

# ÖLÇME VE DEĞERLENDİRME

## I- TEMEL KAVRAMLAR

### Ölçme:

Ölçme bir betimleme işlemidir. Belli bir nesnenin ya da nesnelerin belirli bir özelliğe sahip olup olmadığını, sahipse oluş miktarını sayı ve sembollerle ifade etme işlemidir.

Geniş anlamıyla, herhangi bir niteliği gözlemek ve gözlem sonucunu sayılarla ya da sembollerle ifade etmektir.

Ölçme başka bir ifadeyle nitelikleri nicelendirme işlemidir.

Ölçme bireyin ya da bireylerin belli bir özelliğe sahip olma derecesinin sayısal olarak betimlenmesidir.

Benim boyum 1.85’dir.

Bugün havanın sıcaklığı 25 derecedir.

Ertuğrul Tarih dersi sınavından 8 puan aldı.

Standart ölçme araçları: metre, kg, standart testler, litre.

Bir öğretmenin sınıfta sorduğu sorulara doğru cevap veren her öğrenciye bir artı vermesi.

Bisikletin değeri 300 YTL’dir.

Yasemin çok güzel bir kızdır.

İbrahim iyi bir insandır.

2008- 2009 Bank Asya 1.lig sezonunu Samsun spor birincilikle tamamladı.

Örnekleri incelediğimizde ölçme işleminin yapılabilmesi için her şeyden önce ölçme işlemine ilişkin özelliklerin belirlenmesi ,bunlara ait gözlem yapılması ve gözlem sonuçlarının sayı ve sembollerle ifade edilmesi gerekir.

Ölçme işleminin temeli anahtar kelime **FARK** tır.

### **Ölçme işleminin aşamaları:**

- Ölçülecek bir değişkeni ya da özelliği belirleme.
- Özelliği ya da değişkeni gözlemlenme.
- Gözlemi uygun zamanda gerçekleştirme.

Sonucu sayı ya da sembolle ifade etme

### **Ölçme ve değerlendirme ;**

\*Eğitim ve öğretim hizmetlerinin daha nitelikli yapılmasını sağlar

\*Öğretmenin öğrenciyi tanımasını sağlar

\*Öğrenciye güçlü ve zayıf olduğu alanlar konusunda geribildirim sağlar.

\*Öğrenciye,davranışını nasıl değiştireceği veya geliştireceği konusunda geri bildirim sağlar.

\*Öğretmene kendini tanıması ve öğretim yöntemlerinin ne derece yeterli olduğu konusunda geri bildirim sağlar.

\*Öğrencinin durumu ve gelişimi hakkında velilerin bilgilendirilmesine olanak sağlar.Böylece velilerin öğretim sürecine katılmalarına yardımcı olur.

## **Ölçme Türleri**

1. **1)Doğrudan (Temel/Dolaysız) Ölçme:** Herhangi bir niteliği başka nitelikle ilişkilendirmeden gözlemek doğrudan ölçme olarak adlandırılır. Ölçülen özellik kullanılan ölçme aracının niteliği benzerdir (aynıdır). Ölçülecek değişken dolaysız (doğrudan) ölçülür. Özellik kendisine ilişkin ölçme aracı ile ölçülür. Geçerliği ve güvenilirliği yüksek bir ölçme elde edilir. Ağırlığı başka bir ağırlık ölçüsü ile ölçmek, uzunluğu bir başka uzunluk ölçüsü ile ölçmek.
  - Bir masanın boyunun cetvel ile ölçülmesi.
  - Bir paketin kefeli terazi ile ölçülmesi.
  - Bir sınıftaki öğrenci sayısının belirlenmesi.
  - Öğrencilerin boy uzunluklarının belirlenmesi.
  - Bir öğrencinin sahip olduğu kitap sayısının belirlenmesi.
  - Sezgin'in kaç kardeşi olduğunun belirlenmesi.
2. **2)Dolaylı (Göstergeyle) Ölçme:** Herhangi bir değişkene ait niteliği başka değişkenler veya araçların niteliğinden faydalanarak ölçmektir.(termometre ile ısı ölçme veya yaylı kantarla ağırlık ölçme: birinci durumda cıvanın genişmesi yardımıyla ısı hakkında fikir sahibi olurken, ikinci durumda yayın uzamasına bağlı olarak ağırlık hakkında bir şeyler söyleyebiliriz.) Ölçülen özellik ile kullanılan ölçme aracının niteliği benzer değildir (farklıdır), ölçülecek değişkenin dolaysız (doğrudan) bir şekilde gözlenmediği durumlarda başvurulur. Özellik kendisine ait olmayan başka bir araçla ölçülür. Geçerlik ve güvenilirlik düşüktür. Eğitimdeki ölçmeler tamamıyla dolaylı ölçmelerdir. Her türlü sınav dolaylı ölçmedir. Bir değişkenin (o değişkenin bir göstergesi olduğu kabul edilen) başka bir değişken ya da tanımlanan bir bağıntı yardımıyla ölçülmesidir.
  - Bir öğrencinin bilgisinin sınav ile ölçülmesi.
  - Resim öğretmeninin öğrencinin yaptığı resme bakarak not vermesi.
  - Öğrencinin zeka düzeyinin ölçülmesi.
  - Öğrencilerin yeteneklerinin ve kişilik özelliklerinin ölçülmesi.
  - Bir dersten geçme notunun belirlenmesi.
  - Bir paketin yaylı terazi ile ölçülmesi.

**3)Türetilmiş Ölçme:** İki ayrı ölçülen nitelik üzerinde yapılan işlemler yardımıyla yeni bir ölçme sonucuna ulaşmadır.

Örneğin Hız=km/saat (yol / zaman)

Yoğunluk = Kütle/hacim gibi)

Karadeniz bölgesinin aritmetik nüfus yoğunluğu 175'tir.(Aritmetik nüfus yoğunluğu=Bölgede yaşayan kişi sayısı/Bölgenin yüzölçümü)

Anahtar kelime=Formül

## **Ölçme Kuralı:**

Ölçme işlemini yaparken,ölçmeye konu olan özelliğin hangi miktarına ne kadar verileceğinin belirlenmesidir.

\*Doğru olarak cevaplandırılacak her soruya 10 puan verilecektir.

\*7 nolu soruyu doğru cevap verenler 25 tam puan alacaktır.

Kısacası ölçülecek özellik ile sayı ya da sembollerin eşleşme yollarının belirtilmesidir.

## **Ölçüm:**

Ölçme işlemi yapıldıktan sonra elde edilen sayı veya semboldür.

Kürşat Ölçme dersinden 82 aldı ifadesindeki 82 puan bir ölçümdür.

### **Ölçmede birimler**

Ölçme işleminde doğal ve yapay (tanımlanmış) birimlerin olduğundan söz edilebilir. Örneğin, bir işyerinde çalışan işçilerin sayısını bulmak için birim olarak işyerinde çalışan tek bir işçiyi almak doğru olacaktır ve bu ölçmenin doğal birimi olduğu söylenebilir. Ancak, uzunluk, ağırlık, zaman gibi bazı değişkenlerle ilgili birimler tanımlanmış (yapay) birimlerdir. Örneğin, metre ve kilogram üzerinde anlaşılmış veya tanımlanmış birimlerdenidir. Birimlerin tanımlanmış olmasının bir sakıncası yoktur ve birçok doğal birimden daha güvenilir ölçme sonuçları elde edilebilir. Zaten birçok birimi tanımlama gereği de üzerinde anlaşmaya varılamayan veya tartışmalara neden olan doğal birimler nedeniyle ortaya çıkmıştır.

Birimlerde aranan üç özellik vardır. Bunlar;

**Eşitlik** :Bir birimin,her zaman ve her yerde aynı büyüklüğü temsil etmesidir.

Örneğin metre her zaman aynı büyüklüğü ifade eder.

**Genellik**: Birimlerin herkes tarafından aynı anlamda kullanılması veya birimlerin çoğunluk tarafından kabul edilmesidir.Ana hatlarıyla bir ülkede aynı özellik için aynı birimlerin kullanılmasıdır.

Ülkemizde sıcaklık ölçümlerinde santigrat,Amerika’da ise fahrenheit kullanılmaktadır.

**Kullanışlılık** :Bir ölçme aracının ve ya yönteminin,çoğaltılma,uygulanma ve puanlama yönünden kolaylık,nesnellik ve ekonomik özelliklere sahip olması anlamına gelir.

İki şehir arasındaki mesafe ölçülürken kullanılacak birim km iken, boy uzunluğu ölçülürken kullanılacak birim cm dir.

\*Altın bileziğin gram ile ifade edilmesi...

Anahtar kelime PRATİK

### **Ölçmede Araçların Rolü**

1-Ölçme sonuçlarının duyarlılığını ve güvenilirliğini dolayısıyla kesinliğini artırır. Ölçme sonuçlarına karışan hata miktarının azalmasını sağlar.,

2-Duyu organlarımızla ölçme yapamadığımız durumları gözlememizi kolaylaştırır. Böylece duyu organlarını kullanamadığımız durumlarda da ölçme yapma imkanı buluruz.

3-Herkesin kabul ettiği, üzerinde anlaştığı standart ölçme sonuçları elde etmemize olanak sağlar.

**Ölçek**: Nesnelere verilen sayıların anlamlarını ya da nesnelere sayı vermede ve de bunların kullanılmasında uyulması gereken kuralları ve kısıtlamaları belirtmek için kullanılan araca denir.

Diğer bir tanım olarak;Belli bir başlangıç noktasından itibaren eşit birimlerle bölünmüş ölçme aracına denir

**Ölçekleme**: Eğitim ve psikolojide ölçek veya ölçme aracı geliştirme olarak tanımlanabilir.

## Ölçmede Kullanılan Ölçek Türleri

### 1. 1)Adlandırma (Sınıflama,Nominal )Ölçeği:

Üst düzey bilgiler vermezler. Sadece **frekans ve mod** gibi basit işlemler yapar. Gözlemlerin en basit biçimi olup taşıdığı bilgiler en az olan ölçeklerdir. Bireylerin ya da objelerin belirli özellikleri doğrultusunda benzeyip benzemedikleri biçimde yapılandırmasıdır. Özelliklerine bakarsak:

- Nesnelerin belirli yönlerinden benzeyip, benzemediklerine göre sınıflamaktır.
- Nesnelerin farklılık ve benzerliklerini belirlenir.
- Nesnelerin sınıf, tip ve kategorilerine göre sınıflanır.
- İllere plaka numarası verme,35 İzmir 55 Samsun,69 Bayburt
- Kız – erkek öğrencileri ayırma,
- kısa – uzun boylu öğrencileri ayırma.
- Basit istatistik işlemler yapma.
- Değişkenler taşıdıkları özelliğe göre sınıflara tasnif edilirler.
- Sınıflama ölçeğinde sınıflar oluşturulduğundan dolayı, bu sınıflardan herhangi birine giren tüm elemanlar eşit, farklı sınıflardan alınan elemanlar ise farklı olurlar.
- Eğitimde nadiren kullanılır; bazı tanımlayıcı nitelikler (göz ve saç rengi gibi değişkenler), insanların yaşadığı coğrafi bölgeler gibi değişkenler bu ölçeğe örneklerdir.
- Bu ölçek değişkenleri sadece gruplara tasnif ettiğinden matematiksel olarak elimizde sadece frekanslar vardır. Dolayısıyla, frekanslara dayalı analizler dışında bir çok matematiksel işlemi yapmak anlamsızdır. Örneğin elmalarla armutların ortalamasını alamayız.

Matematiksel işlem olarak sadece sınıf içindeki elemanların sayısı yani frekanslar belirlenir.

- Bu ölçekte bazen isimler yerine sayılar kullanılabilir. Mesela bir ankette kız öğrencilere “1” ile, erkeklere ise “2” ile kodlanması. Futbolda kalecilere “1” numarasının verilmesi, okul kimlik numaralarında sayıların kullanılması durumlarında, kullanılan numaraların sayısal bir anlamı yoktur, bunlar sadece koddurlar.

Beşiktaş futbol takımının 10 numarası Delgado,20 numaralı oyuncusu Ricardinyo

20 numaralı Ricardinyo,10 numaralı Delgado’dan 2 kat daha iyi oynar denilemez.

- Sınıftaki öğrencilerin velilerinin mesleklerini belirlemek için bir anket hazırlanmış ve bu ankete memur, esnaf, işçi gibi ifadeler yazılmıştır. Bu ifadeler sınıflama ölçeğinde yer alır.

Farklılıkları birim cinsinden belirleyemeyiz. Ölçmenin basit, ilkel bir türüdür. Ölçme sonuçları sembol, sayı veya sıfatlara dönüştürülür.

İstatistiksel olarak frekans tablosu yapılabilir ve **mod** (tepe değer =en çok tekrar eden ölçüm) bulunabilir.

### **Sınıflama Ölçeğinin Özellikleri:**

**a)-Simetriklik:**Melikşah,Ayşe’nin kardeşi;Ayşe’de Melikşah’ ın kardeşidir.Bu ilişkiye simetriklik denilir ve sembolle  $a=b$  ise  $b=a$  şeklinde gösterilir.

**b)-Geçişlilik:**Yusuf,Oktay’ın kardeşi;Oktay da Muharrem’in kardeşi ise Yusuf ,Muharrem’ in kardeşidir.Bu ilişkiye geçişlilik denir ve sembolle  $a=b$  ,  $b=c$  ise  $a=c$  şeklinde gösterilir

- ### 2. 2)Sıralama Ölçeği:
- Sınıflama ölçeğine göre bir üst düzeyde ölçme sonuçları veren ölçek türüdür. Belirli bir nitelik bakımından büyükten küçüğe ya da küçükten büyüğe sıraya koyma işidir. Bu sıraya uygun sayı veya sembol kullanılırsa sıralama ölçümleri yapılabilir. Kullanılan sayılar sıra sayılarıdır. Bu ölçekler büyüklük küçüklük bağıntısı ve bu bağıntının özelliklerini sağlar. Sıralama ölçeklerinin de birim ve başlangıç noktası olmadığı söylenebilir. Yani büyüklükler hakkında farklılığın ne kadar olduğu yönünde bilgi sahibi olamayız. İstatistiksel olarak frekans tablosu, mod, medyan (ortanca) ve yüzdellikler hesaplanabilir. İlişki ölçülerinden sıra farkları korelasyonu bulunabilir.



Bireylerin ya da objelerin ölçülen özellikleri doğrultusunda sıraya konması ve sıradaki yerine göre bir sayı verilmesi durumu söz konusudur. Özellikler;

- Nesneleri belli özelliğine sahip oluş miktarı bakımında sıralamaktır.
- Öğrencileri boy sırasına, sınav notlarına göre sıralama.
- **Ortanca ve yüzdellikler** belirlenir.
- Niteliklerin belli bir boyutta, büyükten küçüğe, küçükten büyüğe, iyiden kötüye veya kötüden iyiye doğru sıralanmasıyla elde edilir. Bu sıralar sayı ile ifade edileceği gibi bazı sıfatlarla da ifade edilebilirler; pekiyi-iyi-orta-geçer-kalır, en uzun-uzun-kısa gibi. Bu tür ölçeklerde öğrencilerin tüm grup içindeki yeri belirlenmektedir.
- Sıralama ölçeğinde, sınıflamadan fazla olarak sıralar arasındaki büyüklük küçüklük ilişkisi vardır.
- Bu ölçekte belli bir boyutta sıralar arasındaki büyüklük, küçüklük anlamlıdır. Buna rağmen, iki sıra arasındaki farkın miktarını söylemenin bir anlamı yoktur. Başarı sıralamasına göre veya boy uzunluğu sıralamasına göre, ikinci öğrencinin dördüncü öğrenciden iki birim daha üstün olduğu söylenemez. Her bir ardışık sıra numarası alan öğrenciler arasındaki fark bir derece de olsa da bu farkın aynı miktarı ifade ettiği söylenemez. Bu ölçekte toplama çıkarma yapılamaz.
- Sınıftaki en başarılı ilk beş öğrencinin belirlenmesi.
- Sıralama ölçeğindeki sıfır anlamsızdır,yani herhangi bir özelliği yoktur.
- İçimizde aylık geliri en yüksek olan Burhan Bey’dir.
- Bu kitapların en kalın olanı Ölçme Değerlendirmedir.

### **Sınıflama Ölçeğinin Özellikleri:**

**a)-Asimetriklik:**Hanife Elif’ten daha başarılı ise Elif Hanife’den daha başarılıdır denilemez.Sınıflama ölçeğinden farklıdır.Bu özelliğine asimetriklik denir ve sembolle  $a > b$  ise  $b > a$  **değildir** şeklinde gösterilir.

**b)-Geçişlilik:**Ziya Alim’den daha uzun,Tuğrul da Alim’den daha uzun ise Tuğrul,Ziya’dan daha uzundur.Bu özelliğe geçişlilik denir ve sembolle  $a < b$  ve  $b < c$  ise  $a < c$  şeklinde gösterilir.

**3)Eşit Aralık Ölçeği:** Birimi ve izafi (tanımlı, yapay, göreceli, görelî) bir sıfır noktası vardır. Gerçekte “0” noktası matematikte yokluk anlamındadır. Yani bir ölçüm sonucunda 0 değeri elde etmek ölçtüğümüz değişkenin hiç olmadığı noktadır. Uzunluk 0 dediğimizde uzunluk yoktur. Aralık ölçeğinde ise başlangıç noktası ölçme yapan kişiye, bir araçtan diğerine veya bir uygulamadan diğerine değişir. Eşit aralık ölçeklerinde “0” noktası görelidir. Örneğin sıcaklık ölçeğinden  $^{\circ}\text{C}$  ve Fahrenheit bu tür ölçeklerdendir. Bunların sıfır noktaları görece (bağıl) sıfırdır. Yani aslında bu ölçeklerdeki “0” noktaları ölçülen özellik olarak sıcaklığın olmadığı bir miktarı göstermez. Yapay bir başlangıç noktasına işaret ederler. Ancak birimler arasındaki miktar eşittir. Sıralama ölçeğinde birimler arasındaki miktarı belirleyemezken aralık ölçeğinde belirleyebiliriz. Bu nedenle eşit aralık ölçeği adı da verilmektedir. Öğrencilerin testlerden aldıkları standart test puanları aralık ölçeğindedir. Eğitim ve psikolojideki ölçme araçları ve genelde sosyal bilimlerdeki ölçme araçlarının büyük çoğunluğu aralık ölçeğinde ölçeklerdir. Eşit aralık ölçeklerinde doğrusal dönüştürme yapılabilir. Yani gözlem sonuçları matematiksel işlemlerle yine eşit aralıklı fakat başlangıç noktası farklı ölçme sonuçlarına dönüştürülebilir. Bu ölçeklerde ortalama, standart sapma, korelasyon vb istatistik işlemler yapılabilir. Farkların miktarı belirlenebilir. Sadece çarpma, bölme diğer bir deyişle oran işlemleri yapılamaz.

Bireyler ya da objelerin ölçme konusu olan özellikleri reel sayılar kümesinin elemanlarıyla bire bir olarak eşlenirler. İzafi 0 ise; öğrencinin sınavda 0 alması. Bir öğrenci sınavdan 0 alıyorsa bu onun hiçbir şey bilmediği anlamına gelmez. Saatimizin 00:00

göstermesi ya da termometrenin 0 dereceyi göstermesi ortamda hala sıcaklığın olduğuna işarettir. Özellikleri:

- Nesnelerin belirli bir başlangıç noktasına göre ve belli bir özelliğe sahip oluş derecesi bakımından eşit aralıklarla sıralanmasıdır.
- **Başlangıç noktası (0 noktası) tanımlanmıştır.** 0 noktası keyfidir ve eşit aralıklarla bölümlenmiştir.
- **Termometre, takvim, , sınavlar.**
- **Aritmetik ortalama ve standart sapma** bulunur.
- Türkçe dersinin sınavında 90 alan bir öğrencinin zekası 45 alan bir öğrencinin zekasından tam iki kat fazladır anlamı çıkmaz.
- Hava sıcaklığının 18 derece olarak belirlenmesi.
- Bir öğrenciye uygulanan zeka testi sonucu.
- Bir öğrencinin Matematik dersine olan tutumu.
- Nurşat'ın karnedeki notlarının hepsi 5
- Matematik dersi 90 olan bir bireyin notu matematik dersi 45 olan bireyin notunun tam iki katıdır anlamı çıkmaz. Çünkü 0 noktası görecelidir yani izafidir.
- Buraya kadar eğitimde uygulanabilir.
- T ve Z puanları
- Anahtar kelimeler...**Sıfır=Biz,Eğitim,Her iki yön**

**4)Oran Ölçeği:** Başlangıç noktası gerçek (mutlak) sıfırdır. Eşit bir birimi vardır. Bu ölçeklerle elde edilen verilerde her türlü matematiksel ve istatistiksel işlemler yapılabilir.

En üst düzeydeki ölçektir. Aralık ölçeğinden ayrıldığı nokta “0” noktasıdır zaten eğitimde de kullanılmamasının nedeni de budur. Oran ölçeğinde “0” gerçek (mutlak) “0”dır. Oran ölçeğinde elde edilen ölçme sonuçları çarpım ve bölme yapılabildiğinden matematiksel 4 işlem ve bu işlemlere dayanan tüm istatistiksel işlemler yapılabilir ve bu da anlamlı olur.

Özelliklerine bakarsak:

- Başlangıç noktası olan 0 noktası gerçek bir sıfır noktasıdır.
- 0 noktası ölçülen özelliğin hiçliğini, yokluğunu gösterir.
- **Uzunluğun ya da ağırlığın ölçülmesidir.**
- **Her türlü istatistik işlemi yapılır.**
- Sınıftaki öğrenci sayısının belirlenmesi.
- Rabia'nın ağırlığı 49 kg dır.
- 304 no lu sınıfın öğrenci sayısı 20 dir.
- Benim hiç kardeşim yok.
- Basketbol takımına girecek öğrenciler için boy uzunluğu sınırı belirleme.

Boy 90 cm olan bireyin boyu 45 cm olan başka bir bireyin boyunun tam 2 katıdır anlamı çıkar.

**Anahtar Kelimeler:**Sıfır Yokluk,Tek yön,Çarpma-Bölme

### Ölçeklerle elde edilen bilgilerin anlamlılıkları

	Ölçeğin yapısal özellikleri		Elde edilecek Olan ölçülerin anlamlılık dereceleri			
Ölçek tipi	Başlangıç noktası	Birimi (Aralığı)	Nitelik gösterme gücü	Nicelik gösterme gücü		
				Sıra	Fark	Oran
Sınıflama	Yok	Yok	Var	-	-	-
Sıralama	Değişken	Değişken	Var	Var	-	-
E. Aralıklı	Keyfi	Keyfi	Var	Var	Var	-
Oranlı	Mutlak	Mutlak	Var	Var	Var	Var

1. Ölçekler yukarıdan aşağıya en basitinden en karmaşığına bir sıra göstermektedir (Basitlik ve karmaşıklık, elde edilen ölçümlerin bize verdiği bilgi, bu bilgilerin güvenilirliği ve ölçümler üzerinde yapılabilecek işlemlerin çokluğuna dayalı olarak belirlenmektedir).

2. Karmaşık ölçekler basit ölçeklere dönüştürülebilirler. ( Aksı çoğu zaman -test standart puanlarını elde etme dışında- mümkün değildir.)

Davranışların ölçülmesi, ya o davranışların ya da o davranışlar sonunda ortaya çıkan ürünün veya ikisinin birden gözlenmesi şeklinde gerçekleşir.

**Tablo 1. Ölçek türleri özet tablosu**

ÖLÇEKLER	TEMEL İŞLEMLER	ÖRNEKLER	İST. TEKNİKLER
SINIFLAMA (ADLANDIRMA)	Farklılık ya da Benzerliği belirlemek	-kişileri saç ve ten rengine göre sarışın esmer diye gruplama -evlere kapı numarası, sokaklara sokak adı verme	-frekans -tepe değer (mod) -yüzde
SIRALAMA	Azı – çoğu veya küçüğü-büyüğü belirlemek	-kitapları kalınlık sırasına göre dizme -öğrencileri boy veya başarı sırasına koyma	-ortanca (medyan) -yüzdelik -sıra farkları korelasyon katsayısı
(EŞİT) ARALIK	Farkları belirlemek	-termometreyle sıcaklık ölçme -kullanılan takvimler -standartlaştırılmış test puanları	-ortalama -standart sapma -Pearson Moment Çarpımları korelasyon katsayısı
ORAN ÖLÇEĞİ	Oranları belirlemek	-hacim, uzunluk, ağırlık vb. ölçüleri	-matematiksel ve istatistiksel her türlü işlem.

### DEĞERLENİRME

Değerlendirme, ölçme sonuçlarının aynı alana ait bir kriter ile kıyaslanarak bir değer yargısına ve oradan da bir karara ulaşma sürecidir. Öğretmen öğrencisine: “Bu dersten tekrara kaldın” demiştir. Öğretmenin bu ifadesi **Karar**’dır.. Ölçme ve değerlendirme etkinliklerinin eğitimin sürecindeki **temel işlevi**:

- Davranışların kazanılma düzeyini belirleme.
- Öğrencilerin hazır bulunuşluk seviyelerini belirleme ve bilgi edinme.
- Öğrenciyi motive etme.
- Öğrencilerin geleceğine yönelik karar alma.

Ölçme Sonucu (ölçüm) → Kriter (ölçüt) → Değer Yargısı → KARAR  
Mine 163 cm boyundadır.)(Polis okuluna ön kayıt(Mine'nin boyu kısadır.(Mine polis okuluna yapmak için 165 cm boya sahip olmak gerekmektedir.) giremez.)

Eğitimde değerlendirme, sürecin tamamında yapılmaktadır. En belirgin ve bilinenleri;

- 1-Hazır bulunuşluluk yada düzey belirleme amacıyla,
- 2-Öğrenme eksiklerini saptamak amacıyla=biçimlendirmeye yönelik,
- 3-Öğrenci başarısını saptamak amacıyla=değer biçmeye yönelik; not verme,
- 4-Öğretimin etkililiği hakkında bir karar vermek amacıyla,
- 5-Öğretim programının sağlamlığı hakkında karar vermek amacıyla,
- 6-Öğretmen ve diğer eğitim görevlileri hakkında karar vermek amacıyla,
- 7-Değerlendirme yöntem ve ölçütleri hakkında karar vermek amacıyla,
- 8-Sistem hakkında karar vermek vb. amacıyla yapılan değerlendirmeler ve bunlara eklenebilecek olanlardan söz edilebilir. Gözden kaçmaması gereken değerlendirme hangi amaç için yapılırsa yapılsın sistemin onarılması anlamını kapsamaktadır.

### **ÖLÇÜT:**

Ölçme sonuçlarını karşılaştırdığımız ya da kıyasladığımız sabit değere kriter (ölçüt) denir. Ölçüt olarak; programın hedefleri, öğrencinin yeteneği, grup başarısı, öğrenme eksiklikleri kullanılabilir. Öğretmenin öğrenciye ilişkim kanaati kullanılmaz. Bir öğretmen, dersinde 3 kez artı alana, sınavda 5 puan fazla vermektedir. 3 kez artı alma ölçüttür.

Örneğin, 'İki sınavın ortalaması 70 ve üzerinde olanlar geçer.' dediğimizde, öğrencinin 'Geçti' ya da 'Kaldı' şeklinde değerlendirilebilmesi, bu iki sınavının ortalamasının 70 ya da daha fazla olmasına bağlıdır.

- Yasemin birinci sınavdan 80 puan aldı.(ÖLÇME)
- Yasemin ikinci sınavdan 60 puan aldı.(ÖLÇME)
- Yasemin'in iki sınavının ortalaması  $(x)=(80+60)/2=70$  puan (Aritmetik Ortalama)
- İki sınavın ortalaması 75 puan ve üzerinde olanlar geçer. (ÖLÇÜT)
- Yasemin kaldı.(DEĞERLENDİRME)
- İki sınavın ortalaması 65 puan ve üzerinde olanlar geçer. (ÖLÇÜT)
- Yasemin geçti.(DEĞERLENDİRME)

Ölçüt :Mutlak Ölçüt ve Bağıl Ölçüt diye 2 ye ayrılır.

### **Kullanılan Ölçüte Göre**

#### **1.Kriter Dayanlı (Mutlak):**

Bireyin başarısını daha önceden konulmuş bir kuralı yerine getirip getirmediğine bakılır. Bireyin başarısı tamamıyla kendi çabasıyla alakalı. Diğer bireylerden bağımsızdır.Yani ölçüt tektir ve önceden belirlenmiştir.

Matematik dersinde öğrencilerin öğrenme eksikliklerini belirleme.

Beden dersinde 100 metreyi 12 sn. de koşan öğrencileri başarılı kabul etme.

Matematik dersinde başarılı olmak için en az 60 puan almak gerekmektedir

Bu ifadede yer alan "60" puan mutlak ölçüttür.

İçerikte yer alan konuların en az %80'nini öğrencilere kazandırma.

1.sınıftan dersi kalanlar 3.sınıfa geçemez.

Hedef dayanaklı değerlendirme de denir. Sistem ve yetişiğin etkinliğini değerlendirmede işe koşulmalıdır. Ayrıca ünite ve hedef analizlerinde, sıkı aşamalılık ilişkisi bulunduğu derslerde, yarım, eksik öğrenmelerin hayati tehlike doğurduğu kurslarda

ve bir alana belli bir beceri, bilgi üzerindeki kişilerin alınacağı durumlarda bu değerlendirme türüne başvurulabilir.

Öğrenciler mutlak ölçütlere göre değerlendirildiklerinde Mutlak Başarı Yüzdesi hesaplanır. Bu formül aşağıdaki gibidir.

$$MBY = \frac{\text{BireyinPuanı}}{\text{TesttenAlınabilecekEnYüksekPuan}} \times 100 = \frac{X_b}{X_{mak}} \times 100$$

MBY= Mutlak başarı yüzdesi

X<sub>b</sub> = Öğrencinin (Bireyin) puanı

X<sub>mak</sub> = Testten alınabilecek en yüksek puan

Anahtar kelime: **Kriter ÖNCE**

## 2. Norm Dayanlı (Bağıl):

Bireyin başarısının içinde bulunduğu gruba göre değerlendirmedir. **Çan eğrisi** buna örnektir. Grup olarak ele alınır. Başka bir deyişle, bu ölçüte göre yapılacak değerlendirme, diğer öğrencilerden bağımsız değildir. Tam tersine değerlendirmede temel etken, grubu oluşturan diğer öğrencilerdir.

Türkçe dersinde öğrencilerin dönem sonunda alacakları notu belirlenen sınıfın aritmetik ortalamasını dikkate alma.

İngilizce bilgisi en iyi olan ilk 5 öğrenciyi belirleme.

