

**Ölçme ve Değerlendirmeye ilgili olarak sınavda çıkan ve çıkabilecek olan bütün konular eklenmeye çalışılmıştır. Konu başlıkları aşağıda olmakla beraber, bu konuların ders videolarına <https://www.youtube.com/watch?v=qswox4MXgds&list=PLbmkRvdEZ-IS4yFLGXwlodNQJDA1qIeb> adresinden ulaşabilirsiniz.**

**Konu videolarından sonra soru çözümleri eklenmeye çalışılacaktır. Konu ve soru çözümü videolarını kaçırmamak için kanala abone olabilirsiniz ve bildirimlerinizi açabilirsiniz.**

**Her türlü sorularınız için [phd.ibrahim@outlook.com](mailto:phd.ibrahim@outlook.com) adresine mail atabilirsiniz.**

## **ÖLÇME VE DEĞERLENDİRME KONU BAŞLIKLARI**

- Sabit ve değişken kavramları (nicel, nitel, sürekli, süreksiz, bağımlı, bağımsız)
- Ölçme, Ölçme Sonucu, Ölçme Kuralı, Ölçüt
- Ölçme Türleri (doğrudan, dolaylı, türetilmiş)
- Ölçmede Birim (doğal, yapay)
- Birimin Sahip Olması Gereken Özellikler (genellik, eşitlik, kullanışlılık)
- Ölçmede Sıfır (gerçek, izafi)
- Ölçek Türleri (sınıflama, sıralama, eşit aralıklı, eşit oranlı)
- Değerlendirme ve Değerlendirme Türleri
- Ölçmede Hata (sabit, sistematik, tesadüfî)
- Hata Kaynakları (ölçmeciden, ölçme aracından, ölçülen kişiden, ölçme yönteminden vb.)
- Korelasyon (pozitif, negatif, nötr)
- Bir Ölçme Aracında Olması Gereken Özellikler (geçerlik, güvenirlik, kullanışlılık)
- Test ve Test Hazırlama Süreci
- Geleneksel Ölçme Araçları ve Özellikleri (yazılı yoklamalar, sözlü sınavlar, eşleştirmeli testler vb.)
- Çağdaş Ölçme Araçları ve Özellikleri (kontrol listesi, portfolyo, dereceleme ölçeği vb.)
- Test İstatistikleri
- Merkezi Yığılım Ölçüleri (aritmetik ortalama, ağırlıklı ortalama, mod, medyan, vb.)
- Merkezi Dağılım Ölçüleri (standart sapma, varyans, ranj, çeyrekler açıklığı vb.)
- Standart Puanlar (z ve t puanları)
- Madde İstatistikleri (madde güçlüğü ve madde ayırt ediciliği, madde varyansı)
- Dağılım Grafikleri

## SABİT VE DEĞİŞKEN KAVRAMLARI

**Sabit:** Kişiden kişiye, durumdan duruma, nesneden nesneye değişmeyen özelliklere denir. Genellikle matematik ve fizik kuralları sabit kavramı içerisinde değerlendirilir. Örneğin Pi sayısı,  $mc^2$  gibi. Ancak yapılan çalışmanın içerisinde sadece belirli bir özellik dikkate alınıyorsa bu da bir sabit olabilir. Örneğin yapılacak bir çalışmada sadece öğretmenlerin dikkate alınması ya da sadece üniversite öğrencilerinin dikkate alınması bu özelliklerin sabit olmasını sağlayabilir.

**Değişken:** Kişiden kişiye, durumdan duruma, nesneden nesneye değişen özelliklere denir. Örneğin: Boy uzunlukları, kilo, zekâ, cinsiyet gibi kavramlar her zaman değişebilir kavramlar arasındadır. Değişkenlerin çeşitleri vardır. Bunları aşağıdaki tablodaki gibi düzenleyebiliriz:

Değişken Türleri		
Nicel Değişken	Sürekli Değişken	Bağımlı Değişken
Nitel Değişken	Süreksiz Değişken	Bağımsız Değişken
		Karıştırmacı Değişken

**Nicel Değişken:** Bu tür değişkenler sayısal değişkenlerdir. Diğer bir deyişle sayılarla ifade edebilmektir. Örneğin: Boy uzunluğu, ağırlık, zekâ düzeyi gibi. Melike 64 kg ağırlığındadır. Kenan Ölçme ve Değerlendirme dersinden 85 aldı. Buradaki 64 ve 85 sayıları nicel birer örnektir.

