması 60 ve standart sapması 5'tir.

5. Didem'in kimya sınavından aldığı ham puanı kaçtır?

A) 80 B) 75 C) 65 D) 61,5 E) 55

6. Bu iki sınav sonucu elde edilen puanlar için aşağıdakilerden hangisi kesinlikle söylenebilir?

A) Kimya sınavı biyoloji sınavından daha zordur.
B) Biyoloji puanlarının dağılımı sola çarpıktır.
C) Öğrencilerin biyoloji dersindeki puanları daha heterojendir.
D) Kimya puanlarının dağılımı sağa çarpıktır.
E) Her iki sınavdaki puanların değişkenliği aynıdır.

7. Bir sınıfa uygulanan fizik sınavı puanlarının aritmetik ortalaması, modu ve medyanı 65 olarak elde edilmiştir.

Buna göre bu sınıfın puan dağılımı nasıldır?

A) Simetrik
B) Çok modlu
C) Sağa çarpık
D) Sola çarpık
E) Asimetrik

KPSS'DE ÇIKMIŞ SORULAR

1. Aşağıdaki tabloda beş dersin sınavlarının puan dağılımlarına ilişkin aritmetik ortalama ile standart sapmalar ve Ali'nin bu sınavlardan elde ettiği puanlar verilmiştir.

Ders	Aritmetik ortalama	Standart sapma	Ali'nin puanı
Matematik	50	5	65
Türkçe	50	8	65
Fizik	50	10	65
Kimya	50	15	65
Biyoloji	50	20	65

Ali'nin sınıfına göre en başarılı ve en başarısız olduğu dersler hangileridir? (2007)

	En başarılı	En başarısız
A)	Matematik	Türkçe
B)	Matematik	Biyoloji
C)	Fizik	Kimya
D)	Fizik	Biyoloji
E)	Kimya	Biyoloji

2. Aynı öğrencilere uygulanan ve bağıl değerlendirme yapılan iki sınavla ilgili aritmetik ortalama ve standart sapma değerleri aşağıda verilmiştir.

Sınav	Aritmetik ortalama	Standart sapma	Ali'nin puanı	Ayşe'nin puanı
I	55	15	70	?
II	60	5	?	68

Buna göre, aşağıdaki yorumlardan hangisi kesinlikle doğrudur? (2008)

A) Ali birinci sınavdaki soruların hepsini doğru yanıtlamıştır.
B) Ali ikinci sınavdan da 70 alırsa ilk sınava göre daha başarısız olmuş olur.
C) Birinci ve ikinci sınav sonuçları arasındaki korelasyon yüksektir.
D) Ayşe'nin ikinci sınavdaki başarısı Ali'nin ilk sınavındaki başarısından daha iyidir.
E) İkinci sınavda puanların ranjı ilk sınava göre daha büyüktür.

egitim@hotmail.es

3. Bir öğretim programına öğrenci seçmek için mülakat yapılacaktır. Bu mülakata bir genel yetenek testinden 50 ve daha üstü standart T puanı alanlar başvurabilecektir. Yetenek testinden elde edilen puanlar normal dağılım göstermektedir.

Bu ön koşul işlemiyle ilgili olarak,

I. Bağıl ölçüt kullanılmıştır.

II. Adayların % 50'si ön koşulu sağlamaktadır.

III. Soruların en az yarısını doğru cevaplayanlar ön koşulu sağlamaktadır.

ifadelerinden hangileri doğrudur? (2010)

A) Yalnız I B) I ve II C) I ve III

D) II ve III E) I, II ve III

4. Bir öğretmen, 100 öğrenciye uygulamış olduğu testten elde edilen ham puanları z standart puanlarına önüştürmüştür. Aşağıdaki tabloda, beş öğrencinin ham puanları ve z puanları verilmiştir.