Bağıl değerlendirmede, sınıftaki diğer öğrencilerin başarısı, öğrencinin yeteneği, ülke çapındaki normlar, erişim düzeyi ölçüt olarak kullanabiliriz.

Matematik olimpiyatlarına katılacak öğrencilere uygulanan bir testin sonucuna bağlı olarak en iyi üç öğrenciyi seçmek.

Sınıfın öğrenme düzeyine göre geri kalan öğrencileri belirleme.

Aritmetik ortalamanın altında kalan öğrencileri başarısız sayma.

Oyunu kazanabilmeniz için, parkuru en kısa sürede tamamlamanız gerekir.

## Kullanılma Amaçlarına Göre

### 1. Tanıma – Yerleştirmeye göre Değerlendirme:

- Öğrenme sürecinin başında ya da bir dersin giriş bölümünde yapılan değerlendirmedir.
- Öğrencilerin belli bir kurs, ders ya da ünitenin ön koşulu niteliğindeki giriş davranışlarına sahip oluş derecelerini belirlemek.
- Öğrencileri düzeylerine göre gruplama.
- Öğrencilerin giriş davranışlarını belirleme.
- Öğrencileri bir programa yerleştirme.
- Öğrencilerin hazır bulunuşluk düzeylerini belirleme.
- Üniversiteye yeni başlayan öğrencilere İngilizce dersinde yapılan bir testin sonucuna dayalı olarak öğrencilerin %30'u dersten muaf sayılmışlardır. Bu değerlendirmenin amacı tanıma yerleştirmeye yönelik değerlendirmedir.

**Not:** Dershanelerin sene başında yaptığı sınavlar tanımlayıcı değerlendirme türleridir.

**2. Biçimlendirici/İzlemeye Değerlendirme (Formatif):** Eğitim – öğretim sürecinde devam etmekte yapılan sınavların hepsi bu değerlendirmeye girer. Her konu sonu testleri, deneme sınavları. Üniteler arasındaki aşamalılığı ve geçişi sağlar, öğrenme yetersizlikleri belirlenir ve önlem alınır. ünite içindeki öğrenme eksikleri belirlenir. Ünitenin hedef – davranışlarında gerçekleşme düzeyi belirlenir. Öğrencilere ara sınav yapılarak not verilir. Öğrencilerin öğrenme süreci takip edilir. (örneğin, konu testleri, quizler, dersane tarama testleri, yaprak testler gibi). Bu sınavlardaki değerlendirme mutlaklıdır.

**3.Düzyer Belirleyici (Summatif):** Eğitim –öğretim sonunda yapılır. Hedeflere ne derecede ulaşıldığını gösterir. Vizeler (biçimlendirici), final sınavları ise (düzyer) belirleyicidir. KPSS sınavı. Öğretim – öğrenme etkinliklerinin sonunda öğrencilerin son durumlarını belirlemek amacıyla yapılan düzyer belirleyici değerdendirme ölçme araçları olarak: Başarı sınavı, bitirme sınavı (final), yeterlilik sınavı, uzmanlık sınavı, erişi testi.

- Öğretim programının hedeflerine ulaşma düzyeri, öğrenme düzyeri ve öğrencileri başarı düzyerini belirlemek için yapılır.

## **SÜREÇ TEMELLİ DEĞERLENDİRMELER**

### **1-Öğrenci Ürün Dosyası (portfolyo):**

Öğrencilerin öğrenme sürecindeki performansının ve başarısının kaydedildiği, öğrencilerin yıl boyunca gerçekleştirdiği tüm çalışmaların örneklerinin bulunduğı, böylelikle çok yönlü (bir bütün olarak) gelişiminin izlendiği bireysel dosyalara, öğrenci gelişim dosyası (portfolyo) denir.

Öğrencilerin aktif olduğu öğrenme-öğretim sürecinde öğretmenler, öğrencileri değerdendirirken sadece sınav sonuçlarına göre karar vermezler.

Karar aşamasında , öğrenci gelişim dosyasında bulunan

- Öğrencilerin sınıf içi etkinliklere katılma düzeyleri.
- Sorumluluk alma düzeyleri
- Bilimsel tutum ve davranışları
- Gözlem, inceleme ve araştırma yapabilme becerileri
- Ekiple çalışabilmeleri

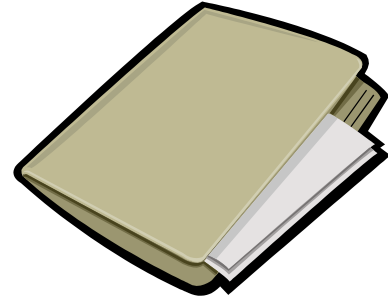
Özelliklerini de dikkate alarak başarıları hakkında karar vermektedir.

Temel amaç, öğrenciyi **tüm yönleriyle** tanımak ve öğrencinin **kendisini tanımasını** sağlamaktır.

Portfolyo öğrencilerin çalışmalarını, harcadıkları çabayı ve gelişmelerini gösterebilir. Öğrencinin çalışmalarını, çabalarını ve gelişimini, velisinin, öğretmenlerinin ve diğer ilgililerin izlemesini kolaylaştırır.

*Portfolyonun amaçları:*

- 1-Öğrencinin öz disiplin ve sorumluluk bilincini geliştirir
- 2-Öğrencilere kendini değerdendirme becerisi kazandırır.
- 3-Ürün değerdendirme araçlarının yanında süreç (performans) değerdendirme araçlarının da kullanılmasına olanak tanır.
- 4-Öğrencinin gelişimini izleyebilme olanağı sağlar.
- 5-Öğrencinin gelecekteki öğrenmelerine yol gösterir.
- 6-Öğrencilerin yeteneklerini sergilemelerine ve ilgi alanlarını geliştirmelerine fırsat verir.
- 7-Öğrencilere arkadaşlarının gelişimlerini izleme ve değerdendirme şansı verir.
- 8-Grup çalışmalarına katılımı ve öğrencilerin birbirlerine yardımcı olmalarını kolaylaştırır.
- 9-Öğrencilerin çalışmalarını kendilerinin ve arkadaşlarının değerdendirmelerine olanak vererek öğrencinin değerdendirme sürecine aktif katılımını sağlar.
- 10-Öğretime öğrenciyi tüm yönleriyle (bilgi, beceri, ilgi, kişilik vb.) tanıma olanağı verir.
- 11-Aile ve öğrenci arasında iletişim güçlendirir.
- 12-Öğrencini bir bütün olarak görülmesi ve değerdendirilmesi amacına hizmet eder.



### *Portfolyo Oluşturma Sürecinin Temel Aşamaları*

- 1-Öğretimin genel hedefleri ve kazanımları belirlenir.
- 2-Öğrencilerle birlikte, öğrencilerin kazanması istenen kazanımlara veya davranışlara karar verilir.
- 3-Portfolyonun ne olduğu, hangi amaçlarla oluşturulduğu ve nasıl oluşturulduğu öğrencilere açıklanır.
- 4-Öğrencilere ürün dosyalarını oluşturmada rehber olması açısından örnek bir portfolyo gösterilir.
- 5-Ürün dosyalarına hangi çalışmaların (ödevler, resimler, video ve bant ses kayıtları, çeşitli sınavların sonuçları, projeler, çeşitli ölçme araçlarından elde edilen sonuçlar, mektuplar, öğrencinin yazıları, günlükler vb.) ne amaçlarla alınacağı belirtilir.
- 6-Ürün dosyasındaki çalışmaların hangi ölçütlere göre nasıl değerlendirileceği öğrencilerle birlikte belirlenerek, bu ölçütler bir yönergeyle karara bağlanır ve bu yönerge öğrencilere dağıtılır.
- 7-Ürün dosyasına öğrencilerin süreç değerlendirme araçlarından ve ürün değerlendirme araçlarından elde ettikleri sonuçlar eklenir. Aynı zamanda öğrenciyi tüm yönleriyle (kişilik, yetenek, ilgi vb.) tanımamızı sağlayacak belge ve bilgiler de portfolyona alınır.
- 8-Ürün dosyası değerlendirilerek öğrencilere dönüt (geri bildirim) sağlanır. Öğrencilerin güçlü ve zayıf yönleri belirlenir. Gerekirse tümel bir değerlendirme yapılarak öğrencilere not verilebilir.

Tablo 2. Geleneksel –Ürün Dosyası değerlendirme karşılaştırması

Geleneksel Değerlendirme	Ürün Dosyası Değerlendirme
<ul style="list-style-type: none"><li>• Öğrencinin yeteneği, tutumu, bilgisi vb. tek bir zamanda ölçülür</li><li>• Öğrenci ölçütlerden habersizdir Değerlendirmeyi öğretmen yapar</li><li>• Öğrenciler içinde bulundukları öğrenci grubuna göre değerlendirilir</li><li>• Öğrenciye değerlendirme konusunda aktiflik ve sorumluluk vermez,</li><li>• Değerlendirmeler öğrenme sürecinde bağımsız olarak yapılır</li></ul>	<ul style="list-style-type: none"><li>• Öğrencinin yeteneği, tutumu, bilgisi vb. süreç içinde ölçülür</li><li>• Öğrenci ölçütlerden haberdardır ve değerlendirme sürecinde aktiftir</li><li>• Öğrencilerin bireysel olarak değerlendirilmesi söz konusudur,</li><li>• Öğrenci aktiftir, sorumluluk verilir ve sorumluluk almayı öğrenir.</li><li>• Öğrenme ve değerlendirme süreci bir bütün olarak görülür.</li></ul>

## **2-PERFORMANS DEĞERLENDİRME**

Öğrencilerin bilgi ve becerilerini ortaya koyarak oluşturdukları çalışma,ürün ya da etkinliklerin değerlendirilmesi süreci performans değerlendirme olarak ifade edilebilir.Bunun yanında performans değerlendirme,öğrencilerin gerçek yaşam problemlerine akademik bilgilerini uygulayabilme ve bunu problem üzerinde gösterebilmeleri ile ilgilenir,öğrencilerin öğrendiklerini gerçek durumlarda göstermelerini sağlar.

Performans değerlendirmenin belli aşamaları vardır.Bu aşamalar:

\*Amaçın Belirlenmesi: Performans değerlendirmede sürecin mi,sonucun mu yoksa her ikisinin birlikte mi değerlendirileceğine karar verilmelidir.

\*Performans Ölçütlerinin Belirlenmesi: Performansın ölçütleri, öğrencilerin bir etkinliği tam ve doğru bir şekilde yapması için göstermesi gereken belli davranışları tanımlar. Performans ölçütlerinin belirlenmesi aşamasında, önce değerlendirilecek performansın belirlenmesi gerekmektedir. Ardından, belirlenen performansın özellikleri tanımlanmalıdır.

Bu ölçütlerin sayısının çok fazla olmaması (10-15) gerekmektedir. Bunlar, gözlenebilir davranış veya ortaya çıkacak ürün şeklinde açık ifadelerle belirtilmelidir.

\*Performansın ya da ürünün gözlemlenebileceği ortam oluşturma: Öğrencinin performansını doğru ve güvenilir biçimde değerlendirmek için gözlemlerin birden fazla tekrarlanması gerekmektedir. Gözlem sayısı, yapılacak değerlendirmenin önemine ve gözlem için gereken süreye göre belirlenmelidir.

\*Performansın Puanlanması: Performansın değerlendirilmesi, dereceli puanlama anahtarı, kontrol listesi ile yapılır.

**Performans değerlendirmenin bir çok yararı vardır. Bunlar:**

\*İnsanların bilgiyi kullanması ve gerçek yaşam durumlarına yakın ürünler ortaya koymasını sağlar

\*Performans değerlendirmede kaynakları tekrar tekrar kullanılabilir. Örneğin, öğretmenin konuşma becerisinin ölçülmesi için geliştirdiği konuşma esnasında öğrencilerin uyması gereken ölçütlerin listesi şeklinde hazırlanmış olan bir performans değerlendirme formu; ölçütleri öğrenciden öğrenciye, sınıftan sınıfa veya yıldan yıla değişmeyeceğinden her öğrencinin performansını ölçmek için kullanılabilir.

\*Performansın belirli bölümlerine odaklanıldığından öğretmenin her bir parçayı gözlemleyebilmesini ve değerlendirebilmesini sağlar.

\*Aynı değerlendirme aracı, öğrencinin zaman içindeki gelişimini bir çizelge olarak kullanarak izleyebilir.

Performansın değerlendirilmesinde aşağıdaki noktalara dikkat edilmelidir.

1-Performans değerlendirmede, öğrencilere ödevi tamamlamaları için verilen süre ödevin niteliğine göre ayarlanmalıdır.

2-Ödev, birçok beceriyi kapsamalıdır.

3-Ödevlerin bazıları bireysel ödevler, bazıları ise grup ödevleri şeklinde verilmelidir.

4-Ödev hem ürüne hem de sürece odaklanmalıdır.

**Performans Ödevleri:**

Öğrencilerin sahip olduğu bilgi ve becerileri günlük yaşamla da ilişkilendirerek ortaya koymasını gerektiren kısa dönemli çalışmalardır. Performans ödevi , öğrencilerin bilgi ve becerilerini gerçek yaşam durumlarına uygun olarak kullanmalarını gerektirir.Çok çeşitli konularda performans ödevi verilebilir.Aşağıda bunlara bazı örnekler verilmiştir:

\*Aynı grafikleri farklı ölçülerde çizme

\*Sergi oluşturma

\*Bir konu hakkında yazı yazma

\*Başka zamanlarla, yerlerle veya kültürlerle ilgili hayali bir mizansen oluşturma

\*Bir haritadan sonuçlar çıkarma

\*Bilimsel gözlemlerini tablo oluşturarak belirtme

\*Deney yapma

\*Bir tiyatro oyununa eleştiri yazma

\*Bir konuyla ilgili afiş, poster, broşür vb. hazırlama

**3-DERECELEME ÖLÇEKLERİ (RUBRİK) PUANLAMA YÖNERGESİ**

Dereceleme ölçekleri performans değerlendirmede öğrencileri puanlama amacı ile kullanılan ölçme araçlarından biridir.. Dereceleme ölçekleri, bir ölçme veya etkinliğin tamamının ya da bölümlerinin nasıl puanlanacağını ana hatlarıyla gösterilmesidir. Hangi durum için kaç puan verileceği önceden belirlenir.



Başka bir tanımıyla dereceleme ölçekleri; kabul edilebilir ya da edilemez performans sınırlarının öğretmen ve öğrenciler için açıkça belirtildiği kriter grubu ölçeğidir. En kullanışlı dereceleme ölçekleri ölçme ve değerlendirme yapacak kişinin kendi yaptığı dereceleme ölçeğidir. Dereceleme ölçekleri bir ölçme sürecinin tamamının ya da bölümlerinin nasıl puanlanacağını ana hatlarıyla gösterebilir. Öğrenciden beklenen performansın tüm boyutlarını dikkate alarak puanlama yapmak amacı ile kullanılırlar. Dereceleme ölçekleri, öğrencinin bir kavrama ilişkin bilgisini, bir ürünü belirlenen ölçütlere uygun olarak ortaya koyma becerisini, verilen bir görevi yapma sürecinde uyulması gereken aşamalı adımları yerine getirmedeki ustalığını veya verilen bir ödevi yapmadaki yeterlik düzeyini belirlemede kullanılabilirler. Dereceleme ölçekleri kullanılırken, öğrencilerden beklenen performansların iyi tanımlanması veya sınırlarının iyi çizilmesi gereklidir. Bu amaçla öğrencilerden beklenen performans belli kategoriye ayrılabilir ve her kategorideki ölçütler ayrıntılı olarak belirlenebilir.

Dereceleme ölçeklerinin bütüncül (holistic) ve analitik (analytical) olmak üzere iki biçimi vardır.

Bütüncül dereceleme ölçekleriyle performansın genel sürecinin veya ürünün bir bütün olarak (parçalarını dikkate almadan) puanlaması, analitik dereceleme ölçekleriyle önce performansın genel sürecinin veya ürünün parçalarının ayrı ayrı puanlaması ve sonra da bu kısmî puanların toplanarak toplam puana ulaşılması söz konusudur.

Analitik dereceleme ölçeklerini kullananlar öğrencilerin verdiği bütün cevapları incelemeli ve belirlenmiş ölçütlere göre puanlamalıdır. Böylece yanlılığın az olması beklenebilir. Analitik dereceleme ölçekleri üründen çok süreçle ilgilenmektedir. Bu nedenle öğrenme eksikliği ya da davranış bozukluğu olan öğrencilerde kullanılması iyi sonuç vermeyebilir.

Dereceleme ölçekleri yazılı sınavların puanlanmasında kullanılan anahtarlara ya da ürün ve süreç değerlendirilmesi için kullanılan gözlem ölçekleriyle kontrol listelerine benzer. Dereceleme ölçeklerinde her bir kazanım veya davranış için kriterlerin mükemmelden kötüye kadar derecelenmesi söz konusudur ve her derece için performans ölçütlerinin ayrıntılı tanımlanması gerekir.

#### *Dereceleme ölçeklerini kullanma nedenleri*

- 1-İyi bir ürünün nasıl olacağı konusunda öğretmen ve öğrenci için açık ve anlaşılır bir tanım verir.
- 2-Öğrencilerin üstlendikleri görevlerde, görevleri yerine getirmede, ortaya koydukları ürünlerde ve bunların değerlendirilmesinde daha fazla sorumluluk duymalarını veya almalarını sağlar.
- 3-Öğretmenin öğrenci çalışmalarını veya ürünlerini değerlendirmelerini kolaylaştırır.
- 4-Öğrencilere, kendi performanslarını değerlendirebilecekleri ölçütler sağlar.
- 5-İyi ürün veya performansla ilgili ölçütlerin velilere bildirilmesi, çocuklarına yardımcı olmak isteyen velilere kolaylık sağlayabilir.

#### *Dereceleme Ölçekleri Geliştirme Aşamaları*

- 1-Ölçeğin amacını belirleme: Ölçekle elde edilen sonuçların ne amaçla kullanılacağı belirlenmelidir.
- 2-Değerlendirilecek öğeleri belirleme: Ürün veya görevlerde nelere puan verileceği belirlenmelidir. Ölçekle ölçülmeye çalışılan davranışlar, ürünler ya da becerilerin anahtarı oluşturulur.
- 3-Yeterlik düzeylerine karar verme: Ürün veya görevlerde değerlendirilecek öğelerin gerçekleşme düzeylerine göre nasıl puanlanacağı belirlenmelidir. Ölçülecek davranış, ürün, görev ya da becerinin yeterlik düzeyleri ile ilgili ölçütler belirlenir. Örneğin en düşük düzey için “yetersiz veya 1 puan”; orta düzey için “kısmen yeterli veya 2 puan”; yüksek düzeydeki performans için “yeterli veya 3 puan” gibi kısa ölçütler yazılabildiği gibi değerlendirilecek

öğeler ve yeterlik düzeyleri ile ilgili ayrıntılı açıklamalar da yapılabilir. Burada önemli olan düzeyler arasında ölçütleri iyi ayırt edebilmektir.

4-Ölçeğin hazırlanması: Yukarıdaki aşamalardan sonra ölçek taslak olarak hazırlanır. Gözden geçirme ve düzeltmeler yapıldıktan sonra öğrencileri değerlendirmede kullanılabilir. Ölçeği kullanma sürecinde karşılaşılan sorunlara göre ölçek yetkinleştirilmelidir.

Rubriklerde neyin, kaç puanla değerlendirileceği önceden belli olduğu için değerlendirme esnasında önceden belirlenmiş bu ölçüt dışında hiçbir şey dikkate alınmaz. Böylelikle değerlendirmenin geçerlik ve güvenirliği de yükselir.

PUAN	ÖLÇÜTLER
4	Tablo titizlikle hazırlanmış, görünümü temiz ve düzenli, ilgi çekici
3	Tablo titizlikle hazırlanmış, görünümü temiz ve düzenli, ancak ilgi çekici değil
2	Tablonun yapımı için uğraşmış ancak görünümü temiz ve düzenli değil ilgi çekici değil
1	Tablonun yapımı için uğraşılmamış görünümü temiz ve düzenli değil ilgi çekici değil
0	Öğrenci tablo yapmamış

#### **4-Öz değerlendirme (kendini değerlendirme)**

Belli bir konuda bireyin kendini değerlendirmesine öz değerlendirme denir. Öz değerlendirme, bireylerin yeteneklerini kendilerinin keşfetmelerine yardımcı olan bir yaklaşımdır. Öz değerlendirme, öğrencilerin okulda yaptıkları çalışmaları, nasıl düşündüklerini ve nasıl yaptıklarını değerlendirmelerini gerektirir. Öz değerlendirme, öğrencilerin güçlü ve zayıf yönlerini tanımlarına yardım eder. Performansının düzeyi hakkında karar vermek için kişisel ya da kişiler arası ölçüt koymada öğrencilere fırsatlar sunar. Öz değerlendirme öğrencilerin güdülenme düzeylerinin yükselmesine yol açar. Öğrencilerin değişik durumlarda davranışlarını kontrol altına almalarını sağlar.



Öz değerlendirmenin olumlu ve olumsuz yönleri vardır. İnsanların genellikle kendi performanslarını değerlendirirken yanlış davranma olasılıkları göz ardı edilmemelidir. Öz değerlendirme, ilk başlarda öğrencilerin deneyimsizliği nedeniyle yanılgılara neden olabilir. Öğrenciler deneyim kazandıkça daha doğru kararlar alacaklardır.

## **II- ÖLÇME ARACINDA ARANAN NİTELİKLER**

Ölçme aracı üç özelliğin yeterli düzeyde bulunması istenir. Bunlar kısaca aşağıda tanımlanmıştır.

**1-Güvenirlik:** Ölçme aracının hatalardan arınık olma derecesi, ölçtüğü şeyi tutarlı ölçebilme derecesi.

**2-Geçerlik:** Ölçme aracının kullanım amacına hizmet etme derecesi, ölçtüğü şeyi, diğer değişkenleri karıştırmadan ölçebilme derecesi.



**3-Kullanışlılık:** Ölçme aracının hazırlama, uygulama, puanlama ve tüm bu süreçlerde ekonomiklik derecesi olarak adlandırılabilir. Ölçme aracının emek, maliyet, zaman, hazırlanışı ve uygulama koşulları gibi yönlerden uygun olmasıdır.

### **Bir Ölçme Aracında Bulunması Gereken Nitelikler**

#### **Güvenirlik:**

Bir ölçme aracının ölçümlerine birçok kaynaktan hata karışabilir. Güvenirlik ölçme aracının hatalardan arınlık düzeyidir. Yani ölçme sonuçlarına hata karıştırmadan ölçme yapabilmesidir. Ölçme sonucundaki hata miktarı düştükçe güvenilirlik artar. Bir ölçme işleminde hata miktarı arttıkça güvenilirlik de düşer.

Güvenirlik ölçme aracının kararlılığı ve tutarlılığı ile ilgilidir. Ölçme aracının birkaç kez uygulandığında benzer (birbirine yakın) sonuçları vermesi düzeyidir. Teknik anlamda ise güvenilirlik gerçek puanlar ile gözlenen puanlar arasındaki korelasyonun karesidir.

- Ölçme araçlarının kararlılığı, tutarlılığı (aynı ve benzer sonuçları vermesi) ve duyarlılığıdır.
- Ölçme aracının hatalardan arınık olma.
- Bir ölçme aracının birden fazla uygulandığında benzer ya da yakın sonuçların alınması (tutarlılığı) güvenilir olduğunu gösterir.
- Bir ölçme işlemi ne kadar az hata yaparsa güvenilirliği o kadar yüksek olur.
- Ölçme aracı gerçeği yansıtır derecesidir.
- Bir testin güvenilirliğini doğrudan artıran faktör “cevap anahtarı” hazırlamadır.

**Not:** Bir ölçme aracı geçerliyse mutlaka güvenilirliklidir. Ancak güvenilir olan her ölçme aracı geçerli olmayabilir. Bir ölçme aracı önce güvenilir ve geçerli arkasında da kullanışlı olmalıdır.

**Ölçme ve değerlendirme sürecine karışan hata kaynakları:** Ölçmeyi yapan kişi, ölçmede kullanılan araçlar, ölçülen özellikler, ölçmenin yapıldığı ortam.

Gerçek Puan Kuramı Temel Denklemi:  $X = T + E$

X: Gözlenen puan      T: Gerçek Puan      E: Ölçmeye Karışan Hata Puanı

Örnek bir öğrenci 20 soruluk bir testten 17 puan almış olsun. Bu 17 puanı gözlenen puandır. Ancak öğrencinin sahip olduğu bilgi düzeyi ile 14 puan alabileceğini varsayalım. Bu 14 puan gerçek puandır. Öğrencinin gerçek performansını ya da başarısını yansıtan puandır. Öğrenci geriye kalan 3 puanı da kopya çekerek ve şansla kazanmıştır. İşte bu 3 puana, ya da gözlenen puanla gerçek puan arasındaki farka hata puanı diyebiliriz. X bilinir, çünkü gözlenen puandır. T hipotetiktir (teoriktir), bilinemez. E eğer yönü ve miktarı belli ise bilinebilir. Ancak tesadüfi ise yönü ve miktarı bilinemez. Dolaylı yollardan hesaplanır.

Hiçbir ölçme hatasız değildir. Hemen her ölçmeye milyarda bir oranında da olsa hata karışır. Hataların çeşitli kaynakları vardır. Hatalar: Ölçmeci, ölçme aracı, ölçme yöntemi, ölçme ortamı veya ölçülen kişiden gelebilir.

#### **2.1 Hata Türleri**

**1-Sabit Hatalar:** Her ölçme sonucuna aynı miktarda ve aynı yönde karışan; her ölçme sonucu için aynı yönde etkili olan ve her bir ölçmede miktarı değişmeyen hatalardır. (Ör: Her öğrenciye 10 puan fazla verilmesi veya bir soruyu bütün sınıfın doğru cevaplandırması sabit hatalardandır.)

1. Hata miktarı ölçümden ölçüme değişmez. Hata miktarı her ölçme işleminde aynı orandadır. Daha çok ölçmenin güvenilirliği ile ilgilidir. Sınav aracı, yanlış cevap anahtarı,

soruların sınavda yoklanacak hedef davranışları uygun örneklendirmemesi durumlarında ortaya çıkar.

- Terazinin her ölçtüğü nesneyi 10 gr. fazla ölçmesi.
- Bir öğretmenin okuduğu her sınav kağıdına 10 puan fazla vermesi.
- 25 soruluk bir testte 13.sorunun okunmaması nedeniyle hiç kimse tarafından cevaplandırılmaması
- Sabit hata özellikle **geçerliliği** düşürür.

Dikkat edilirse hata işlemi hep aynı nesnelere veriliyor

**2-Sistematiik Hatalar:** Ölçme sonuçlarına belirli bir kurala göre karışan ve yönü ile miktarı tespit edilebilen hatalara sistematiik hatalar denir. Çoğunlukla her bir ölçme için sabit olmayan, puanlayıcı yanlılığını yansıtan hatalardır. (Ör: Matematik ya da Fen dersinde öğrencilerin güzel yazı yazma, düzen vb.ne puan verme veya sadece erkek / kız öğrencilere fazladan puan verme sistematiik hatalardandır.)

2. Hata miktarı ölçümden ölçüme belli oranda artar ya da azalır. Ölçme aracı belirli şartlar gerçekleştiği zaman hata yaparsa. Öğretmenin (ölçmeyi yapanın) yanlılığından kaynaklanır. Ölçülen büyüklüğe, ölçmeciye ve ölçme koşullarına bağlı olarak miktarı değişen hatadır. Puanlayıcının subjektif olduğu (tarafalı – yanlı) hatalardır. Daha çok ölçmenin geçerliliği ile ilgilidir. Bir testin sistematiik hatalardan arınıklılığının artması testin geçerliliğini artırır. Diğer hata türleriyle birlikte ele alındığında geçerliliği düşürür.
- Öğretmenin sadece erkek öğrencilere 10 puan fazla vermesi.
  - Yazısı güzel olan öğrencilere 5 puan fazla verilmesi.
  - Terazinin 2 kg. ve üstü ağırlık ölçümlerinde 10 gr. eksik tartması.
  - Öğretmenin 50 puanlık sınav kağıdına 5, 60 puanlığa 10, 70 puanlığa 15 puan fazla vermesi.

Terazinin 1kg'lık nesneyi 100 gr, 2 kg'lık nesneyi 200 gr, 3 kg'lık nesneyi 300 gr eksik göstermesi.

Eğitim sistemimizdeki hatalarda öğretmenlerin YANLILIĞI söz konusudur.

**3-Random (tesadüfi, gelişigüzel, rasgele) Hatalar:** Kaynağı, yönü ve miktarı bilinmeyen, ölçme sonuçlarına tesadüfen karışan hatalardır. Bu tür hatalar ölçme sonuçlarına nerede, ne zaman, nasıl, ne kadar karıştı tespit edilemezler. (Ör: Sınav koşullarının öğrencilere etkisi, öğretmenin puanlama dikkatsizliği gibi etkenlerden ölçmelere karışan hatalar bu türdendir.) Ancak bu tür hataların ortalamasının sıfıra yaklaştığı ve normal dağıldığı kabul edilir.

Ölçmede asıl ilgilenilen hata türü tesadüfi hatalardır. Sabit hatalar standart sapma ve varyansa etki etmezler. Sistematiik ve sabit hatalar kontrol edilebilirler. Ölçme sonuçlarına hangi yönde ve hangi miktarda karıştığı tespit edilerek ölçme sonuçlarından çıkarılabilirler. Böylece de ölçme sonuçları hatalardan arındırılmış olur. Dolayısıyla sabit ve sistematiik hataların ölçme sonuçlarının güvenilirliğine etkileri azaltılabilir veya tamamen ortadan kaldırılabilir. Bu nedenle güvenilirlik söz konusu olduğunda ölçmecileri asıl ilgilendiren hata türü olarak tesadüfi hatalara vurgu yapılır. Ancak random hataları kontrol etme imkânı yok gibidir. Bu nedenle yapılabilecek en doğru işlem random hataları azaltmak için gerekli önlemleri alabilmektir. Hata miktarı kuralsız olarak artar ya da azalır. Hatanın kaynağı bilinmez. Daha çok ölçmenin güvenilirliği ile ilgilidir. Kaynağı bilinmeyen ve ölçme sonuçlarına gelişigüzel karışan hatalardır. Örneğin; öğrencinin sınav günü hasta olması, sınav salonuna güneş vurması, şans başarısı, kopya çeken öğrencinin aldığı puan.

- Terazinin bir nesneyi 100 gr. eksik, başka bir nesneyi 50 gr. fazla, başka bir nesneyi 15 gr. eksik tartması.
- Öğretmenin bir sınav kağıdına 10 puan eksik, başka bir kağıda 15 puan fazla, bir diğer kağıda 20 puan eksik vermesi.