**Not:** Nicel değişkenler daha geçerlidir.

**Nitel Değişken (Kategorik Değişken):** Sıfatsal değişkenlerdir. Sıfatlarla veya sembollerle ifade edilebilirler. Örneğin: Cinsiyet, renkler, medeni durum, din, dil, inanç, millet. Matematiksel olarak anlam ifade etmezler. Kızlar “1”, Erkekler “2”.

**Not:** Nicel değişkenler nitel değişkenlere dönüştürülebilir. Örneğin: Kenan 196 cm uzunluğundadır. Kenan uzundur. Melih KPSS’den 91 aldı. Melih KPSS’de başarılı oldu. Buradaki başarılı kelimesi sıfattır, semboldür. Ayrıca bir sınıfta 15 erkek 10 kız öğrenci vardır dendiğinde ise buradaki durumda değişkenimiz *niteldir*. Çünkü buradaki sayılar sadece adet belirtmektedir.

**Not 1:** Zekâ, ilgi, başarı, sıcaklık gibi değişkenler *nicel* birer değişkendirler. Çünkü hepsi de ölçülebilir özelliktedir. Ancak cinsiyet, göz rengi, medeni durum gibi değişkenler *nitel*dirler. Çünkü bunlar ölçülebilir bir özelliğe sahip değildir. Hava sıcaklığı 35 derecedir diyebiliriz ancak Ahmet 0,5 evlidir diyemeyiz.

**Sürekli Değişken:** Reel sayılarla eşleştirme yapabilen değişkenlerdir. Yani iki değişken arasında sonsuz sayıda başka değişken yer alabiliyorsa bu tür değişkenlere denir. Örneğin: uzunluk, zekâ, sıcaklık gibi. İki sıcaklık derecesi arasında sonsuz sayıda değer vardır ve sıcaklıkta sürekli olarak değişim gösterir. Ayrıca sürekli değişkenler sayılarla ifade edildikleri için aynı zamanda niceldirler.

**Not:** Reel (gerçek) sayı, sayı doğrusundaki tüm noktalara karşılık gelen sayılardır.

**Süreksiz Değişken:** Tam sayılarla eşleştirme yapabilen değişkenlerdir. İki değer arasına ya hiç değer yazılmaz ya da çok az yazılır. Örneğin konuşulan dil, göz rengi gibi. İnsanın bildiği dil sayısı zamanla değişebilir ancak bu durum çok az olarak değişebilir. Süreksiz değişken aynı zamanda niteldir.

**Bağımsız Değişken:** Araştırmacının kontrolü altında olan, düzenleyip değiştirebildiği değişkendir. Bağımlı değişken üzerinde etkisi araştırılan değişkendir. Yani araştırma da etkisinin olup olmadığı araştırılan değişkendir.

**Bağımlı Değişken:** Araştırmada diğer değişkene (bağımsız değişken) ya da değişkenlere karşı tepkisinin araştırıldığı değişkenlerdir. Yani bağımsız değişkene bağlı olarak değişen değişkenlerdir. Örneğin düzenli uykunun insanlar üzerinde etkisinin araştırılması gibi.

**Not:** Bağımsız değişken olayın nedenidir. Bağımlı değişken ise olayın sonucudur, olaydan etkilenendir. Örneğin yağmur yağınca şemsiye satışları artar (Neyin neye etkisinin olduğuna bakılır). Reklamın insanların alışveriş yapmaları üzerindeki etkisi. Bilgisayar programlarının matematik öğrenme üzerindeki etkisi. Güdülenmenin başarı üzerindeki etkisi gibi.

**Not 1:** Bağımlı ve bağımsız değişken durumu göre değişebilir. Örneğin bir durumda bağımlı değişken olan değişken başka bir durumda bağımsız değişken olabilir.

**Karıştırıcı Değişken:** Araştırma da doğrudan gözlenmeyen ancak bağımsız değişken ile bağımlı değişken arasındaki ilişkiyi etkileme olasılığı olan, kontrol altına alınamayan

değişkendir. Örneğin reklamın alış-veriş üzerindeki etkisi incelenirken, çok fazla reklam izlenmesi daha fazla alış-veriş yapılmasına sebep olabilir ancak bireyin sosyo-ekonomik durumunun kötü olması sebebiyle birey ne kadar reklam izlerse izlesin alış-veriş yapmayabilir. Burada sözü edilen sosyo-ekonomik durum gibi; bireyin eğitim düzeyi, cinsiyeti, o anki psikolojik durumu elde edilecek sonuçlar üzerinde beklenmeyen sonuçların oluşmasına neden olabilir.