Öğrenci	Ham puan	z puan
Kemal	100	1,5
Nuri	90	1
Fatma	70	-0,5
Zeynep	50	-1
Mehmet	30	-2

Öğretmen bir öğrencinin puan dönüşümünde hata yaptığına göre, hatalı puan hangi öğrenciye aittir? (2012)

A) Kemal B) Nuri C) Fatma D) Zeynep E) Mehmet

CEVAP ANAHTARI

DEĞERLENDİRME SORULARI

1. BÖLÜM	
1.	E
2.	B
3.	A
4.	E
5.	A
6.	C
7.	B
8.	A
9.	B
10.	C

2. BÖLÜM	
1.	E
2.	A
3.	D

3. BÖLÜM	
1.	D
2.	A
3.	D

4. BÖLÜM	
1.	E
2.	D
3.	E
4.	B
5.	C
6.	B
7.	C
8.	A
9.	A
10.	D

5. BÖLÜM	
1.	A
2.	D
3.	C
4.	D
5.	A
6.	E
7.	B
8.	A
9.	B
10.	D

6. BÖLÜM	
1.	B
2.	A
3.	E
4.	D
5.	A
6.	B
7.	D
8.	B
9.	B
10.	C

7. BÖLÜM	
1.	C
2.	D
3.	D
4.	B
5.	C
6.	E
7.	B
8.	B
9.	D
10.	B

8. BÖLÜM	
1.	B
2.	C
3.	B
4.	A
5.	A
6.	C
7.	A

egitim@hotmail.es

CEVAP ANAHTARI

KPSS'DE ÇIKMIŞ SORULAR

1. BÖLÜM					
1.	D	16.	B	31.	A
2.	A	17.	D	32.	B
3.	B	18.	E	33.	E
4.	A	19.	B	34.	C
5.	A	20.	C	35.	A
6.	D	21.	A	36.	D
7.	C	22.	B	37.	D
8.	A	23.	A	38.	D
9.	B	24.	E	39.	B
10.	E	25.	B	40.	O
11.	A	26.	A	41.	E
12.	C	27.	A	42.	E
13.	A	28.	C	43.	A
14.	E	29.	D	44.	A
15.	E	30.	D		

2. BÖLÜM	
1.	C
2.	E

3. BÖLÜM	
1.	E
2.	A

4. BÖLÜM							
1.	D	13.	B	25.	A	37.	C
2.	E	14.	B	26.	E	38.	C
3.	E	15.	D	27.	E	39.	D
4.	C	16.	C	28.	C	40.	C
5.	D	17.	E	29.	A	41.	B
6.	B	18.	A	30.	B	42.	E
7.	D	19.	D	31.	C		
8.	D	20.	E	32.	A		
9.	A	21.	B	33.	C		
10.	A	22.	A	34.	C		
11.	B	23.	C	35.	D		
12.	C	24.	D	36.	B		

5. BÖLÜM											
1.	E	11.	D	21.	D	31.	B	41.	D	51.	B
2.	B	12.	D	22.	D	32.	D	42.	A	52.	B
3.	A	13.	A	23.	C	33.	B	43.	E	53.	D
4.	E	14.	C	24.	A	34.	E	44.	D	54.	D
5.	D	15.	A	25.	B	35.	C	45.	B	55.	E
6.	C	16.	A	26.	E	36.	A	46.	C	56.	A
7.	A	17.	B	27.	C	37.	E	47.	E	57.	B
8.	C	18.	D	28.	E	38.	C	48.	A	58.	C
9.	E	19.	E	29.	E	39.	E	49.	D	59.	D
10.	E	20.	E	30.	D	40.	A	50.	D	60.	C

6. BÖLÜM			
1.	B	12.	B
2.	A	13.	D
3.	E	14.	E
4.	B	15.	D
5.	A	16.	C
6.	A	17.	E
7.	D	18.	D
8.	C	19.	B
9.	B		
10.	E		
11.	C		

7. BÖLÜM							
1.	C	12.	B	23.	D	34.	A
2.	B	13.	A	24.	A	35.	A
3.	E	14.	D	25.	A	36.	A
4.	E	15.	C	26.	D	37.	C
5.	E	16.	D	27.	B	38.	D
6.	E	17.	C	28.	B	39.	C
7.	D	18.	D	29.	D	40.	B
8.	B	19.	B	30.	E	41.	A
9.	A	20.	A	31.	C	42.	B
10.	C	21.	D	32.	E		
11.	A	22.	B	33.	B		

8. BÖLÜM	
1.	B
2.	D
3.	B
4.	C

egitim@hotmail.es

KAYNAKÇA

Atılgan, H., Kan, A. ve Doğan, N. (2006). Eğitimde Ölçme ve Değerlendirme. Anı Yayıncılık, Ankara.