- Cevap kağıdına kodlama yapılırken kaydırma yapılması.
- Birkaç öğrencinin sınav sonuçlarının fazla gelmesi.
- Öğrencinin sınav günü hastalanması.
- Şans başarısı(atıp-tutma)
- Bir öğretmenin sınav kağıtlarını okurken zaman zaman açık olan TV'den gelen sesler dikkatini çekmiş ve bazı kağıtları puanlamada hatalar yapmıştır bu hata tesadüfi hatadır.

**Not:** Sabit ve sistematik hatalar düzeltilebilir. Tesadüfi hatalar nerden geldiği bilinmediği için düzeltilemez. En çok tesadüfi (Random) hata güvenilirliği düşürüyor. Eğitimde kullanılan ölçme araçlarında her zaman bir miktar hata payı vardır. Bunun olası nedeni ise, ölçülen özelliklerin dolaylı olarak ölçülmesinden kaynaklanır

Ölçme aracının duyarlılığı hataları azaltır. Ölçmecinin yeterli oluşu, ölçme aracının iyi kullanılması, puanlamanın objektifliği, ölçme yapılan ortamın amaca uygunluğu ve ölçülen kişinin performansını yansıtacak durumda olması hataları azaltır.

**Ölçme hatası:** Gerçek ölçme sonucu ile yapılan ölçme sonuçları arasında meydana çıkan farktır.

<b>Ölçmeciden kaynaklanan hatalar</b>	<ul style="list-style-type: none"> <li>➤ Ölçmecinin taraflı davranması – dikkatsiz davranması.</li> <li>➤ Kız öğrencilere ya da yazısı güzel öğrencilere 5 puan fazla not vermek.</li> </ul>
<b>Ölçme aracından kaynaklanan hatalar</b>	<ul style="list-style-type: none"> <li>➤ Soruların anlaşılammaması, ölçme aracının birden fazla davranışı karmaşık ölçmesi, soruların iyi düzenlenmemiş olması, soruların okunamıyor olması gibi hatalar.</li> </ul>
<b>Ölçmenin yapıldığı ortamdan kaynaklanan hatalar</b>	<ul style="list-style-type: none"> <li>➤ Sınav yerinin ısı, ışık ve havalandırmasının yetersiz olması, ses ve gürültünün rahatsız edici olması.</li> </ul>
<b>Ölçülen kişiden kaynaklanan hatalar</b>	<ul style="list-style-type: none"> <li>➤ Öğrencinin derse ve sınava tutumu, sorumluluğu, kişiliği, yorgunluğu, isteksizliği.</li> </ul>
<b>Ölçme aracının çeşidinden kaynaklanan hatalar</b>	<ul style="list-style-type: none"> <li>➤ Sözlü sınav yerine, yazılı sınavın kullanılması.</li> </ul>

### **Ölçmenin Standart Hatası**

Ölçmelerin tutarlılığı ile güvenilirlikten sonraki ikinci yaklaşım hataların büyüklüğünü belirlemektir. Tek bir gözleme bakılarak hata konusunda bir şey söylenemez. Aynı kişiye ait birden fazla ölçme yapılmalıdır. Bu ölçmeler bir dağılım gösterir. Bu dağılımın ortalaması kişinin gerçek puanının tahminidir. Yani gerçek puana en yakın puandır. Bu dağılımın standart kayması ölçmenin standart hatası olarak adlandırılır.

Ancak bireylerle ilgili tekrarlı ölçmeler yapmak zaman alıcıdır. Bu nedenle çok sayıda kişiye test uygulamak, bu testten elde edilen puanların standart kaymasını ve güvenilirliğini kullanmak yoluyla bireylerin puanlarına karışan hata (ölçmenin standart hatası) hesaplanabilir. Bu yolla hesaplanan ölçmenin standart hatası formülü aşağıdaki gibidir:

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

$S_x$ =Testin standart sapması

$r_x$ = Testin güvenilirliği

Örnek1: Oktay öğretmen güvenilirliği 0,84, standart sapması 5 olan Kimya testinin standart hatasını hesaplamaya çalışmaktadır.

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

$$Se = 5\sqrt{1-0,84}$$

$$Se = 5\sqrt{0,16}$$

$$Se = 5.(0,4)$$

$$Se = 2$$

Örnek2: Oktay öğretmen güvenilirliği 0,75 standart sapması 3 olan Biyoloji testinin standart hatasını hesaplamaya çalışmaktadır

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

$$Se = 3\sqrt{1-0,75}$$

$$Se = 3\sqrt{0,25}$$

$$Se = 3.(0,5)$$

$$Se = 1,5$$

Oktay öğretmen bu iki testi karşılaştırdığında Kimya testinin standart hatasını 2 biyoloji testinin standart hatasını 1,5 olduğunu görür.

Bir testin standart hatası ne kadar düşükse güvenilirliği de o derece yüksektir.

O halde Biyoloji testinin güvenilirliği, Kimya testinden daha fazladır.

Standart hata grubun değişirliğinden az etkilenir. Güvenirlik ise grubun değişirliğine bağlıdır. Bu nedenle bazen güvenirlilik yerine kullanılması önerilmiştir. Özellikle testteki puanların ve puanlar arasındaki farkların güvenirliliği ile ilgili yargılar için çok kullanışlıdır. (Ör: Güvenirliliği .64 olan ve standart kayması 5.03 olan testten bir öğrencinin 48 puan aldığını varsayalım. Geçme notu 50 olsun. Bu durumda öğrenci sınıfını geçmeli midir? Nasıl karar vereceğiz? Kararı, standart hatayı hesaplayıp öğrencinin puanına ekleme ve çıkarma yapma yoluyla öğrencinin puan aralığını bulduktan sonra vermek yararlıdır. Bu test için ölçmenin standart hatasını hesaplayacak olursak ilgili formülden 3,02 olarak elde edilir. Bu puanı öğrencinin puanına ekleyip çıkarınca öğrencinin puanı 44,98 – 51,02 arasında olacaktır. Öğrencinin puanı 50 geçme notunu kapsadığından geçer not vermek doğru olur.

### **Güvenirliliği Belirleme Yollarından Bazıları:**

**1-Test Tekrar Test Güvenirliliği:** Aynı testi aynı gruba iki kez uygulama yoluyla elde edilen güvenirlilik katsayısıdır. İki uygulamadan elde edilen puanlar arasındaki korelasyon katsayısıdır. İki uygulama arasından belirli bir zaman geçmelidir. Aradan geçen zaman cevaplayıcıların ölçülen özellik bakımından değişmeyecekleri, aynı zamanda ilk uygulamada verdikleri cevabı hatırlamayacakları bir süre olmalıdır. Bu güvenirlilik katsayısına Kararlılık Katsayısı da denir.

**2-Paralel Test Güvenirliliği:** Test-tekrar test yönteminin dezavantajını gidermek amacıyla. En az iki eşdeğer form geliştirilir. İki test madde sayısı, niteliği ve ölçtükleri davranışlar bakımından denk olmalıdır. İki test aynı öğrenci grubuna uygulanır. İki testten elde edilen cevaplayıcı puanları arasında korelasyon katsayısı hesaplanır. Uygulama aynı öğrenci grubuna aynı zamanda yapılırsa hem kararlılık anlamında güvenirlilik hem de formların eşdeğerliği hakkında bilgi verir.

**3-İki Yarı Güvenirliliği:** Tek uygulama gerektirdiğinden çok kullanılır. Test iki eşdeğer yarıya bölünür (Yarisından veya tek-çift sorular). Öğrencilerin bu iki yarılarından elde ettikleri puanlar arasında korelasyon elde edilerek testin bir yarısı için güvenirlilik bulunur. Bu korelasyon kullanılarak testin tümü için güvenirlilik hesap edilir. Testin tümü için güvenirlilik katsayısının formülü şöyledir.

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

Bu yöntemin kullanılmasında tek boyutlu; aynı niteliği ölçen maddelerden oluşan bir test olmasına özen gösterilir. Aksi durumda yorumlamada sıkıntı yaratacağından kullanılmaz.

**4-KR-20 Güvenirliği:** Bu formül için madde güçlük indeksi (p) bilinmelidir. İkinci olarak testteki her bir maddenin aynı değişkeni ölçtüğü yani testin ölçmek istediği niteliği ölçtüğü sayılışı vardır. Formül madde güçlüklerini ve madde sayısını dikkate alarak hesaplanan bir güvenirlilik katsayısıdır.

$$KR - 20 = \frac{K}{K - 1} \left( 1 - \frac{\sum pq}{S_x^2} \right)$$

K: Madde sayısı

p: Maddeyi doğru cevaplandıranların tüm cevaplayıcılara oranı

q: Maddeyi yanlış cevaplandıranların tüm cevaplayıcılara oranı

$S_x^2$ : Tüm testin varyansı

**5-KR-21 Güvenirliği:** Testteki maddelerin eşit ya da yakın güçlükte olduğu varsayılıyorsa ve madde istatistikleri (güçlükleri veya varyansları = pq) bulunmamışsa test puanları ortalaması ve varyansı kullanılarak KR-21 güvenirlilik katsayısı hesaplanabilir. Maddelerin eşit güçlükte varsayılmasından dolayı KR-21 genellikler KR-20'den küçük çıkar.

$$KR - 21 = \frac{K}{K - 1} \left( 1 - \frac{K \bar{X} - \bar{X}^2}{S_x^2} \right)$$

Eğer maddeler güçlük bakımından oldukça farklı ise yani birbirine yakın veya eşit değilse bu formül güvenirliliğin en alt sınırını verir. Bu durumda elde edilen güvenirlilik katsayısı yeterli görülüyorsa (.75'in üstünde ise) testin güvenilir olduğunu söylemek mümkündür.

KR-20 ve KR-21 formülleri sadece 1-0 puanlamasında kullanılır. Eğer puanlama farklı ise bu formül kullanılmaz. Cronbach alfa güvenirlilik katsayısı kullanılır.

**6- Cronbach alfa:** Maddeler (sorular) 1-0'dan farklı puanlandığı zaman veya bir testin alt testleri varsa bu tür testlerin güvenirliliğini KR-20'nin özel bir hali olan Cronbach Alfa tekniği ile hesaplarız. KR-20'den farklı eğer testin alt testleri var ise madde varyansları toplamı yerine alt test varyansları toplamı alınır. 1-0'dan farklı puanlandığında KR-20'de olduğu gibi madde varyansları kullanılır.

$$Cronbach \alpha = \frac{K}{K - 1} \left( 1 - \frac{\sum S_j^2}{S_x^2} \right)$$

$\sum S_j^2$ : Madde varyansları toplamı veya alt test varyansları toplamı

**7-Puanlayıcı Güvenirliği:** Cevap kâğıtlarına birden fazla puanlayıcının verdiği puanlar arasındaki korelasyondur. Puanlayıcıların yansızlığı ve uyumu hakkında bilgi verir. Daha çok yazılı sınavlar, sözlü sınavlar, performans testleri gibi durumlarda kullanılır.

**8-Puanlama Güvenirliği:** Bir puanlayıcının bir sınav kağıdına farklı zamanlarda (en az iki kez) verdiği puanlar arasındaki korelasyondur. Puanlayıcının kararlılığı ve yansızlığı hakkında fikir verebilir. Daha çok yazılı sınavlarda, ödev-projelerde ve kısmen de kısa cevaplı testlerde kullanılabilir.

**Tablo 1. Güvenirlik Belirleme Tekniklerine İlişkin Özet Bilgiler**

Güvenirlik Katsayısı Tipi	Anlamı	Gereken form sayısı	Uygulama sayısı	Hata varyansı kaynağı
<b>KUDER-RICHARSON KR - 20 ve KR - 21</b>	İç tutarlık	Bir	Bir	Kapsam örnekleme ve heterojenliği
<b>CRONBACH ALFA</b>	İç tutarlık	Bir	Bir	Kapsam örnekleme ve heterojenliği
<b>İKİ YARI</b>	İç tutarlık	Bir	Bir	Kapsam örnekleme
<b>TEST-TEKRAR TEST</b>	Kararlılık	Bir	İki	Zaman örnekleme
<b>PARALEL FORMLAR Birlikte uygulama</b>	Tutarlık	İki	Bir	Kapsam örnekleme
<b>PARALEL FORMLAR Aralıklı uygulama</b>	Tutarlık	İki	İki	Zaman ve kapsam
<b>PUANLAYICI</b>	Tutarlık	Bir	Bir	Puanlayıcılar
<b>PUANLAMA</b>	Kararlılık	Bir	Bir	Puanlama

Açıklanan güvenirlik belirleme tekniklerine ilişkin özet bilgiler tablo 4'te verilmektedir. Tablo 4 güvenirliklerin hangi anlama geldiği, kullanılması gereken form sayısı, uygulama sayısı ve hesaplanan güvenirlik değerlerine karışan hata kaynaklarını göstermektedir.

### Güvenirliği Artırma Yöntemleri ve Dikkat Edilmesi Gerekenler

Soru sayısını artırmak güvenirliği artırır. Güvenirlikle soru sayısı arasında matematiksel bir ilişki vardır. Soru sayısı arttıkça duyarlı ölçme yapılacağı için güvenirlik de artar. Ancak soru sayısını bir noktadan sonra artırmanın getirişi çok fazla değildir. Bu ilişki aşağıdaki şekilde görülebilir. Güvenirliği tam (1) yapabilmek için eklenecek soru sayısı sonsuzdur. Monoton artan bir fonksiyondur. Bir noktadan sonra eklenecek soru sayısı ile artacak güvenirlik miktarı önemsiz bir düzeye gelir. Bunu bir örnekle açıklayacak olursak madde sayısı 4 güvenirliği 10 olan bir teste sürekli soru ekleme durumunda güvenirlik artış miktarı şöyle olacaktır.

Aşağıdaki şekilde örüldüğü gibi bir noktadan sonra meydana gelen güvenirlik artışı eklenen soru sayısı karşısında değmeyecek bir miktardır. Soru sayısı ile güvenirlik ilişkisi kapsamında ne kadar soru eklenince istenilen güvenirliği elde edeceğimizi gösteren formül vardır. Bu formül aşağıdaki gibidir.

$$r_n = \frac{nr_s}{(n-1)r_s + 1}$$

formülde geçen terimlerin açıklamaları şöyledir.

$r_s$ =Soru eklenecek testin (eldeki testin) güvenirliği

$r_n$ =Testin soru eklendikten sonra beklenen güvenirliği (istenilen güvenirlik)

$n$ =Testin kaç kat uzatılacağı

Bir sınavda soru sayısı arttıkça güvenirlik artar. Soru sayısının artmasıyla öğrencilerin sallama şansı azalacağı için güvenirliği artırır. Soru sayısı fazla olan çoktan seçmeli testlerin güvenirliliği, soru sayısı az olan yazılı yoklamalardan daha yüksektir.

1. Soruların öğrenciler tarafından açıkça anlaşılır olması ve kesin yanıtının olması güvenirliği artırır. Bu durum öğrencilerin gelişi güzel cevaplarını engellediği için güvenirliği artırır



2. Sınav için yeterli sürenin verilmesi güvenilirliği artırır. Fazla süre verilmesi nedeniyle kopya çekilmesi nedeniyle geçerlilik düşer.
  3. Ölçmede duyarlılığı yüksek olan araçların (daha küçük birimleri ölçebilen) kullanılması güvenilirliği artırır. Bir ölçme işleminde hata oranının az olması duyarlılığı yüksek olduğunu gösterir.
  4. Ölçme aracındaki soruların güçlüğü “orta düzey” olmalıdır. Soruların çok kolay ya da çok zor olması güvenilirliği düşürür. Sorular, cevaplayacak öğrencilerin yaklaşık yarısı tarafından doğru cevaplandırılacak güçlükte olmalıdır. Soruların çok kolay veya çok zor olması bilenle bilmeyen öğrencilerin birbirinden ayrılmasına neden olduğu için güvenilirliği düşürür.
  5. Sınavın puanlanmasının nesnelliği güvenilirliği artırır. Objektif puanlanabilen sınavların puanlanmasında hatalar az olacağı için güvenilirlik yüksektir. Bir testin değişik kişilerce ya da aynı kişi tarafından farklı zamanlarda puanlanması ile elde edilen sonuca “**Puanlama Güvenirliği**” denir. Sınavın objektif puanlanması güvenilirliği artırır. Objektif olma puanlama biçimiyle alakalıdır.
  6. Öğrencilerin sınavlar ile ilgili fikirleri ve ön yargıları da güvenilirliği etkiler. Öğrencinin sınavı bir ölüm – kalım meselesi haline getirmesi güvenilirliği düşürür.
  7. Bir ölçme aracının güvenilir olması öncelikle, test puanlarına karışan hata miktarının az olmasıyla gerçekleşmiş olur.
  8. Öğrenciler sınava motive edilmeleri soruları dikkatli yanıtlamalarını sağlayacağından güvenilirliği artırır.
- Güvenirlik katsayısı asla eksi olmaz.

#### **2.4 GEÇERLİK**

Bir ölçme aracının sadece o araçla ölçülmek istenen değişkeni ölçme, başka değişkenlerle karıştırmama veya amaca hizmet etme derecesidir. Testin sahip olması gereken en önemli niteliktir. Buna göre *geçerlik, genel anlamda, bir ölçme aracının ya da ölçme yönteminin ölçmeyi amaçladığı özelliği, başka herhangi bir özelliklerle karıştırmadan, doğru olarak ölçebilme derecesi* olarak tanımlanabilir. Başka bir tanım ise geçerliği, maddenin/testin bilenle bilmeyeni veya bir davranışa sahip olanla olmayanı ayırma derecesi olarak ifade eder. Bu tanım eğitimde kullanılan araçların geçerliğini belirlemede, madde analizlerinde çok kullanılır. Bu terim, testlerin sağladığı sonuçların kullanılma amaçları için geçerli olup olmadığını, geçerli oluş derecesini ve gerçekliğini belirtmektedir.

Bir ölçme aracının geçerliği yüksek ise güvenilirliği de yüksektir. Bunun tersi doğru değildir. Güvenirliği yüksek olan testin geçerliği düşük olabilir.

Geçerlikte iki önemli soru vardır. 1) Ölçme aracı hangi amaç ya da amaçlar için kullanılabilir. 2) Ölçme aracı bu amaç veya amaçlara ne derece de hizmet edebilir, uygundur?

Bir testin geçerliği bir derece sorunudur. Ya hep ya hiç sorunu değil. Hatalardan tamamen kurtulmak mümkün değil ancak minimum yapmaya çalışmak gerekir.

Geçerliği birden fazla yolla tanımlamak mümkündür. Tek bir tanım vermek yerine kullanım amacına uygun olduğuna ilişkin kanıtlar bulmak daha doğrudur. Çok bilinen geçerlik türleri kısaca şöyle özetlenebilir.

**1-Kapsam Geçerliği:** Ölçülmek istenen kapsamla geçerliği araştırılan testin ölçtüğü kapsamın örtüşme derecesidir. Bir ölçme aracının ölçülmek istenen davranışları (özellikleri) ne derece kapsadığının belirlenmesidir. Kapsam geçerliğinin belirlenmesi için ölçme konusu olan özelliklerin ya da davranışların önceden belirlenmesi gerekir.

Eğitimde hedefler geneldir, gözlenemezler. Oysa ölçülecek nitelik ya da davranış doğrudan ya da dolaylı olarak gözlenebilmelidir. Bu nedenle hedefler yerine davranışlar ölçme konusu yapılır.

Ör: Hedef 1; "Ölçmeyle ilgili belli başlı kavramların bilgisi." Geneldir. Gözlenemez.

Davranış 1; Güvenirlilik kavramının tanımını yazma, söyleme veya verilen seçenekler arasından işaretleme. Özeldir. Gözlenebilir.

#### **Kapsam Geçerliğini Belirleme Yolları**

İki temel yol olduğu söylenebilir. Birincisi *mantıksal (rasyonel)*, ikincisi *istatistikî yoldur*.

**1-a- Belirtke Tablosu Hazırlamak.** Belirtke tablosu, sütun başlarında hedefler, satır başlarında konular ve hücrelerde hedefler ve konularla ilgili davranışların numara ya da sayılarla gösterildiği tablodur. Yani davranışların konu ve hedeflere göre dağılımını gösteren çizelgelerdir.

**Tablo 2. Örnek Belirtke Tablou**

Hedef-dav.							
Konular	Bilgi	Kavrama	Uygulama	Analiz	Sentez	Değerlendirme	Toplam (%)
1.Ölçmenin Temel	4	3	2	2	1	4	16 (%33)
2. Ölçme Aracının	3	5	8	3	2	3	24 (%50)
3.Yazılı Yoklamalar	-	3	1	2	1	1	8 (%17)
Toplam (%)	7 (%15)	11 (%23)	11 (%23)	7 (%15)	4 (%7)	8 (%17)	48 (%100)

Örnek belirtke tablosunda (Tablo 2)görüldüğü gibi davranışların konu ve hedeflere göre dağılımı belirlenir. Sınav sorularının bu dağılımı ne derece yansıttığı incelenir. Bu inceleme sonucunda kapsam geçerliği olup olmadığına karar verilir. Eğer sorular bu dağılıma uygun olarak ölçme yapıyorsa, yani davranışları konu ve hedef ağırlıklarını da göz

önünde bulundurmak koşuluyla yeterince kapsıyorsa kapsam geçerliğinden söz edilir.

**1-b-Uzman kanısına başvurma:** Sorular uzmanlara gösterilerek ilgili davranışları ne derecede ölçtüğünü belirlemeleri istenir. Bu amaçla "evet, hayır"lı, davranışlarla "ilgili, ilgisiz" veya davranışları "ölçüyor, ölçmüyor" gibi cevapların verilebileceği ölçekler kullanıldığı gibi soruların davranışları ölçme düzeyini gösteren 3 ya da 5'li likert tipi ölçekler de kullanılabilir. Elde edilen sonuçlar yüzde-frekans veya korelasyon katsayısı hesap edilerek değerlendirilebilir. Bu tekniğe *Bir uzmanın soruları belirtke tablosuyla karşılaştırma tekniği de denebilir*. Bu teknik belirtke tablosu hazırlamanın devamı olan bir tekniktir. Bu tekniğin yukarıdaki teknikten tek farkı hazırlanan belirtke tablosunu hazırlayan kişinin değil bağımsız bir uzmanın incelemesidir.

**2-İstatistiksel Yöntem:** Daha önceden hazırlanmış ve kapsam geçerliliğinin yüksek olduğu kabul edilen test ile bizim geliştirdiğimiz test aynı öğrenci grubuna uygulanır ve öğrencilerin iki testten aldıkları puanlar arasındaki korelasyon katsayısı hesaplanır. Korelasyon katsayısı ne kadar yüksekse, testin kapsam geçerliliği de o kadar yüksektir.

Özet olarak kapsam geçerliğini, sağlamak için uzman kişiye danışma, belirtke tablosunu dikkate alma ve konu hedef ağırlığına bakmak gerekir. Kapsam geçerliği için kullanacağımız testin nitelikleri kullanma amaç ve durumlarıyla mümkün olduğunca eşlenmelidir. Kapsam geçerliği alan uzmanlarının testle ilgili görüşleri olarak kabul edilebilir.

Örnek: Dört işlemle ilgili bir test hazırlayalım. Testin amacı öğrencilerin dört işlemle ilgili becerilerini ölçmektir. Ancak testi hazırlarken çıkarma işlemiyle ilgili

sorular yazmazsak kapsam geçerliği düşecektir. Çıkarma işleminin yanı sıra bölme işlemi ile ilgili soruları da yazmayı ihmal edersek kapsam geçerliği daha da düşecektir. Aynı örneği şu şekilde de inceleyebiliriz. Dört işlem becerisi ile ilgili davranışların ağırlıkları (çarpma, bölme, çıkarma ve toplama işlemlerinin ağırlıkları) farklı olsun. Bu konuların ağırlığını belirlemede, çözülen örnek sayısı, harcanan ders saati kitapta ayrılan sayfa sayısı ve esas olarak belirtke tablosundaki davranışların sayısı olacaktır. Kullanılacak testte bu ağırlıkla dengeli olacak şekilde soru yoksa kapsam geçerliği düşük olur.

- En önemli sayılabilecek geçerlilik türüdür.
- Bir ölçme aracının ölçme yapacağı hedef - davranışları (konuları) yeterli olarak ölçebilme derecesidir. Yani ölçme aracının ölçmek istediği davranışları ne derecede kapsadığı geçerliğini gösterir.
- 30 soruluk program geliştirme testi hazırlanırken tüm sorular bilişsel hedeflere ayrılmışsa o ölçme aracının kapsam geçerliliğinden söz etmek mümkün değildir.
- Kapsam geçerliliğini artırmak için **belirtke tablosundan** yararlanılır. Bir de uzman görüşüne başvurulmalıdır.
- Ölçülmesi hedeflenen evreni (hedefler – konular) yeterli ve dengeli olarak örnekleyen ve kapsadığı maddelerin (soruların) her birinin ölçmeyi hedeflediği, hedef – konuyu gerçekten ölçebildiği bir test kapsam geçerliğine sahiptir.
- Örnek, bir derste toplam 10 ünite işlenmiş, sınavında ise 5 soru sorulmuştur. Sorular 2. 5. 7. 8. 9. ünitelerden sorulmuş olsun. Böyle bir sınavda kullanılan sorular işlenen tüm üniteleri yeteri kadar temsil edemediği için sınavın kapsam geçerliği düşüktür. Ya da testteki bir madde (soru) ilgili hedef – davranış yeterince kapsamıyorsa kapsam geçerliği düşüktür.
- Eğer ölçme aracı (sınav) öğrencinin dersin hedeflerini gerçekleştirme derecesi yerine, kişiliği ya da görünümü gibi (ölçmek istediği özelliklerin dışındaki özellikler) faktörlere dayalı olarak yapılması durumunda yapılan ölçmenin geçerliği yoktur.

Kapsam geçerliğinin aşamaları: ölçülecek nitelik ile ilgili davranışların belirlenmesi, ölçme aracını oluşturan davranışların (örneklemenin) belirlenmesi, ölçme aracını oluşturan davranışlar örneklemenin evreni temsil etme düzeyinin belirlenmesi , ölçme aracındaki soruların ilgili olduğu davranış ölçmeye uygun olup olmadığının kontrol edilmesi

**Anahtar Kelimeler:**Belirtke Tablosu,Hedef Davranış,İçerik,Bütün

**2-Ölçüte Dayalı Geçerlik (Ölçüt Geçerliği):** Test puanlarıyla ölçüt yada ölçütler arasındaki korelasyon katsayısıdır. Ölçütler genelde "dış ölçüt"tür. Elde ettiğimiz geçerlik ise bir ölçüte göre geçerliktir. Geçerlik katsayısı ölçüt alınan puanlara göre değişecektir. Bu nedenle öncelikle doğru ölçüt bulmak gerekir. Yani geçerliği araştırılan konuyla ilgili bir ölçüt olmalıdır. Aynı zamanda ölçütün geçerliği ve güvenilirliği yüksek olmalıdır. Ölçüt geçerliği ölçütün elde edildiği zamana göre ikiye ayrılarak incelenir.

**a-Yordama Geçerliği:** Ölçüt puanların yordayıcı (geçerliği araştırılan) puanlardan sonra elde edilmesi durumundaki geçerliktir, Bu geçerlik türünde yordayıcı puanları daha sonraki bir performansı yordamak (kestirmek-tahmin etmek) amacıyla kullanılır. Eğer iyi yorduyorsa (yani yordayıcı ve ölçüt puanları arasındaki ilişki yüksekse) yordama geçerliğinden söz edilir. Yordama bir tahmindir. Ancak her tahmin bir yordama değildir. Bir tahminin yordama olabilmesi için geleceğe dönük olması belli bilgiler temelinde bazı teknikler kullanılarak yapılması gerekir. Yordama işinde (geçerliğinde) kullanılan istatistiksel teknik regresyondur.

Yordama geçerliğinde en önemli sorun uygun bir ölçüt, bulmaktır. İyi bir ölçüt şu özelliklere sahip olmalıdır. 1-Ölçüt öncelikle yordanmaya çalışılan değişkenle ilişkili olmalıdır. (Örneğin müzik yeteneğini ölçen bir araçla kişinin okuduğunu anlama yeteneğini yordayamazsınız.) 2-Ölçüt güvenilir olmalıdır. Ölçüte ilişkin puanlar ne kadar çok hata

içerirse yordama geçerliği o derecede düşük olacaktır. 3-Ölçüt geçerli olmalıdır. Zaten geçerlik çalışması bir aracın amaca uygunluk derecesini ortaya koyma çalışmasıdır. Bu nedenle amaca uygunluğu kanıtlanmamış ya da amaca uygun olmadığı bilinen bir ölçütle bu çalışma başarısız olur. 4-Kolay ve pratik olmalıdır. Elde edilmesi ve uygulanması kolay ve pratik olmalıdır.

**b)Uyum (Uygunluk) Geçerliği:** Ölçüt puanların yordayıcı puanlarla aynı zamanda veya daha önce elde edilmesi durumunda, ölçüt ile yordayıcı arasındaki korelasyon katsayısına uyum geçerliği adı verilir. Uyum geçerliğine hâlihazır geçerlik, mevcut durum geçerliği, eş-zaman geçerliği ve uygunluk geçerliği de denilmektedir. Yordama geçerliğinden farkı; testin ölçülmek istenen davranış veya özellikleri hangi düzeyde ölçebileceğini belirlemektir. Yordamada ise kestirme vardır. Özellikle testin amacı tanı koymak ise uyum (eş-zaman) geçerliğine başvurmak yararlıdır. Örneğin üniversite sınavı ile üniversite başarısı; işe giriş sınavı ile iş performansı arasındaki korelasyon katsayıları yordama geçerliği olurken; üniversite giriş sınavı ile ortaöğretim başarı puanı (ölçüt olarak) veya matematik dersi başarısı ile aynı zamanda elde edilen matematik yeteneği testi puanları arasındaki korelasyon katsayıları uyum geçerliği olur.

Bir testin bir ölçüte göre geçerliği bu testin güvenirliğinin karekökünden büyük olamaz.

Ölçüt geçerliğini yordayıcı ve ölçütün güvenirlikleri, yordayıcı ve ölçütün birbirine etkisi (bağımlı-bağımsız olması), grubun homojenliği ve puanların ranjı (genişliği) etkiler.

