

ÖLÇME VE DEĞERLENDİRME

ÖLÇME NEDİR?

**Herhangi bir niteliği
gözlemek, gözlem sonuçlarını
sayı ya da sembollerle ifade
etmektir.**





Örnek:

Ayşe 80 puan aldı.

Mehmet'in boyu 1.75 m'dir.

Hüseyin sınavda 1. oldu.

Fatma uzun boylu bir kızıdır.

Ölçek Türleri (Ölçme Düzeyleri)

Sınıflama
ölçekleri

Sıralama
ölçekleri

Eşit aralıklı
ölçekler

Eşit oranlı
ölçekler



En az bilgi verenden en fazla bilgi verene doğru;

Sınıflama
ölçekleri

Sıralama
ölçekleri

Eşit
aralıklı
ölçekler

Eşit oranlı
ölçekler



Sınıflama (adlandırma) Ölçekleri

- ➔ En kalitesiz ölçek türüdür.
- ➔ Belli bir özellik yönünden benzer olanların aynı gruba alınmasıdır.
- ➔ Farklar belli değildir.
- ➔ Belli bir başlangıç noktası ve birimi yoktur.
- ➔ Sayısal hiçbir ifade içermez.
(6 Erkek+ 5 Kız= ?)
- ➔ Hiçbir matematiksel işlem yapılamaz.
- ➔ Frekans, yüzde ve mod değerleri bulunabilir.



Örnek:

- ➔ İnsanları saç renklerine göre;
sarışınlar, esmerler, kumrallar ve kızılar
- ➔ Öğrencinin kütüphanesindeki kitapları ansiklopediler, şiir kitapları, romanlar diye ayırması
- ➔ İnsanları cinsiyetlerine göre kadınlar-erkekler şeklinde gruplandırmak

Sıralama Ölçekleri

➡ Azlık ve çoklukların belirtilmesinde kullanılır.

➡ Ölçülen özelliğe sahip oluş miktarı bakımından nesneler büyüklük sırasına konulur.

➡ Büyüklük küçüklüğü gösteren sıra sayıları vardır.

➡ Farklar belli değildir.

➡ Sayısal hiçbir ifade içermez.

➡ Hiçbir matematiksel işlem yapılamaz.

➡ Frekans, yüzde, medyan değerleri bulunabilir. Frekans her zaman 1 çıkar. Bu nedenle mod bulunmaz.

➡ Başlangıç noktası ve sıralar sabit değildir.

Örnek:



Yüzme yarışmasında birinci, ikinci ve üçüncünün belirlenmesi



Boy uzunluk sırasının değerleri tespit edilmeden belirlenmesi



KPSS tercihlerini sıralamak



Eşit Oranlı Ölçekler

- ➔ En kaliteli ölçek türüdür.
- ➔ Sıfır gerçek sıfırdır yani yokluğu, hiçliği ifade eder.
- ➔ Sayısal ifadeler vardır.
- ➔ Başlangıç noktası vardır.
- ➔ Sıfır gerçek sıfır olduğu için çarpma ve bölme de yapılabilir.
- ➔ Bütün matematiksel işlemler yapılabilir.
- ➔ Frekans, yüzde, mod, medyan, ortalama, standart sapma hesaplanabilir.

Örnek:



Sınıfta 3 öğrenci vardır.



Cebindeki parayla 5 kg patates alabildi.

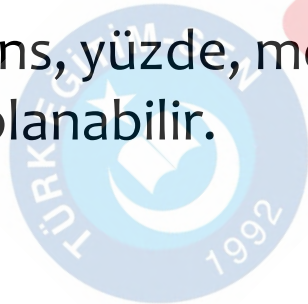


Testte 27 adet soru bulunmaktadır.



Eşit Aralıklı Ölçekler

- ➔ Başlangıç noktası olarak belirlenmiş bir sıfır noktası vardır; fakat keyfi olarak belirlenmiştir. Gerçek yokluğu ifade etmez.
- ➔ Sayısal ifadeler vardır.
- ➔ Toplama çıkarma işlemi yapılırken, çarpma ve bölme işlemi yapılamaz.
- ➔ Frekans, yüzde, mod, medyan, ortalama, standart sapma hesaplanabilir.



273°K.....	0°C
274°K.....	1°C
293°K.....	20°C
313 °K.....	40°C



Örnek:



Takvim



Meridyen



Deniz seviyesinden yükseklik (Rakım)



Eğitimde yararlanılan sınav türleri



Dönüşme var

Eşit Oranlı Ölçek	800 g	600 g	100 g	50 g	10 g
Eşit Aralıklı Ölçek	700 g	500 g	0 g	-50 g	-90 g
Sıralama Ölçeği	1	2	3	4	5
Sınıflama Ölçeği	Ağır	Ağır	Hafif	Hafif	Hafif

Dönüşme yok

Ölçme Türleri

```
graph TD; A[Ölçme Türleri] --- B[Doğrudan (temel) ölçme]; A --- C[Dolaylı (göstergeyle) ölçme]; A --- D[Türetilmiş ölçme]
```

**Doğrudan (temel)
ölçme**

**Dolaylı (göstergeyle)
ölçme**

Türetilmiş ölçme

Doğrudan Ölçme



Ölçülecek özelliğin araya başka değişken girmeden ölçülebilmesidir.



Ölçülecek özellik bu özelliği ölçmek için kullanılacak aracın özellikleri aynı olmalıdır.





➔ **Yüzme hocasının öğrenci performansını değerlendirirken gözlem yapması**



➔ **Pazardan alınan ürünlerin kefeli terazi ile tartılması**



➔ **Öğrencilerin boyunu mezura ile ölçmek**



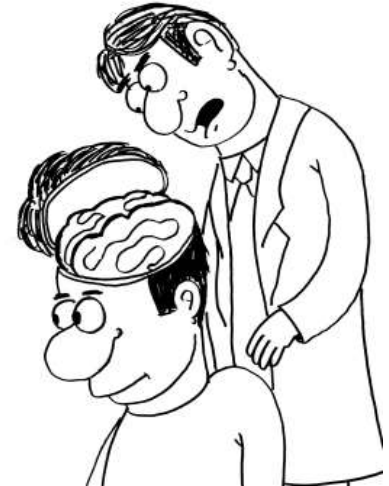
Dolaylı Ölçme



Doğrudan gözlenmesi mümkün olmayan durumlarda ölçülmesi gereken özelliği temsil eden başka değişkenin gözlenerek ölçülmesidir.



Eğitimde ölçülmek istenen özelliklerin çoğu dolaylı olarak ölçülebilen özelliklerdir.



Örnek:

- ➔ Bebeğin ateşini ölçerken termometre kullanma
- ➔ Öğrencileri değerlendirirken yapılan sınavlar
- ➔ Kilosunu merak eden bir bayanın baskülde tartılması
- ➔ Cisim kütlesinin yaylı kantarla ölçülmesi



Türetilmiş Ölçme

- ➔ İki yada daha fazla değişken arasında yapılan matematiksel işlem sonucu elde edilen ölçmelerdir.
- ➔ Değişkenlerin hepsi ayrı ayrı ölçülür.
- ➔ Örneğin, hızın ölçülmesi





Örnek:

- ➔ Özkütlenin hesaplanması
- ➔ IQ'nun hesaplanması
- ➔ Bir ilin aritmetik nüfus yoğunluğunun hesaplanması
- ➔ Bir aracın 2 saatte gittiği yolun hesaplanması
- ➔ Silindirin hacminin hesaplanması

ÖLÇMEDE BİRİM

Birim, bir ölçme işleminin üzerinde, değer

- Hava sıcaklığı 20 °C'dir.

°C

- Ali'nin boyu 172 cm'dir.

cm

- Kitabın ağırlığı 2 kg'dır.

kg

Birimlerde Bulunması Gereken Özellikler



Eşitlik

Genellik

Kullanışlılık

Eşitlik

- Birimlerin birbirine eşit olması
- Metreyi oluşturan her santim

Genellik

- Birimin herkes tarafından aynı şekilde anlaşılması
- Ağırlık ölçmede kullanılan kilogram

Kullanışlılık

- Birimlerin kullanışlı olması
- Yol uzunluğunun “m” yerine “km” ile ifade edilmesi

Ölçme ile İlgili Kavramlar

Ölçme kuralı

- Her soru 5 puandır.
- Not baremi

Not baremi:
Her soru 5 puandır.



Ölçme

- Bugün hava güzeldir.
- Ayşe sınavdan 100 aldı.



Ölçüm

- Güzel
- 100



Değerlendirme Nedir?

➔ Ölçme sonuçlarını bir ölçüte vurarak, ölçülen nitelik hakkında bir değer yargısına varma sürecidir.





- ➔ Ayşe sınıfını geçti.
- ➔ Mehmet basketbol takımına seçildi.
- ➔ Fırat KPSS'den 83 puan alarak atandı.



Değerlendirme ile İlgili Kavramlar

Ölçüt

- KPSS'den 85 alan matematik öğretmenleri başvuruda bulunabilir.
- Matematik öğretmenliği bölümünden KPSS göre ilk 100 'e girenler atanabilir.

Değerlendirme

- KPSS ' den derece yaparak atandı.
- Seviyesi düşük olan öğrenci sınıfta kaldı.

Ölçüt

Neye göre geçilip
neye göre
kalinacağını
belirlerken
kullanılan kriterdir.

Ölçüt

Mutlak

Bağıl



Mutlak Ölçüt Nedir?

Mutlak Ölçüt



Ölçmeci tarafından belirlenir.



Kişiden kişiye değişiklik göstermez.



Kesin, net ve standarttır.



Bağıl Ölçüt Nedir?

Bağıl ölçüt



Öğrenci belirler.



Öğrenciler birbirini etkiler.



Kesinlik, netlik ve standartlıktan söz edilemez.





- ➔ Sınıf **ortalamasının altında** puan alanlar sınıfta kalmış sayılır.
- ➔ Öğrenciler gelirlerine göre sıralanıp **en fakir** 5 öğrenciye burs verilecektir.
- ➔ Koşu yarışında **ilk üçe** girenler atletizm takımına girecektir.



**Mutlak ölçüt kullanılarak yapılan değerlendirme
Mutlak değerlendirmedir !**

(Ölçüt Dayanaklı, Kriter Referanslı)

**Bağıl ölçüt kullanılarak yapılan değerlendirme
Bağıl değerlendirmedir !**

(Norm Dayanaklı, Norm Referanslı)



Değerlendirme Türleri



Örnek:

➔ Boyu 1.80m ve üzerinde olan kız öğrenciler basketbol takımına seçildi.



➔ En az 5 yıl deneyimi olanlar işe alındı.



➔ Yüzme yarışmasında ilk 10'a girenler kulübe girmeye hak kazandı.



➔ Matematik dersinde belirlenen hedeflerin %70'ini kazanan öğrenciler sınıfı geçti.



Tanımaya ve Yerleřtirmeye Dönük Değerlendirme (Hazırbulunuşluğun değerlendirilmesi, Tanı koymaya dönük, Diagnostik)

- ➔ Öğretim sürecinin başında öğrencileri tanımak, hazırbulunuşlukları hakkında bilgi sahibi olmak.
- ➔ Öğrencilerin öğrenme ihtiyaçlarını tespit etmek.
- ➔ Ön koşul davranışların öğrencilerde bulunup bulunmadıklarını görmek.
- ➔ Not vermek amaç değildir.



Uygulama Şekilleri



Hazırbulunuşluk testleri



Seviye belirleme sınavları



Seçme testleri



Tanıma testleri



Özel ve genel yetenek testleri



Muafiyet sınavları

Biçimlendirme ve Yetiştirmeye Dönük (Formatif/İzlemeye dönük) Değerlendirme



Öğretim süreci devam ederken gerçekleşir.