Baykul, Y., Gelbal, S. ve Kelecioğlu, H. (2003). Anadolu Öğretmen Liseleri İçin Eğitimde Ölçme ve Değerlendirme. Devlet Kitapları, Milli Eğitim Basımevi,İstanbul.

Baykul, Y. (2000). Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması. ÖSYM Yayınları, Ankara.

Baykul, Y. (1999). İstatistik: Metodlar ve Uygulamalar. Anı Yayıncılık (3. Baskı), Ankara.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2008). Bilimsel Araştırma Yöntemleri.(Genişletilmiş 2. Baskı).Pegem Akademi, Ankara.

Büyüköztürk, Ş., Çokluk, Ö. ve Köklü, N. (2010). Sosyal Bilimler için İstatistik. PegemA Yayıncılık, Ankara.

Güler, N. (2008). Klasik Test Kuramı Genellenebilirlik Kuramı ve Rasch Modeli Üzerine Bir Araştırma. Yayınlanmamış doktora tezi. Sosyal Bilimler Enstitüsü, Hacettepe Üniversitesi,Ankara.

Linn, R. L. ve Gronlund, N. E. (1995). Measurement and Assessment in Teaching. (Eighth edition).Prentice Hall, Columbus, Ohio. USA

Özçelik, D. A. (1998). Ölçme ve Değerlendirme.(Genişletilmiş 2. Baskı).ÖSYM Yayınları, Ankara.

Özçelik, D. A. (1989). Test Hazırlama Kılavuzu.(Genişletilmiş 3. Baskı).ÖSYM Yayınları, Ankara.

Shavelson, R. S. (1988). Statistical Reasoning for the Behavioral Sciences. (Second edition). Allyn and Bacon, Inc. USA

Turgut, M. F., (1987). Eğitimde Ölçme ve Değerlendirme Metotları, Saydam Matbaacılık (5.Baskı), Ankara.

DİZİN

A

alt-üst grup yöntemi, 111, 113,

aritmetik ortalama, 13, 46, 62, 122, 123, 133, 138, 141, 142, 143, 146, 147, 151, 152, 153, 154, 158, 162, 166, 167, 169, 170 ,171

B

bağıl değerlendirme, 13, 20, 22, 103, 163, 171

bar grafiği, 137

birim, 1, 6, 7, 8, 9, 10 ,12, 15, 23, 40, 113, 146, 168

Ç

çeyrek sapma, 133, 138, 144, 145, 146

çoktan seçmeli testler, 29, 69, 70, 79, 80, 82, 83, 93, 95

D

değerlendirme, 1, 2, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 40, 60, 65, 66, 69, 70, 74, 75, 77, 83, 84, 85, 86, 87, 88, 89,92, 93, 94, 97, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 158, 163, 166, 171

dereceli puanlama anahtarı (rubrik), 69, 89, 106

doğru-yanlış testleri, 69, 78, 79, 93, 113

E

eşit aralıklı ölçek, 10, 12

F

formatif değerlendirme, 14

frekans, 8, 9, 134, 135, 136, 138, 139, 140, 141, 145, 146, 147, 150, 152, 153, 154, 156, 157, 158, 160, 161

G

geçerlik, 30, 39, 40, 49, 51, 52, 54, 56, 60, 62, 72, 80

güven aralığı, 39, 48

güvenirlik, 30, 39, 40, 42, 44, 46, 47, 50, 53, 54, 56, 62, 63, 72, 73, 75, 76, 80, 82, 94, 113, 124, 162

H

hata, 2, 6, 14, 21, 27, 28, 29, 30, 31, 39, 40, 41, 42, 46, 47, 48, 49, 56, 57, 58, 59, 60, 61, 62, 64, 70, 74, 76, 81, 90, 96, 97, 103, 106, 112, 119, 124, 125, 144, 162, 166

Henrysson yöntemi, 111, 115

histogram, 136, 137

İ

iç tutarlılık, 56, 59

K

kapsam geçerliği, 47, 49, 50, 53, 56, 59, 60, 61, 63, 65, 66, 67, 72, 73, 74, 75, 77, 78, 82, 93, 95, 96, 99, 108, 130

kısa cevaplı testler, 69, 75, 76, 77, 93

korelasyon, 33, 34, 35, 36, 37, 38, 42, 43, 44, 45, 46, 47, 50, 51, 52, 53, 60, 61, 62, 64, 67, 155, 171

egitim@hotmail.es

egitim@hotmail.es