**3-Yapı Geçerliği:** Bireylerde var olduğu kabul edilen özelliklere psikolojide yapı denir. "Zeka, tutum, ilgi, kaygı, depresyon, atılganlık, içedönüklük, dışa dönüklük, güvensizlik vb." gibi kişilik özellikleri psikolojik yapıya örnek olarak verilebilir. Yapı, birbiriyle ilişkili olan öğeler ya da öğelerin oluşturduğu örüntü olarak tanımlanır. Bu yapılar hipotetiktir (teoriktir, kuramsaldır). Yani var oldukları kabul edilir. Doğrudan ispatlanamaz. Ancak gözlemlerle varlıkları hissedilir. Örneğin günlük dilde "zeki adam" dediğimizde bunu ispatlayanlarız. Yani zekânın varlığını ispatlamaya kalkmayız. O kişinin bizim kabul ettiğimiz bir yapıyla ilgili davranışları gösterdiğini düşünürüz.

Yapı geçerliği bir araçla ölçülmek istenen yapının o araçla ortaya konulma derecesi olarak tanımlanabilir.

Evinize misafir geldiğini düşünün.Evde yemek yok ve siz de pilav yapmaya karar verdiniz. Prinçleri tabağa döktünüz; ve teker teker içerisindeki siyah taşları ayıklıyorsunuz.Beyaz olan prinç taneleri bir tarafa,siyah olan taşlara diğer tarafa... Aynı örneği eğitim sistemimiz için düşünürsek,bir eğitimcinin sorduğu 30 soru içerisinde amacına hizmet eden ayırt edici ve güvenirliliği yüksek olan soruları alıp;diğerlerini testin dışında bırakmasıdır.

Yapı geçerliği, yapının tanımlanmasındaki doğruluğa, yapıdan çıkarılan gözlenebilir davranışların yapıyla ilgili olmasına, davranışların yapının tamamını (kapsamını) temsil etmesine, gözlenebilir davranışları ölçmek için yazılan maddelerin (soruların) onları ölçebilme derecesine bağlıdır.

Bir testin, veya ölçme işleminin teorik bir yapıyı ölçüp ölçmediğinin belirlenmesidir. Başka bir ifadeyle bir ölçme aracının teorik olarak ölçmek istediği yapıyı ölçmeye uygun olup olmadığına dair kanıt toplama çalışmasıdır. Yapı geçerliği çalışmalarında bir grubun ve grupta yer alan bireylerin sahip olduğu teorik yapıyı (kişilik, yaratıcılık, ilgi, zeka, tutum) belirlemeye çalışılır. Yapı geçerliğinin temel mantığı; ölçülmesi planlanan teorik yapıya yüksek derecede sahip olan bir gruba düşük derecede sahip olan bir gruba testi uygulayıp, testin bu grupları ne derece uygun olarak ayırıp, ayırmadığına bakmaktır

**Anahtar Kelimeler:**Prinç,örüntü,Faktör Analizi

**4-Görünüş Geçerliği:** Bir testin ne ölçtüğü ile ilgili değil ne ölçüyor görüldüğü ile ilgilidir. Testin ölçülmek istenen değişkeni ölçüyor görünmesidir. Bu çalışma testin tamamı için "olduğu kadar maddeler için de yapılır. Yani maddelerde ölçülmek istenen davranışı ölçüyor görünmelidir. Örneğin "14/4 işlemi sonucunda kalan kaç olur." ile "On dört'ü dörde böldüğümüzde kalan kaç çıkar." Soruları matematik davranışını aynı geçerlik düzeyinde ölçüyor görünmemektedir. Birincisi daha iyi bir matematik sorusu gibi görünmektedir.

Bir ölçme aracının, ölçmek istediği değişkeni ne derece ölçebilir görüldüğüdür. Genelde ölçme aracının görünüş geçerliği uzman incelemesiyle belirlenir. Testin bütünü için gerekli olan görünüş geçerliği her bir soru içinde geçerlidir. Testteki her bir soru ölçmek istediği özelliği ölçüyor görülmelidir. Test kitapçığının üzerinde ne yazıyorsa testte de o alanla ilgili sorular yer almalıdır.

#### **Geçerliği Artırmanın Yolları**

- 1-Her ölçme aracının geçerliği o araçtaki maddelerin teker teker geçerliğine bağlıdır. Her soru araçla ölçülmek istenen davranışların en az birini ölçmek zorundadır.
- 2-Aracın kapsamı geçerliği etkiler. Hem çok dar hem de çok geniş bir kapsam geçerliği düşürür.
- 3-Test güvenilirliğini artıran tüm etmenler geçerliği artırır.
- 4-Cevaplamaya verilen süre (cevaplanamayan soruların kapsamı daraltmasından dolayı) geçerliği etkiler.
- 5-Puanlama yanlışlıkları geçerliğin etkiler.
- 6-Ölçüt puanların geçerlik ve güvenilirlikleri geçerliği etkiler.
- 7-Ölçme aracının gücü ve grubun yetenek düzeyi geçerliği etkiler. Örneğin okul birincisi, ikincisi seçilecekse zor bir sınav yapılabilir. Kolay bir sınav yapılırsa araç amaca hizmet etmeyeceğinden geçerlik düşük olacaktır.
- 8-Soruların yıllarca değiştirilmeden kullanılması geçerliği düşürür.
- 9-Soruların sınavdan önce açıklanması geçerliği düşürür.
- 10-Kopya çekilmesi geçerliği düşürür.
- 11-Testin kullanım sayısı geçerliği etkiler. Her kullanımın geçerliği bir miktar düşürdüğü söylenebilir.
- 12-Testin kullanım amacı ve uygulama gurubu geçerliği etkiler.

#### **Güvenirlik Ve Geçerlik Arasındaki Temel İlişkiler**

1- Bir ölçme aracının geçerli olabilmesinin ön koşulu güvenilir olmasıdır. Başka bir ifadeyle geçerliğin ön koşulu güvenirliliktir. Güvenilir olmayan bir ölçme aracı geçerli olamaz

2-Bir test güvenilirliği sağlıyorsa, geçerli olabilmenin ilk şartını yerine getirmiş demektir. Buna karşın bir testin güvenilirliğinin olması aynı zamanda geçerli olduğu anlamına gelmez. Yani bir testin güvenilirliğinin yüksek olması, geçerli bir test elde edebilmek için yeterli değildir. Çünkü güvenilirlik sadece tesadüfi hatalardan etkilenirken, geçerlik sabit ve sistematik hatalardan etkilenmektedir.

3-Ölçme aracının geçerli olması aynı zamanda güvenilir olduğunu belirtir.

4- Test aracının en önemli özelliği geçerli olmasıdır. Buna karşın geçerlik ve güvenilirlik birbirleri yerine tercih edilebilecek iki özellik değildir.

5-Birkaç istisna durum dışında, güvenilirliği artırmaya yönelik çabaların tümü, geçerliği de artırır.

6-Bir aracın ölçüt geçerliği o aracın güvenilirliğinin karekökünden büyük olamaz.

$$Max.r_{xy} = \sqrt{r_{xx}} ;$$

Formülde max.  $r_{xy}$  geçerliği,  $r_{xx}$  güvenirliliği göstermektedir..

## **2.5 Kullanışlılık**

Araçların hazırlanış, uygulanış ve puanlanış bakımından ekonomik ve kolay olması aynı zamanda sınırlılıklarının az olması anlamına gelmektedir. Şu alt başlıklar altında incelenebilir:

- 1-Ekonomiklik: Araç, öğretmen, yönetim, öğrenci vb. için ekonomik olmalıdır.
- 2-Hazırlama süresinin azlığı: Aracı geliştirmenin ne kadar zaman aldığı ile ilgilidir. Diğer koşullar eşit olduğunda en kısa zamanda hazırlanan testler tercih edilmelidir.
- 3-Uygulama süresinin azlığı: Aracın kullanılışlılığı, verdiği puanın önemine oranla uygulama gerektirdiği zaman azaldıkça artar.
- 4-Hazırlayıcı ve uygulayıcının niteliklerine bağlı olarak sınırlı olması: Araçların hazırlanması ve uygulanması özel beceriler gerektiriyorsa kullanılışlılığı azalır.
- 5-Cevaplayıcının niteliklerine bağlılık: Cevaplayıcıların niteliklerine uygun araçlar kullanılışlıdır.
- 6-Uygulama kolaylıklarına bağlılık: Sayfa düzeni, baskı kalitesi vb. özellikler uygulamayı kolaylaştırır. Araçları kullanılışlı yapar.
- 7-Puanlama kolaylığına bağlılık: Objektif olarak ve çabuk puanlanan araçlar kullanılışlıdır.
- 8-Yorumlama kolaylığına bağlılık: Elde edilen puanlar anlamlı biçimde yorumlanamaz ise değerlendirme ve karar verme de kullanılamaz. Araç yorumlamaya yardımcı, yorumlamayı kolaylaştırıcı ve yorumda hatayı azaltıcı ise kullanılışlıdır.

# EĞİTİM ÖĞRETİM SÜRECİNDE KULLANILAN ÖLÇME ARAÇLARI

## 3.1 Test Planı (Test Geliştirme)

Bir sınavın amacını belirleme adımından not verme adımına kadar olan süreci kapsayan işlemleri içerir. Aşamalar şunlardır.

*1-Testin amacını belirlemek:* Testle hangi değişkenlerin/özelliklerin ölçüleceği (örneğin matematik yeteneği, matematik başarısı, matematiğe karşı tutum gibi) sonuçların hangi değerlendirme türü için kullanılacağı (öğrenme eksikleri, hazırbulunuşluk düzeyi vb.) ve sonuçlar üzerinde hangi işlemlerin yapılacağını belirleme işlemlerini kapsar.

Eğitimde yapılan süreçlerin birçoğunda öğrenciler hakkında bilgi toplama amacı vardır. Bu amaçla çeşitli yollardan ve araçlardan faydalanılabilir. Ölçme araçları yardımıyla öğrenciler üzerinde yapılan ölçme işlemleri aşağıdaki amaçları kapsayabilir.

- 1) Başarı ya da performansın yordanması
  - 2) Öğrenme eksik veya yanlışlarının ortaya çıkarılması,
  - 3) Öğrenci ihtiyaçlarını belirlemek
  - 4) Öğrencilerin performans ya da başarısını değerlendirmek
  - 5) Öğrencileri yönlendirmek veya onlara rehberlik yapmak
- Öğrenciler arasında seçme yapmak için.
  - Öğrencileri tanıma ve bir programa – alana yerleştirmek, öğrenme düzeyini ve öğrenme eksikliklerini belirlemek için.
  - Öğrencilerin öğrenmelerine ilişkin sonuç belirlemek için yapılıyor olabilir.
- 2.-Ölçülecek davranışların belirlenmesi.
  - 3-Belirtke tablosunun hazırlanması.
- Dersin içeriğinin (işlenecek konuların) hangi hedef – davranış alanlarına göre gerçekleştirilebileceğine bakılarak uygun soru geliştirilir.
  - Sorulacak soruların evrenini temsil gücü yüksek ve geçerli sorular belirlenir.
- 4-.Soru sayısını belirleme.
  - 5.Soru tipini belirleme.
- Kullanılacak soru tipi belirlemede ölçülmesi istenilen hedef davranışların alanı (bilişsel, psikomotor, duyuşsal) ve konunun özelliği dikkate alınır.
- 6.Sınavın güçlük derecesini belirleme.
- Öğrenme düzeyini ve başarıyı ölçmeyi amaçlayan bir sınavın güçlük derecesi “orta düzey” (0,50) olmalıdır. Yani sınav ne çok zor, ne de çok kolay olmalıdır.
- 7.Sınavın süresini belirleme.
- Bir sınavın amacı, soru sayısı, zorluk düzeyi, soru tipi öncelikle ve sınava giren öğrencilerden en zayıf olan öğrencinin durumu dikkate alınarak bir düzenleme yapılır.
- 8-Soruların yazımı.

Puanlamanın nasıl olacağını belirleme

## 3.2 Birey Hakkında Bilgi Toplama Yolları

Bireyler hakkında bilgi toplamanın çeşitli yolları vardır. Bunlar genel olarak ikiye ayrılır.

*1-Test dışı teknikler:* Gözlem, vak’a kayıtları, kontrol listeleri, dereceleme ölçekleri, biyografi, otobiyografi vb. yöntemler.

*2-Test teknikleri:* Öğrencileri yeri ve zamanı önceden belirlenmiş sınama durumları içine sokarak tepkilerini almak amacıyla geliştirilen araçlar kullanarak bilgi toplamaktır. Örnek olarak zekâ, başarı, yetenek vb. testleri verilebilir.

Testlerle yapılacak deęerlendirmelerde, yer ve zaman herkes için aynı olduęundan bireyleri birbiriyle kıyaslamak daha objektif görülebilir. Çünkü test koşulları herkes için aynıdır, eşittir. Farklı deęişkenlerin etkisinin minimuma indięi kabul edilir.

Test uygulaması birey hakkında bilgi toplama yollarından birisidir. Test kelimesi sınama, yoklama anlamındadır. Türkiye’de ise sınav anlamında, özellikle çoktan seçmeli sınav anlamında kullanılır..

Psikolojik testler sınıflandırılırken bazı ölçütler kullanılır. Bu ölçütler çerçevesinde birçok testin hangi kategoriye alınacaęı şüphelidir. Birçok testin sınıflanması binişiklik gösterir. Yani testlerin çoęu aynı zamanda birden çok sınıflamanın içine girebilmektedir. Sınıflamaları tek başına doęru ya da yanlış diye belirlemek mümkün deęildir. İşe yarayıp yaramadıklarına bakılır. Sınıflama belirli ölçütlere göre yapıldıęından çeşitli sınıflamalara rastlanmaktadır. Aşaęıda bazı ölçütlere göre testlerin nasıl sınıflanabileceęine ilişkin kısa bilgiler bulunmaktadır.

## **A – Davranışların Doğası Bakımından**

### **1 – Maksimum Yerlik Testleri**

Bireyin belli bir işi, belli bir durumda, ne derece iyi ya da doęru yapabildięini belirlemede kullanılır. Bireyin zihinsel, duyuşsal, devinişsel ya da akademik yetenekleri hakkında bilgi toplamak amacı ile hazırlanan araçlardır.

**a) Zeka Testleri:** Kişinin takvim yaşı ile zihinsel gelişimi arasındaki uyumu ölçen testlerdir. Çocukları için öğrenme gücünü, yetenekler içinse karşılaşılan zihinsel sorunlarla başa çıkabilme yeteneęini ölçer. Bazı ölçmecilerce 3 grupta incelenir.

**1. Genel Yetenek Testleri;** Belirli düzeylerde genel zihin yeteneklerini ölçer. Doğuştan getirildięine, çevre etkisiyle çok az deęiştiiğine inanılan zihin gücünü ölçer. Genellikle kişinin ileride, kendisine olanak verildiğinde, belirli bir şeyi ne kadar öğrenebileceęini veya yapabileceęini yordamada, öğrenme gücünü (neyi ne kadar öğrenebileceęini) ölçmede kullanılır.

**2. Özel Yetenek Testleri;** Sınırlı ve dar bir alandaki zihin gücünü ölçer. Örneğin, müzik, resim ve sanat dalarında bireyin sahip olduęu özel zihin yeteneklerini ölçen testler bunlar arasındadır.

**3. Farklı yetenek testleri:** Dil yeteneęi, soyut kavrama yeteneęi, mekanik kavrama yeteneęi gibi farklı zihin yeteneklerini ölçen testlerdir.

**b) Başarı Testleri:** Kişinin eğitim süreci içinde veya belirli çevre koşulları altında ne kadar öğrendięini ölçen testlerdir. Bu testler, bireylerin geçmişte ne kadar öğrendięini ortaya çıkarmak için kullanılabilir. Esas olarak bireyin ileride ne kadar öğrenebileceęinin deęil geçmişte ne kadar öğrendięinin belirlenmesinde kullanılır. Üç tür başarı testi yaygın kullanılır.

**1.Eriş i testleri;** amaçlanan davranışlardan hangilerinin ne ölçüde öğrenildięini/kazanıldıęını ölçmeyi amaçlar.

**2. İzleme testleri;** amaçlanan davranıştan her birinin ne ölçüde kazanıldıęını ve hangi düzeyde öğrenme eksik ve güçlükleriyle karşılaşıldıęını belirlemeyi amaçlar.

**3. Hazırbulunuşluk (düzey belirleme) testleri:** Bir ders ya da kursa başlamadan önce öğrenenlerin sahip olması gereken kazanım veya davranışlardan ne kadarına sahip olduklarını belirlemek amacıyla kullanılan testlerdir.

### **2 – Kişilik Testleri**

Yansıtıcı türden ve kişisel yoruma açık olan bu test grubunda amaç; danışanın psikolojik danışma yardımı almasını gerektirecek bir kişilik bozukluęunun olup olmadıęını anlamaktır. Bireyin psikolojik ihtiyaç ve deęerlerini, normal ve normal olmayan kişilik özelliklerini görmemize yardımcı olur.



**a) Projektif kişilik testleri:** Açık olmayan uyarıcılara karşı bireyin gösterdiği kişisel ve sübjektif tepkilere dayanır. Projektif (yansıtma) testlerde resimler, mürekkep lekeleri, tamamlanmamış cümleler kullanılır.

**b) Kişilik envanterleri:** Bireylerin psikolojik ihtiyaç ve değerlerini ölçer. Bunun yanında bireylerin normal ve normal olmayan kişilik yönlerini belirlemeye de yardımcı olur.

**c. Uyum envanterleri:** Bireyin çevresine veya çevresindeki insanlara uyum düzeyini belirlemek üzere geliştirilmiştir.

**d. İlgi envanterleri:** Bireylerin hoşlandıklarını ve hoşlanmadıklarını tanıyarak ileride doyum sağlayacağı alanları meslek olarak seçmelerine yardımcı olmak amacı ile kullanılan ölçme araçlarıdır.

### **B – Puanlama Biçim Bakımından**

#### **1 – Objektif Testler**

Herkesin puanlayabileceği, anahtarı daha önceden belirlenmiş testlerdir. Çoktan seçmeli, doğru-yanlış ve eşleştirmeli testler objektif testler için örnek oluşturabilir.

#### **2 – Subjektif Testler**

Puanın, puanlayıcının öznel kanı ve yargısından etkilendiği testlerdir. Sözlü ve yazılı testler bu türden testlere örnek verilebilir.

#### **1 – Öğretmen Yapımı Testler**

Öğrencilerin veya öğrenenlerin öğrenme düzeyini belirlemek için, sınıf atmosferi gibi faktörleri göz önüne alarak öğretmenlerin kendi sınıf veya okulları için hazırladıkları testlerdir. Sınıf öğretmenin sınıfındaki öğretimin etkililiğini ölçmek ya da artırmak için geliştirdiği testlerdir.

#### **2 – Standart Testler**

Belli bir alanda öğretilenlerin tümünü diğer değişkenlerden arınık olarak ölçmeyi amaçlayan testlerdir. Daha kapsamlı bir kullanım için uzmanlarca hazırlanır. İçerik, puanlama, yönergeler ve yorumlama özgülleşmiştir.

Unutulmamalıdır ki, başka ölçütler kullanarak testleri farklı şekilde gruplamak da mümkündür.

### **D – Uygulama Süresi Bakımından Açısından**

**1 – Hız Testleri:** Verilen süre içerisinde öğrencilerin, kendilerine sorulan sorulardan ne kadarını cevapladığına bakılarak, öğrencilerin soruları cevaplama hızlarının ölçüldüğü testlerdir. Örnek ...Pasa parola yarışması

**2 – Güç Testleri:** Öğrencilerin başarı düzeylerini ölçmek amacıyla, süre sınırlaması yapılmadan uygulanan ve gittikçe zorlaşan sorulardan oluşan testlerdir.

Örnek...Kim 500 Milyar ister yarışması.

### **3.3 Değerlendirmede Kullanılan Ölçme Araçları**

<b><u>Bilişsel</u></b>	<b><u>Duyusal</u></b>	<b><u>Psiko-Motor</u></b>
Yazılı yoklama	İlgi Testi	Derecelendirme Ölçeği
Sözlü yoklama	Tutum Testi	İş-Performans Testi
Çoktan Seçmeli	Gözlem Tekniği	Gözlem Fişi
Kısa Cevaplı		
Eşleştirmeli		
Yetenek Testi		
Hazırbulunuşluk Testi		
Ödev ve Projeler		

## Yazılı Yoklamalar

**Tanım:** Cevaplayıcının sorulan soruların cevaplarını düşünüp hatırlayarak ve hatırladığı cevabı organize edip yazılı olarak sunduğu sınav türüdür. Kâğıt kalem veya essey tipi sınavlar da denir. Yazılı sınavlar kişinin özgün ve yaratıcı düşünce gücünü, yazılı anlatım becerisini, belli konulardaki görüşünü, ilgisini ve tutumunu ölçmede kullanışlıdır. Yazılı sınavı uygulamaları farklı şekillerde yapılabilir.

Bunlar;

**1-Klasik yazılılar:** Tanıma uygun olarak, günümüzde birçok öğretmenin uyguladığı sınav tarzıdır. Az sayıda soru içerir ve öğrenciler tüm soruları cevaplamak zorundadır.

**2-Tercihli sınavlar:** Sınav soruları arasında tercihli sorular varsa bu tür sınavlara tercihli sınavlar denmektedir. Tercihli sınavlar birkaç biçimde uygulanabilir. Örneğin, sorular konulara göre gruplandırılarak her gruptan belirli sayıda soru seçilip cevaplandırılması istenebilir; belirli soruları zorunlu tutarak yanında seçimlik sorular verilerek belirli sayıda da seçimlik sorunun cevaplandırılması istenebilir ya da tüm sorular verilerek bunlardan seçilecek belirli sayıda sorunun cevaplandırılması istenebilir.

**3-Ad çekmeli Yazılı Yoklama :** Öğretmen önceden çeşitli konularla ilgili soruları hazırlayarak bir kutuya koyar ve öğrenciler bu kutuda kura ile bir soru çekerek kendilerine çıkan soruyu cevaplar.

**4-Açık Kitap yazılıları:** Özellikle kullanılacak formül sayısının çok fazla olduğu derslerde öğrencilerin sınav esnasında diledikleri kaynaklardan faydalanmasına olanak sağlayan yazılı yoklamalardır. Bu tip sınavlarda öğretmenlerin hazırladıkları soruların orijinal olmasına dikkat etmeleri gerekmektedir. Aksi halde, herhangi bir kaynakta sorunun aynen bulunması ve öğrencilerinde bu kaynakla sınava girme riski vardır.

Yazılı yoklamalarda kullanılan çeşitli soru tipleri vardır. Bunlar;

**a-Sınırlı cevap soruları (Kısa cevap gerektiren sorular):** Cevapları çoğunlukla sınırlı veya kısa olan sorulardır. Verilecek cevabın niteliğine, uzunluğuna ve örgütlenmesine ilişkin bazı sınırlamalar getirilir (Listeleyin, tanımlayın, nedenlerini sıralayın gibi).

**b-Serbest cevap soruları (Uzun cevap gerektiren sorular veya açık uçlu sorular):** Cevapları sınırlı olmayan, kompozisyon tipi sorulardır. Cevabın içerik, nitelik ve uzunluğunu belirlemede cevaplayıcıya belli ölçülerde özgürlük verilir (Yaratıcı düşünme gücü, bilgiyi örgütleme yeteneği, düşüncelerin iç tutarlılığını ve dış tutarlılığını değerlendirme gücü gibi özelliklerin ölçülmesinde kullanılabilirler). Bu tip sorulardan az sayıda sorulabildiğinden kapsam geçerliğini düşürücü etkileri vardır. Ancak, kompozisyon becerileri, sentez ve değerlendirme basamağındaki davranışlar, yaratıcılık ve özgünlük gibi değişkenler ölçüleceği zaman zorunlu olarak kullanılması gereken sorulardır.

**c-Zorunlu cevaplı sorular:** Bütün cevaplayıcılar tarafından cevap verilmesi zorunlu olan sorulardır.

**d-Seçimlik sorular:** Cevaplandırılması cevaplayıcıların seçimine bırakılan sorulardır. Bazı durumlarda verilen bütün sorular arasından belirlenen kadarının cevaplandırılması istenir. Bazı durumlarda da zorunlu cevaplandırılacak soruların yanında seçimlik sorularda verilerek, zorunlu sorular yanında seçimlik sorulardan belirlenen kadarının seçilip cevaplandırılması istenir.

### Yazılı sınavların özellikleri

1-Öğrenciler cevaplarını belirli bir sürede yazmak zorunda olduklarından çok sayıda soru sorulamaz. Bu nedenle kapsam geçerliği düşüktür. Kapsam geçerliği önemli olduğu durumlarda yazılı sınav uygulamak doğru olmaz.

2-Yazma hızı öğrenciler arasında farklı olduğundan "yazı yazma hızı" gibi bir değişken ölçülecek değişkene karışarak geçerliği düşürebilir.

3- Soru sayısı az ve soruları yazma zamanı kısa olduğundan hazırlanması kolaydır.

4-Puanlama süresi zaman aldığından özellikle çok sayıda cevaplayıcının olduğu durumlarda, puanlayıcının kâğıtları okumaya zamanı azsa veya sonuçların çabuk bildirilmesi gerekiyorsa yazılı sınavların kullanışlılığı düşer.

5-Özellikle serbest cevaplı sorularda kesin doğru, kesin yanlış ayrımı zor olduğundan kâğıdın okunması ve puanlanması zaman alır. Bu durum yansız puanlama yapmayı da zorlaştırır. Aynı zamanda öğrencileri birbiriyle karşılaştırmak da zorlaşır. Bu nedenle güvenilirlik düşer.

6-Öğrenciler cevaplarını düşünüp hatırlayarak yazdıklarından, bazı ölçmecilerce yazılı yoklamaların davranışları daha geçerli ölçtüğü iddia edilmektedir.

7-Puanlama güvenilirliği düşüktür.

8-Analiz, sentez ve değerlendirme gibi üst düzey basamaklardaki davranışların ölçülmesine uygundur.

9-Cevaplamada özgürlük bulunduğundan, cevaplayıcılar cevapları ve soruları isteyerek veya istemeden başka tarafa çekerlerse bu durum puanlama farklılıklarına, dolayısıyla güvenilirliğin düşmesine yol açar.

10-Şans başarısı yoktur. Bu durum güvenilirliği dolayısıyla geçerliği artırır.

11-Soruların güçlüğü, testin geçerliği ve güvenilirliği önceden kestirilemez.

12-Anlatımın duru, yazının güzel, kâğıdın düzenli olması ve kompozisyon yeteneği puanlama yanlışlıklarına neden olarak güvenilirliği düşürür. Aynı zamanda bu değişkenler ölçülecek değişkene karışarak geçerliği de düşürür.

13-Soruların belirginliği ve anlaşılabilirliği azaldıkça öğrenciler tarafından istenilen yere çekilme olasılığı artar.

14-Yazılı sınav soruları şişirme yolla cevaplamaya elverişlidir.

#### **Yazılı Sınav Hazırlarken Dikkat Edilmesi Gereken Noktalar:**

1-Sorular açık, anlaşılır olmalı; cevaplayıcılar farklı şekilde algılanmamalı.

2-Mümkünse veya amaca uygunsa açık uçlu sorular yerine kısa cevap gerektiren sorular sorulmalı.

3-Sorular birbirinden bağımsız olarak cevaplandırılabilir. Her biri belirli bir davranışlar zincirini yoklamalı.

4-Soru ders kitabından veya bilinen bir kaynaktan olduğu gibi alınmamalı (Ezberi yoklar ve geçerliği düşürür).

5-Sorularda ifade ve yazım hataları bulunmamalıdır. Bu hatalar soruların yanlış anlaşılmasına sebep olabilir. "Adamın, eskiden ne yaptığı ile ilgili bir kaygısı yoktu" ile "Adamın eskiden, ne yaptığı ile ilgili bir kaygısı yoktu" farklı anlamlarda cümlelerdir.

6-Gereğinden çok soru yazılmalıdır. Böylece sorular incelenerek düzeltilir ve uygun sorular seçilir.

Yazılı sınavlarının puanlanmasında değişik tekniklerden yararlanılabilir.

#### **Yazılı Sınavlarını Puanlama Teknikleri**

**1-Genel İzlenimle Puanlama:** Puanlayıcının cevaplayıcının kâğıdını baştan sona okunduktan sonra genel izlenimine göre bir puan vermesidir. Özellikle kompozisyon kâğıtları genelde böyle puanlanır. Bu tekniğin puanlama güvenilirliği düşüktür.

**2-Sınıflama:** Cevaplayıcıların kâğıtlarına hızlıca göz atılarak "iyi, orta, zayıf" gibi sınıflara ayrılması ve daha sonra bu sınıflardaki kâğıtların puanlanmasıdır. İstenirse bu sınıflar da kendi içinde yeniden sınıflandırılabilir. Genel izlenime göre puanlama güvenilirliği yüksektir ancak yine de puanlama güvenilirliğinin düşük olduğu söylenebilir.

**3-Sıralama:** Tüm cevaplayıcıların kâğıtları puanlayıcı tarafından birbiriyle karşılaştırılarak en iyiden en zayıfa doğru sıraya konur. Bütün kâğıtlar bu şekilde sıralandıktan sonra en iyi kâğıda en yüksek not, diğerlerine de buna göre azalan puanlar verilir. Puanlar ilk iki yöntemden daha güvenilirdir ancak yeterli puanlama güvenilirliği olduğu söylenemez.

**4-Anahtarla Puanlama:** Cevaplayıcılardan beklenen cevaplar listelenir. Bu listeye cevap anahtarı denir. Sonra bu listedeki cevaplara göre puanlama cetveli oluşturulur. Puanlama

cevvel ve cevap anahtarı kullanılarak yapılan puanlamaya anahtarla puanlama denir. Puanlar diğ er puanlama tekniklerine göre daha güvenilir dir yani daha az hatalıdır.

### **Puanlamada Dikkat Edilecek Noktalar**

- 1-Ölçme konusu olmayan etkenlere puan verilmemelidir (Yazı güzelliğ i, anlatım gücü vb).
- 2-Cevaplayıcıyı tanımaktan ileri gelen puanlama yanlılığ ını önlemek için isimler kapatılmalıdır.
- 3-Bir sorunun iyi cevaplanmış olması diğ er sorulara verilecek puanı etkileyebileceğ inden puanlama soru-soru yapılmalı. Her kâğıtta aynı soru okunmalı ve puanlanmalıdır.
- 4-İyi bir öğrenciden sonra orta bir öğrenci zayıf, zayıf bir öğrenciden sonra ise aksine iyi görünebilir. Bu etkiyi gidermek için kâğıtlar sınıflandırılabilir.
- 5-Mümkünse birden fazla puanlayıcı kullanılmalıdır.
- 6-Yorgunluk ve dikkat eksikliğ i başladığına okuma ve puanlama bırakılmalıdır
- 7-Anahtarla puanlama imkânı yaratılmalıdır.