Ünite ya da konu bitiminden sonra yapılır.



Öğrenme eksiklerini ve yanlışlıkları ortaya çıkarmak ve bunları ortadan kaldırmak için etkinlikler yapılır.



Not vermek amaç değildir.



Uygulama Şekilleri

- ➔ İzleme testleri
- ➔ Formatif testler
- ➔ Ünite testleri
- ➔ Quizler



Düzey Belirlemeye Dönük Değerlendirme (Summatif değerlendirme-Değer biçmeye yönelik değerlendirme-Durum muhasebesine yönelik değerlendirme-Öğrenci başarısının değerlendirilmesi)

- ➔ Öğretim süreci sonunda (dönem sonunda) gerçekleştirilir.
- ➔ Öğrencilerin başarı seviyeleri belirlenir.
- ➔ Öğrencilerin hangi hedef davranışları kazandıklarını ya da hedeflerin ne kadarını kazanamadıklarını belirlemeye yarar.
- ➔ Geçti-kaldı, başarılı oldu yada başarısız oldu yorumları yapılır.
- ➔ Not vermek amaçtır.



Uygulama şekilleri



Summatif (dönem sonu) testler



Final sınavları



Vize sınavları



Eriş testleri



Başarı testleri



Düzey belirleme sınavları

Ünitemiz Sona Ermiştir



Ölçmede Hata ve Ölçme Aracında Bulunması Gereken Teknik Özellikler



Ölçmede Hata



Hata : Bir özelliğin gerçek değeri ile ölçülen değeri arasındaki farktır.

Hatasız, mükemmel bir ölçme aracı yoktur. Bütün ölçme sonuçlarına az ya da çok hata karışır!

Hata Türleri

```
graph TD; A[Hata Türleri] --> B[Sabit Hatalar]; A --> C[Sistematik Hatalar]; A --> D[Tesadüfi Hatalar]
```

**Sabit
Hatalar**

**Sistematik
Hatalar**

**Tesadüfi
Hatalar**



Sabit Hatalar

Gerçek Değer	Ölçülen Değer
5 kg	7 kg
12 kg	14 kg
35 kg	37 kg
23 kg	25 kg
32 kg	34 kg
3 kg	5 kg
8 kg	10 kg



Sabit miktarda
karışan hata



Ölçülen özelliğin
miktarı değıştikçe
hata miktarının da
değişmesi sonucu
oluşan hataya
“sistematiik hata”
denir.





Örnek:



Her öğrencinin puanından 10 puan düşürmek.



Yapılan sınavda, dördüncü sonunun okunamaması nedeniyle *hiçbir öğrenci* tarafından cevaplandırılmaması.



2 cm'lik kısmı kırık olan 50cm'lik bir cetvel ile *bir kez* kullanarak ölçme yapmak.

Sistemik Hata

Gerçek Değer	Ölçülen Değer
10 kg	11 kg
20 kg	22 kg
30 kg	33 kg
40 kg	44 kg
50 kg	55 kg



Sanki belli bir
kural varmış gibi





Ölçülen özelliğin
miktarı değıştikçe
hata miktarının da
değişmesi sonucu
oluşan hataya
“sistematiik hata”
denir.





Hap Bilgi

Eğitimdeki sistematik hatada öğretmenin **yanlılığı** ve **taraf tutması** söz konusudur!





Örnek:

- ➔ Öğretmenin yaptığı sınavı değerlendirirken hemşehrilerine fazladan puan vermesi
- ➔ Sınav sonuçlarının düşük olması üzerine öğretmenin her öğrencinin sınav sonuçlarına %10 puan eklemesi
- ➔ Müzik öğretmeninin müzik korosunda olan öğrencilerine 10 puan fazla vermesi

Tesadüfi (random/rastgele) Hata

Gerçek Değer	Ölçülen Değer
5 kg	5 kg
12 kg	12,5 kg
20 kg	23 kg
8 kg	8 kg
11 kg	10,6 kg



Ne sabit
ne sistematik



Ölçme sonuçlarına tesadüfen karışan; yönü miktarı ve nedeni bir başka deyişle kaynağı belli olmayan hatalar **“tesadüfi hata”** denir.



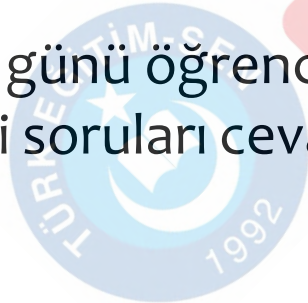
Hem arttırıcı hem de azaltıcı yönde karışabilmektedir.





Örnek:

- ➔ Dikkatsizliği nedeniyle öğretmenin yaptığı sınavı puanlarken olması gerekenden farklı puanlar vermesi
- ➔ Öğretmenin sınavlara verdiği puanları not defterine işlerken farkında olmadan yanlış yazması
- ➔ Sınav günü öğrencinin hasta ya da çok kaygılı olması nedeniyle bildiği soruları cevaplayamaması



Hata Kaynakları

ÖLÇMECİ

ÖLÇME
ARACI

ÖLÇÜLEN
BİREY

ÖLÇME
ORTAMI



Bir Ölçme Aracında Bulunması Gereken Teknik Özellikler

GEÇERLİK

GÜVENİRLİK

KULLANIŞLILIK



Güvenirlik

- ➔ Ölçme sonuçlarının tesadüfi hatalardan arınık olması, yani ölçme sonuçlarına hata karışmadan ölçme işleminin yapılabilmesidir.
- ➔ Ölçme sonuçlarında karışan hata miktarı ne kadar yüksek yüksek olursa güvenirlik o derece düşük olacaktır.



**Güvenirlik, genellikle paralel testler arasındaki
korelasyon olarak tanımlanır!**



Bir anlamda güvenirlik katsayısı korelasyon katsayısıdır.

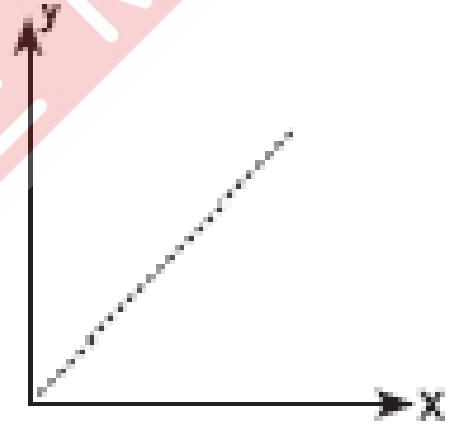


Korelasyon

- ➔ İki veya daha fazla değişken arasındaki ilişkinin miktarını ve yönünü sayısal olarak belirlememizi sağlayan istatistiksel bir tekniktir.
- ➔ Korelasyon katsayısı $+1$ ile -1 arasında değişen değerler alır.
- ➔ Korelasyon katsayısı $0,00$ ise iki değişken arasında bir ilişki yoktur.

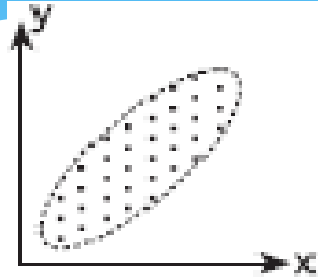


Öğrencinin Adı	1. Ölçüm	2. Ölçüm
Ali	80	80
Ayşe	45	45
Ahmet	65	65
Veli	32	32
Osman	98	98
Hüseyin	100	100
Mehmet	74	74

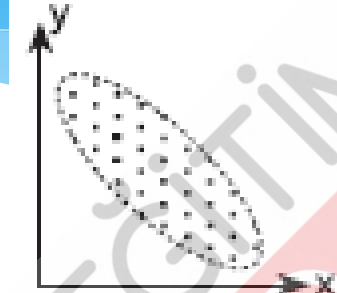


$$r_{xy} = 1,00$$

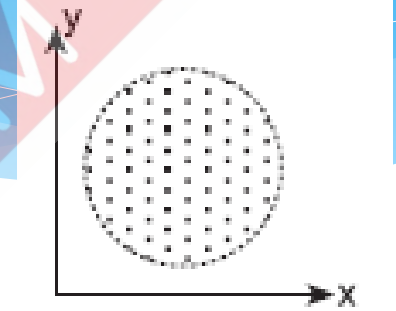




$$r_{xy} = 0,70$$



$$r_{xy} = -0,70$$



$$r_{xy} = 0,00$$



Pozitif yönde ilişki

- Zeka ile başarı
- Yaş ile eğitim düzeyi



Negatif yönde ilişki

- Sigara ile sağlık
- Yorgunluk seviyesi ile anlama düzeyi



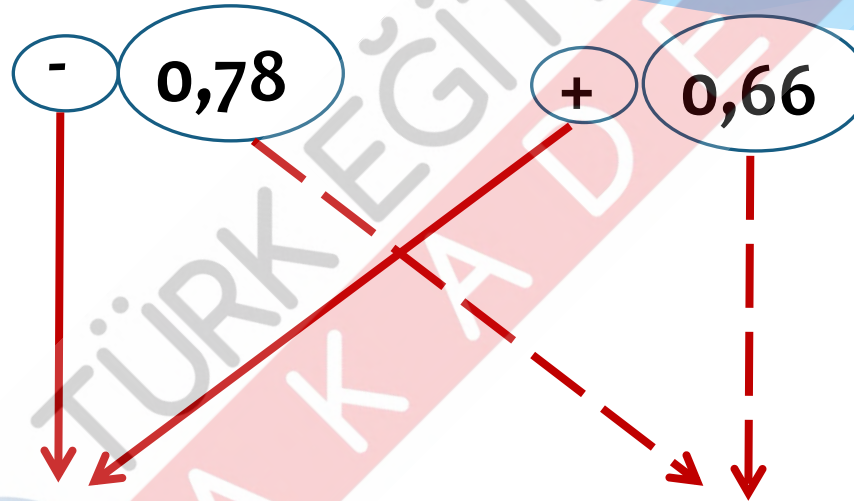
İlişki yok

- Boy ile zeka
- Ayakkabı no ile dersteki başarı





Uyarı !



İlişkinin Yönü

İlişkinin Miktarı



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?

- 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.



Hap Bilgi

Katsayısı -1 ile +1 arasındadır!

Güvenirlilik katsayısı 0 ile +1 arasındadır!



Güvenirlik Belirleme Yöntemleri

Birden Fazla Uygulamayla

Test Tekrar Test

Paralel (Eşdeğer) Test

Tek Uygulamayla

İki Yarı (Eşdeğer) Yöntemi

Maddenin tutarlılığına bağlı

*KR20/21

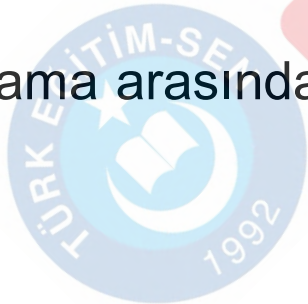
*Cronbach-Alfa

Testteki Hatayı Puan Cinsinden İfade Etme

Ölçmenin Standart Hatası

Test–Tekrar Test Yöntemi

- ➔ Bir test belli bir zaman aralığıyla aynı öğrenci grubuna iki kez uygulanır.
- ➔ İki uygulama sonucunda öğrencilerin aldıkları puanlar arasındaki korelasyon katsayısı hesaplanır.
- ➔ Korelasyon katsayısı güvenirlik katsayısı olarak kabul edilir ve bu güvenirlik **“kararlılık”** anlamında güvenirlik olarak bilinir.
- ➔ İki uygulama arasında geçen zaman önemlidir (3-4 hafta).