## **ÖLÇME**

Bir değişkenin (ölçülmek istenen bir özelliğin) gözlenerek, gözlem sonucunun sayı ya da sembolle eşleştirilme sürecidir. Bir durum tespittir. Gözlenen özelliğin sayı ya da sembolle ifadesidir. Ölçmede dört temel öge bulunur. Bu ögeler aşağıda sıralanmıştır:

1. Özelliğin belirlenmesi (ne ölçülecek),
2. Ölçme aracı (gözlem),
3. Ölçme kuralı,
4. Sayı ya da sembol.

## **Ölçüm (Ölçme Sonucu)**

Bu sürecin sonunda elde edilen sayı ya da sembole ölçme sonucu ya da ölçüm denir. Ölçülebilecek bazı durumları şu şekilde sıralayabiliriz: Uzunluk, yükseklik, sıcaklık, ısı, nem, deprem şiddeti, hız, alan, hacim, bilgi, zekâ, mevcut, öğrenme vb.

1. Mehmet'in boyunun uzunluğu 174 santimetredir.
2. Erciyes Dağı'nın yüksekliği 3917 metredir.
3. Ali sınavdan 71 aldı.

## **Ölçme Kuralı**

Ölçülen özelliğin hangi sayı ya da sembolle eşleştirileceğine karar verme işlemine de ölçme kuralı denir.

1. Sınavdaki her soru 10 puan değerindedir.
2. 11. soruyu yapamayan bu sınavdan geçemez.
3. Sınavda dört yanlış bir doğruyu götürür.

**Not:** Ölçme kuralı her zaman için puanlamayla ilgilidir. Puanlamayla ilgili olmayan ancak sınav esnasında uygulanan başka kurallara *sınav kuralı* denir.

1. Sınav esnasında hesap makinesi kullanmak yasaktır.
2. Sınavımız toplam 30 dakikadır.
3. Sınava 5 dakika geç kalanlar, sınava alınmayacaklardır.

## **Ölçüt**

Ölçme sonuçlarının kıyaslamaya tabi tutulduğu değerdir. Ölçüt değiştikçe değerlendirme de değişir. Ölçüt olmaksızın değerlendirme yapılamaz. Örneğin Matematik dersini geçme puanı 50'dir. Aritmetik ortalamanın üstünde olanlar sınıfı geçer. Ölçüt ikiye ayrılır. Bunlar *mutlak ölçüt* ve *bağıl ölçüt*tür.

### **Mutlak Ölçüt**

Ölçme sonuçlarının hangi kritere göre değerlendirileceğinin ölçme işlemi yapılmadan önce belirlenmesidir.

1. Sınavdan 70 alan herkes geçer. (Buradaki 70 rakamı daha önceden belirlenmiştir.)
2. Yüksek lisans yapabilmek okul ortalamasının en az 2,50 olması gerekir.
3. Mesleğinde 3 yıl deneyimi olan herkes, uzman olarak atanacaktır.

### **Bağıl (Norm) Ölçüt**

Ölçme sonuçlarının hangi kritere göre belirleneceğinin sınav sonuçlarına göre belirlenmesidir.

1. Sınav sonucuna göre ilk 10'a giren herkes bu dersi geçer.
2. %10'luk başarı dilimine girenler, ücret ödemek zorunda değildir.
3. Mesleğinde en deneyimli olan 5 kişi, uzman olarak atanacaktır.

**Not:** Bağıl ölçütümüzde henüz ölçme işlemine başlamadan belirlenebilir. Ancak mutlak ölçütteki gibi net bir sayı ortaya konulamaz.

# ÖLÇME TÜRLERİ

## Doğrudan (Temel) Ölçme

Ölçülecek özelliklerin, araya başka bir değişken girmeden doğrudan doğruya gözlenmesi sonucu yapılan ölçme türüne “doğrudan ölçme” denir. Eğitimde doğrudan ölçülecek özellikler çok azdır, el ile yapılan işler dışında, doğru ve güzel okuyabilme, iyi konuşabilme gibi özellikler, eğitimde doğrudan gözlenebilen özelliklerdir. Ölçülen özellik ile ölçme aracının özelliğinin aynı olmasıdır. Özelliğin kendisi ölçülür ve özellik var olduğu şekliyle ölçülür. Doğrudan doğruya karşılaştırma şansı vardır.

1. Bir tahtanın uzunluğunu metre ile ölçmek doğrudan ölçmedir.
2. Eşit kollu teraziyle ağırlık ölçmek doğrudan ölçmedir.
3. Varlıkların sayısını söylemek doğrudan ölçmedir.
4. Mehmet’in kaç kardeşinin belirlenmesi doğrudan ölçmedir.
5. Bir kişinin kaç adet kitabının olduğunun belirlenmesi doğrudan ölçmedir.