### **Sözlü Sınavlar**

**Tanım:** Soruların sözlü olarak sorulup, cevaplayıcının cevaplarını düşünüp hatırlayarak sözlü olarak sunduğ u sınav türleridir. Sözlü yoklamalar da denir. Birçok soru tipi kullanılabilir. Uygulanmasında çeşitlilik görünmekte belirli standartlardan uzak sınavlardır. Bu nedenle sınav türlerini sınıflamak zor olmaktadır.

### **Kullanıldığı alanlar:**

- 1-Okul öncesi ve ilköğ retimin ilk yıllarında yararlanılmalıdır.
- 2-Konuşma becerileri, sözlü anlatım gücü gibi özellikleri ölçmede kullanılabilir.
- 3-Müzik alanındaki becerileri ölçmede de kullanılabilir.
- 4-Yabancı dil konuşma becerisini belirlemede kullanılır.

### **Özellikleri:**

- 1-Sorular ve cevaplar sözlü olarak sunulur.
- 2-Cevaplayıcı cevabını düşünüp bulmak zorundadır. Şans başarısı olmaması güvenilirliğ i artırır.
- 3-Cevaplayıcının cevabında belirsiz kalan noktaların deş ilmesi yoluyla bilginin derinlemesine ve genişlemesine ölçülmesi mümkündür.
- 4-Soruların hazırlanması diğ er sınavlara göre daha kolaydır ve daha az zaman alır. Bu nedenlerle kullanış lılığ ı yükselir.
- 5-Sınav yapanla cevaplayıcı arasında diğ er sınavlara göre daha fazla etkileş im vardır. Sınav yapanın sınav anındaki durumu öğrenciyi etkilediğ inden geçerliğ i ve güvenilirliğ i düşüktür.
- 6-Cevaplayıcının cevaplama anındaki durumu (ses tonu, cevaplama tarzı, duruş u vb.) puanlayıcıyı olumlu ya da olumsuz etkileyerek puanlama güvenilirliğ ini düşürür.
- 7-Sözlü sınavlar bireyseldir. Her öğrenciye ayrı soru sorulur. Soruların kapsamı ve güçlükleri farklı olacağından puanların karşılaştırılabilirliğ i, güvenilirliğ i ve geçerliğ i düşer.
- 8-Cevaplama, sorudan hemen sonra yapıldığından öğrencinin düşünme ve cevabını gözden geçirme şansı çoğunlukla yoktur. Bu durum geçerliğ i ve güvenilirliğ i düşürür.
- 9-Az soru sorulduğ undan kapsam geçerliğ i düşüktür.
- 10-Her cevaplayıcıya ayrı soru sorma zorunluluğ u vardır. Bu sebeple kullanış lı değildir.
- 11-Cevapların doğruluk derecesini puanlayıcı belirler. Bu da puanlama güçlüğü ve yanlılığ ını getirir. Bu da güvenilirliğ i düşürür.
- 12-Sınav yapan sevmediklerine kasıtlı davranabilir. Güvenirlik düşer.

### **Aksak Yanları:**

- 1-Bir sorunun ancak bir cevaplayıcıya sorulması; cevap yanlış da olsa başka cevaplayıcıya sorulamaması,
- 2-Az sayıda soru sorulmasının güvenilirliğ i ve geçerliğ i düşürmesi,

- 3-Puanlama güvenilirliğinin genelde düşük olması,
- 4-Özellikle ölçülmediği durumlarda konuşma diline hâkimiyet, sözlü anlatım gücü gibi değişkenlerin ölçülecek değişkene karışarak geçerliği düşürmesi,
- 5-Cevaplayıcıya fazla zaman verilmemesi,
- 6-Soruyu soranla cevaplayıcının etkileşim içinde olması,
- 7-Heyecan-sıkılma-korku gibi etmenlerle soruyu soranın ses tonu-jest ve mimikleri veya cevaplayıcıya karşı takındığı tavrın geçerliği düşürmesi olarak sayılabilir

### **Doğru Yanlış Testler**

**Tanım:** Cevaplayıcılardan soruları / ifadeleri / maddeleri belli ölçüt(ler)e göre sınıflamasının istendiği sınav türleridir. En çok uygulananı (bilineni) doğru yanlış testleridir. Bu testlerde sorular doğru-yanlış önermeler olarak verilir. Cevaplayıcılar önermeleri/ifadeleri doğru-yanlış olarak sınıflar. İki seçenekli çoktan seçmeli testlere benzerler.



**Madde Tipleri:** 1-Doğru-yanlış önermelerden oluşan madde tipleri. 2-Önermeyi yanlış yapan ifadeyi işaretleten madde tipleri. 3-Önermelerin yanlış kısmını düzelttiren madde tipleri (Yanlış önermelerin doğrusunu yazması istenir). 4-İkiden fazla sınıflama gerektiren madde tipleri. Aslında bu madde tipleri doğru-yanlış testlerindeki şans başarısının azaltmak amacıyla geliştirilmiştir. Bu nedenle madde tipleri olarak almak yerine doğru-yanlış testlerinde şans başarısını azaltmada alınacak önlemler olarak kabul edilebilir.

### **Özellikleri:**

- 1- Cevaplar bir harf veya işaretle verildiğinden puanlama hem kolaydır hem objektiftir.
- 2-Cevaplama zamanı hemen hemen ifadeleri okumaya ayrıldığından (yani yazmaya ayrılan zaman az olduğundan) çok sayıda soru sorulabilir. Kapsam geçerliği yüksektir.
- 3-Madde yapısı basit olmasına rağmen uygulama basamağına kadar çeşitli düzeylerde davranışları ölçebildiği söylenir.
- 4-Sınavda cevaplama işlemi, ölçülen bilgi veya yeteneğin dışında bir zorluk getirmez. Bu nedenle ilköğretimin alt basamaklarında da uygulanabilir.
- 5-Şans başarısı vardır ve diğer seçmeli testlere göre en yüksek şans başarısına sahiptir
- 6-Öğrenme eksiklerini belirleme ve tanı koymaya elverişli değildir (Örneğin yanlış olarak verilmiş bir ifadeyi yanlış diye işaretleyen bir öğrencinin ifadenin doğrusunu bilip bilmediğinin kanıtı yoktur).
- 7-Bazı maddelerin yanlış olarak verilmesi test yapısı için zorunlu fakat eğitim açısından sakıncalıdır. Eğitimde yanlış cümleleri vermek ve onu vurgulamak öğrenme açısından sakıncalıdır. Önemli olan doğru olanı vurgulamak ve onun öğrenilmesini sağlamaktır.
- 8-İfade mutlaka doğru ya da yanlış olmalıdır. Oysa bazı özelliklerin ölçülmesinde kesin doğru ya da kesin yanlış yargılar geliştirmek zordur. Bu güçlük doğru yanlış testi uygulayanları sadece bilgi basamağındaki davranışları ölçmeye yöneltir.
- 9-Sorulan hazırlamak ve puanlamak kolaydır.
- 10-Yazı güzelliği, düzen gibi değişkenler işe karışmadığından geçerlik yüksektir.

### **Soru Yazarken Dikkat Edilecek Noktalar**

- 1-Her madde bir davranışı yoklamalıdır (her ifadede bir ana fikir bulunmalıdır). Örneğin: "Güvenirlik madde sayısından etkilenmez."
- 2-Her madde kesinlikle doğru ya da yanlış olmalıdır. Cevaplayıcıların ifadenin doğruluğu ve yanlışlığı hakkında karar verebilmesi için gerekli minimum bilgi verilmelidir. İfadelerde (birçok, birkaç, bazen vb.) belgisiz sıfatlardan kaçınılmalıdır.
- 3-Önermenin doğru-yanlış olduğunu açığa vuracak ipuçlarından sakınılmalıdır.
- 4-Bir maddede birden fazla yanlış ya da doğru fikir bulunmamalıdır. Bu durumda madde kolaylaşır. Bunun yerine ifade iki madde haline getirilmelidir.
- 5-Doğru-yanlış ifadelerin uzunlukları birbirine yakın olmalıdır.

6-Doğru-yanlış maddelerin sayıları birbirine yakın olmalıdır.

7-Doğru-yanlış maddeler testlere belirli bir örüntü halinde yerleştirilmemelidir (İki doğru, bir yanlış vb. gibi).

8-Bir maddenin yanlışlığı önemsiz veya aldatıcı bir noktada olmamalıdır.

Örneğin, "Fransız ihtilali 1879 tarihinde gerçekleşmiştir." Bu tip maddeler dikkati ölçtüğünden geçerliği düşürür.

9-Her madde mümkün olduğunca kısa ve açık yazılmalıdır. Gereksiz ayrıntılarla şişirilmemelidir. Ana fikrin doğruluğuna etkisi olmayan, edebi olsun diye açıklığı kaybedilen maddeler iyi madde değildir.

10-Çift olumsuzlama kullanılmamalıdır. İfadenin anlaşılmasını zorlaştırır. Bunlardan kurtulmanın yolu cümleyi olumlu ifade etmektir. Örnek:"Geçerliğin düşüklüğü testin güvenilir olmaması değildir."

11-Doğru bir önermeden yanlış bir önerme elde etmek için cümlelerin fiilinin olumsuz hale getirme yoluna gidilmemelidir. Böyle maddelerde "me, ma" olumsuzluk ekleri gözden kaçabilir. Üstelik böyle maddelerin yanlış olduğu, zorla yanlış hale getirildiği anlaşılabilir.

12-Maddelerdeki ifadeler bilinen bir kaynaktan alınmamalıdır.

13-Kanı ifadeleri bir kaynağa yâda otoriteye dayandırılarak verilmelidir.

14-Birbiriyle ilgili maddeler örgütlenerek sunulmalıdır. Örneğin: "Aşağıda doğru yanlış maddeleri ile ilgili olarak verilen ifadeleri okuyarak karşılıklarına doğru ise 'D', yanlış ise 'Y' harfi koyunuz.

—Puanlama nesneldir.

—Geniş kapsamda soru sormaya elverişlidir.

—Sentez ve değerlendirme basamağında davranış ölçmek için çok kullanılır.)

### **Kısa Cevaplı Testler**

Tanım: Bir kelime, bir rakam, bir cümle veya kısa bir liste ile cevaplandırılabilir sorulardan oluşan, cevaplayıcının cevabı düşünüp hatırlayarak yazılı olarak sunduğu ölçme araçlarına kısa cevaplı testler denir.

Cevapların düşünüp hatırlanması ve yazılması bakımından yazılı sınavlara; puanlama kolaylığı ve objektifliği bakımından çoktan seçmeli testlere benzer.

### Soru Türleri:

1-Soru Kipinde Sorular: Madde soru cümlesi biçimindedir. "Kim, nasıl, niçin, nerede vb." soru kipleri ile biten cümlelerden oluşur.

ÖR: "Bir ölçme aracında aranması gereken en önemli nitelik hangisidir?"

2-Eksik Cümle Tipinde (Boşluk Doldurma) Sorular: Öğrencilerden ifadelerde, cümlelerde veya paragraflarda bırakılan boşluk ya da boşluklara uyan kelimeyi veya cümleyi (kelimeler veya cümleleri, sayıları vb.) yazması beklenen sınav türüdür.

ÖR: "Ölçek türlerinden, .... her işlemi yapmaya uygundur." Cümledeki boşluğa uygun kavramı yazınız.

3-Tanımlamayı Gerektiren Sorular: Bu tip sorularda bir kavramın tanımı, bir tanımı belirten kavram veya bir olgunun belirgin özelliklerinin yazılması istenir.

ÖR: "Random (tesadüfi) hatayı tanımlayınız."

4-Tanımayı Gerektiren Sorular: Bu tip sorularda bir bilginin, kavramın, tanımın, özelliklerin veya işlem basamaklarının tanınması istenir.

ÖR: "Bu çalışmada, testin tanımlanan yapıya uygun, soruların birbiriyle ve test puanlarıyla ilişkilerinin yüksek olup olmadığına bakılmıştır."



Yukarıda yapılan çalışmada çalışılan geçerlik türleri nelerdir?

Özellikler:

- 1-Cevaplayıcı cevabını düşünüp bularak yazmak zorundadır. Şans başarısı olmadığından güvenilirlik yükselir.
- 2-Cevaplayıcı istediği cevabı verme özgürlüğüne sahiptir ancak yazılı sınav sorularına verilen cevaplarda olduğu gibi şişirme yapamaz.
- 3-Uygulama ve analiz düzeyine kadar davranışlar yoklanabilir. Bu nedenle kullanışlıdır.
- 4-Her eğitim düzeyine uygun düşen sınav tipidir. Her yaştaki öğrenciye uygulanabilir. Bu sebeple kullanışlıdır.
- 5-Puanlama göreceli kolaydır ama tamamen objektif (nesnel) olmayabilir. Kullanışlıdır ve kısmen de güvenilirirdir.
- 6-Çok soru sorulabileceğinden kapsam geçerliği yazılı yoklamalardan yüksektir.
- 7-Cevaplar kısa ve belirgin olduğundan puanlama güvenilirliği yazılı yoklamalara göre yüksektir. Bu sebeple güvenilirlik artar.
- 8-Fikirleri kompoze etme, yazı güzelliği, kâğıdın düzeni gibi etkenler ölçmeye daha az karışacağından geçerliği yazılı yoklamalardan yüksektir.
- 9-Soruları hazırlama ve puanlama kolaydır. Bu nedenle kullanışlıdır.
- 10-Şans başarısı yoktur. Bu nedenle güvenilirlik artar.
- 11-Testi tekrar kullanma imkânı ayrı bir cevap kâğıdı kullanılarak yaratılabilir. Bu durumda kullanışlılık artar.
- 12-Sorularda şekil, harita ve grafiklerden faydalanılabilir. Bu sebeple testin kullanışlılığı yükselir.

Soru Yazarken Dikkat Edilecek Noktalar:

- 1-Yoklanmak istenen davranış önemli bir davranış olmalıdır. Kıyıda köşede kalmış bilgiler yoklanmamalıdır.
- 2-Maddenin cevaplanabilmesi, bu madde ile yoklanan davranışın beklenen derecede öğrenilmiş olmasını gerektirmeli ve sadece buna bağlı olmalıdır.
- 3-Cevaplayıcıların cevabı bulabilmesi için gerekli bilgiler eksiksiz verilmiş olmalı, cevaplama için gereksiz ve şaşırtıcı bilgiler bulunmamalıdır.
- 4-Cevaplayıcıların cevabı bulmalarına yardımcı olabilecek ipuçları bulunmamalıdır.
- 5-Maddenin tek ve kesin doğru cevabı olmalıdır. Birden fazla cevap var ise ayrı puanlanmalıdır.
- 6-Madde açık, seçik ve kolay anlaşılır bir ifadeyle sunulmalıdır.
- 7-Eksik cümle (Boşluklu) maddelerinde, eksik olanın ne olduğu açıkça anlaşılmalıdır. Boş bırakılan yer/ler belirsizliğe yol açmamalıdır.
- 8-Maddeyi oluşturan cümle cevaplayıcıların bildiği bir kaynaktan aynen alınmamalıdır.
- 9-Testte yer alan maddeler birbirini cevaplandırmayı kolaylaştırıcı veya zorlaştırıcı olmamalıdır. Maddeler birbirinden bağımsız olmalıdır.
- 10-Her bir madde ile sadece bir davranış yoklanmalıdır.
- 11-Eksik cümle maddelerinde bırakılan boşluklar aynı boyutta olmalıdır.
- 12-Eksik cümle maddelerinde bırakılan boşluk sayısı tercihen bir, en fazla iki olmalıdır.
- 13-Cevap yeri olarak bırakılan boşluklar cevaplamayı ve puanlamayı kolaylaştıracak biçimde düzenlenmelidir.
- 14-Cevap yerleri aynı boyutta, tek tip olmalıdır.

### Çoktan Seçmeli Testler

Tanım: Cevaplayıcının cevabını, (sorunun doğru cevabı olasılığı olan) birtakım seçenekler arasından seçerek vermesini gerektiren sorulardan oluşan testlerdir. Doğru yanlış testleri seçenek sayısı iki olan çoktan seçmeli test olarak kabul edilir. Seçmeli testlerde sorulara madde denir. Madde, testin puanlanabilen en küçük birimidir.

Bir madde, kök ve seçeneklerden oluşur. Kök sorunun sorulduğu ve çözüm için gerekli bilgilerin verildiği kısımdır. Seçenekler ise öğrenci için olası cevaplardır. Doğru seçeneğe doğru cevap, yanlış seçeneklere ise çeldirici adı verilir. Çeldiriciler, madde ile ölçülmek istenen davranışa sahip olmayanların veya doğru cevabı bilmeyenlerin aklını çelmek, onları yanıltmak amacıyla yazılmış, doğru olmayan seçeneklerdir.



#### Madde Türleri:

##### 1)Doğru Cevaba Göre:

a)*Doğru cevabı kesin ve bir tane olan maddeler:* Maddenin seçeneklerinden sadece biri doğru diğerleri kesinlikle yanlış olan maddelerdir. Sıkça kullanılan madde tiplerindendir.

b)*Anahtarlanmış cevabı en doğru olan maddeler:* Seçeneklerin tümü kısmen doğru olan fakat biri diğerlerinden daha doğru olan madde tipidir. Bu tip maddelerle daha ileri düzey davranışlar daha kolay ölçülebilir. Bu tip maddelerde çeldiriciler doğru cevaba yakındır.

c)*Bileşik cevap gerektiren maddeler:* Bir maddenin birden fazla doğru cevabı olduğunda kullanılan madde türüdür. Bu tip maddeler çok kullanışlıdır. Hem ileri düzey davranışların ölçülmesinde hem de seçenek bulma zorlaştığında kullanılabilir.

d)*Doğru cevabı gizli maddeler:* Çoktan seçmeli maddelerin en zayıf yönlerinden biri, kişinin doğru cevabı bilmediği halde verilen seçenekler arasında onu gördüğünde tanınmasıdır. Bazı hatırlamaların bilgiyi akla getirmesi veya davranışın gerektirdiği işlem ve düşünceden farklı bir yolla akla getirmeyi önlemek için maddenin cevabı gizlenir.

##### 2-Madde Köküne Göre:

a)*Kökü soru kipinde olan maddeler:* Madde kökü soru kelimeleri ile ifade edilmektedir.

b)*Kökü eksik cümle tipinde olan maddeler:* Kökte bir harf, kelime, sembol veya cümle yazılmayarak boş bırakılır. Verilmesi gereken cevabın, yani bu boşluğa yazılması gereken cevabın, cevaplayıcının verilenler arasından seçmesi istenir. Bırakılacak boşluğun o soru ile ölçülmek istenen davranışa yoklayıcı nitelikte olması gerekir.

c)*Kökü olumsuz maddeler:* Madde kökünün olumsuz olarak ifade edildiği madde türüdür. Olumsuz madde bir sorunun birkaç doğru cevabı bulunduğu durumlarda kullanılır. Olumsuz madde olumlu maddeler içinde yer aldığı olumlu madde gibi görülebilir. Bu nedenle olumsuz ifadenin koyu, altı çizili vb. yazılması yoluyla olumsuzluğa dikkat çekerek, dikkatsizlikten kaynaklanacak hataların önüne geçilir.

##### 3-Maddelerin Gruplanışına Göre:

a)*Ortak köklü maddeler:* Bir paragraf, bir tablo, bir grafik veya harita gibi bir materyalden iki ya da daha fazla sorunun sorulduğu maddelerdir. Bu tip maddeler, okuduğunu anlama, sayılar arasındaki ilişkiler ve bilginin ötelenmesi(transferi)ni içeren davranışların yoklanmasında çok kullanılır. Bu tip maddeleri yazarken dikkate alınacak en önemli noktalar, soruların madde kökündeki bilgiye dayalı olması ve birbirinden bağımsız olması olarak düşünülebilir. Soruların bağımsız olmasının sağlanmadığı sık sık görülür. Bu şans başarısını artıran bir hatadır. Ortak köke dayalı olarak sorulan madde sayısı çok



olmamalıdır. Bu durumda soruların bağımsızlığını sağlamak zorlaşır. Sorulacak soru sayısı konusundan bir kural olmamasına rağmen beşi geçmemesi önerilir.

b) *Ortak seçenekli maddeler*: Bazı hallerde birden çok davranışın yoklanmasında aynı seçeneklerin kullanılması gerekebilir. Böyle durumda seçenekler başa yazılır. Madde köklerindeki sorular bu seçeneklere göre cevaplandırılır. Bu tip maddelere ortak seçenekli maddeler denir. Bu tip maddeler yazma-okuma zamanı ve basılı kâğıt miktarında ekonomi sağlar.-Bu tip maddeleri yazarken seçenek sayısı ile soru sayısının eşit tutulmamasında yarar vardır. Bu durumda sonuncu sorunun cevabı geriye kalan seçenek olacaktır. Genel olarak soru sayısının seçenek sayısından az olması veya bir seçeneğin doğru cevap olarak birden fazla kullanılması önerilir.

### Özellikleri:

1 -Doğru cevap maddenin içinde verilir. Diğer sınav türlerinde böyle bir durum yoktur. Doğru cevabın madde içinde verilmesi, ölçülen davranışa yeterince sahip olmayanların doğru cevabı görünce tanınmalarını sağladığından; bazı uzmanlar, çoktan seçmeli soruların ileri düzeydeki karmaşık davranışları ölçmede yetersiz kaldığını ileri sürerler. Yapılandırılmış olduklarından özellikle yaratıcılık gibi davranışları ölçemezler. Bu durum kullanışlılığı azaltır.

2-Sınav süresinin önemli bir kısmı maddeyi okumaya ve doğru cevabı bulmaya ayrılır. Yazma süresi az olduğundan çok soru sorulabilir. Bu durum kapsam geçerliliğini ve dolaylı olarak güvenilirliği artırır.

3-Yazı güzelliği, kompozisyon yeteneği gibi değişkenler işe karışmadığından geçerliği yüksektir.

4-Puanlama objektif ve kolaydır. Objektif puanlama, bir puanlayıcının cevap kâğıtlarını farklı zamanlarda ve farklı puanlayıcıların da farklı zamanlarda aynı cevap kâğıtlarını okumaları durumunda puanların değişmemesidir. Çoktan seçmeli testlerde cevaplayıcının cevabı ve doğru cevap okuyan kişiye göre değişmez. Bu nedenle puanlama güvenilirliği tamdır.

5-Şans başarısı vardır. Bu durum güvenilirliği düşürür. Ancak seçenek sayısının çokluğu şans başarısını azaltır. Bu nedenle az seçenek yerine çok seçenek yazmak güvenilirliği artırır. Ancak zorlama yazılan seçenekler yazmak bu amaca hizmet edemez.

6-Her ne kadar bazı ölçmeciler aksini iddia etse de çoktan seçmeli testlerle ileri düzey davranışların ölçülmesi zordur. Özellikle sentez basamağındaki davranışlar ölçülemez.

7-Uygulanması kolaydır. Bu nedenle, eğitim basamaklarının çoğunda (basit madde formu kullanmak şartıyla ilköğretim 3. sınıftan itibaren) uygulanabilir. İlköğretimdeki kullanımlarında ayrı cevap kâğıdı kullanılmaması önerilir (Dikkatsizliğin getireceği hataları önlemek amacıyla).

8-Hazırlanması zordur. *Soru* yazma tekniklerinin bilinmesi ve deneyimi gerektirir. Kullanışlık düşer.

9-Geniş gruplara uygulanabilir ve makine ile puanlanabilir. Kullanışlık yükselir.

10-Cevaplama sadece işaret ile yapılmaktadır. Bu uygulama zamanını kısaltır. Uygulama zamanının kısa oluşu cevaplayıcının dikkatinin dağılmadan cevaplamayı tamamlamasına olanak verir.

11-Önceden saptanan özelliklerde sınav yapılması mümkündür. İstatistiksel temellere dayanır. Bu temeller, üzerinde çok çalışılmış, formülleri geliştirilmiş ilkelerdir. Maddelerle testin bütününe birbirine bağlayan istatistiksel eşitlikler kurulmuştur. Bu eşitlikler sayesinde maddeleri önceden deneyerek test puanlarının ortalaması, değişkenliği, güvenilirliği vb. yaklaşık olarak hesaplanabilir.

12-Özelliklerinin çoğu geçerliği, güvenilirliği ve kullanışlılığı artırıcı yöndedir.

13-Çoktan seçmeli testlerde okuduğunun anlama ve okuma hızı gibi bir değişken de işe karışır. Yeterli zaman verilerek bu değişkenlerin etkisi giderilebilir.  
14-Maddeler açık, yalın ve anlaşılır bir dille yazılmalıdır. (İstisnası okuduğunu anlama ve hız testlerinde mümkündür.)

#### **Soru Yazmada Dikkat Edilecek Noktalar**

- 1-Madde bir davranışı yoklamalı ve yoklayacağı davranışa sahip olanların doğru cevabı verebilmeleri için gerekli bilgileri kapsamalıdır.
- 2 Madde kökü olabildiğince açık-seçik ifade edilmiş olmalıdır. Kök hiçbir yoruma meydan vermeyecek açıklıkta yazılmalıdır. Bu durum maddenin güvenilirliğini düşürür. Soruyu zorlaştırmak için kökte belirsizlikler bırakılmamalı veya şaşırtma yoluna gidilmemelidir. Soru zorlaştırılacaksa seçenekler doğru cevaba yakınlaştırılmalıdır.
- 3-Madde kökünde gereksiz ifade ve açıklamalardan (anlaşılabilirliğe katkısı olmayan kelimelerden) kaçınılmalıdır. Madde kökünde ölçülecek davranışı ortaya koyan ifadeler bulunmalıdır.
- 4-Madde kökünde seçeneklerden herhangi birine ipucu veren ifadeler kullanılmamalıdır.
- 5-Seçeneklerde tekrar eden kelimeler madde köküne alınmalı ve tekrar edilmemelidir. Yani seçenekler mümkün olduğunca kısa olmalıdır.
- 6-Her test maddesi diğerlerinden bağımsız cevaplandırılacak şekilde yazılmalıdır.
- 7-Madde noktalama ve yazım kurallarına uygun yazılmalıdır.
- 8-Seçeneklerin uzunlukları birbirine yakın olmalıdır. Seçenekler ifade tarzı ve kapsam bakımından birbirine yakın olmalıdır.
- 9-Çeldiricilerin doğru cevaba yakınlığı maddenin düşünülen güçlüğüne paralel olmalıdır.
- 10-"Yukarıdakilerin hepsi" seçeneği dikkatle kullanılmalıdır. Seçeneklerin birbirine çok yakın olduğu hallerde kullanılabilir. Seçeneklerden birinin yanlışlığının açık olması (bu durumda yukarıdakilerin hepsi seçeneği elenir) veya seçeneklerden birinin doğru olduğunun görülmesi (bu durumda diğer seçenekler elenir) hallerinde güvenilirliği düşürücü etki yapar.
- 11-"Yukarıdakilerin hiçbiri" seçeneği de dikkatli kullanılmalıdır. Bu seçenek doğru cevap olarak kullanılmamalıdır. Doğru cevap olarak kullanılırsa, cevaplayıcının davranışa sahip olup olmadığı anlaşılamaz. Kesin bir doğru cevabı olan maddelerde çeldirici olarak kullanılabilir.
- 12-Birçok madde türünde "Yukarıdakilerin hepsi" ve "yukarıdakilerin hiçbiri" aynı madde içinde seçenek olarak kullanılmamalıdır. Kullanılırsa ikisi birden elenir. Mantıksal olarak ikisinin aynı anda doğru olması mümkün değildir.
- 13-Seçenekler arasında büyüklük sırası varsa bunlar sıralı olarak verilmelidir.
- 14-Seçenek sayısı uygulama yapılan cevaplayıcıların düzeyine uygun olmalı, ilköğretim ikinci sınıfta uygulanırsa (ki önerilmez) iki seçenekli, ilköğretim 3, 4, 5 üç seçenekli, 6, 7, 8 dört seçenekli ve ortaöğretimle yüksek öğretimde beş seçenekli testler kullanılabilir.
- 15-Seçenekler birbirinden bağımsız olmalı, biri diğerini içermemeli, ipucu olmamalıdır. Böyle bir zorunluluk varsa bileşik cevap gerektiren soru tipi kullanılmalıdır.
- 16-Testteki tüm maddelerin seçenek sayıları aynı olmalıdır.
- 17-İyi bir test maddesinin madde kökü okunduğunda seçeneklere bakılmaksızın davranışa sahip olan cevaplayıcı doğru cevabı verebilmelidir.
- 18-Çeldirici bulmada güçlük çekildiğinde madde formu değiştirilmelidir.
- 19-Doğru cevaplar test içine belli bir örüntüde yerleştirilmemelidir. Maddelerin doğru cevaplarının yeri belirli bir düzene bağlı olmamalıdır.
- 20-Doğru cevaplar seçeneklere yaklaşık olarak eşit dağıtılmalıdır.
- 21-Olumsuz maddeler olumlu algılanma riskine karşı, olumsuzluk ifadesinin altını çizme, koyu yazma vb. yoluyla belirginleştirilmelidir.
- 22-Maddeler arasında ve madde kökü ile seçenekler arasında belirli bir boşluk bırakılmalıdır.

- 23-Ortak köklü ve ortak seçenekli maddeler başlangıç ve bitimlerinde diğer sorularla karıştırılmaması için çizgi ile ayrılabilir. Ayrıca bu tip soruların ortak olan kök ve seçenekler ile aynı sayfada olması dikkatin dağılmasını önleyebilir.
- 24-Testin ilk sorularının kolay olması cevaplayıcıyı güdüler.
- 25-Yazı büyüklüğü uygulanan kişilerin düzeyine uygun olmalıdır.
- 26-Aynı seçeneğe art arda çok sayıda doğru cevap getirilmemelidir.
- 27-Maddelerdeki ifadeler bilinen bir kaynaktan olduğu gibi alınmamalıdır.
- 28-Her test maddesi önemli bir davranışı yoklamalıdır.
- 29-Çeldiriciler davranışa sahip olmayanları yanıltmalı fakat sahip olanları yanıltmamalıdır.
- 30-Birbirinin zıddı olan seçeneklerin kullanılmaması gerekir.

#### *Puanlama*

- a) 1-0 puanlaması, b) Maddelerin kolaylık ve zorluğuna göre puan vermek, c) Maddelerin ölçtüğü yeteneğe göre puan vermek, d) Madde seçeneklerine doğruluk derecesine göre ağırlık vererek puanlamak.