Hap Bilgi

de edilen korelasyon katsayısının
yüksek olduğunu söyleyebilmemiz için korelasyon
katsayısının **0,70**'ten yüksek olması gerekir!



Paralel (Eşdeğer) Formlar Yöntemi

- ➔ Birbirine paralel yani aynı davranışları ölçen, fakat farklı sorulardan oluşan iki test hazırlanır.
- ➔ İki test aynı öğrencilere uygulanır ve öğrencilerin aldıkları puanlar arasındaki korelasyon katsayısı hesaplanır.
- ➔ Korelasyon katsayısı güvenirlik katsayısı olarak kabul edilir.
- ➔ Bu yöntemle elde edilen güvenirlik katsayısı **“eşdeğerlik katsayısı”** olarak da ifade edilir.
- ➔ Korelasyon katsayısının yüksek olması iki testin birbiriyle **tutarlı**, aynı zamanda ölçümlerin de **kararlı** olduğunu bildirir.

İki Yarı (Eşdeğer Yarılar) Yöntemi

- ➔ Bir test hazırlanır ve öğrencilere uygulanır.
- ➔ Uygulama sonuçları yani test maddeleri iki yarıya ayrılır.
- ➔ İki yarıdan elde edilen puan dağılımları arasındaki korelasyon hesaplanır.
- ➔ Hesaplanan korelasyon katsayısının yüksek olması , iki yarıdan elde edilen puanların **tutarlı** olduğu, yani testin güvenilir olduğu anlamına gelir.





Elde edilen korelasyon katsayısı testin yarısına ait korelasyon katsayısıdır. Tüm testin korelasyon katsayısı aşağıdaki formülle hesaplanır.

$$\text{Testin güvenirliği} = \frac{2 \times \text{Testin yarısından elde edilen korelasyon}}{1 + \text{Testin yarısından elde edilen korelasyon}}$$



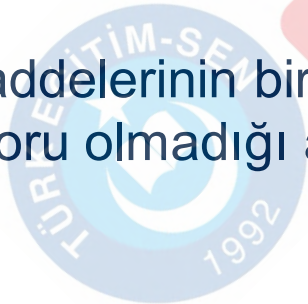


Tüm testin güvenilirlik katsayısı kesinlikle testin yarısından elde edilen korelasyon katsayısından büyüktür !



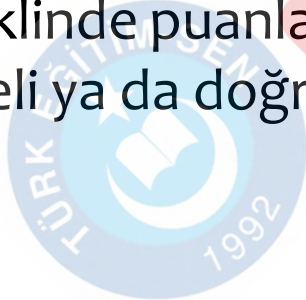
İç Tutarlılık Güvenirliği Hesaplama Yöntemi

- ➔ Bir testin, bir defa uygulanması ve testte yer alan maddelerin analiz edilmesine dayanır.
- ➔ İç tutarlılık güvenirlği bir testteki maddelerin birbirleriyle ne derece tutarlı olduğunun belirlenmesi için kullanılan bir güvenirlık hesaplama yöntemidir.
- ➔ Test maddelerinin birbirleriyle tutarlı olması, testte ilgisiz ya da hatalı soru olmadığı anlamına gelir.



KR-20 / KR-21

- ➔ Test bir kez uygulanır.
- ➔ İç tutarlılık katsayısı olarak adlandırılır.
- ➔ Testteki maddelerin aynı yeterliği ölçtüğü anlamına gelir (testin tek boyutlu olduğu).
- ➔ 1-0 şeklinde puanlanan maddeli testlerde uygulanabilir (çoktan seçmeli ya da doğru-yanlış).





Uyarı !

KR-20: Testteki her bir maddenin güçlük derecelerinin hesaplanmış olması gerekir.
(Madde analizi yapılmış testler)

KR-21: Testteki her bir maddenin güçlük derecelerinin eşit olduğu varsayılır.
(Madde analizi yapılmamış testler)

KR-20 ve KR-21 Arasındaki İlişki

- ➔ $KR-21 < KR-20$
- ➔ KR-21 güvenirliğin alt sınırıdır.
- ➔ Güvenirlik ister KR-21 ister KR-20 ile hesaplansın yapılacak yorum değişmez.



KR-20 ya da KR-21 yüksek ise yapılabilecek yorumlar

- ➔ Uygulanan test hatalardan arındırılmıştır.
- ➔ Ölçülen değişken tek boyutludur. (Farklı bir değişken, amaç dışı bir soru karışmamıştır.)
- ➔ Testi oluşturan maddeler homojendir.



Cronbach Alfa: KR yöntemlerinden tek farkı 1-0 şeklinde puanlanmayan testlerde maddeler arasındaki tutarlılığı bulmak için kullanılmasıdır. (Likert tipi tutum ölçekleri ve klasik yazılılar gibi)



Ölçmenin Standart Hatası

- ➔ Güvenirlik için hatasızlığın bir ölçüsü olduğunu söylemiştik. Güvenirlik katsayısı ölçme işleminde yapılan hatanın miktarını belirtmez.
- ➔ Standart hata bize hatanın miktarını söyler.
- ➔ Öğrencilerin gerçek puan aralıklarını bulmamızı sağlar.



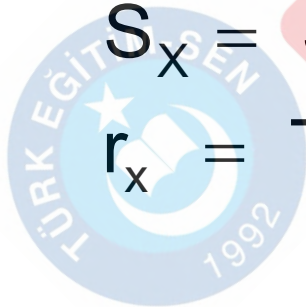
Ölçmenin Standart Hatası

$$S_e = S_x \sqrt{1 - r_x}$$

S_e = Ölçmenin standart hatası

S_x = Standart sapma

r_x = Testin güvenirliği





Örnek:

➔ Standart sapması 4, güvenirlik katsayısı 0,75 olan bir testin standart hatası kaçtır?



Çözüm:

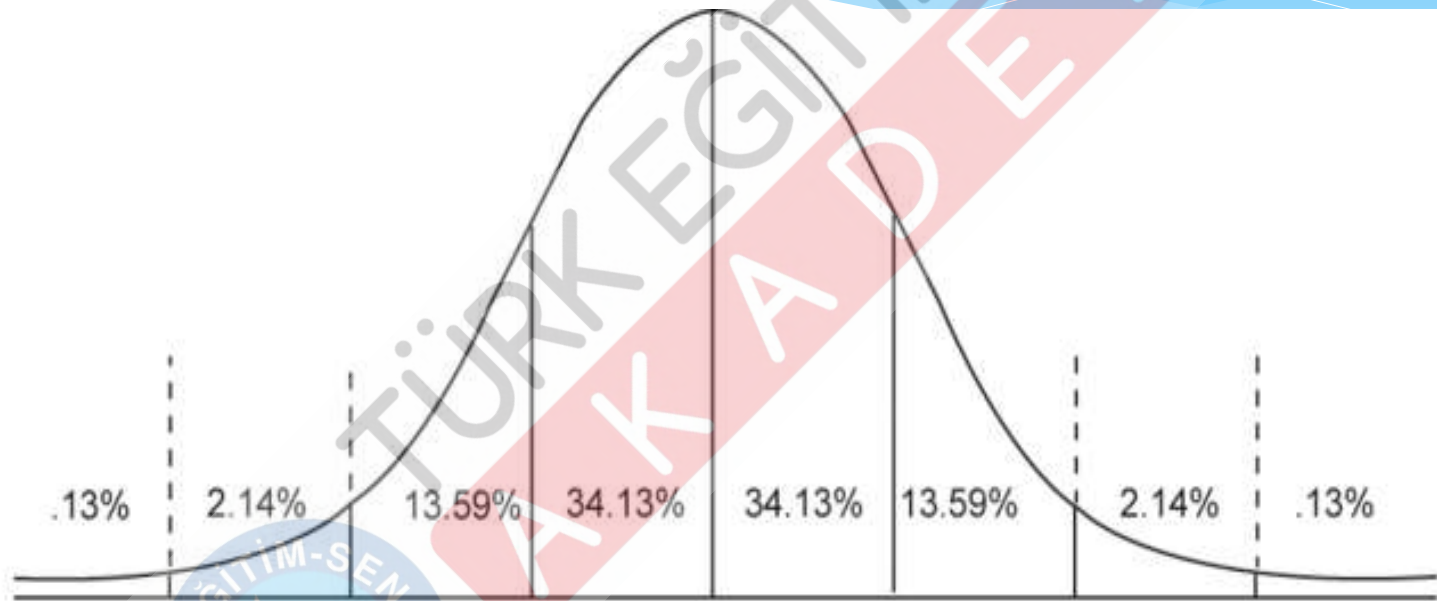
$$S_e = S_x \sqrt{1 - r_x}$$

$$S_e = 4 \sqrt{1 - 0,75}$$

$$S_e = 2$$



NORMAL DAĞILIM EĞRİSİ



SS

$-3Ss$

$-2Ss$

$-1Ss$

$1Ss$

$2Ss$

$3Ss$

Yüzdelik

0.13

2,27

15,86

50

84,12

97,71

99,85

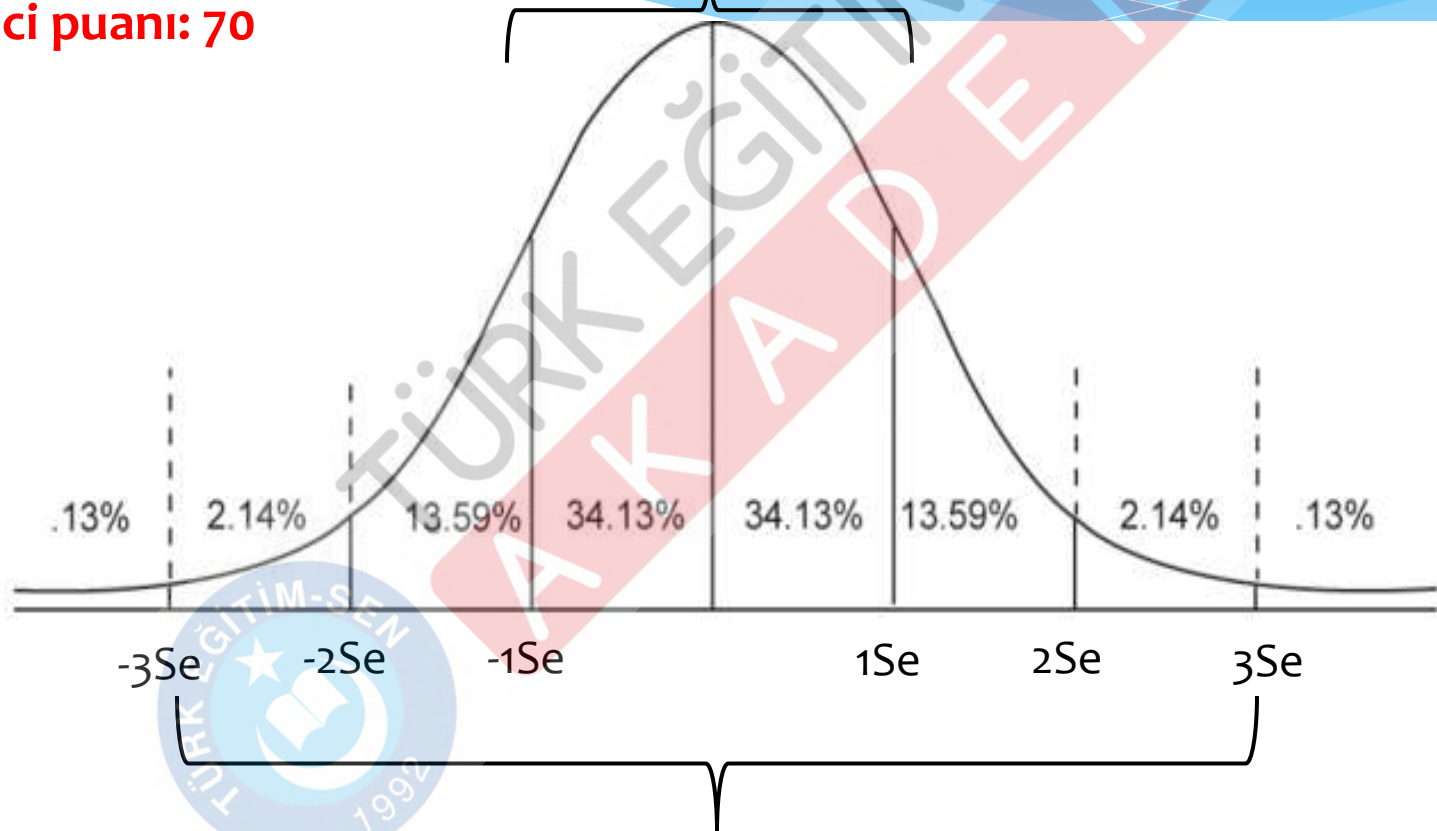


%95 olasılıkla 66-74 arasında

%68 olasılıkla 68-72 arasında

Öğrenci puanı: 70

Se = 2



%99 olasılıkla 64-76 arasında

Güvenirlilik

KARARLILIK

TUTARLILIK

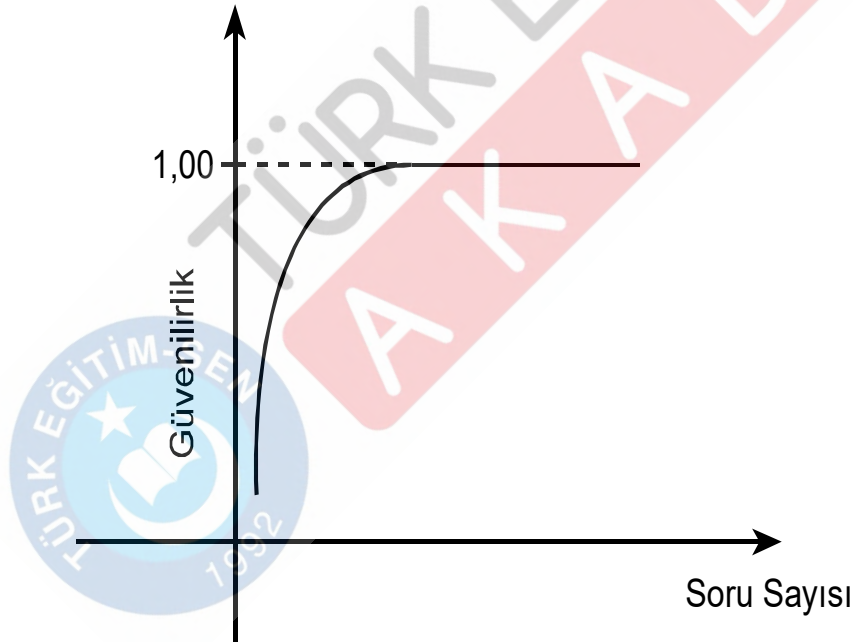
GÜVENİRLİK

OBJEKTİFLİK

DUYARLILIK

Güvenirligi Arttıran Faktörler

- ➔ Soruların açık, net ve anlaşılır olması
- ➔ Sınavdaki soru sayısının arttırılması



Güvenirliği Etkileyen Faktörler

- ➔ Sınavdaki soruların, öğrencilerin yaklaşık yarısı tarafından yanıtlanabilir olması
- ➔ Şans başarısının düşük olması
- ➔ Öğrencilerin psikolojik özelliklerinin olumlu olması
- ➔ Testin uygulandığı grubun heterojen olması
- ➔ Sınav süresinin yeterli olması
- ➔ Sınav sonuçlarının objektif puanlanması
- ➔ Yönergenin bulunması

Geçerlik



Bir ölçme aracının ölçmek istediği özelliği başka değişkenler ya da özelliklerle karıştırmadan tam ve doğru ölçmesidir.



Bir ölçme aracının kullanılış amacına hizmet etme derecesidir.



**SABİT
HATA**

**D
O
Ğ
R
U
D
A
N**

GEÇERLİK

SİSTEMATİK HATA

**D
O
Ğ
R
U
D
A
N**

GEÇERLİK

TESADÜFİ HATA

**- L
Y
A
D
O**

**D
O
Ğ
R
U
D
A
N**

GÜVENİRLİK

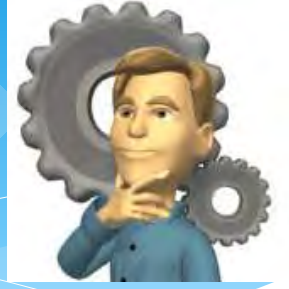
GEÇERLİK TÜRLERİ



Kapsam Geçerliği

- ➔ Kapsam, bir testin ölçmek istediği konulardır.
- ➔ Bir ölçme aracının ölçülmek istenen hedef davranışları (kazanımları) ne derecede örneklendirdiğidir.
- ➔ Belirtke tablosu ile test uyumuna bakmak.
- ➔ Uzman görüşüne başvurmak.

Yordama Geçerlik



- ➔ Yordama, istatistiksel yöntemlerle bilinen bir durumdan yola çıkarak geleceğe yönelik tahminde bulunmaktır.
- ➔ Yordama geçerliği, önceden uygulanmış bir testin sonuçlarının gelecekteki bir başarı hakkında önceden bilgi sağlayabilme derecesidir.
- ➔ Yordama geçerliğinde, ölçüt puan sonradan elde edilir.





Örnek:

YKS

**Üniversitedeki
akademik
başarı**



**Üniversite giriş
sınavındaki puan**

Ölçüt



2007

KPSS

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?

- ☒ A) Yordama geçerliği
- B) Yapı geçerliği
- C) Kapsam geçerliği
- D) İç tutarlılığı
- E) Puanlayıcılar arası güvenirliği

Uygunluk Geçerliği

- ➔ Bir diğer adı da **benzer ölçekler geçerliği**dir.
- ➔ Bir testin uygunluk geçerliğini bulurken testin aynı değişkeni ölçtüğünü bildiğimiz geçerli bir testten elde edilen puanların ilişkisine bakarız.
- ➔ Uygunluk geçerliğinde ölçüt puanlar ile sınav sonucu elde edilen puanlar arasında korelasyon katsayısı hesaplanır.



Örnek:

YKS Deneme sınavı

YKS

Ölçüt puan önce

**Sınav sonucu
sonra**



Yapı Geçerliđi



İnsan beyinde var olduđu düşünölen ama doğrudan gözlenemeyen psikolojik özelliklere yapı denir.(Zeka, matematik bilgisi, kişilik v.s.). Her psikolojik yapı kendi içinde bölümlere ayrılır.



Bu yapıların ölçölmesi söz konusu olduğunda testlerin de ölçölecek yapının bölümleri kaç taneyse o kadar sayıda bölüm barındırması ve bölümlerin her birini birbirine karıştırmadan ölçmesi gerekmektedir.



Görünüş Geçerliği

- ➔ Bir testin ya da bir maddenin ölçmek istediği değişkeni ölçüyor gibi görünmesidir.
- ➔ **Örneğin**, bir testin yönergesinde “Fizik soruları” yazıyorsa, ilk bakışta o test fizik konularıyla ilgili gibi görünmelidir.



Geçerliğı Artırma Yolları

- ➔ Her sorunun belirtke tablosunda yer alan bir davranışı ölçmesi kapsam geçerliğini dolayısıyla da geçerliğı artırır.
- ➔ Sınavda kopya çekme, ipuçlarından yararlanma, tahminle doğru yanıtlama gibi geçerli olmayan davranışlar engellenmelidir.
- ➔ Sınavdan önce soruların verilmesi, aynı soruların tekrar tekrar kullanılması vs. geçerliğı düşüreceğinden test hazırlarken ve uygularken bunlara dikkat edilmelidir.

Geçerlik- Güvenirlik

- ➔ Güvenirlik, geçerlik için ön koşuldur.
- ➔ Güvenirliği etkileyen her şey geçerliği de etkiler.
- ➔ Geçerli olan bir test güvenilirdir. Fakat güvenilir bir test her zaman geçerli olmayabilir.
- ➔ Şans başarısı ve kopya hem güvenirliği hem de geçerliği düşürür.
- ➔ Öğretmenin her yıl aynı soruları sorması öncelikle geçerliği düşürür.

GÜVENİRLİK	GEÇERLİK
Tesadüfi hatalardan etkilenir.	Doğrudan sabit ve sistematik hatadan dolayı olarak tesadüfi hatadan etkilenir.
Ölçme aracı ile ilgilidir.	Ölçme işleminin amacı ile ilgilidir.
Hatalardan arınlık derecesidir.	Amaca hizmet etme derecesidir.
Duyarlık, tutarlılık, objektiflik, kararlılık	Amaç, amaca uygunluk, hedef davranış
Güvenilir her test geçerli değildir.	Güvenirliği içine alır.
Geçerlik için ön koşuldur.	Güvenirliği içine alır.
0 ile +1 arasındadır.	-1 ile +1 arasındadır.

Kullanışlılık



➔ Ekonomik olmalı

➔ Uygulaması ve puanlaması pratik olmalı

➔ Uygulama süresi kısa olmalı

➔ Ölçme aracı yapan kişiden istenen becerinin az olması

➔ Puanların yorumlanmasının kolay olması

Ünitemiz Sona Ermiştir



Ölçme Sonuçları Üzerine İstatistiksel İşlemler



TEST İSTATİSTİKLERİ

ARİTMETİK ORTALAMA

- ➔ Bir öğrenci grubunun bir testten aldıkları puanların toplamının (ya da verilerin toplamının), öğrenci sayısına yada veri sayısına bölünmesi ile hesaplanan merkezi eğilim ölçüsüne “**aritmetik ortalama**” denir.
- ➔ Aritmetik ortalama, öğrenci grubunun (sınıfın) başarısı hakkında bilgi verir.

Bir grup verinin aritmetik ortalamasının hesaplanması

$$\bar{X} = \frac{\sum X}{N}$$

\bar{X} : Aritmetik ortalama

$\sum X$: Verilerin ya da ölçümlerin toplamı

N: Veri ya da kişi sayısı



Bir grup verinin aritmetik ortalamasının hesaplanması



Bir sınıftaki 4 öğrenci İngilizce dersinden 20, 40, 36 ve 52 almıştır. Bu öğrencilerin aldıkları puanların aritmetik ortalaması kaçtır?



$$\bar{X} = \frac{\sum X}{N} = \frac{20 + 40 + 36 + 52}{4} = \frac{148}{4} = 37$$

Frekans tablolarında aritmetik ortalamasının hesaplanması

$$\bar{X} = \frac{\sum f.x}{N}$$

\bar{X} : Aritmetik ortalama

$\sum f.x$: Verilerin ya da ölçümlerin toplamı

N: Veri ya da kişi sayısı



Frekans tablolarında aritmetik ortalamasının hesaplanması

Örnek:

Puan (x)	Frekans (f)
80	3
70	2
60	2
45	5
30	8

Verilen tabloda 20 öğrencinin tarih dersinden aldığı puan dağılımı verilmiştir.

Bu verilere göre puan dağılımının aritmetik ortalaması kaçtır?



Çözüm:

Puan (x)	Frekans (f)	f.x
80	3	240
70	2	140
60	2	120
45	5	225
30	8	240

$$\sum f.X = 240 + 140 + 120 + 225 + 240$$

$$\sum f.X = 965$$

$$\bar{X} = \frac{\sum f.X}{N} = \frac{965}{20} = 48.25$$



Gruplandırılmış verilerin aritmetik ortalamasının hesaplanması

$$\bar{X} = \frac{\sum f.X_o}{N}$$

\bar{X} : Aritmetik ortalama

$\sum f.X_o$: $f.X_o$ değerlerinin toplamı

N: Veri ya da kişi sayısı



Gruplandırılmış verilerin aritmetik ortalamasının hesaplanması



Örnek:

Puan Aralığı (x)	Frekans (f)
45 – 49	2
40 – 44	3
35 – 39	5
30 – 34	4
25 – 29	2
20 – 24	4

Verilen tabloda 20 öğrencinin tarih dersinden aldığı puan dağılımı verilmiştir.

Bu verilere göre puan dağılımının aritmetik ortalaması kaçtır?



Çözüm:

Puan Aralığı	f (frekans)	X_o (Orta Değer)	$f \cdot X_o$
45 – 49	2	47	94 (47 x 2)
40 – 44	3	42	126 (42 x 3)
35 – 39	4	37	148 (37 x 4)
30 – 34	4	32	128 (32 x 4)
25 – 29	2	27	54 (24 x 2)
20 – 24	5	22	110 (22 x 5)
	$N = 20$		$\Sigma fX_o = 660$

$$\bar{X} = \frac{\sum f \cdot X_o}{N}$$

$$\bar{X} : \frac{660}{20} = 33$$

AĞIRLIKLI ORTALAMA

Puanların ortalamaya olan katkılarına farklı ağırlıklar verilerek hesaplanan ortalamaya denir.



Puan (x)	Kredi	Not
Tarih	4	4
Kimya	4	4
Biyoloji	3	2
Matematik	3	2

Yandaki tabloda 4 dersin kredileri ve bir öğrencinin bu derslerden aldığı notlar gösterilmiştir.

Öğrencinin ağırlıklı ortalaması kaçtır?



Çözüm:

Puan (x)	Kredi	Not	<u>NotxKredi</u>
Fizik	4	4	16
Kimya	4	4	16
Biyoloji	3	2	6
Matematik	3	2	6
Toplam	14	12	44

$$\bar{X}_{\text{ağırlıklı}} = \frac{\sum (\text{Kredi} \times \text{Not})}{\text{Toplam Kredi}} = \frac{44}{12} = 3.66$$

Aritmetik ortalama olan tablo sorularında,

- ➔ Grubun başarı düzeyinin
- ➔ Grubun mutlak başarı düzeyinin
- ➔ Öğrencilerin ortalama başarı düzeyinin
- ➔ Öğrencilerin öğrenme düzeylerinin en düşük veya en yüksek olduğu ders ya da test hangisidir.

Sorularının cevapları bulunur.

Örnek:

	Aritmetik Ortalama	Mod (Tepe Değer)	Ortanca Medyan	Standart Sapma
İşletme	70	50	65	4
İktisat	74	88	80	2
Tarih	68	56	60	3
Anayasa	65	65	65	5
Matematik	57	68	46	2

Tabloya göre sınıfın en başarılı olduğu ders hangisidir?

Cevap : İktisat

MEDYAN (ORTANCA)



Sıralanmış puan dağılımında tam ortada bulunan puana (değere) **medyan(ortanca)** denir.



Bir grup veri dağılımında veri sayısı
tek olduğunda ortanca hesaplama:

$$X_{\text{ort}} = \frac{n+1}{2}$$

X_{ort} : Ortanca

n : Veri sayısı





12, 14, 10, 23, 35, 42, 38, 68, 78 veri grubunun ortancası kaçtır?



10, 12, 14, 23, 35, 38, 42, 68, 78

~~10~~, ~~12~~, ~~14~~, ~~23~~, 35, ~~38~~, ~~42~~, ~~68~~, ~~78~~

$$X_{\text{ort}} = \frac{n+1}{2} = \frac{9+1}{2} = 5$$

Bir grup veri dağılımında veri sayısı
çift olduğunda ortanca hesaplama:

$$X_{\text{ort}} = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$$

X_{ort} : Ortanca

n : Veri sayısı

Örnek:

12, 14, 10, 23, 36, 42, 38, 68, 78, 89 veri grubunun ortancası kaçtır?

Çözüm:

10, 12, 14, 23, 36, 38, 42, 68, 78, 89

$$\frac{n}{2} = \frac{10}{2} = 5 \text{ ve } \frac{n}{2} + 1 = \frac{10}{2} + 1 = 6$$

~~10, 12, 14, 23, 36, 38, 42, 68, 78, 89~~

$$\frac{36 + 38}{2} = 37$$



Örnek:

Puan (x)	Frekans (f)
67	6
56	8
37	5
30	2
23	5
10	2

Tablodaki verilere göre bu dağılımın ortancası kaçtır?



Çözüm:

$$\sum f = 28$$

14. ve 15. ortancadır?

$$\text{medyan} = \frac{37 + 56}{2} = 46.5$$

Puan (x)	Frekans (f)	
67	6	23. 24. 25. 26. 27. 28. 29.
56	8	15. 16. 17. 18. 19. 20. 21. 22.
37	5	10. 11. 12. 13. 14.
30	2	8. 9.
23	5	3. 4. 5. 6. 7.
10	2	1. 2.

MOD (TEPE DEĞER)

➔ Bir puan ya da veri dağılımında en çok tekrar edilen, yani frekansı en yüksek olan değere **mod** denir.





Örnek:

12, 12, 13, 14, 23, 23, 23, 32, 45, 56, 56, veri grubunun modü kaçtır?



Çözüm:

12, 12, 13, 14, 23, 23, 23, 32, 45, 56, 56

mod: 23





Örnek:

Puan (x)	Frekans (f)
67	6
56	14
37	2
30	2
23	5
10	8

Tablodaki verilere göre bu dağılımın modu kaçtır?



Çözüm:

Tabloda verilen puanlardan frekansı en yüksek olan puan **56**'dır.



Örnek:

Puan (x)	Frekans (f)
45-51	1
38-44	7
31-37	9
24-30	2
17-23	5
10-16	8

Yandaki tabloda verilen verilerin
modu kaçtır?



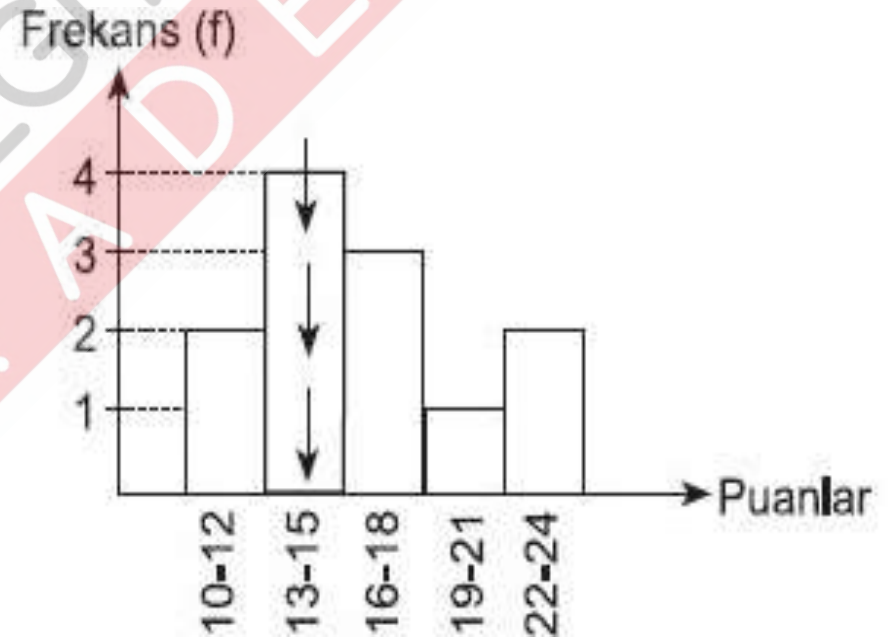
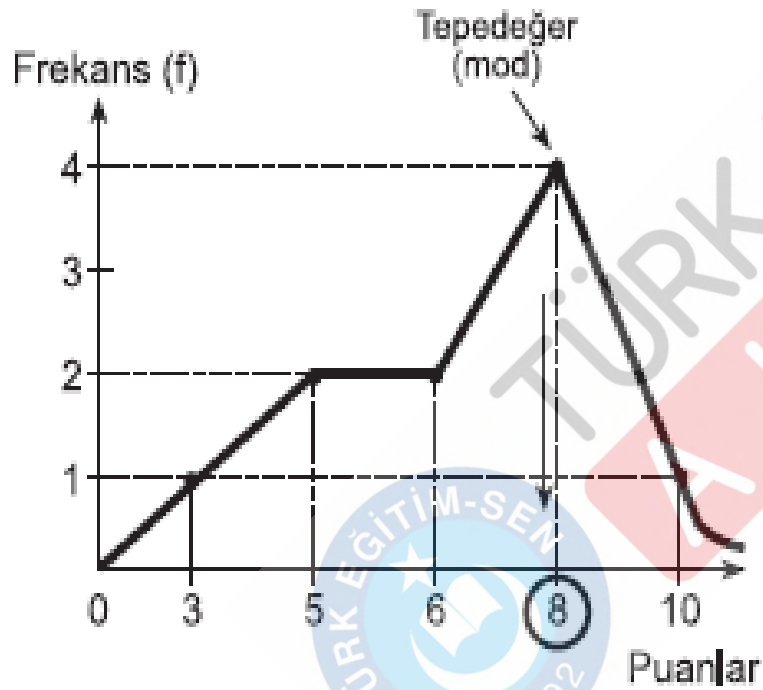


Çözüm:

Puan (x)	Frekans (f)	<u>X_o</u>
45-51	1	49
38-44	7	41
31-37	9	34
24-30	2	27
17-23	5	20
10-16	8	13

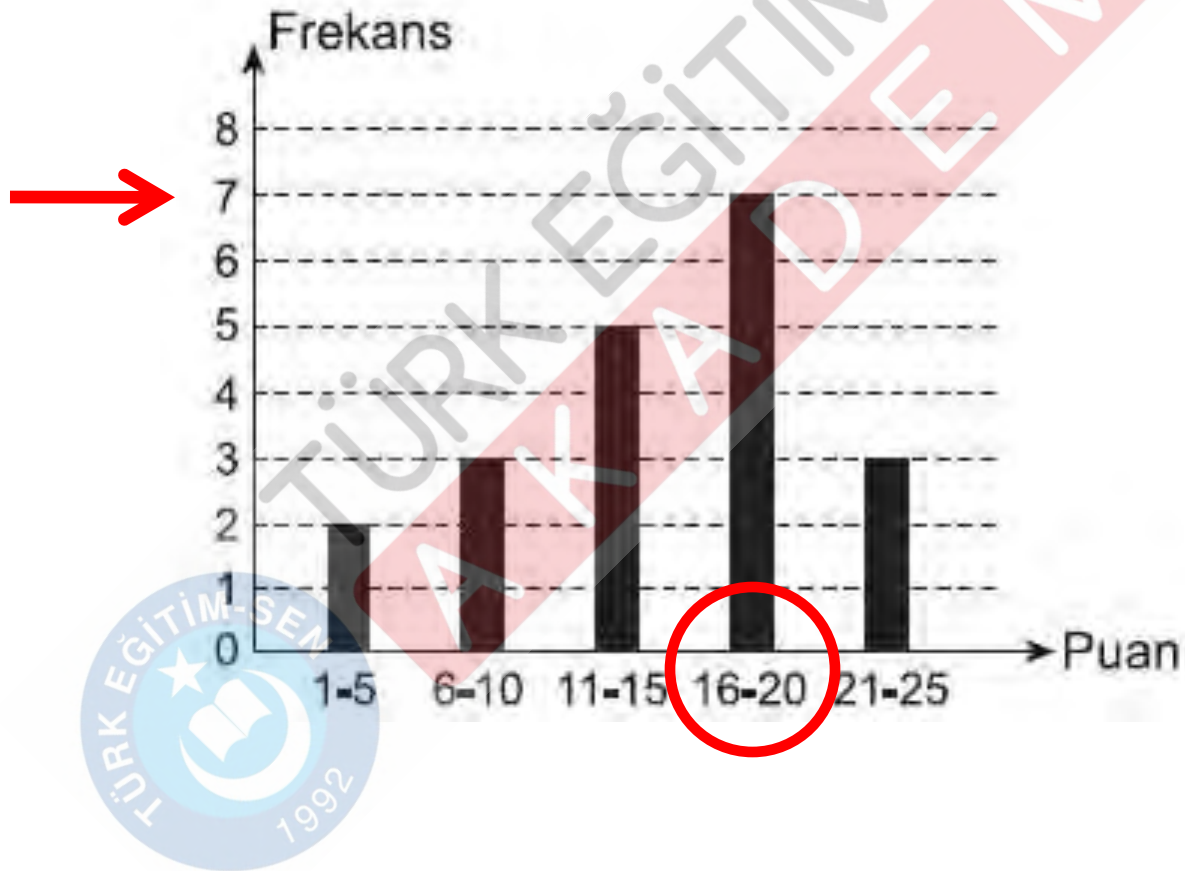
$$\frac{31 + 37}{2} = 34$$

Grafiklerde mod bulma





Örnek:





- ➔ Bir puan dağılımında, bütün puanların frekansları aynı ise bu puan dağılımının modu yoktur.
(1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3)
- ➔ Bir puan dağılımında ardışık iki değer en büyük ve eşit frekansa sahipse, mod bu iki değer ortalamasına eşittir.
(1, 2, 2, 3, 5, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 9)
- ➔ Bir puan dağılımında ardışık olmayan iki değer en büyük ve eşit frekansa sahipse bu dağılımın iki farklı modu olur.
(1, 2, 3, 5, 5, 5, 7, 8, 9, 9, 9)



10, 10, 20, 30, 30, 30, 40, 40, 40, 50, 50

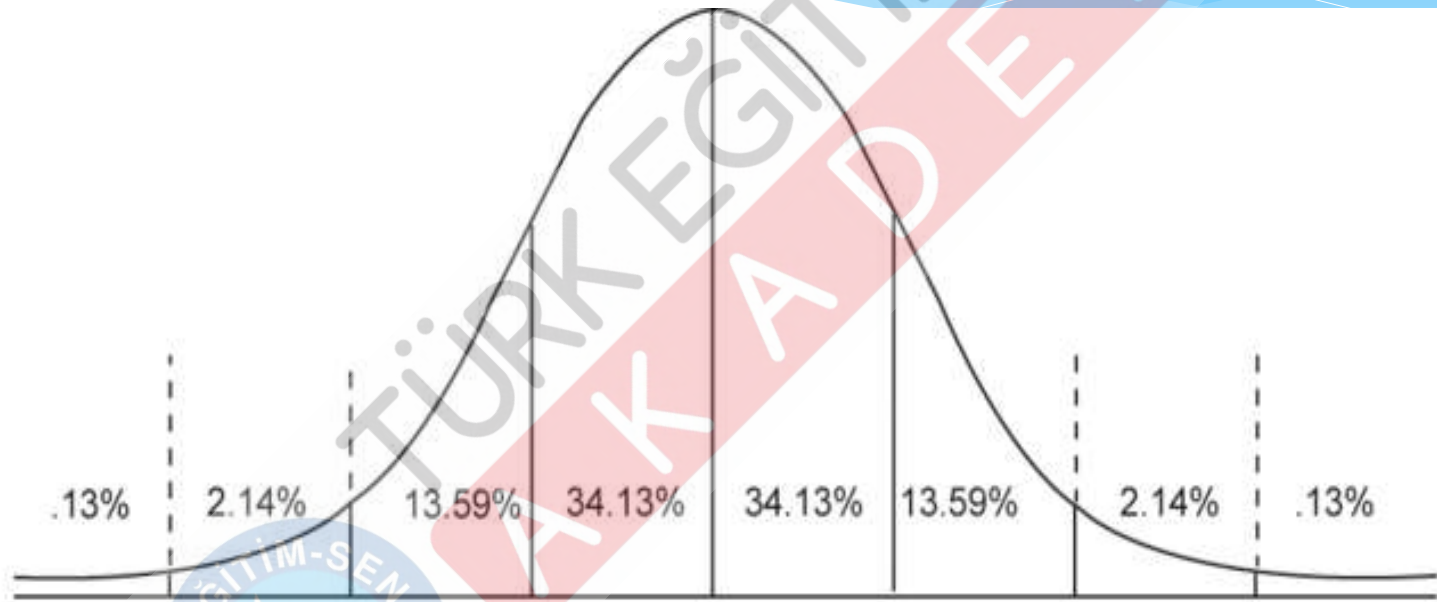
Puan dağılımının modu kaçtır?

CEVAP: 30 ve 40



DAĞILIMLAR

NORMAL DAĞILIMLAR (ÇAN EĞRİSİ)



SS

-3σ

-2σ

-1σ

1σ

2σ

3σ

Yüzdelik

0.13

2,27

15,86

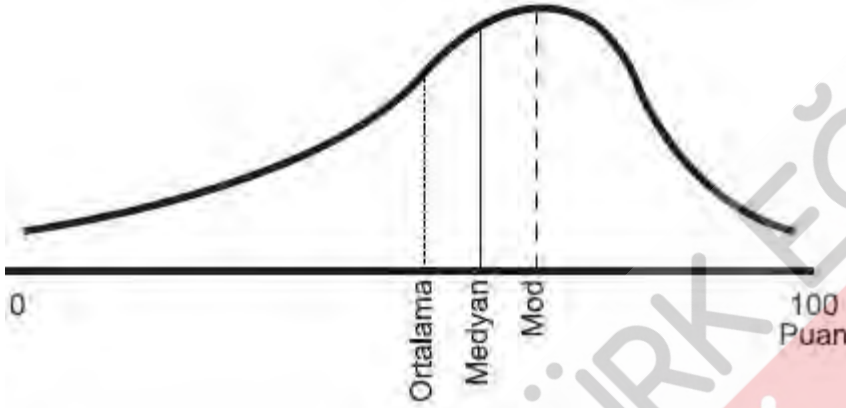
50

84,12

97,71

99,85

SOLA ÇARPIK DAĞILIM



$X_{ort} < Medyan < Mod$

Negatif kayışlıdır.

Öğretim yeterlidir.

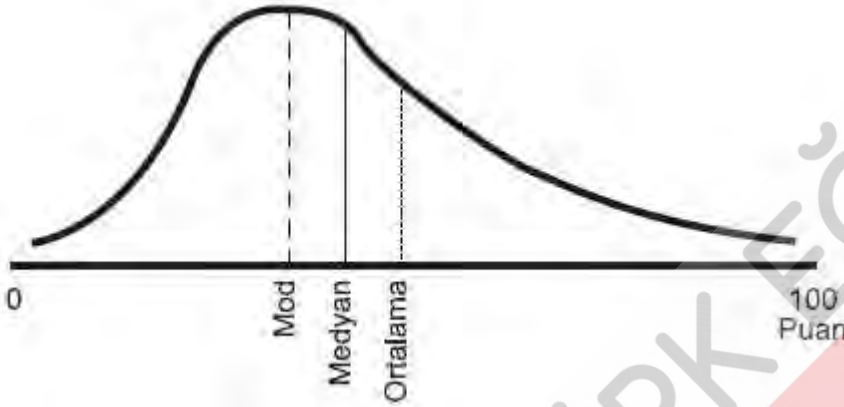
Test kolaydır.

Öğrencilerin başarıları yüksektir.

Öğrencilerin öğrenme düzeyi yüksektir.

Öğrenciler hedef-davranışları kazanmışlardır.

SAĞA ÇARPIK DAĞILIM



Mod < Medyan < Xort

Pozitif kayışlıdır.

Öğretim yetersizdir.

Test zordur.

Öğrencilerin başarıları düşüktür.

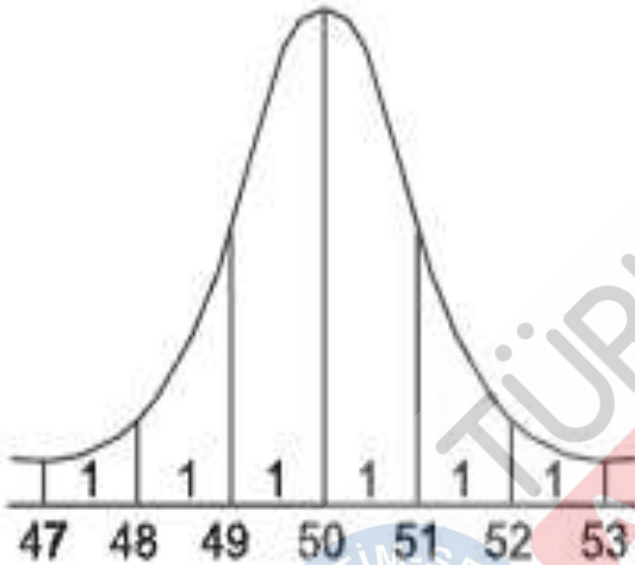
Öğrencilerin öğrenme düzeyi düşüktür.

Öğrenciler hedef davranışları kazanamamışlardır.

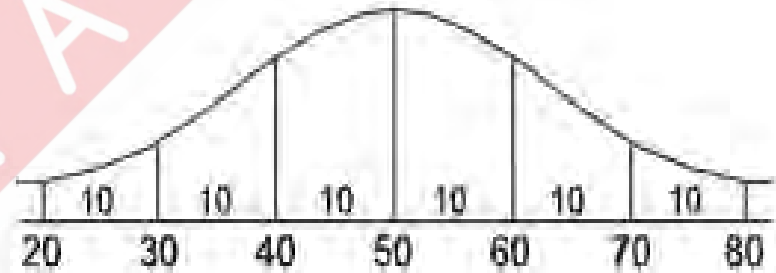
STANDART SAPMA

- ➔ Bir dağılıdaki verilerin ortalamadan ne kadar uzaklaştığının ortalamasını veren bir yayılma ölçüsüdür.
- ➔ Eğer standart sapma büyükse gruptaki puanlar aritmetik ortalamadan uzaklaşmış demektir. Bu durumda o grup için **heterojendir** denilebilir.
- ➔ Eğer standart sapma küçükse gruptaki puanlar ortalamadan fazla uzaklaşmamış demektir. Bu durumda o grup için **homojendir** denebilir.

Standart Sapma Küçükse;



Standart Sapma Büyükse;



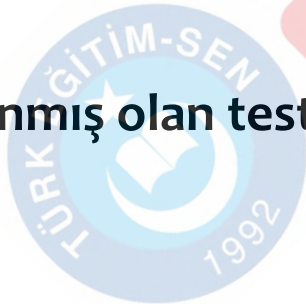
Standart sapma büyük olduğunda yapılacak yorumlar

- ➔ Testin uygulandığı grup heterojendir.
- ➔ Öğrenciler arasındaki farklılaşma fazladır.
- ➔ Bilen öğrenciler ile bilmeyen öğrenciler birbirinden ayrılmıştır.
- ➔ Uygulanmış olan testin ayırt ediciliği yüksektir.