#### *Bütün ölçme araçlarında uyulması gereken kurallar*

- 1-Sorular açık ve anlaşılır olmalıdır.
- 2-Sorularda bilimsel hata bulunmamalıdır.
- 3-Sorularda dil ve anlatım bozukluğu olmamalıdır.
- 4-Soruların cevaplandırılması için herkese aynı süre verilmelidir.
- 5-Cevaplamayla ilgili ve gerektiği durumlarda puanlama ile ilgili bir yönerge verilmelidir.
- 6-Sorular birbirinden bağımsız olarak cevaplandırılabilir bir yapıda olmalıdır. Bir sorunun doğru ya da yanlış cevaplandırılması başka bir sorunun doğru ya da yanlış cevaplandırılmasını etkilememelidir.
- 7-Her soru önemli bir davranışı yoklamalı, önemsiz ayrıntılar üzerinde durulmamalıdır.
- 8-Soruların kökünde davranışa sahip olmayan cevaplayıcıları doğru cevaba götürecek ipuçları verilmemelidir.
- 9-Aksi gerekmedikçe her soru bir davranışı yoklamalıdır.

#### **Ödev Ve Projeler**

Tanım: Öğrencilerin soru veya konuları ders dışında kendi çabalarıyla çözmesi / araştırması; bir inceleme yaparak sonucunu rapor halinde yazabilmesi; bir konuyu derinlemesine araştırması; araştırma, bağımsız çalışma, yaratıcı düşünme, farklı kaynaklar kullanabilme, yardım alma becerilerinin geliştirilmesi; uygulama ve üst düzeydeki davranışların geliştirilmesi ve ölçülmesi vb. amacıyla kullanılan araçlardır.

Ödevler-projeler bazı ölçmecilerce bir ölçme yöntemi olarak kabul edilmez. Ancak ödev ve projelerdeki başarı genel başarının önemli bir yönünü yansıtır ve tanımdaki amaçlar eğitimin amaçları arasındadır. Bazı ölçmeciler ise yukarıdaki görüşün aksine ödev ve proje uygulamaları sonucunda sınav kaygısından uzak, gerçeğe daha yakın bir davranış ölçüsünün elde edileceğini savunur. Bu görüşte de gerçek payı vardır. Ancak bu uygulamalar gözetimden uzak yapıldığından puanların geçerliği şüphelidir. Ödev öğrencinin kendisi tarafından yapılmamış olabilir. Bu tür uygulamalar iyi bir ölçme yöntemi olmasa da iyi bir kendi kendine öğrenme yöntemidir. Puan verilmediği zaman yeterli etkiyi gösterememe riskini önlemek için puan verilmeli ama sakıncalarından dolayı nota katılacak ağırlığı düşük olmalıdır.

#### *Türleri:*

- 1-Bireyse ödev veya projeler.
- 2-Grup ödevleri veya projeleri

#### *Uygulama Şekilleri:*

- 1-Herkese aynı konu ya da sorunun verilmesi,
- 2-Kişilere veya gruplara farklı konu ya da soruların verilmesi.

#### Özellikleri

- 1-Tanımda sayılan becerilerin geliştirilmesi ve ölçülmesi amaçlarına uygundur. Bu amaçların ölçülmesi ve geliştirilmesi için faydalı hatta zorunludur.
- 2-Özellikle grup ödevleri görev bilgiyi paylaşma, görev alma, yardımlaşma-dayanışma, sorumluluk alma, eleştiri ve özeleştirme yapma, empatik düşünme, hesap verme vb. davranışların geliştirilmesi ve ölçülmesi için uygundur.
- 3-Ödev-projelerin öğrencilerin kendileri tarafından yapılıp yapılmadığının kontrolü zordur. Bu nedenle önlemler alarak kullanmak, kontrolü sağlamak yararlı olabilir.
- 4-Verilen puanların geçerliği düşüktür.
- 5-Puanlama güvenilirliği düşüktür. Öğretmen kanıları işe diğer sınav türlerinde daha fazla karışmaktadır. Bu yönüyle yazılı sınavlara benzer.
- 6-Öğrencilere verilen konu veya soruların güçlüklerin ayarlamak zordur. Bu nedenle verilen puanlarla öğrencileri karşılaştırmak da zordur.
- 7-Derste öğrenilen-öğretilen davranışların ölçülmesi için uygun bir yöntem değildir.

#### Uygulamada Dikkat Edilecek Noktalar

- 1-Her öğrenciye ya da gruba farklı konu veya soruların verilmesi amaca daha fazla hizmet edebilir.
- 2-Ödev veya projeler öğrencilerin ilgilerine göre dağıtılırsa veya kendilerinin seçmelerine izin verilirse beklenen faydanın sağlanma olasılığı artar.
- 3-Verilen konular öğrencilerin üstesinden kalkabilecekleri bir güçlükte olmalıdır.
- 4-İnceleme/araştırma amacıyla yararlanabilecek muhtemel kaynaklar ve bulunabilecekleri yerler belirtilmelidir.
- 5-Konunun kapsamı açıkça belirtilmeli, sınırları çizilmelidir.
- 6-Puanlama bir anahtara göre yapılmalıdır. Bu nedenle bir anahtar hazırlanmalıdır.

## IV- ÖLÇME SONUÇLARI ÜZERİNDE YAPILAN İSTATİSTİKSEL İŞLEMLER

### Madde Analizi

Madde analizi, bir testteki maddelerin iyi işleyip işlemediğini ortaya koymak amacıyla uygulanan bir tekniktir. Çeşitli yolları vardır. En çok kullanılanı Henrysson ve alt-üst grup yöntemidir. Henryson yöntemi özellikle gözlem sayısının ve madde sayısının çok olduğu durumlarda yorucu olmaktadır. Bu durumda alt – üst grup tekniği ile elde edilen madde analizi sonuçları ile Henryson analizi sonuçlarının birbirine yakınlığı söylenebilir. Bu tekniklerden alt üst grup tekniği ile yapılan madde analizi üzerinde durulacaktır.

Madde analizi yapmadan önce öğrencilerin verdiği cevaplardan doğru olanlar 1, yanlış olanlar 0 olarak puanlanır. Bu puanlamaya dayanarak öğrencinin test puanı bulunur.

Test puanları hesaplanırken cevap kağıtları en yüksek puandan, en düşük puana doğru sıralanarak test puanı en yüksek kısmın % 27'lik kısmı ile en düşük kısmın %27'lik kısmı alınır. (Arada kalan % 46'lık dilim dikkate alınmaz)

Soru No	A	B	C	D	E	Toplam
Üst Grup	n(A <sub>ü</sub> )	n (B <sub>ü</sub> )	n (C <sub>ü</sub> )	n (D <sub>ü</sub> )	n (E <sub>ü</sub> )	N (ü)
Alt Grup	n(A <sub>a</sub> )	n (B <sub>a</sub> )	n (C <sub>a</sub> )	n (D <sub>a</sub> )	n (E <sub>a</sub> )	N (a)
Toplam	n(A)	n (B)	n (C)	n (D)	n (E)	N

### Madde Güçlüğü:

Maddeye doğru cevap verenlerin maddeyi cevaplayan tüm cevaplayıcılara oranı olarak tanımlanır. Orta güçlükteki madde iyi maddedir.

Bir maddenin güçlük indeksi ( $p_j$ ) olarak ifade edilir ve o maddeyi (soruyu) doğru cevaplayan öğrencilerin yüzdesini verir.

Madde güçlük indeksi 0 ile 1 arasında değer alır.

0,00-0,25	0,25-0,39	0,40-0,60	0,60-0,75	0,75-1,00
Çok zor	Zor	Normal Güçlükte	Kolay	Çok Kolay

1'e doğru yaklaştıkça soruyu bilen öğrenci sayısı artar, soru kolaylaşır.

0'a doğru yaklaştıkça soruyu bilen öğrenci sayısı azalır, soru zorlaşır.

**Örnek:** 60 kişilik bir sınıfta Vatandaşlık dersinin 4. sorusuna öğrencilerin vermiş olduğu cevap dağılımı aşağıda verilmiştir. Bu sorunun madde güçlük indeksini ( $P_j$ ) bulunuz?

$$\frac{A}{3} \quad \frac{B^*}{25} \quad \frac{C}{7} \quad \frac{D}{4} \quad \frac{E}{21}$$

**Çözüm:**  $p_j = \frac{25}{60} = 0,42$  (Normal güçlükte bir sorudur.)

Alt üst grup tekniğinde aşağıdaki formülle hesaplanır.

$$p_j = \frac{n(d, \bar{u}) + n(d, a)}{N}$$

$p_j$ : Maddenin güçlük indeksi

$n(d, \bar{u})$ : Maddeyi üst grupta doğru cevaplandıranların sayısı

$n(d, a)$ : Maddeyi alt grupta doğru cevaplandıranların sayısı

$N$ : Üst ve alt gruptaki toplam öğrenci sayısı

**Madde Geçerliliği:** Bir maddenin bilenle bilmeyeni, bir davranışa sahip olanla olmayanı ortaya koyabilme işlevine denir. -1 ile +1 arasında değerler alır. Eğer değeri eksi ise ölçülmek istenen değişkenin zıttı bir değişkeni ölçtüğü, sıfır ile 0,20 arasında geçersiz bir madde olduğu, 0,20 – 0,30 aralığında düzeltilmesi gereken bir madde olduğu ve 0,30’u aşır 1’e yaklaştıkça geçerliliği artan bir madde olduğu söylenebilir.

$$r_{jX} = \frac{n(d, \bar{u}) - n(d, a)}{n}$$

$r_{jX}$  : Madde ayırtıcılık gücü indeksi

$n(d, \bar{u})$ : Maddeyi üst grupta doğru cevaplandıranların sayısı

$n(d, a)$ : Maddeyi alt grupta doğru cevaplandıranların sayısı

$n$ : üst ya da alt gruptan herhangi birisinde yer alan birey sayısı.

Ayırt Edicilik İndeksi	Yorum
0,40 ve üstü	Oldukça ayırt edici bir maddedir
0,30-0,39	Ayırt edici fakat düzeltilmesi gerekir.
0,20-0,29	Kesinlikle düzeltilmesi gerekir.
0,19 ve altı	Testten çıkarılmalıdır.

Bir sorunun düzeltilmesi demek;

\*Bilmeyeni yeterince ayırt edemeyen seçeneklerin güçlendirilmesi(Yani seçeneklerin doğru cevaba yakınlaştırılması)

\*Bileni çeldiren seçeneklerin zayıflatılması , şeklinde gerçekleştirilir.

Bilmeyenlerin de soruları doğru cevaplaması halinde,maddenin ayırt ediciliği negatif bir değere doğru ilerler.

### **ÖRNEK**

Aşağıda 30 soruluk bir test maddesine cevap veren gruptan 200 öğrencinin cevap dağılımı görülmektedir. Bu dağılıma göre 20. sorunun güçlük ve ayırt edicilik değerlerinin hesaplayalım.

Soru no: 20	A	B*	C	D	E	Boş	Toplam
Üst grup	5	70	5	10	5	5	100
Alt grup	20	30	15	20	10	5	100
Toplam	25	100	20	30	15	10	200

### **Madde Güçlülüğünün Hesabı**

Madde güçlük hesabında hem alt grupta maddeye doğru cevap verenleri hem de üst grupta maddeye doğru cevap verenleri dikkate alıyoruz. İki gruptan doğru cevap verenlerin sayısını tüm cevaplayıcı sayısına oranlıyoruz. Böylece maddenin gruba zor mu yoksa kolay mı geldiğini anlıyoruz.

$$p_{20} = \frac{n(d, \ddot{u}) + n(d, a)}{N} = \frac{(70 + 30)}{200} = 0,50$$

Elde edilen madde güçlük değerine göre 20. madde orta güçlükte bir maddedir. Grubun hemen yarısı civarında öğrenci doğru cevap vermiş, diğer yarısı ise doğru cevap verememiştir.

### **Madde Geçerliği Hesabı**

Maddenin bilenle bilmeyeni ayırt etmesi olarak tanımlanmıştır. Bunun için maddeye üst gruptan doğru cevap verenlerle alt gruptan doğru cevap verenler arasındaki fark büyüdükçe madde geçerli hale gelecektir. Aksi durumda madde geçerli olmayacaktır.

$$r_{jX} = \frac{n(d, \ddot{u}) - n(d, a)}{n} = \frac{70 - 30}{100} = 0,40$$

Elde edilen madde geçerlik değerine göre maddenin yeterince geçerli olduğu, maddenin sorunun cevabını bilenle bilmeyeni ayırabildiğini söyleyebiliriz. Bu anlamda madde bilenle bilmeyeni ayırabilmektedir.

Aynı zamanda iki hesaptaki sonuçlar birbirini desteklemektedir. Madde orta güçlükte olduğu zaman genellikle maddenin ayırıcı olması beklenebilir. 20. madde de bu beklentiye uygun bir sonuç vermiştir.

### **Madde Varyansı ve Standart Sapması:**

İki çıktısı olan dağılımlarda varyans ilgilenile olayın olasılığı ile diğer olasılıkların çarpımıdır. Bir maddede ilgilenilen olayı doğru cevaplandırılması olarak ele alabiliriz. Yanlış seçenekleri işaretlemek ise diğer olayların olasılıklarıdır. Madde varyansı aşağıdaki formülle hesaplanır.

$$s_j^2 = p_j * q_j = p_j * (1 - p_j)$$

$$s_j^2 : \text{Maddenin varyansı}$$

$$p_j : \text{Maddenin güçlük indeksi}$$

Madde varyansı ve standart sapması madde güçlük indeksine bağlıdır. Madde güçlük indeksinin 0 ya da 1'e eşit olduğu durumlarda madde varyansı 0 değerini alır. Madde güçlük indeksi 0.50 olan maddelerin madde varyansları en yüksek değeri alır. Bu sebeple, güçlük indeksi 0.50 değerini alan maddelerin en ayırıcı maddeler olduğu söylenebilir.

### **4.1 Verilerin Düzenlenmesi**

Bireyler ya da gruplar üzerinde yapılan gözlem sonuçları işlenmeden önce kolay yorumlanamazlar. Ölçme sonuçlarını yorumlamak, özetlemek, betimlemek veya karşılaştırma yapmaya uygun hale getirmek için yapılması gereken ilk işlem verileri düzenlemektir. Gözlem veya ölçme sonuçlarına, yani işlenmemiş verilere **ham veri** denir. Üzerinde matematiksel ve istatistiksel hiçbir işlem yapılmayan verilere ham veri adı verilir. Verileri düzenlemede kullanılan işlem basamakları aşağıdaki gibi gösterilebilir.

- 1-Verileri büyükten küçüğe ya da küçükten büyüğe sıraya koymak
- 2-Frekans tablosunu oluşturmak
- 3-Verileri gruplamak
- 4-Verileri uygun grafiklerden biriyle göstermek.

Verileri düzenlemekle ilgili aşamaları örnek üzerinden devam ettirmek faydalı olabilir. Öncelikle veri toplamamız, gözlem yapmamız gerekir. Tablo 11 50 öğrenciye uygulanan 10 soruluk testten elde edilen sonuçları göstermektedir. Bu sonuçlar görüldüğü gibi sadece öğrenciler hakkında bilgi vermekte ve bu bilgileri yorumlamak, uygulama grubu hakkında fikir yürütmek zor görünmektedir. Bunun için öncelikle elimizdeki verileri düzenlememiz gerekir.

**Tablo 3. 50 Öğrencinin 10 soruluk testte aldığı puanlar**

Öğrenci	Puan	Öğrenci	Puan	Öğrenci	Puan	Öğrenci	Puan	Öğrenci	Puan
1	6	11	8	21	6	31	9	41	5
2	6	12	1	22	7	32	5	42	4
3	3	13	5	23	8	33	4	43	9
4	3	14	7	24	10	34	1	44	6
5	4	15	3	25	7	35	7	45	9
6	2	16	10	26	8	36	6	46	9
7	4	17	5	27	7	37	9	47	8
8	6	18	5	28	10	38	2	48	5
9	7	19	2	29	8	39	3	49	4
10	7	20	7	30	8	40	6	50	8

#### *1-Verileri Sıralamak*

Bu adımda elimizdeki verileri büyükten küçüğe veya küçükten büyüğe sıralamak gereklidir. Tablo 4’de 50 öğrenciye ait veriler (tablo 3) sıralanmış olarak görünmektedir.

**Tablo 4. 50 Öğrencinin 10 soruluk testte aldığı sıralanmış puanlar**

Öğrenci	Puan	Öğrenci	Puan	Öğrenci	Puan	Öğrenci	Puan	Öğrenci	Puan
1	1	11	4	21	6	31	7	41	8
2	1	12	4	22	6	32	7	42	8
3	2	13	4	23	6	33	7	43	9
4	2	14	4	24	6	34	7	44	9
5	2	15	5	25	6	35	7	45	9
6	3	16	5	26	6	36	8	46	9
7	3	17	5	27	6	37	8	47	9
8	3	18	5	28	7	38	8	48	10
9	3	19	5	29	7	39	8	49	10
10	4	20	5	30	7	40	8	50	10

Verileri sıralamak bize fazla bir yorum şansı vermemekle beraber sonraki adımlar için büyük kolaylık getirir. Sıralanmış veriler, en büyük ve en küçük değeri hemen görmemizi, verilerin değişim aralığını hesaplamamızı sağlayabilir.

#### *2-Frekans Tablosu Oluşturmak*

Frekans sıklık demektir. İstatistikteki anlamı ise bir gözlemin verinin tekrar sayısı veya bir verinin kaç kez gözlemlendiğini gösteren sayı olarak tanımlanabilir. Kısaca verilerin gözlenme sıklığı da denilebilir. Tablo 5’de 50 öğrenciye ait verilerin frekans tablosu görünmektedir.



**Tablo 5. Frekans Tablosu**

X <sub>i</sub> (Puanlar)	f (frekans)	tf	%	t %
1	2	2	4	4
2	3	5	6	10
3	4	9	8	18
4	5	14	10	28
5	6	20	12	40
6	7	27	14	54
7	8	35	16	70
8	7	42	14	84
9	5	47	10	94
10	3	50	6	100
	50		100	

*Frekans tablosunun özellikleri*

1) Özel bir gruplamadır. Grup aralığı 1 kabul edilir. Yaptığımız işlem 1 den 40 kadar her değerin kaç kez tekrar ettiğini bulmaktır. Ancak tekrarı (frekansı) sıfır olan değerleri yazmayız. Bu nedenle sabit bir aralığı dikkate almadığımızı söylenebilir.

2) En çok tekrar eden değer ya da değerleri gösterir.

3) Tablodan faydalanarak puan dağılımının grafiği çizilebilir. Puanların dağılımı hakkında bilgi verir.

4) Herhangi bir kişinin yeri konusunda bilgi verir. Frekans bir verinin tekrar sayısı hakkında, toplam frekans ise bir puanın (kişinin) yeri hakkında bilgi verir. % bir verinin toplam veri içindeki yüzdesini, yığılmalı yüzde ise %'lik dilimler içindeki yerini (sırasını) verir.

Frekans ve yüzde hacim, tekrar ve alan çağrıştırırken, yığılmalı frekans ve yüzde sıra çağrıştırmaktadır.

Frekans tablosunun ilk sütunu gözlenen puanları (bireylerin sınavdan aldıkları puanları) göstermektedir. Gözlenme olasılığı olduğu halde gözlenmeyen puan bu tabloda gösterilmez. Çünkü gözlenmeyen puanın frekansı sıfır olduğundan yapılacak hesaplamalara önemli bir etkisi olmayacaktır. Örneğin sıfır (0) puan gözlenme olasılığı olan ama gözlenmediği için bu tabloda gösterilmeyen puandır.

Tablonun ikinci sütununu frekanslar oluşturur. Frekans sütunu bir puanın kaç kez gözlendiğini ya da aynı puandan kaç tane olduğunu göstermektedir. Frekans sütununun son satırı toplam gözlem sayısına denk olmalıdır.

Tablonun üçüncü sütunu yığılmalı frekansları göstermektedir. Yığılmalı frekans sütununa toplam frekans sütunu da denmektedir. Yığılmalı frekans bir puanın kendi frekansı ile kendinden küçük puanların frekansları toplamıdır. Yığılmalı frekans sütunu bize puanların sırasını verir. Bunun anlamı yığılmalı frekans bize bireyin grup içindeki yerini söyler.

Örneğin 6 puan alan bir birey sınıfta kaç kişiden başarılı veya başarısızdır sorularının cevabını yığılmalı frekans sütunundan görebiliriz. 6 puan alan bireyin sınıftaki kaç kişiden daha başarılı olduğunu görmek için 6 puandan bir küçük değerin yığılmalı frekansına bakarız. Bu durumda 6 puan alan birey 20 kişiden daha başarılı görünmektedir. 6 puan alan bireyin kaç kişiden daha başarısız olduğunu belirlemek için ise 6 puanın yığılmalı frekansını toplam frekanstan çıkarmak gereklidir. Bu durumda 6 puan alan bireyin yığılmalı frekansı 27, toplam frekans 50 ve  $50 - 27 = 23$  olduğundan, 6 puan alan birey sınıftaki 23 kişiden daha başarısızdır diyebiliriz.

Tablonun dördüncü sütunu yüzdeleri göstermektedir. Yüzde sütununa bağlı frekans sütunu da denmektedir. Bir gözlemin % sütunundaki değeri o gözlemin frekans sayısının tüm gözlemlere oranını 100 ile çarparak elde edilebilir. Örneğin 3 puanı 4 kez gözlenmiştir, yani frekansı 4'tür. 3 puanının tüm gözlemlere oranı  $4 / 50 = 0,08$ 'dir. Bu değeri 100 ile çarparsak 8 değerine ulaşırız. Yüzde sütunu frekans sütununa benzer, Frekans sütunun verdiği bilgiyi % cinsinden verir.

Tablonun beşinci sütunu yığılmalı ya da toplamlı yüzde sütunudur. Yığılmalı yüzde, bir gözlemin kendi yüzdesi ile kendinden küçük gözlemlerin yüzdeleri toplamından oluşur. Yığılmalı frekans sütununun verdiği bilgiyi yüzde cinsinden verir. Yani bireyin grup içindeki yerini yüzde cinsinden verir. Zaten yığılmalı yüzdelere, yığılmalı frekansların yüzdeleridir.

Örneğin, altı puan alan bireyin sınıfın yüzde kaçından daha başarılı veya başarısız olduğu sorularının cevabı yığılmalı yüzde sütunundadır. 6 puan alan bireyin sınıfın yüzde kaçından daha başarılı olduğunu bulmak için 6 puandan bir küçük değerin yığılmalı yüzdesine bakarız. 6 puandan küçük değer 5 puandır ve yığılmalı yüzde değeri 40'tır. Demek ki 6 puan alan birey sınıfın yüzde 40'ından daha başarılıdır. 6 puan alan bireyin kaç kişiden başarısız olduğunu bulmak için, 6 puan alan bireyin yığılmalı yüzde değerini 100'den çıkarmak gerekir. Bu durumda 6 puan alan bireyin yığılmalı yüzde değeri 54 ve  $100 - 54 = 46$  olduğuna göre, 6 puan alan birey sınıfın yüzde 46'sından daha başarısızdır.

Frekans tablosunda dikkat edilmesi gereken durumlardan biri de, frekansların toplamıyla yığılmalı frekansların son satırı; yüzdelerin toplamıyla yığılmalı yüzdelerin son satırı birbirine eşit veya çok yakın çıkmasıdır. Bazen hesaplanan yüzdeleri yuvarlamadan dolayı bu değerler arasında çok küçük farklılıklar çıkabilir. Bu farklılığın yuvarlamalardan geldiğinden emin olmak gerekir.

Verileri düzenlemenin üçüncü adımı verilerin gruplandırılmasıdır.

### **3-Verileri Gruplama**

Verileri gruplama, verileri eşit aralıklarla düzenleyerek ve bu aralıklara düşen frekansları saptayarak özetleme işlemidir. Verileri gruplamanın temel gerekçesi şöyle açıklanabilir. Veriler çok geniş bir aralıkta (ranjda) dağıldığında frekans tablosundan yorum yapılması zorlaşır. Bu zorluğu gidermenin yollarından biri verileri gruplamaktır. Ayrıca verileri gruplamak istatistiğin gruplarla ilgilenme mantığına da uygundur. Verileri gruplamak üç adımda gerçekleşen bir işlemdir.

1. adım, grup sayısını belirlemek: Keyfidir. Araştırmacı karar verir. Verilere uygun, bilgi kaybını en aza indirecek şekilde saptanmalıdır. En büyük-en küçük veriye ve toplam frekansa bakılarak karar verilir. Tek sayı olması yararlıdır. Eğitimde 5-7 grup oluşturmak yaygındır.

2.adım grup aralık katsayısını hesaplamak: Grup aralıklarının eşit olması sağlayacak bir sayı elde etme işidir. Eşit olmayan aralıklarla da gruplama yapılabilir. Ancak bu durumda veri analizi yapılamaz. Grupların genişliği grup aralık katsayısı denilen sayı ile saptanır. Bu katsayı şöyle hesaplanır:

$$a = (X_{eb} - X_{ek}) / bgs$$

a: grup aralık katsayısı: gereğinden küçük (grup sayısı artacağından gruplama anlamsız olur) ve büyük (bilgi kaybı fazla olacağından) olmamalı.

X<sub>eb</sub>: Gözlenen en büyük değer (en büyük ölçüm-veri)

X<sub>ek</sub>: Gözlenen en küçük değer (en küçük ölçüm-veri)

bgs: belirlenen grup sayısı

3. adım, grupları oluşturmak: Grup sayısı ve grup aralıkları dikkate alınarak grupları oluşturma basamağıdır. Grupları oluştururken aşağıdaki koşulların sağlanmasına dikkat edilmelidir.

### **Grupları oluştururken sağlanması gereken koşullar:**

1.Grup aralıklarının eşit olması gereklidir.: gruplar arası karşılaştırma, grupları sürekli hale getirme, verilerin tamamını temsil edecek istatistiklerin hesaplanabilmesi için gereklidir.

2.Verilerin birden çok gruba girmemesigereklidir. her verinin temsil edilmesinin sağlanması, frekans hesabında hatanın önlenmesi ve verilerin hangi gruba ait olduğunun tartışılmaması için gereklidir. Bir veri yalnız bir grupta olmalı.

3.İlk veri ilk grupta son veri son grupta olmalıdır.

Verileri grupladığımızda, istatistikleri ve analizleri gruplar üzerinden yürüttüğümüz için verilerle ilgili bilgi kaybı olur. Ancak işleri kolaylaştırması bakımından bir miktar bilgi kaybını göze alırız. Kurallara uyulursa bilgi kaybı az olur.

**Tablo 6. Gruplanmış veriler**

Gruplar	Kesikli Grup Aralıkları	Sürekli Grup Aralıkları	Grup Orta Noktaları ( $X_0$ )	Frekanslar (f)	Toplamlı Frekanslar ( $t_f$ )	Yüzde (%)	Toplamlı Yüzde ( $t_{\%}$ )
1. Grup	9 – 11	8,5 – 11,5	10	8	50	16	100
2. Grup	6 – 8	5,5 – 8,5	7	22	42	44	84
3. Grup	3 – 5	2,5 – 5,5	4	15	20	30	40
4. Grup	0 – 2	-0,50 – 2,5	1	5	5	10	10

Gözlem sonuçları ve kesikli grup aralıkları doğal sayılarla ifade edilmişlerdir. Kesikli (süreksiz) sayılar üzerinde bölme işlemi yapılamadığından grup aralıkları sürekli hale getirilir. Kesikli grup aralıklarının sürekli grup aralıklarına çevirme, grup aralıklarının başlangıç ve bitim değerlerini yarımsar puan sağa ve sola kaydırılarak yapılır.

Verileri sürekli hale getirirken, verilerin ifade ettiği değişkenin sürekli sayılarla ifade edilip edilmeyeceği dikkate alınmalıdır. Ölçülen değişken yetenek, başarı veya tutum ise sürekli kabul edilebilir. Ama sınıf mevcudu, illerin nüfusları kesiklidir. Veriler sürekli olarak kabul edilmesi mümkün değilse sürekli hale getirme işlemi yapılmamalıdır. Kesikli grup aralıklarının ifade ettiği bilgilerle yetinilmelidir.

Grup orta noktaları( $X_0$ ); grupları temsil amacıyla hesaplanır. Grubun başlangıç değeri ve bitiş değerinin toplamının ikiye bölünmesi ile elde edilir.

$$X_{01}=(0+2)/2= 1 \text{ veya } X_{01}=(-0,50 +2,5)/2= 1 \text{ vb.}$$

#### 4- Verileri grafikte gösterme

Grafik: bir değişkene veya birbiriyle ilgili değişkenlere ait verilerin şekille gösterilmesidir. Ölçme sonuçlarının özetlenmesi ve düzenlenmesinde başvurulan yollardan biridir. Göze hitap ettiklerinden tablolardan daha kolay anlaşılır ve yorumlanır. Veriler farklı ölçeklerden elde edildiğinden farklı grafiklerle gösterilirler. Yani verileri elde ettiğimiz ölçeğe göre grafik kullanılır.

1) Sütun Grafiği: Verilerin dikdörtgenel bölgelerle gösterildiği grafiklerdir. Kesikli değişkenler için kullanılabilecek çubuk grafiği sürekli veriler için kullanılabilecek histogram denir.

a) *Çubuk grafiği*

\*Kesikli değişken için kullanılır.

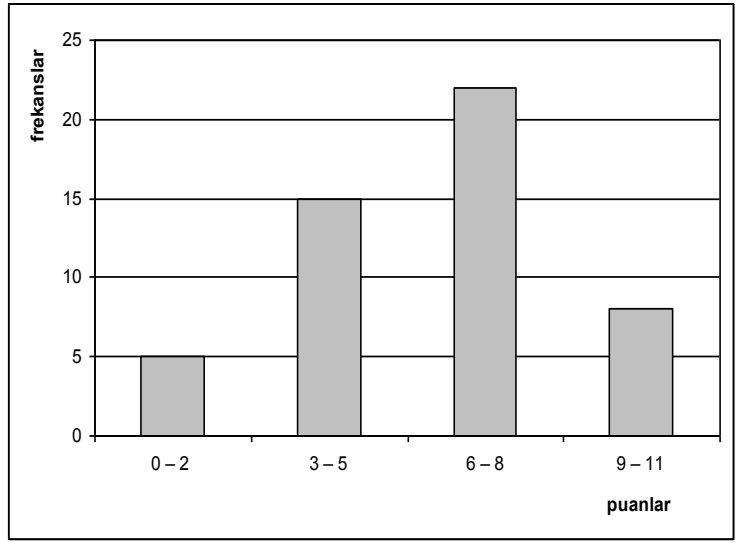
\*Adlandırma (sınıflama veya sıralama ölçekleriyle elde edilmiş veriler) için kullanılır.

\*Frekansların sayılabildiği veya frekansların toplam frekansa oranının hesaplanabildiği durumlar için kullanılır.

\*Değişkenler eksenlere anlamayı kolaylaştıracak biçimde yerleştirilir.

\*Grupların (sınıfların) başlangıç ve bitişleri çıkarılmaz.

Şekil 7. Çubuk Grafiği



\*Frekanslar alanla ifade edilmektedir. Her çubuğun alanı o grubun frekansını, çubukların alan toplamı da toplam frekans verir.

\*Frekans yerine yüzdeler de kullanılabilir.

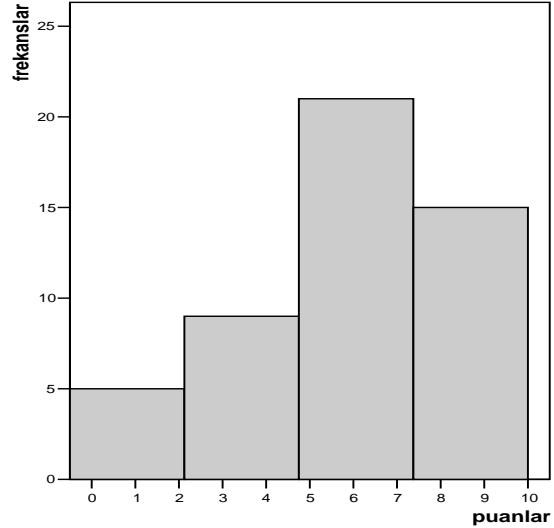
Histogram:

\*Çubuk grafiğe benzer ancak çubuk grafiğinin aksine sürekli grup aralıkları ile çizilir. \*Sürekli olan ya da sürekli hale getirilebilen gruplandırılmış veriler için kullanılır. Başka deyişle sürekli grup aralıkları, sürekli olan ya da sürekli hale getirilebilen değişkenler için kullanılır.

\*Genellikle gruplar yatay, frekanslar-yüzdeler düşey ekseninde gösterilmektedir.

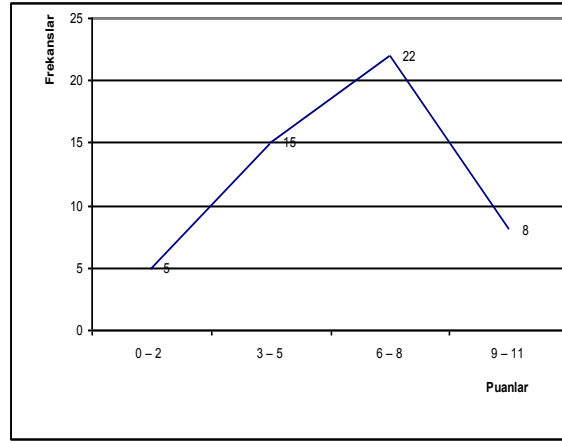
\*Aralıklar sürekli olduğundan grup başlangıç ve bitişlerinde boşluk bırakılmaz.

Şekil 8. Histogram



\*Frekanslar veya yüzdelere alanla temsil edilirler, her çubuğun alanı o grubun frekansını; çubukların alan toplamı toplam frekansı verir.

**Şekil 9 Çizgi grafiği**



**2)Çizgi Grafiği:** Gruplanmış verilerde grup orta noktalarının, gruplandırılmamış verilerde frekans alanlarının uç noktalarının birleştirilmesiyle elde edilen, kırık çizgilerden oluşmuş grafik türüdür. Sadece gruplandırılmış verileri değil, veriler sürekli olmak koşuluyla her türlü frekans dağılımını göstermek amacıyla kullanılabilir.

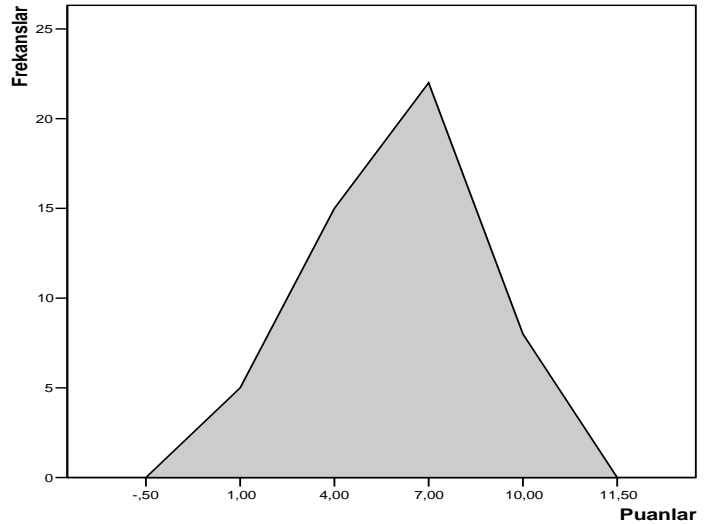
\*Gruplanmış verilerde grup orta noktaları dikkate alınarak çizilir.

\*Frekans ve yüzdelere için çizilebilir.

\*Gruplandırılmış veriler için çizginin altında kalan alan (gruplandırılmamış veriler için doğru parçalarının uzunlukları toplamı) toplam frekansı verir.

\*Yığılmalı frekans tablosunu göstermede de kullanılabilir.

**Şekil 10. Frekans Poligonu**



**3) Frekans Poligonu (çokgeni):** Süreklileştirilmiş verilerde ilk değerle son değer X eksenine birleştirilmesi sonucu elde edilir.

\*Süreklili veriler için çizilir.

\*Yüzde ve frekanslar için çizilebilir.

\*Frekans çok geninin alanı toplam frekansa eşittir.

\*Gözlemlerin dağılımını alanlarla göstermeye yarar.

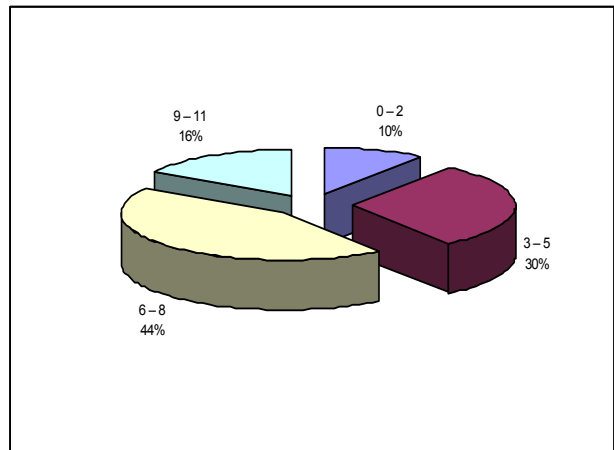
**4)Pasta Grafiği:** Verilere ait frekans veya yüzdelerin dairenin parçalarıyla (açılarla) gösterilmesine dayanan tekniktir.

\*Hem sürekli hem de süreksiz değişkenleri göstermede kullanıldığına rastlanmaktadır.

**Şekil 11 Pasta grafiği**

\*Her bir daire parçası gözlemlerin frekans veya yüzdesini, daire parçalarının toplamı toplam frekansı (yüzdeyi) verir.

Verileri sıralama, gruplandırma ve grafiklerle gösterme veriler hakkında genel bilgiler verir. Bu işlemler verilerin tümü hakkında bir bilgi vermez. Verilerin yorumlanabilmesi için daha ileri düzeyde işlemlere gerek vardır. Verilerin bütününe ait özellikleri belirtecek bilgilere ihtiyaç vardır. Bunlardan biri merkezi yığılma (eğilim) ölçüleri, diğeri de değişme ölçüleridir.



### 4.3 Merkezi Eğilim Ölçüleri

#### MOD (Tepe Değer)

Ölçümlerde en çok tekrar eden değere mod denir. Kaba bir ölçüdür. Ortalamanın ve ortancanın hesap edilmediği durumlarda kullanılır.

- Örnek:1-) 2,3,3,4,7,9 Mod=3  
2-) 2,3,3,4,5,5,7,8 Mod=3 ve 5  
3-) 2,3,3,5,5,7,9 Mod=4  
4-) 2,2,2,2 Modu yoktur.

Özellikleri:

- 1-Gözlenen frekanslar birbirine eşit ise mod olmaz yani dağılım modsuzdur.
- 2-Ardışık iki veya daha çok ölçüm birbirine eşit sayıda ve diğer ölçümlerden daha çok tekrar etmişse bu durumda mod ardışık ölçümlerin orta noktasıdır.
- 3-Ardışık olmayan iki ya da daha çok ölçüm eşit sayıda ve diğer ölçümlerden daha çok tekrar etmiş ise bu durumda dağılım çok modludur.
- 4-Gruplanmış verilerde frekansı en çok olan aralığın orta noktası mod olarak alınır. Bunun dışında üsteki açıklamalar gruplanmış veriler içinde geçerlidir.
- 5-Sadece bir değer diğerlerinden fazla tekrar etmişse dağılım tek modludur.
- 6-Eğer 2 değer diğer değerlerden fazla tekrar etmiş (gözlenmiş) ve eşit sayıda tekrar etmiş ise bu durumda dağılım çift modludur.

Örnek:1-) 

<u>Puan (x)</u>	<u>frekans (f)</u>
60	4
50	7
40	6
30	4
20	5

Bu dağılımın Modu kaçtır?

**Cözüm:** En çok tekrarlanan puan 50 olduğu için (7 kere tekrar etmiş) dağılımın modu 50 dir.

Örnek:2-) 

<u>Puan (x)</u>	<u>frekans (f)</u>
80	4
70	8
60	6
50	4
40	8

Bu dağılımın Modu kaçtır?

**Cözüm:** En çok tekrarlanan puan 70 ve 40 olduğu için (8 kere tekrar etmiş) dağılımın modu 70 ve 40 dur.Dağılım çift modludur.(Bimodal)

Örnek:3-) 

<u>Puan (x)</u>	<u>frekans (f)</u>
90	4
80	8
70	8
60	4
50	5

Bu dağılımın Modu kaçtır?

**Cözüm:** En çok tekrarlanan puan 80 ve 70 (8 kere tekrar etmiş)dir.Oysaki en yüksek puana sahip bu iki değer ardışık olarak sıralandığından dağılımın modu hesaplanırken her iki değer aritmetik ortalamasını alırız. Mod=75

Örnek:4-) Puan (x)      frekans (f)

75-77	4
72-74	9
69-71	8
66-68	7
63-65	<b>5</b>

Bu dağılımın Modu kaçtır?

**Çözüm:** En çok tekrarlanan puan aralığı (75-77) olduğu için (9 kere tekrar etmiş) dağılımın modu (75-77) aralığıdır. Bu aralıkta 75,76,77 sayıları vardır. Aralığın orta noktası baz alınır.

Mod=76

### **MEDYAN (ORTANCA)**

Sıralı verilerde % 50 sıraya gelen puan değerine denir. Dağılımı iki eşit parçaya ayırır. Uç değerlerden etkilenmez. Ortancanın hesap edilmesi için aşağıdaki aşamalı işlemleri yapmak gerekir.

a) Ölçümler büyüklük sırasına konur (Eğer veriler sıralı verilmiş ise bu işlem adımı geçilir).

b)  $n+1/2$  yardımıyla ortancanın dağılımdaki sırası belirlenir.

c) Duruma göre aşağıdaki formüllerden biri kullanılarak ortanca hesaplanır.

1-Veri sayısı tekse (yani ölçüm sayısı 3, 5, 7 gibi tek sayı ise) ortanca  $n+1/2$  formülü ile bulunur.

Örnek: 1, 7, 13, 16, 19, 22, 24, 27, 26 gibi değerler verildiğinde  $(n+1)/2 = (9+1)/2 = 5$  bulunur ki bu değer ortancanın sırasını verecektir. Sağdan ve soldan sayıldığında veri grubu için ortancanın 19 değerine (gözlemine) karşılık geldiği söylenebilir.

2-Veri sayısı çift ise ortancanın sırası belirlendikten sonra  $X_a + (X_b - X_a)/2$  formülü kullanılarak ortancayı hesaplamak doğru olacaktır. Burada  $X_a$  çift sayıda olan gözlemi ikiye ayırdıktan sonra elde edilen ilk grubun son değeri,  $X_b$  ise ikinci grubun ilk değeridir.

Örnek 1: 5, 7, 8, 10, 14, 16, 17, 18 verilerini ikiye bölüyoruz.

$X_a$   $X_b$  olur.

Ortancayı hesaplarsak  $10 + (14-10) / 2 = 12$  elde ederiz.

Dikkatli bakıldığında ortancanın denk geldiği aralığın sağında ve solunda bulunan değerlerin ortalaması alındığında ortancanın bulunabildiği görülecektir.

Örnek:1-) 7,4,6,9,8 medyanı kaçtır?

**Çözüm:** Dağılım ilk başta büyüklük sırasına göre dizilir.

4,6,7,8,9 Tam ortaya denk gelen değer 7 dir. O halde medyan=7

Örnek:2-) Puan (x)      frekans (f)      Yığılmalı frekans(tf)

60	4	26
50	7	22
40	6	15
30	4	9
20	<b>5</b>	<b>5</b>

İlk başta yığılmalı frekanslar toplanır. Sınıf mevcudu 26 kişidir. % 50 si 13'tür. Yığılmalı frekansta 15'in içerisinde 15'i içine alan puan değeri 40'tır

Medyan:40

3-Gruplanmış verilerde aşağıdaki işlemler gerçekleştirilir.

a)Yığılmalı frekans hesap edilir.

b)n/2 bulunur.

c)Yığılmalı frekanslar sütununun bir ucundan başlanarak n/2'nin olduğu gruba kadar frekanslar sayılır.

d)n/2 değerinin rastladığı grup (aralık) bulunur.

e)Aralığın frekansa bölünmesi gerçekleştirilir.

Bu işlemler yapıldıktan sonra elde edilenler aşağıdaki formüle yerleştirilir.

$$Ortanca = L + \left( \frac{\frac{n}{2} - t_{fa}}{f_b} \right) a$$

Formüldeki simgelerin anlamları şöyledir;

L:Ortancanın içine rastladığı aralığın alt sınırı

T<sub>fa</sub>: “ “ “ aralıktan önceki (düşük puanlı gruplardaki) toplamli frekans

F<sub>b</sub>: “ “ “ aralıktaki ölçüm sayısı (gözlenen frekans)

a: aralık katsayısı

Örnek:3-)

Puan (x)	frekans (f)	Yığılmalı Frekans(tf)
75-77	4	40
72-74	9	36
69-71	11	27
66-68	8	16
63-65	8	8

**Cözüm:**

L:Ortancanın içine rastladığı aralığın alt sınırı

= 68,5

T<sub>fa</sub>: “ “ “ aralıktan önceki (düşük puanlı gruplardaki) toplamli frekans =16

F<sub>b</sub>: “ “ “ aralıktaki ölçüm sayısı (gözlenen frekans) =11

a: aralık katsayısı =3

$$Ortanca = 68,5 + \left( \frac{\frac{40}{2} - 16}{11} \right) 3$$

Medyan=69,6

**Ortancanın Özellikleri:**

1-Verileri tam ortadan ikiye ayırır.

2-Gruba eklenecek herhangi bir veriden,anında etkilenir ve değişiklik gösterir.

3-En büyük ve en küçük değerlerin büyüklüklerinden (uç değerlerden) etkilenmez.



### ARİTMETİK ORTALAMA

Ölçümlerin (verilerin) ağırlık merkezidir. Verilerin toplamının veri sayısına bölümü olarak da tanımlanabilir. En çok kullanılan yığılma ölçüsüdür.

Özellikleri:

- 1-Örneklemeden örnekleme en az değişme gösteren dolayısıyla en kararlı yığılma ölçüsüdür.
- 2-Verilerin hemen hepsini kullanır.
- 3-Birçok istatistiksel işlem için uygundur.
- 4-Evren ya da örneklemden hesap edilebilir.
- 5-Verilerin aritmetik ortalamadan farkları 0'a eşittir.
- 6-Aritmetik ortalama alanları dengeleyen bir noktadır.
- 7-Veri sayısı az ise uç değerlerden etkilenir. Bu aritmetik ortalamanın verilerin ağırlık merkezi olmasından kaynaklanır.

**Örnek:1-)**

Puan (x)	frekans (f)
70	5
60	9
50	4
40	7
30	2

Bu dağılımın Aritmetik Ortalaması kaçtır?

**Cözüm:**Aritmetik ortalamayı bulmak için puan ve frekanslarının çarpımına ihtiyaç duyarız.

Puan (x)	frekans (f)	x.f	x.f
70	5	70x5	350
60	9	60x9	540
50	3	50x3	150
40	7	40x7	280
30	6	30x6	180

Kişi Sayısı =30

Toplam=1500

$$\text{Aritmetik Ortalama} = \left( \frac{\sum f \cdot x}{n} \right)$$

$$\text{Aritmetik Ortalama} = \left( \frac{1500}{30} \right) = 30 \quad \text{Aritmetik Ortalama}=30$$

Puan Aralığı Verilen Frekans Tablosundan Aritmetik Ortalama Hesaplama

**Örnek:1-)**

Puan (x)	frekans (f)
71-80	7
61-70	4
51-60	8
41-50	2
31-40	4

Puan (x)	frekans (f)	Yığılmalı Frekans(tf)	Puanların Orta Noktası Xo	Xo.f	Xo.f
71-80	7	25	75	75x7	525
61-70	4	18	65	65x4	260
51-60	8	14	55	55x8	440
41-50	2	6	45	45x2	90
31-40	4	4	35	35x4	140

Kişi Sayısı :25

Toplam:1455

$$\text{Aritmetik Ortalama} = \left( \frac{\sum f \cdot x_o}{n} \right)$$

$$\text{Aritmetik Ortalama} = \left( \frac{1455}{25} \right) = 58,2$$

#### **Aritmetik Ortalamanın Özellikleri:**

- 1-Bir grubun ortalama öğrenme ve başarı düzeyleri hakkında bilgi verir.
- 2-En duyarlı vasat ölçüsüdür.
- 3-Puanların elde edildiği ölçme aracının güçlük düzeyi hakkında bilgi verir.

Not:Eğer tablo sorularında şu özelliklere dikkat çekilmiş ise;

- Grubun **mutlak başarı** düzeyi.
- Grubun **başarı düzeyi**
- Öğrencilerin **öğrenme düzeyinin en düşük** ve **en yüksek** olduğu ders
- Öğrencilerin **ortalama başarı düzeyi**

Tabloda direk aritmetik ortalamanın olduğu sütuna bakılır.

\*Örnek:

Ders	X	ORTANCA	MOD	Sx
Program	34	36	41	4
Öğrenme	48	48	48	5
Gelişim	51	53	56	6
Ölçme	59	58	41	3
Rehberlik	57	58	59	4

\*Mutlak başarı düzeyine göre en başarılı ders hangisidir?

**Çözüm:**Aritmetik ortalamanın en fazla olduğu ders (59) Ölçme'dir

\*Öğrencilerin öğrenme düzeyinin en düşük olduğu ders hangisidir?

**Çözüm:**Aritmetik ortalamanın en az olduğu ders (34) Program'dır

#### **Ortanca, Mod ve Ortalamanın Karşılaştırılması:**

1-Ortalama, Medyan ve moda göre daha fazla bilgiye dayanır. Ortalama hesaplanırken verilerin tümü işleme katılır. Ortancada yarısı, mod da ise biri ya da birkaçı işleme katılmaktadır. Verilerin tümünü dikkate alması ortalamayı daha istikrarlı ve daha güvenilir yapar. Örneğin bir evrenden random olarak seçilecek örneklemelerde ortalama diğerlerine göre daha istikrarlı daha yakın değerler verecektir.

2-Ortalamanın değeri ölçümlerin değerine bağlıdır. Uç değerler ortalamayı etkiler. Ortalama uç değerler doğru çekilir. Dizideki ölçümlerin birinin değeri değişirse ortalamanın değeri de değişir. Medyan ve mod ise uç değerlerden daha az etkilenir. Dizinin tam ortasındaki değer değişmeden ortanca, ölçümlerin dağılışı biçimi (frekansları) değişmeden de modun değeri değişmez.

3-Ortalama diğer matematik işlemler için de uygundur. Diğer yandan ortanca ve mod merkezi eğilim ölçüsü olmanın dışında pek fazla kullanılmaz.

4-Herhangi bir durumda hangi merkezi eğilim ölçüsünün kullanılacağı, incelemeye ya da cevabı aranan soruya bağlıdır. Genel bir kural olarak birden fazla eğilim ölçüsü kullanılabilecekse bunlardan en fazla ölçüm kullanılarak hesaplanan (en fazla bilgi içeren) ve diğer işlemlere de elverişli olan seçilir.

5-Kullanımda tercih sırası ortalama, ortanca ve mod şeklindedir.

6-Uç değerlerin olması durumunda uygun olan ortancadır.

7-Mod ortalama ve ortancanın hesaplanamadığı durumlarda ve en çok tekrarlayan ölçümü belirlemenin önemli olduğu durumlarda kullanılır.

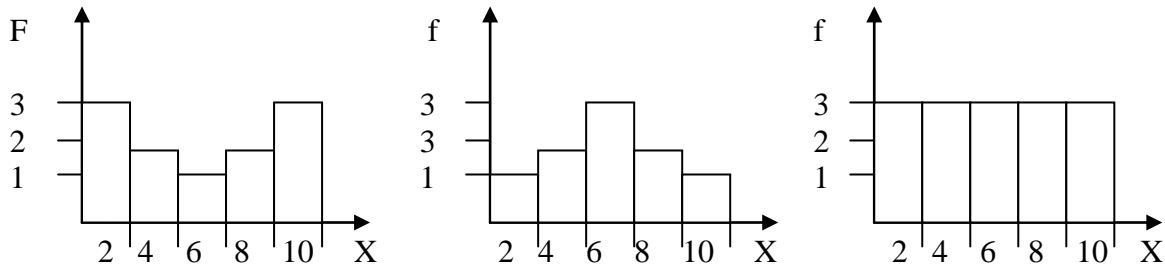
#### 4.4 Dağılım Ölçüleri

Merkezi eğilim ölçüsü bir gruptaki verilerin etrafında toplanma eğilimi gösterdiği bir değerdir. Bu değer veri grubunun bir özelliğini gösterir. Veri grubunu bir yönüyle betimler.

Verilerin hangi değer etrafında yığıldıklarını bilmek önemlidir ancak verileri betimlemeye yetmeyebilir. Çünkü verilerin birbirinden ne kadar uzaklaştıkları, yayılmalarının nasıl olduğu da önemlidir. Ölçme sonuçları birbirine eşit olsaydı, veriler arasında farklılaşma olmazdı. Değişimden söz edilmezdi. Ölçme sonuçlarının yayılması bu değişimin bir sonucudur.

Bir grup verideki değişmelerin ölçüsü “değişme, yayılma veya dağılma” ölçüleri olarak adlandırılır.

Verileri sadece merkezi eğilim ölçüleri ile betimleyemeyiz. Çünkü aynı merkezi eğilim ölçüsü değerine sahip olduğu halde değişik şekillerde dağılan veriler vardır. Aşağıda ortancası ve ortalaması eşit olan 3 veri grubunun dağılımlarının farklılaştığı görülmektedir.



Şekil 13. Ortalamaları ve ortancaları eşit, değişim ölçüleri farklı dağılımlar

Yukarıdaki dağılımlar incelendiğinde ortalamaları ve ortancalarının 3 olduğu görülür. Buna rağmen dağılımları oldukça farklıdır. Merkezi eğilim ölçüleri incelendiğinde benzer grup veya aynı grup denilebilecekken dağılımları incelendiğinde farklı gruplar olduğu görülmektedir. Dolayısıyla verilerin değişim ölçülerini bilmeden onları betimlemek bilgi yetersizliğinden dolayı hata yapmamıza neden olabilir. Sadece merkezi eğilim ölçülerinden hareketle verileri betimlemek onlar arasındaki farkları görmemizi engeller.

#### Ranj

En büyük ölçme (gözlem) sonucu ile en küçük ölçme sonucu arasındaki farka ranj denir.

$$\text{Ranj} = X_{\text{eb}} - X_{\text{ek}}$$

#### **Örnek :**

2,3,7,9,11,13,17

Bu dağılımın Ranjı kaçtır?

**Çözüm:** En büyük değer - En küçük değer

$$17 - 2 = 15$$

**Örnek:2-) Puan (x) frekans (f)**

63	8
52	11
43	6
34	9
25	5

Bu dağılımın Ranjı kaçtır?

**Çözüm:**En büyük değer-En küçük değer

$$63 - 25 = 38$$

**Örnek:3-) Puan (x) frekans (f) Grup Orta No**

79-81	4	80
76-79	9	77
73-75	8	74
70-72	7	71
67-69	5	68

Bu dağılımın Ranjı kaçtır?

**Çözüm:**Grup orta noktaları belirlenir.

En büyük değer-En küçük değer

$$80 - 68 = 12$$

Ranj Değerine bakarak Şu yorumları yapabiliriz.

1-Ranj değeri büyük olan grup heterojen,küçük olan homojendir.

2- Ranj değeri büyük olan grup bilen ile bilmeyeni birbirinden ayırır.Ölçülen özellikler bakımından bireyler farklılık gösterir.

3- Ranj değeri büyük olan grup küçük olana göre daha güvenilirdir.

Sadece veri grubundaki iki değere bağlı olduğundan verilerin dağılımı hakkında fazla bilgi vermez. Çok kullanışlı bir dağılım ölçüsü değildir. Merkeze yığılma ölçüsü olarak modun kullanıldığı veri gruplarında dağılım ölçüsü olarak kullanılır.

**Yüzdelikler ve Çevrekler**

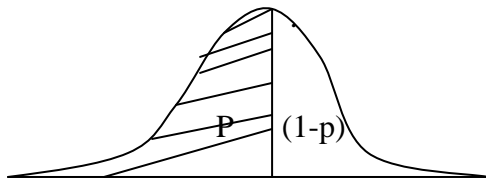
Ortanca sıraya konulmuş verilerin tam ortasını gösteren (bir yer gösteren) değer olarak merkezi yığılma ölçüsüdür. Sıraya konulmuş verilerde tam ortayı gösteren ölçme sonucu yanında başka sıradakilerde dikkate alınabilir. Örneğin verilerin % 25'ini sağında % 75'ini solunda bırakan değere ihtiyaç olabilir. Bir dağılımda bir verinin yerini gösteren ölçüler yüzdeliklerdir. Buna göre ortanca bir yüzdelik ve merkeze yığılma ölçüsüdür ancak diğer yüzdelikler yer bildirmekle birlikte merkezi eğilim ölçüsü değildir.

Yüzdelik: Bir dağılımda belirli bir verinin yerini gösteren ölçülerdir.

P bir yüzde olmak üzere p. Yüzdelik küçükten büyüğe sıralanmış veriler için ölçme sonuçlarının p kadarından büyük, (1-p) kadarından küçük ölçme sonucudur.

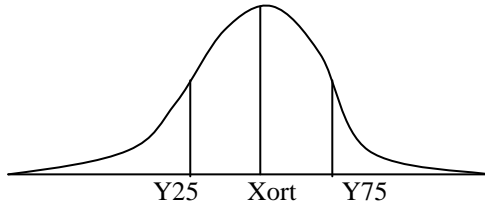
p. yüzdelik ölçme sonuçlarının % p'sini solunda, % (1-p)'sini sağında bırakan değerdir. P. Yüzdeliği  $Y_p$  olarak gösterirsek, ortanca  $Y_{50}$  olarak gösterilebilir.

Bir dağılıma ait p. Yüzdelik aşağıdaki gibi gösterilebilir.



### Çeyrek Kayma

Üçüncü ve birinci çeyreğin farkının yarısıdır.



**Şekil 12. Çeyrek sapmanın  
Grafiksel Gösterimi**

Çeyrek kaymada ölçme sonuçlarından sınırlı sayıda veriye dayandığından çok kullanışlı ve bilgi verici değildir. Ortancanın merkezi eğilim ölçüsü olarak kullanıldığı durumlarda dağılım ölçüsü olarak kullanılabilir.

### Standart Sapma (Kayma)

Veri grubunda merkezi yığılım ölçüsü olarak ortalama kullanılmışsa dağılım ölçüsü olarak kullanılır. Verilerin aritmetik ortalamadan ne kadar uzaklaştığının ortalama bir ölçüsünü verir. Ölçme sonuçlarının aritmetik ortalamadan farklarının karelerinin aritmetik ortalamasının karekökü olarak tanımlanabilir. Yani varyans değerinin kareköküdür.

$$s^2 = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

Standart sapma ölçme sonuçlarının aritmetik ortalamaya göre yayılmalarının bir ölçüsünü verir. Standart kayma (ve varyans) ile dağılımın sivrililiği ve basıklığı arasında bir ilişki vardır. Ölçme sonuçları ortalamaya yaklaştıkça farklar küçülür uzaklaştıkça büyür. Farkların küçülmesi dağılımın sivrileşmesine, büyümesi basıklaşmasına yol açar. Yani dağılım sivrileştikçe standart sapma küçülür, dağılım basıklaştıkça standart sapma büyür.

Standart sapma dağılımın birimi olarak kabul edilir. Bu nedenle varyansın daha çok kullanılır.

Standart sapma da varyans gibi yorumlanır. Standart sapma da ancak bir başka gruba karşılaştırıldığında anlamlı olur. Kendi başına büyük ya da küçük diye yorumlamak mümkün değildir. Standart sapması büyük olan grup heterojen, basık veya değişkenliği fazla olarak; standart sapması küçük olan grup ise homojen, sivri veya benzeşik (değişkenliği küçük-az) olarak yorumlanabilir.

Not:Eğer tablo sorularında şu özelliklere dikkat çekilmiş ise;

- Öğrenciler arası **farklılaşma** en fazla ve en az
  - Öğrencilerin öğrenme düzeyleri benzer ve ya farklı
  - Grubun **homojen** veya **heterojen** olduğu
- Tabloda direk Standart Sapma' nın olduğu sütuna bakılır.

Standart Sapma Küçükse:

- Öğrenciler arası **farklılaşma** en azdır.
- Grup homojendir.
- Öğrencilerin öğrenme düzeyleri benzerdir.

Standart SapmaBüyükse:

- Öğrenciler arası farklılaşma en fazladır.
- Grup heterojendir.
- Öğrencilerin öğrenme düzeyleri farklıdır.

\*Örnek:

Ders	X	ORTANCA	MOD	Sx
Program	39	42	46	5
Öğrenme	40	48	51	3
Gelişim	54	55	56	6
Ölçme	60	54	43	2
Rehberlik	39	47	58	3

En çok hangi derste öğrenciler arası farklılaşma vardır?

**Cözüm:** Standart Sapmanın en fazla olduğu ders (6) Gelişim'dir

Örnek:2-)

Hangi derste daha homojen bir dağılım vardır?

**Cözüm:** Standart Sapmanın en az olduğu ders (2) Ölçme'dir

**Örnek: 3**

Dersler	Aritmetik	Mod	Medyan	Standart sapma	Gülşah'ın notu
Türkçe	68	75	70	5	55
Matematik	70	65	60	10	60
Tarih	80	80	80	7	73
Fizik	65	75	70	4	72
Coğrafya	60	70	65	3	69

1. Hangi derste dağılım normaldir.

- Aritmetik ortalama, mod ve medyanın eşit olduğu dağılımlar normaldir yani simetrikliktir. Tarih dersi'nde dağılım normaldir yani simetrikliktir.