Standart sapma küçük olduğunda yapılacak yorumlar

- ➔ Testin uygulandığı grup homojendir.
- ➔ Öğrenciler arasındaki farklılaşma azdır.
- ➔ Bilen öğrenciler ile bilmeyen öğrenciler birbirinden ayrılmamıştır.
- ➔ Uygulanmış olan testin ayırt ediciliği düşüktür.





Örnek:

	Aritmetik ortalama	Standart Sapma
Coğrafya	50	4
Psikoloji	30	3
Tarih	33	3
Anayasa	22	5
Matematik	50	2

Tabloya göre en homojen ve en heterojen sınıf hangisidir?

	Aritmetik Ortalama	Mod (Tepe Değer)	Ortanca Medyan	Standart Sapma
1. Test	65	75	70	7
2. Test	35	55	40	6
3. Test	40	40	40	8
4. Test	40	25	30	9
5. Test	60	40	45	16

Bir sınıftaki öğrencilerin yanıtları 80 üzerinden puanlanmış ve puanlardan yukarıdaki tabloda bulunan veriler çıkarılmıştır. Buna göre en çok hangi testte öğrenciler arasında farklılaşma en fazladır?



I. grup: 100 üzerinden aritmetik ortalaması 50; standart sapması 5

II. grup: 6 üzerinden aritmetik ortalaması 3; standart sapması 1

Gruplardan hangisi daha heterojendir?



BAĞIL DEĞİŞKENLİK KATSAYISI

$$V = \frac{Sx}{\bar{X}} \cdot 100$$

V: Bağıl değişkenlik katsayısı

Sx: Standart sapma

\bar{X} : Aritmetik ortalama

Bağıl Değişkenlik Katsayısı	Puan Dağılımı	Dağılım Özelliği
20'den yukarı	Heterojen	Basık
20 civarı	Normal	Normal
20'den aşağı	Homojen	Sivri

RANJ (DİZİ GENİŞLİĞİ)

- ➔ Bir puan ya da veri dağılımında en büyük ölçme sonucuyla en küçük ölçme sonucu arasındaki farktır.
- ➔ Ranj, en yüksek puanla en düşük puan arasındaki farktır.
- ➔ Örneğin en yüksek puanın 83, en düşük puanınsa 41 olduğunu düşünelim. Bu durumda ranj $83 - 41 = 42$ olur.



Örnek:

Puan (x)	Frekans (f)
45-51	1
38-44	7
31-37	9
24-30	2
17-23	5
10-16	8

Yandaki tabloda verilen verilerin
ranj kaçtır?





Çözüm:

Puan (x)	Frekans (f)	<u>Xo</u>
45-51	1	49
38-44	7	41
31-37	9	34
24-30	2	27
17-23	5	20
10-16	8	13

$$49 - 13 = 36$$



Ranj deęerine gre yapılacak yorumlar



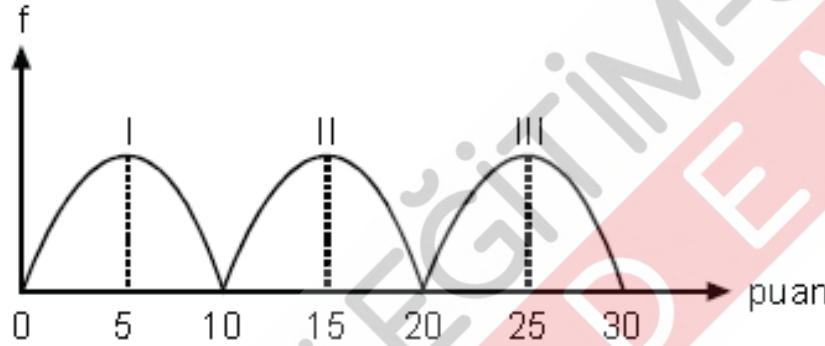
Ranj deęeri byk olan grupların test sonuları ranj deęeri kk olan grupların test sonularından daha gvenilirdir.



Ranj deęeri byk olan grubun heterojen olduęu, bilenle bilmeyenin birbirinden ayrıldıęı, bireylerin llen zellik bakımından farklılık gsterdięi sylenebilir.



Ranj deęeri kk olan bir grubun homojen olduęu, bilenle bilmeyenin birbirinden ayrılmadıęı, bireylerin llen zellik bakımından benzer olduęu sylenebilir.



Grafikte I, II, III numarayla gösterilen üç dağılım için aşağıdaki ifadelerden hangisi doğrudur?

- A) Ranj ve medyan değerleri aynıdır.
- B) Ranj değerleri aynı, mod değerleri farklıdır.
- C) Ranj değerleri farklı, mod değerleri aynıdır.
- D) Ranj değerleri farklı, medyan değerleri aynıdır.
- ✓ E) Ranj ve mod değerleri aynıdır.



Örnek:

Puan (x)	Frekans (f)	Yığılmalı Frekans
25-27	1	30
22-24	1	29
19-21	3	28
16-18	7	25
13-15	4	18
10-12	3	14
7-9	5	11
4-6	4	6
1-3	2	2

Yandaki tabloda verilen verilerin modu kaçtır?

Yandaki tabloda verilen verilerin medyanı kaçtır?

Yandaki tabloda verilen verilerin ranjı kaçtır?

MADDE İSTATİSTİKLERİ

MADDE GÜÇLÜK İNDEKSİ

- ➔ Madde güçlük indeksi, bir maddenin (sorunun) uygulandığı grupta doğru cevaplanma oranıdır.
- ➔ Eğer bir madde grubun çoğunluğu tarafından doğru yanıtlanıyorsa o madde kolay bir maddedir.



$$P = \frac{n(d)}{N}$$

P : Madde güçlük indeksi

n (d) : Maddeyi doğru cevaplayanların sayısı

N : Toplam öğrenci sayısı



40 kişilik bir sınıfta eğer herkes doğru cevaplarsa;

$$P = \frac{n(d)}{N} = \frac{40}{40} = 1$$



40 kişilik bir sınıfta eğer herkes yanlış cevaplarsa;

$$P = \frac{n(d)}{N} = \frac{0}{40} = 0$$



Madde Güçlük İndeksi;

- ➔ 0 - 1 arasında değer alır.
- ➔ 0.00 ile 0.40 arasında ise madde zordur
- ➔ 0.40 ile 0.60 arasında ise madde orta güçlüktedir.
- ➔ 0.60 ile 1.00 arasında ise madde kolaydır.



Aşağıdaki seçeneklerden hangisi bir sorunun madde gücü (P_j =olasılığı) olamaz?

A) $8/9$

B) $2/4$

C) $1/4$

D) $1/3$

☒ E) $4/3$





Örnek:

Öğrenciler	6. madde	X Ham puanı
1. Öğrenci	1	4
2. Öğrenci	0	8
3. Öğrenci	0	6
4. Öğrenci	0	3
5. Öğrenci	1	8
6. Öğrenci	1	4
7. Öğrenci	0	8
8. Öğrenci	0	5
9. Öğrenci	0	4
10. Öğrenci	1	10

10 sorudan oluşan coğrafya sınavının 6. maddesine doğru cevabı veren öğrenciler “1” yanlış veren öğrenciler “0” ile puanlandırılmıştır

Bu sınavdaki 6. maddenin güçlük indeksi kaçtır?



Çözüm:

Doğru cevabı veren 4 öğrenci vardır.
Toplam öğrenci sayısı da 10' dur.

$$P = \frac{n(d)}{N} = \frac{4}{10} = 0.4$$

Yorum : Bu madde zor bir maddedir.

%27'lik Alt ve Üst Gruplar Yöntemi ile Madde Güçlük İndeksi Hesaplama:

$$p_j = \frac{n(d.\ddot{u}.)+n(d.a)}{N}$$

p_j : Madde güçlük indeksi
 $n_{(d,\ddot{u})}$: Maddeyi üst grupta doğru
yanıtlayanların sayısı
 $n_{(d,a)}$: Maddeyi alt grupta doğru
yanıtlayanların sayısı
 N : Üst ve alt gruptaki toplam
öğrenci sayısı



Örnek:

	A	B*	C	D	E	Toplam
Üst Grup	12	45	10	15	18	100
Alt Grup	18	26	20	20	16	100
Toplam	30	71	30	35	34	200

Doğru yanıt “B” seçeneği ise madde güçlük indeksi kaçtır?





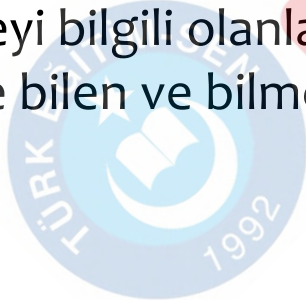
Çözüm:

$$p_j = \frac{n(\text{d.ü.}) + n(\text{d.a})}{N} = \frac{45 + 26}{200} = 0.36$$

- ➔ Bu zor bir maddedir.
- ➔ Grubun %36'sı doğru %64'ü yanlış cevaplamıştır.
- ➔ Bu madde ile ilgili olan hedef davranışa sınıfın %34'ü ulaşmıştır.

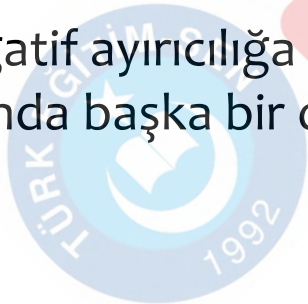
MADDE AYIRT EDİCİLİK GÜCÜ İNDEKSİ

- ➔ Madde ayırt ediciliği, bir maddenin bilen öğrenci ile bilmeyen öğrenciyi birbirinden ayırt etmesidir.
- ➔ Yani bir maddenin görevi bilen (başarılı) öğrenci ile bilmeyen (başarısız) öğrenciyi birbirinden ayırmaktır.
- ➔ Maddeyi bilgili olanlar doğru yapıp bilgisiz olanlar yanlış yapıyorsa o madde bilen ve bilmeyeni birbirinden ayırt ediyor demektir.



Madde Ayırt Edicilik Gücü İndeksi;

- ➔ Maddenin kalitesi hakkında bilgi verir.
- ➔ Maddenin bilen ve bilmeyen öğrencileri birbirinden ayırıp ayırmadığı hakkında bilgi verir.
- ➔ $-1,00$ ile $+1,00$ arasında değer alır.
- ➔ Negatif ayıricılığa sahip maddelerin testte ölçülmek istenen amaç dışında başka bir değişkeni ölçtüğü söylenebilir.



$$r_{jx} = \frac{\bar{X}_{jd} - \bar{X}}{S_x} \sqrt{\frac{P_j}{q_j}}$$

r_{jx} : Madde ayırt edicilik gücü indeksi

\bar{X}_{jd} : Maddeyi doğru cevaplayan öğrencilerin ortalaması

\bar{X} : Testin aritmetik ortalaması

P_j : Madde güçlük indeksi

q_j : Maddeye doğru cevap veremeyenlerin yüzdesi

S_x : Standart sapma



Örnek:

	A	B*	C	D	E	Toplam
Üst Grup	10	60	5	10	15	100
Alt Grup	20	10	30	15	25	100
Toplam	30	70	35	25	40	200

Doğru yanıt “B” seçeneği ise madde ayırt edicilik gücü indeksi kaçtır?





Çözüm:

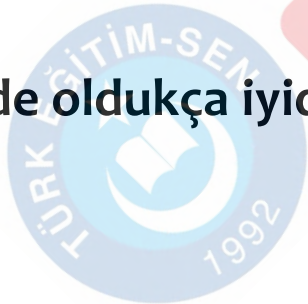
$$r_{jx} = \frac{n(d.ü.) - n(d.a)}{n} = \frac{60 - 10}{100} = 0.50$$



Madde bilen bilmeyenli ayırmıştır.