2. Hangi derste **grup homojendir** ya da **farklılaşma en düşük** ya da **öğrencilerin öğrenme düzeyi birbirine ne yakındır**.

- Bir dağılımda grubun homojen olmasını ve farklılaşma düzeyinin en düşük olmasını ve öğrencilerin öğrenme düzeyinin birbirine en yakın olmasını gösteren en önemli gösterge, **standart sapmanın en küçük** olmasıdır. Tablomuzda Coğrafya dersinin standart sapması en küçük derstir. Grubun homojendir, öğrenciler arasındaki farklılaşma en düşüktür ve öğrencilerin öğrenme düzeyi birbirine en yakındır yorumları yapılabilir.

3. Hangi derste **grup heterojendir** ya da **farklılaşma en yüksek** ya da **öğrencilerin öğrenme düzeyi birbirine en uzaktır**.

- Bir dağılımda grubun homojen olmasını ve farklılaşma düzeyinin en düşük olmasını ve öğrencilerin öğrenme düzeyinin birbirine en yakın olmasını gösteren en önemli gösterge, **standart sapmanın en büyük** olmasıdır. Tablomuzda standart sapmanın en büyük olduğu ders Matematik dersidir. Bu ders için; öğrenciler arasındaki farklılaşma en yüksektir ve öğrencilerin öğrenme düzeyi birbirine en uzaktır yorumları yapılabilir.

4. Ortalama öğrenme düzeyi ya da grup başarı düzeyi en yüksek ve en düşük olan dersler.

- Aritmetik ortalamanın en yüksek olduğu ders, ortalama öğrenme düzeyi en yüksek olan dersi verir. Tarih dersi (80), grup başarı düzeyi ve ortalama öğrenme düzeyi en yüksek olan derstir.

- Aritmetik ortalamasının en düşük olduğu ders bizlere ortalama öğrenme düzeyinin en düşük olduğu dersi verir. Coğrafya (60), grup başarı düzeyi ve ortalama öğrenme düzeyi en düşük olan derstir.
- 5. Gülşah'ın en başarılı ve en başarısız olduğu dersler.
- Z formülünden yararlanılmaktadır.  $Z = (X_i - \bar{X}) / S_x$
- Tek tek bütün dersler için Z puanı bulunduğu zaman en başarılı ders Coğrafya; en başarısız ders ise, Türkçe'dir.  $Z_{Coğrafya} = (69 - 60) / 3 = 3$  çıkar.

**Not:**

Aritmetik ortalaması eşit iki gruptan standart sapması daha küçük olan daha başarılı kabul edilir.

Standart sapmanın artması, ölçülen özelliğin ayırt ediciliğinin yüksek olduğunu gösterir. Ayırt ediciliğin artışı doğal olarak **güvenirliliği** artırır.

**Varyans**

Ölçme sonuçlarının aritmetik ortalamadan farklarının karelerinin aritmetik ortalamasına varyans denir. Verilerin ortalamadan uzaklıklarının bir ölçüsüdür. Bir değişim ölçüsüdür. Grubun değişkenliğini diğer gruplarla karşılaştırmada ve yorumlamada kullanılabilir.

Varyansın büyüklüğü arttıkça puanlar arasındaki fark artar. Bu durumda grup heterojendir denir. Varyans azaldıkça puanlar arasındaki farklar da azalır. Bu durumda da grup homojendir denir. Ancak büyüklüğü ve küçüklüğü için nesnel bir sınır yoktur. Varyans büyüdükçe örneğin bir test puanlarını ele aldığımızda testin ayırt edici olduğu söylenebilir. Yani test davranışlara sahip olanlarla olamayanları ayırıyor denir. Varyansın büyük ya da küçük olarak yorumlanabilmesi için nesnel bir sınır olmadığından en doğru işlem bir grubun varyansını başka grup(lar)la karşılaştırmaktır.

Buna göre varyansı büyük olan grup heterojen, basık veya değişkenliği büyük olarak; varyansı küçük olan grup ise homojen, sivri veya benzeşik (değişkenliği küçük-az) olarak yorumlanabilir.

**DAĞILIMLARIN YORUMLANMASI**

**Bağıl Değişim Katsayısı**

Standart kayma ortalamaya bağlı olarak hesaplandığından belirli sınırlılıklar taşır. En önemlisi de ortalamaya göre yanlılık göstermesidir. Bu nedenle her iki dağılımı da içeren iyi bir dağılım ölçüsü olarak bağıl değişim katsayısı hesaplanabilir. Bağıl değişim katsayısı standart sapmanın aritmetik ortalamaya göre değeri olarak tanımlanabilir. Formülü;

$BDK = \text{standart sapma} / \text{ortalama} \times 100$

$v = \frac{s}{\bar{X}} \times 100$  olarak gösterilir.

Bu formülden elde edilen sonuçlar hem birden fazla grubun değişkenliğini karşılaştırmada hem de bir grubun değişkenliğini yorumlama da kullanılabilir.

Bağıl Değişkenlik 25 den BÜYÜKSE	BASIK	HETEROJEN
Bağıl Değişkenlik 20-25 ARASINDAYSA	NORMAL	SİMETRİK
Bağıl Değişkenlik 20 den KÜÇÜKSE	SİVRİ	HOMOJEN

Eğer bir grubun;

- bağıl değişim katsayısı 25’den büyükse ( $V>25$ ) dağılım basık, heterojen, farklı,
- bağıl değişim katsayısı 20 ile 25 arasında ise ( $20<V<25$ ) dağılım normal, simetrik,
- bağıl değişim katsayısı 20’den küçükse ( $V<20$ ) dağılım sivri, homojen, benzeşik denir.

Birçok durumda varyans veya standart sapmaya bakarak grubun homojen mi yoksa heterojen mi olduğunu yorumlamak yanıltıcı sonuçlara götürebilir. Bu nedenle grupları homojenlik heterojenlik bakımlarından karşılaştırmada hesaplamanın mümkün olduğu her durumda bağıl değişim katsayısını kullanılması önerilebilir.

### **Basıklık (Yayvanlık) Katsayısı**

Standart kayma ve varyans büyüdükçe dağılımın sivrileştiğini, büyüdükçe dağılımın basıklaştığını biliyoruz. Ancak bunlar bize basıklıkla ilgili bir ölçü vermez. Bir dağılımın basıklığının (yayvanlığının) veya sivrililiğinin ölçüsü basıklık katsayısı adını alır.

Genel formülü;

$$Bs = BK = \frac{\sum_{i=1}^n (X - \bar{X})^4}{nS^4} - 3$$

Basıklık katsayısı değeri için aşağıdaki durumlar söz konusudur.

- a)
- b)  $BK < 0$  ise dağılım normalden basık, yayvan ve veriler heterojendir.
- c)  $BK = 0$  ise dağılımın basıklığı normal dağılım kadardır.
- d)  $BK > 0$  ise dağılım normalden sivridir ve veriler homojendir.

Aşağıda merkezi eğilim ve değişme ölçülerine bağlı olarak karşılaşılabilecek tipik dağılımlarla ilgili örnekler verilmiştir.

Aşağıdaki dağılımların en basığı, en heterojeni modsuz dağılımdır. Çift modlu dağılımlarda ise, ölçülen değişken yönünden düzeyleri birbirinden çok farklı iki gruba ait sonuçların grafiksel yapısı görülmektedir. Solu çarpık dağılımlarda ortalama düşük değerler alabilmesine rağmen grup genelde başarılı kabul edilir. Tersine, sağa çarpık dağılımlarda ise ortalama yüksek değer almasına rağmen grup genellikle başarısız kabul edilir.

\*Normal dağılımlarda:

- Ortalama =Medyan = mod.
- Dağılımın her iki tarafı birbirine eşit olduğundan dağılım simetriktir.
- Normal dağılımlarda eğri tabanı birleşmez.
- Test orta güçlüktedir.
- Öğrencilerin başarısı orta seviyededir.
- Normal dağılımlarda



%68'i (+1), (-1) standart sapma içinde  
 %95'i (+2), (-2) standart sapma içinde  
 %99'i (+3), (-3) standart sapma içinde  
 alanı içinde yer alır.

**\*Sağa çarpık dağılımlarda**

Aritmetik ORTALAMA	MEDYAN	MOD
Örnek: 85	79	68
Aritmetik Ortalama > MEDYAN > MOD		

- Test zordur.
- Öğrencilerin çoğu ortalamanın altında puanlar almıştır.
- Öğrencilerin başarısı düşüktür.
- Öğrenciler hedef davranışları kazanamamışlardır.
- Öğretim yetersizdir
- Pozitif kayışlıdır.

**\*Sola çarpık dağılımlarda:**

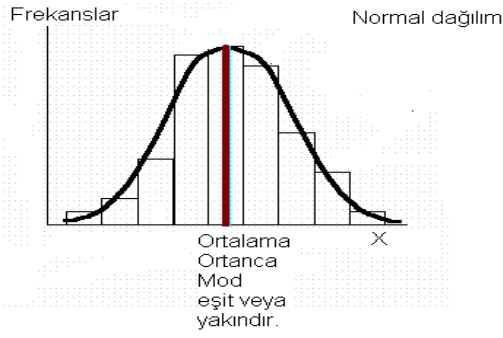
Aritmetik ORTALAMA	MEDYAN	MOD
Örnek:63	68	71
MOD >MEDYAN >Aritmetik Ortalama		

- Test kolaydır.
- Öğrencilerin çoğu ortalamanın üstünde puanlar almıştır.
- Öğrencilerin başarısı yüksektir..
- Öğrenciler hedef davranışları kazanmışlardır.
- Öğretim yeterlidir
- Negatif kayışlıdır.

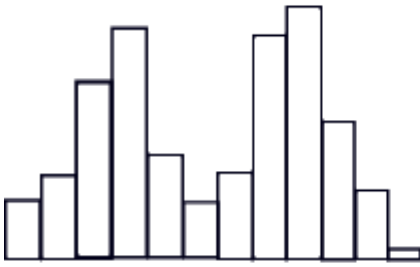
**Örnek:** Hayriye' nin 5 derse ait değerleri aşağıda verilmiştir. Çarpıklık değerlere göre Hayriye'nin en başarılı olduğu ders hangisidir?

Ders	X	ORTANCA	MOD	S <sub>x</sub>
Tarih	79	72	71	4
Öğrenme	58	58	58	5
Vatandaşlık	81	73	66	6
Ölçme	59	58	41	3
Matematik	57	58	59	4

**Çözüm:** Hayriye' nin en başarılı olduğu ders Mod >Medyan >Aritmetik Ortalama olduğu **Matematik**'tir.



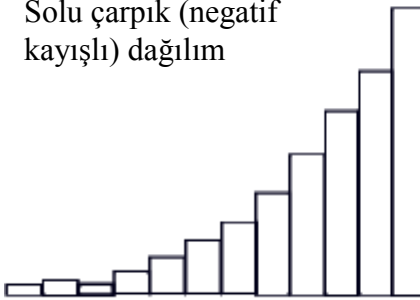
Cift modlu dağılım



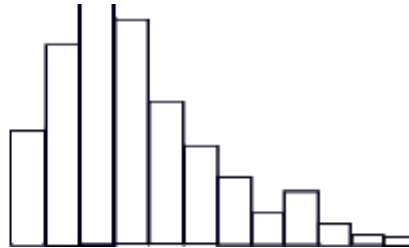
Modszuz dağılım



Solu çarpık (negatif kayışlı) dağılım



Sağı çarpık (pozitif kayışlı) dağılım



Şekil 13. Tipik dağılımlardan örnekler

### **STANDART PUANLAR**

-Ali Ölçme Değerlendirme sınavından 52 aldı.

-İmdat' ın maaşı 1.350 YTL'dir.

-Rahmi Rehberlik sınavından 69,Program sınavından 55 aldı.

Yukarıdaki ifadeler yorumda bulunmak aslında çok da sağlıklı değildir.Çünkü gerçekçi yorumlar için farklı bilgilere ihtiyaç duyulmaktadır.İlk örnekte Ali 'nin aldığı puan düşük,orta ya da yüksek olarak değerlendirmek için sınıf ortalamasını,ikinci örnekte İmdat 'ın aldığı maaşın az,normal ya da çok olarak değerlendirilebilinmesi için çalıştığı yerdeki diğer kişilerin ya da ülke genelindeki diğer insanların aldığı maaşların ortalamasının bilinmesi gerekir.

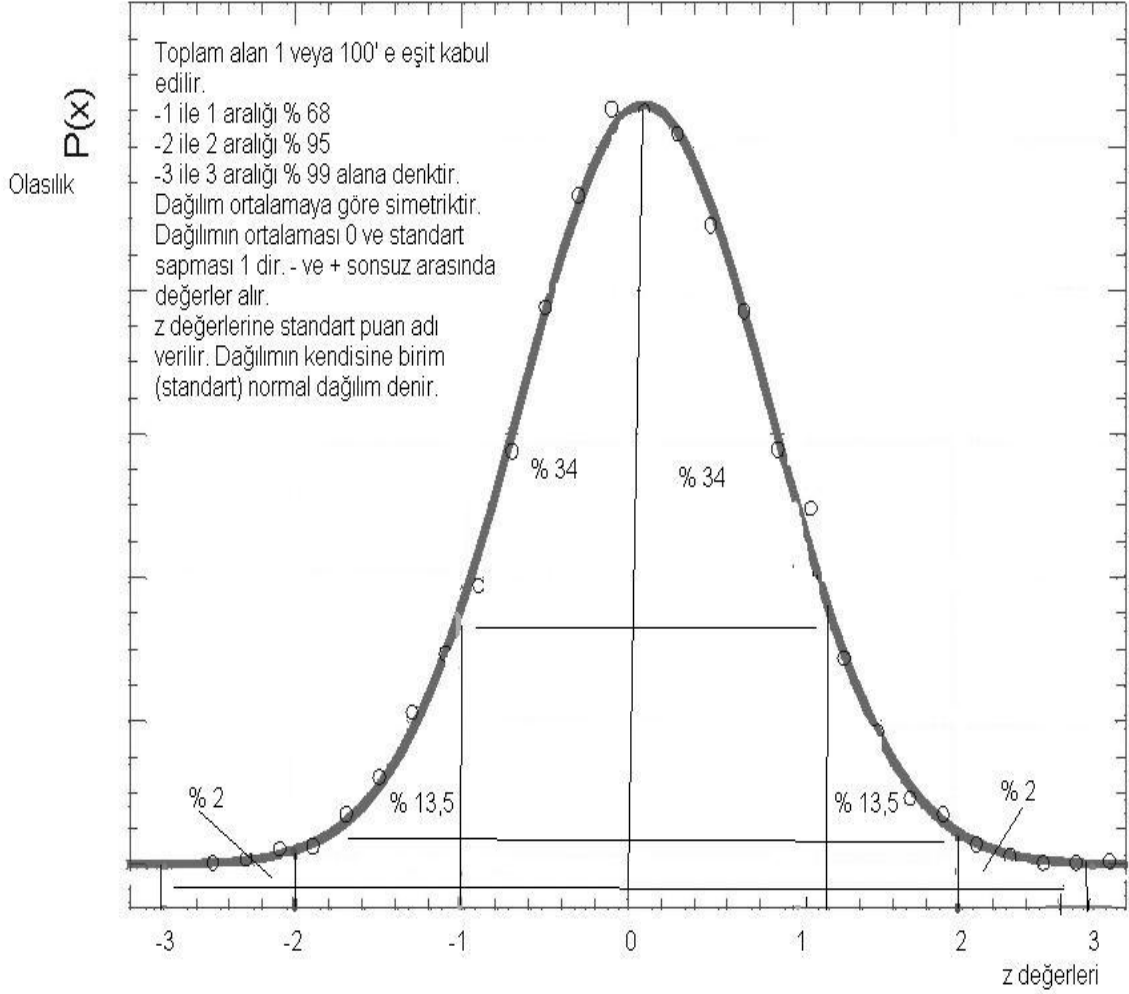
3.örnekte “Rahmi dersinden daha başarılıdır” demek doğru olmayacaktır.Çünkü her iki sınavın da güçlüğü aynı olmayabilir.

Gerçekçi yorumlarda bulunabilmek için bu ve bunun gibi verileri doğrudan kullanmak yerine, başka verilerle karşılaştırabilmek için standart puanlara dönüştürmek gereklidir.

Bu standart puanlara **Z** ve **T** puanları denir

### Z Puanı:

Normal dağılıma sahip bir ölçme kümesi ortalaması 0,00 ve standart kayması 1 olan dönüşüme tabi tutulabilir. Bu dönüştürmeden elde edilen yeni dağılıma Standart Normal Dağılım ya da Birim Normal Dağılım denir. Bu hesaplama yapıldıktan sonra Z puanı yüksek olanın daha yüksek puana sahip olduğu, daha fazla maaş aldığı söylenilebilir. Birim normal dağılım istatistiksel işlemlerde ve çeşitli karşılaştırmalarda kolaylık sağlar. Aşağıda birim normal dağılım grafiği görülmektedir.



### Standart Puan:

Verilerin ortalamadan farklarının standart sapmaya bölümü standart puan veya Z puanı olarak adlandırılır.-

$$Z = \frac{X - \mu}{\sigma} \text{ evrende veya } Z = \frac{X - \bar{X}}{S} \text{ örnekleme de kullanılan formüldür.}$$

**Örnek:**

Dersler	Puan	X	Sx
Coğrafya	70	60	8
Öğrenme	60	50	5
Türkçe	70	60	12
Ölçme	80	70	6
Geometri	70	80	15

Yukarıdaki tabloda Yasemin ‘in beş farklı dersten almış olduğu puanlar, ve bu derslere ait sınıfın aritmetik ortalaması ile standart sapma verilmiştir. Yasemin ‘in derslerdeki başarısını sıralayınız?

**Cözüm:** Derslere ilişkin aritmetik ortalama ve standart sapma hazır olarak verildiği için direkt tablo üzerinde formülde yerine koyalım.

Dersler	Puan	X	Sx	Z	Başarı Sırası
Coğrafya	70	60	8	$(70-60)/8=1,25$	3
Öğrenme	60	50	5	$(60-50)/5=2$	1
Türkçe	70	60	12	$(70-60)/12=0.83$	4
Ölçme	80	70	6	$(80-70)/6=1.66$	2
Geometri	70	80	15	$(70-80)/15=-0,67$	5

(Z puanı daha yüksek olan daha yüksek başarıyı belirtir.)

Yasemin ‘in en başarılı olduğu dersler sırasıyla Öğrenme, Ölçme, Coğrafya, Türkçe ve Geometridir.

**Z Puanının Yorumlanması**

Bu dönüştürmeyle ölçme sonuçları birbiriyle karşılaştırılabilir. Z puanlarının, dağılımların ortalaması 0,00’a getirildiğinden mutlak olmayan, izafi bir başlangıç noktası vardır. Diğer yandan standart puanlar elde edilirken, bir ölçme kümesinde sabit bir değer olan standart sapmaya bölme işlemiyle birim cinsinden ifade edilmiş olurlar. Bu nedenle standart puanların bir birimi vardır ve standart puanlar eşit aralıklı ölçektir. Z puanlarına eşit aralıklı ölçeklerdeki tüm işlemler yapılabilir.

Z puanlarının 2 temel işlevi olduğu söylenebilir. Öncelikle bir bireyin grup içindeki yerini gösterir. Örneğin z puanı 2 olan bir bireyin şekil 17’deki değerlerden hareketle grubun % kaçından daha başarılı veya daha başarısız olduğu belirlenebilir. Buna göre

$Z = 2$  için başarılı olma olasılığı  $= (50+34+13,5) = 97,5$  olur.

Yapılan işlem bireyin z puanının solunda kalan alanı toplamaktır. Birim normal dağılım ortalamaya (0’a) göre simetrik olduğundan eksi sonsuzdan sıfıra kadar olan alan 50 (+) 0 ile 1 arasındaki alan 34 (+) 1 ile 2 arasındaki alan 13,5 = 97,5 olmaktadır. Buna göre birey grubun % 97,5’undan daha başarılı olmuştur. Bu değeri 100’den çıkarırsak bireyin grubun % kaçından başarısız olduğunu bulabiliriz.  $100 - 97,5 = 2,5$ ’tur. Yani birey grubun % 2,5’ğundan başarısız olmuştur.

Bu bilgilerden hareketle özetle:

- Z puanı ne kadar yüksekse başarı o kadar artar.
- Z puanı 0 olan bir öğrenci, sınıftaki öğrencilerin yarısından daha başarılı, yarısından daha başarısızdır. Çünkü Z puanı simetrik bir yapı sergiler. (51 kişilik sınıfta başarı sıralamasında 26.sırada yer alır)
- Z puanı -1 olan bir öğrenci, sınıftaki öğrencilerin yaklaşık %16 ‘sından daha başarılı, % 84 ‘ünden daha başarısızdır.

- Z puanı -2 olan bir öğrenci,sınıftaki öğrencilerin yaklaşık %2 ‘sinden daha başarılı,% 98 ‘inden daha başarısızdır.
- Z puanı +1 olan bir öğrenci,sınıftaki öğrencilerin yaklaşık %84 ‘ünden daha başarılı,% 16 ‘sından daha başarısızdır.
- Z puanı +2 olan bir öğrenci,sınıftaki öğrencilerin yaklaşık %98 ‘inden daha başarılı,% 2 ‘sinden daha başarısızdır.

Z puanının ikinci işlevi farklı gruptaki bireyleri birbiriyle karşılaştırmaktır. Sıralama ölçeğinde elde edilen ham puanlarda farkların görünmemesi nedeniyle karşılaştırma mümkün olamadığından, farklı gruptaki bireyler standart puanlar elde edildikten sonra, yani puanlar eşit aralığa dönüştürüldükten sonra karşılaştırılır. Böyle bir karşılaştırma da z standart puanı büyük olan bireyler başarılı; z standart puanı küçük olan bireyler başarısız kabul edilir. Yani en başarılı birey z standart puanı en büyük olandır.

### **T Puanı:**

Z puanının negatif değerler alabilmesi nedeniyle uygulamada kolaylık sağlaması için T puanı geliştirilmiştir. T puanı , Z puanına dayanarak hesaplanır ve **kesinlikle** eksi değer alamaz.

Z puanı elde edildikten sonra, bu standart puanların ortalaması ve birimi (standart sapması) değişen çeşitli standart puanlara dönüştürülmesi mümkündür. Temel standart puan z puanıdır. Örneğin T standart puanları ortalaması 50, standart sapması 10 olan standart bir puan türüdür. İstenirse bu doğrusal dönüştürme kullanılarak başka standart puanlar da elde edilebilir. Z değerlerinin T değerlerinde dönüşümü şöyledir.

$$T = 50 + (10 \cdot Z) \quad T = \left( \frac{X - \bar{X}}{S} \right) \times 10 + 50$$

T puanı z puanının eksi değerlerinden (eksi değerler bazı durumlarda işlemleri güçleştirdiğinden) kurtulmak için hesaplanan bir standart puandır. Dolayısıyla T puanı da bir standart puandır ve z puanıyla aynı işlevlere sahiptir. Ayrıca z puanı nasıl yorumlanıyorsa aynı şekilde yorumlanır. Yani t puanı en büyük olan birey en başarılı kabul edilir.

**Örnek:** -0,67 olarak hesaplanan Z puanını T puanına çeviriniz.

**Cözüm:**  $T = 50 + (10 \cdot Z_{\text{puanı}})$   
 $= 50 + (10 \cdot (-0,67))$   
 $= 50 + (-6,7)$   
 $= 43,3$

## ***HAM PUANLARIN SINAV NOTUNA ÇEVİRİLMESİ***

Bir öğretmenin, öğrencilerinin sınav kağıtlarını okuyarak puanlaması bir ölçme işlemidir. Bu ölçme işleminden elde edilen puanlara **ham puanlar** denir. Bu puanların sınav notuna çevrilmesinde 2 yol kullanılır:

**a)Bağıl Değerlendirme:**

Ham puanların sınav notuna çevrilmesinde kullanılarak sınav notlarının alt ve üst sınırlarını belirleyecek olan ölçüt, ölçmenin yapıldığı grubun bu sınava ilişkin aritmetik ortalamasına ve standart sapmasına bağlıdır.

Ham puanları sınav notuna çevirmede aşağıdaki tablodan faydalanılır.

Notlar	Alt Sınır	Üst Sınır	İşlem Açıklaması
1	0,00	$X-(0,5.Sx)$	Üst sınırı belirlemek için,standart sapma 0,5 değeri ile çarpılarak elde edilen sonuç, aritmetik ortalama ile çıkarılır.
2	$X-(0,5.Sx)$	$X+(0,5.Sx)$	Üst sınırı belirlemek için,standart sapma 0,5 değeri ile çarpılarak elde edilen sonuç, aritmetik ortalama ile toplanır.
3	$X+(0,5.Sx)$	$X+Sx$	Üst sınırı belirlemek için aritmetik ortalama ile standart sapma toplanır.
4	$X+Sx$	$X+(1,5.Sx)$	Üst sınırı belirlemek için standart sapma 1,5 ile çarpılarak elde edilen sonuç, aritmetik ortalama ile toplanır
5	$X+(1,5.Sx)$	En yüksek puan	Grup içerisindeki en yüksek puan üst sınırı oluşturur.

**Örnek :**

Türkçe dersinden yapılan bir sınavın aritmetik ortalaması 55, standart sapması 15’dir.5’ lik sistemdeki tüm notlar için üst sınırı belirleyiniz.

**Çözüm:**

Notlar	Alt Sınır	Üst Sınır	İşlem Açıklaması
1	0,00	$X-(0,5.Sx)$	$55-(0,5.15) = 47,5$
2	$X-(0,5.Sx)$	$X+(0,5.Sx)$	$55+(0,5.15) = 62,5$
3	$X+(0,5.Sx)$	$X+Sx$	$55 + 15 = 70$
4	$X+Sx$	$X+(1,5.Sx)$	$55+(1,5.15) = 77,5$
5	$X+(1,5.Sx)$	En yüksek puan	Grup içerisindeki en yüksek puan, üst sınırı oluşturur.

**a)Mutlak Değerlendirme:**

Ham puanların sınav notuna çevrilmesinde kullanılacak ölçüt gruba bağımlı değildir.Önceden net olarak belirlenmiştir. Öğrenci notları 100 üzerinden hesaplanır.

Mutlak değerlendirmeye göre hesaplama yapmada aşağıdaki formülden yararlanılır.

$$MBY=(X/K).100$$

MBY=Mutlak Başarı Yüzdesi

X= Öğrencinin Ham Puanı

K = Alınabilecek en yüksek ham puan

**Örnek:** 50 soruluk bir sınavda Hayrettin soruların 15 tanesini doğru olarak cevaplandırmıştır. Hayrettin' in alması gereken sınav notu kaçtır ?

**Çözüm:**

$$MBY=(X/K).100$$

$$= (15/50).100 = 30$$

### Örnek soru ve çözümler

Sağdaki tabloda 5 sınıfa uygulanan 20 soruluk testten elde edilen istatistikler görülmektedir. Aşağıdaki soruları bu tabloya göre cevaplayınız.

sınıflar	$\bar{X}$	s
1	12	3
2	11	3
3	10	4
4	9	2
5	8	5

1-Sınıflarda 15 doğru yapan bireyler karşılaştırılmak istenmektedir. Sınıfların her birinde 15 puan alan bireylerden hangisi diğerlerine göre daha başarılı olmuştur?

2- Sınıflarda 15 puan alan bireylerden hangi sınıftaki birey en başarısızdır?

3-Hangi sınıf en homojendir?

4-Hangi sınıf en heterojendir?

5-3 numaralı sınıfta 15 puan alan bireyin T puanı kaçtır?

6-4 numaralı sınıfta 15 puan alan öğrenci grubun % kaçından daha başarılıdır?

7-En simetrik dağılım gösteren grup hangisidir?

Cevaplar:

İlk iki soru z puanlarının hesaplanmasını gerektirmektedir. Buna göre sınıfların z standart puan formülü kullanılmalıdır.

$$Z_1 = \frac{X - \bar{X}}{S} = \frac{15 - 12}{3} = \frac{3}{3} = 1 \quad Z_2 = \frac{15 - 11}{3} = \frac{4}{3} = 1,33 \quad Z_3 = \frac{15 - 10}{4} = \frac{5}{4} = 1,25$$
$$Z_4 = \frac{15 - 9}{2} = \frac{6}{2} = 3 \quad Z_5 = \frac{15 - 8}{5} = \frac{7}{5} = 1,4$$

Z değerlerini büyükten küçüğe sıraya koyarsak  $Z_1 < Z_3 < Z_2 < Z_5 < Z_4$

1-Hesaplanan z değerlerine göre en büyük z değerine 4. sınıfta 15 puan alan birey sahip olduğundan en başarılı birey 4. sınıftaki öğrencidir.

2. En başarısız birey z değeri en küçük olan 1. sınıftaki öğrencidir.

Üçüncü ve dördüncü sorular için bağıl değişim katsayılarının hesaplanması gereklidir.

$$v_1 = \frac{s}{\bar{X}} \times 100 = \frac{3}{12} \times 100 = 25 \quad v_2 = \frac{3}{11} \times 100 = 27 \quad v_3 = \frac{4}{10} \times 100 = 40$$
$$v_4 = \frac{2}{9} \times 100 = 22,22 \quad v_5 = \frac{5}{8} \times 100 = 62,50$$

3-En homojen grup bağıl değişim katsayısı en küçük olan grup olacağına göre 4 numaralı sınıf en homojen, benzeşik ve sivri gruptur.

4-En heterojen grup bağıl değişim katsayısı en büyük olan grup olacağına göre 5 numaralı grup en heterojen, en farklı ve en basık gruptur.

5- $T_3 = Z_3 \times 10 + 50 = (1,25) \times 10 + 50 = 62,5$  tur.

6-  $Z_4 = 3$  olduğuna göre;

Tablo 17 deki standart normal dağılım alanlarının toplamı =  $50 + 34 + 13,5 + 2 = 99,5$  olacaktır. Bu da 4 numaralı sınıfta 15 puan alan öğrencinin sınıftaki bireylerin % 99,5'undan başarılı olduğunu gösterir. Sınıftaki en başarılı öğrencilerden biri olduğu söylenebilir.

7-Simetriklik elimizdeki verilere göre bağıl değişim katsayısından anlaşılmalıdır. Buna göre bağıl değişim katsayısı 20 ile 25 arasında olan dağılım simetriktir. Buna göre 4 numaralı sınıfın en simetrik dağılım olduğu görülmektedir. 1 numaralı sınıfında belirli bir düzeyde simetrik olduğu söylenebilir.