Madde oldukça iyidir.



Örnek:

Soru No	P_j	R_{jx}
1	0.10	0.60
2	0.52	0.75
3	0.60	-0.19
4	0.38	0.15
5	0.20	0.48

Tabloya göre zor olmasına rağmen ayırt ediciliği yüksek olan soru hangisidir?

Hem ayırt ediciliği hem de güçlüğü bakımından hangisi ideal bir test sorusudur?

	A	B	C	D	E	Boş	N
Üst Grup	18	30	10	10	10	22	100
Alt Grup	2	25	15	14	10	33	100

Yukarıda analizi verilen maddede doğru cevap hangi seçenekte olursa ayırt edicilik en yüksek olur?

A) A

B) B

C) C

D) D

E) E



Çözüm:

	A	B	C	D	E	Boş	N
Üst Grup	18	30	10	10	10	22	100
Alt Grup	2	25	15	14	10	33	100

$$r_{jx} = \frac{n(d.ü.) - n(d.a)}{n}$$

$$\text{A Seçeneği için : } \frac{18 - 2}{100} = 0.16$$

$$\text{D Seçeneği için : } \frac{10 - 14}{100} = -0.06$$

$$\text{B Seçeneği için : } \frac{30 - 25}{100} = 0.05$$

$$\text{E Seçeneği için : } \frac{10 - 10}{100} = 0.00$$

$$\text{C Seçeneği için : } \frac{10 - 15}{100} = -0.05$$



Örnek:

Testler	Ayırt Edicilik	Güçlük İndeksi	Kapsam Geçerliği	Yordama Geçerliği
1	0.80	0.14	0.70	0.60
2	0.90	0.55	0.90	0.85
3	0.70	0.70	0.68	0.74
4	0.28	0.28	0.82	0.75
5	0.30	0.80	0.55	0.40

Hangi test bir okula öğrenci seçiminde verilen kararın başarısını yansıtacak niteliktedir?

Hangi test ayırt edici olmasına rağmen zordur?



Örnek:

Puan (x)	Frekans (f)	Yığılmalı Frekans
25-27	1	30
22-24	1	29
19-21	3	28
16-18	7	25
13-15	4	18
10-12	3	14
7-9	5	11
4-6	4	6
1-3	2	2

Yandaki tabloda verilen verilerin modu kaçtır?

Yandaki tabloda verilen verilerin medyanı kaçtır?

Yandaki tabloda verilen verilerin ranjı kaçtır?

	A*	B	C	D	E	Boş	N
Üst Grup	55	5	10	30	0	0	100
Alt Grup	23	2	30	15	18	12	100

Bu madde için düzeltmeye gerek var mı varsa en önemli değişiklik ne olmalıdır?

- A) Soruda düzeltme yapmaya gerek yoktur
- B) E çeldiricisi alt gruptan fazla öğrenciyi çeldirdiği için değiştirilmesi gerekir.
- C) B çeldiricisi grupta az öğrenciyi çektiğinden güçlendirilmelidir.
- ☒ D) D çeldiricisi üst gruptan fazla öğrenciyi çektiğinden dolayı geliştirilmelidir.
- E) Soru tümüyle testten çıkarılmalıdır.

MADDE İSTATİSTİKLERİNDEN HESAPLANAN TEST İSTATİSTİKLERİ

ARİTMETİK ORTALAMA:



Bir testteki maddelerin güçlük indekslerinin toplamı, testin aritmetik ortalamasına eşittir.

$$\bar{X} = \sum_{j=1}^k p_j$$



TESTİN ORTALAMA GÜÇLÜĞÜ



Testin ortalama güçlüğü, genel olarak testin bütününün ne kadar zor ya da kolay olduğunu gösterir.

$$\bar{P} = \frac{\bar{X}}{K}$$

\bar{P} : Testin ortalama güçlüğü

\bar{X} : Testin aritmetik ortalaması

K: Testte bulunan madde sayısı



STANDART PUANLAR

Standart puanlar aritmetik ortalaması ve standart sapması farklı dağılımların, aynı aritmetik ortalama ve standart sapmaya sahip dağılım haline dönüştürülmesini sağlar.

Z PUANI

$$Z_i = \frac{X_i - \bar{X}}{S_x}$$

Z_i : Herhangi bir i bireyinin standart Z puanı

X_i : i bireyinin ölçme sonucu (puanı)

\bar{X} : Puan dağılımının aritmetik ortalaması

S_x : Puan dağılımının standart sapması

Örnek:

	Aritmetik ortalama	Medyan	Mod	Sx
Matematik	60	60	70	8
Fizik	68	75	80	12
Kimya	70	50	40	5
Biyoloji	65	65	65	10
Edebiyat	90	70	60	5

Bütün derslerden 80 alan bir öğrencinin öğrenme düzeyinin en yüksek ve en düşük olduğu dersler hangileridir?



Çözüm:

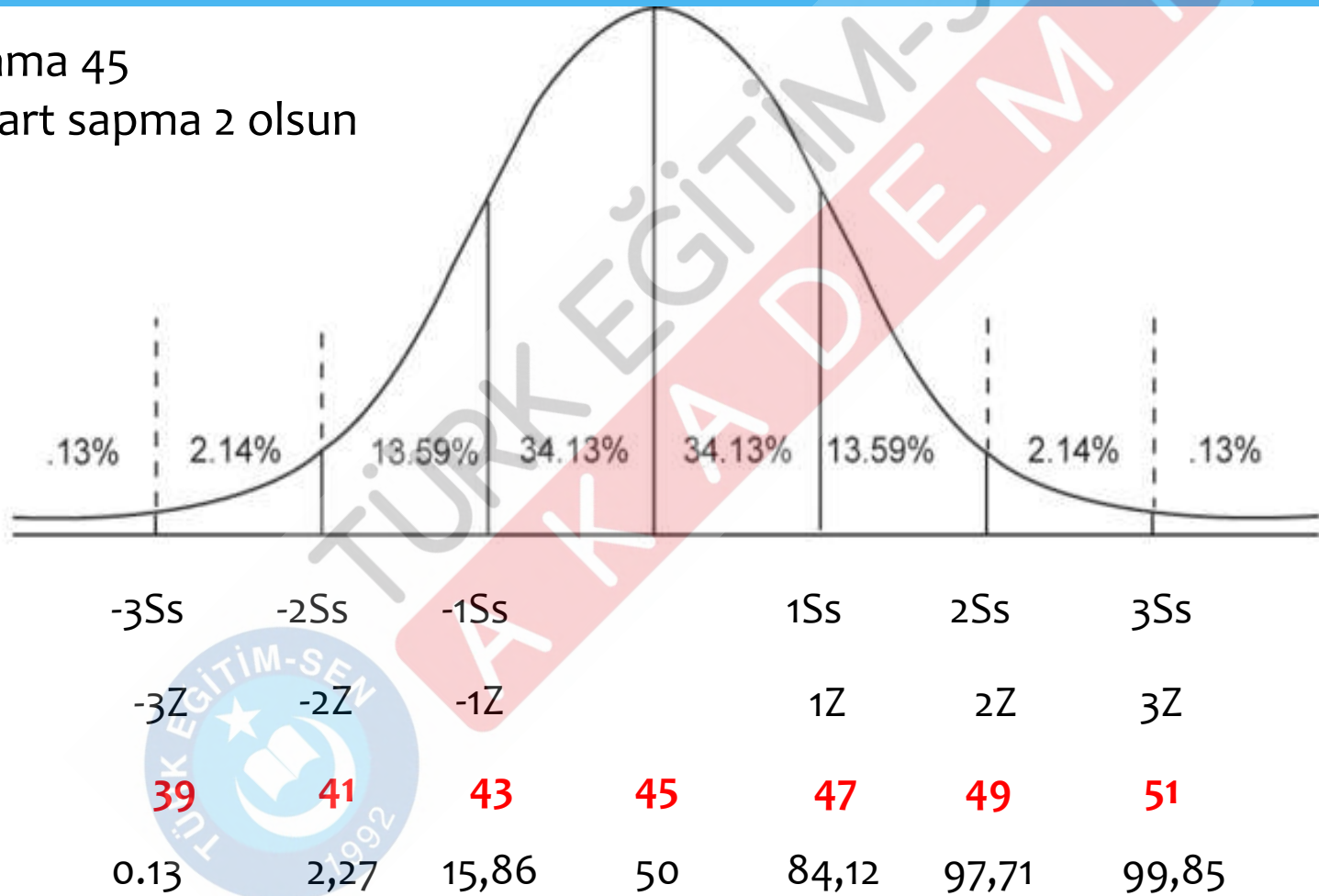
$$Z_{\text{mat.}} = 2.5 \quad Z_{\text{fiz.}} = 1 \quad Z_{\text{kim.}} = 2 \quad Z_{\text{biy.}} = 1.5 \quad Z_{\text{edeb.}} = -2$$

$Z_{\text{matematik}} > Z_{\text{kimya}} > Z_{\text{biyoloji}} > Z_{\text{fizik}} > Z_{\text{edebiyat}}$

Öğrencinin öğrenme düzeyinin en yüksek olduğu ders (en başarılı olduğu ders), Z puanının en yüksek olduğu matematik dersidir.

Öğrencinin öğrenme düzeyinin en düşük olduğu ders (en başarısız olduğu ders) ise, Z puanının en düşük olduğu edebiyat dersidir.

Ortalama 45
Standart sapma 2 olsun





Örnek:

Bir sınavda alınan puanların aritmetik ortalaması $X = 70$ ve standart sapması da $S_s = 3$ olarak belirlenmiştir.

- ➔ Öğrencilerin %68'inin puanı aritmetik ortalamanın 1 standart sapma yukarısında ($70+3$) ve aşağısında ($70-3$) yer alır. Öğrencilerin %68'inin puanı 73-67 arasındadır.
- ➔ Öğrencilerin %95'inin puanı aritmetik ortalamanın 2 standart sapma yukarısında ($70+2.3=76$) ve aşağısında ($70-2.3=64$) yer alır.
- ➔ Öğrencilerin %99'unun puanı aritmetik ortalamanın 3 standart sapma yukarısında ($70+3.3 = 79$) ve aşağısında ($70-3.3 = 61$) yer alır.

T PUANI

$$T_i = 10 \left(\frac{X_i - \bar{X}}{S_x} \right) + 50$$

$$T_i = 10Z_i + 50$$

T_i : Herhangi bir i bireyinin standart T puanı

X_i : i bireyin ölçme sonucu (puanı)

\bar{X} : Puan dağılımının aritmetik ortalaması



	A*	B	C	D	E	Boş	N
Üst Grup	55	5	10	30	0	0	100
Alt Grup	23	2	30	15	18	12	100

Bu madde için düzeltmeye gerek var mı varsa en önemli değişiklik ne olmalıdır?

- A) Soruda düzeltme yapmaya gerek yoktur
- B) E çeldiricisi alt gruptan fazla öğrenciyi çeldirdiği için değiştirilmesi gerekir.
- C) B çeldiricisi grupta az öğrenciyi çektiğinden güçlendirilmelidir.
- ☒ D) D çeldiricisi üst gruptan fazla öğrenciyi çektiğinden dolayı geliştirilmelidir.
- E) Soru tümüyle testten çıkarılmalıdır.

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
Felsefe	40	14	10	7	9
Fizik	50	30	30	30	11

Öğrencilerin en başarılı olduğu ders hangisidir?

- A) Kimya
D) Fizik

- ☒ B) Tarih
E) Felsefe

- C) Türkçe



TÜRK EĞİTİM-SEN
AKADEMİ