**Not:** Doğrudan ölçmede hata yapma şansı çok azdır. Doğrudan ölçmenin geçerliği ve güvenilirliği daha yüksektir.

## Dolaylı (Göstergeli) Ölçme

Ölçülen özellik ile ölçme aracının özelliğinin farklı olduğu durumlarda dolaylı ölçme vardır. Özelliğin etkisi, belirtisi, yansıması ölçülür.

1. Hava sıcaklığını termometre ile ölçmek dolaylı ölçmedir.
2. Yaylı terazi ile ölçme yapmak dolaylı ölçmedir.
3. Elektronik teraziyle ölçme yapmak dolaylı ölçmedir.
4. Öğrencinin derse karşı tutumunu ölçmek dolaylı ölçmedir.
5. Öğrencinin başarı durumunu sınav yaparak ölçmek dolaylı ölçmedir.

**Not:** Dolaylı ölçmede hata yapma ihtimali daha fazladır. Eğitimde kullanılan ölçme araçları dolaylı ölçme işlemini gerçekleştirir.

### Doğrudan ölçme ile dolaylı ölçme arasındaki farklar:

Doğrudan Ölçme	Dolaylı Ölçme
Sıfır noktası gerçektir.	Sıfır noktası tanımlanmıştır.
Eşit kollu terazi ile yapılan ölçümler.	Baskül, yaylı kantar, dinamometre
Eşit oranlı ölçek kullanılır.	Eşit aralıklı ölçek kullanılır.
Metre ile boy ölçümü veya binanın yüksekliğinin ölçümüdür.	Rakım ölçümü
Varlığın kendisi ölçülür.	Varlığın kendisi değil araç üzerindeki etkisi ölçülür.
Geçerlik güvenirliği yüksektir.	Geçerlik güvenirliği düşüktür.
Hata oranı düşüktür.	Hata oranı yüksektir.

### Türetilmiş Ölçme

Bir özelliği birden fazla özellik yardımıyla ölçmektir. Formüllerle yapılan ölçmedir.

1. Arabanın hızını bulmak türetilmiş ölçmedir.  $Hız = Yol/Zaman$ .
2. Bir yerin nüfus yoğunluğunu bulmak türetilmiş ölçmedir.  $Nüfus\ yoğunluğu = Nüfus / Alan$ .
3. Bir ülkenin Gayri Safi Milli Hâsıla (GSMH)'nın hesaplanması da yine türetilmiş ölçmedir.
4. Bir kişinin IQ seviyesini bulmak türetilmiş ölçmedir.  $IQ = Zeka\ yaşı / Gerçek\ yaş$ .
5. Bir maddenin yoğunluğunu bulmak türetilmiş ölçmedir.  $Yoğunluk = Kütle/Hacim$ .
6. Öğrencinin taktir teşekkür hesabı yapması türetilmiş ölçmedir.

**Not:** Bilişsel alan, dolaylı ölçme; duyuşsal alan, dolaylı ölçme; psiko-motor alan ise doğrudan veya dolaylı ölçme kapsamına girer.

## ÖLÇMEDE BİRİM

Ölçme sonuçları çoğu kez bir birimle ifade edilir. Standart bir birimin varlığı ölçme kuralının uygulanmasını kolaylaştırır ve ölçme hatalarını belli bir oranda azaltır (Turgut ve Baykul, 1992:24). Birim, ölçmenin yapılmasını, sonuçlarının anlaşılması ve kullanılışlılığını oluşturan en küçük ölçektir. Ölçmecinin önemli sorumluluklarından biri ölçme amacına uygun düşecek birimi seçmektir. Ölçmede kullanılacak olan birimlerde eşitlik, genellik ve ölçme amacına uygunluk gibi özellikler aranmalıdır (Taşdemir, 2011: 23).

Bir ölçme aracını oluşturan en küçük parçacığa birim denir. Örneğin “Hava sıcaklığı 35 C derece” dendiğinde birim C derecedir. “Sınıfta 25 öğrenci vardır” dendiğinde öğrenci birimdir. “Sınavda 40 soru vardır” dendiğinde her bir soru bir birimdir. “Sınıf 6 metre uzunluğundadır” dendiğinde metre bir birimdir.

**Birim Çeşitleri:** Birimler *doğal (gerçek, tanımlanmamış)* ve *yapay (standart, izafi, tanımlanmış)* birim olmak üzere ikiye ayrılırlar.

### a. Doğal (Gerçek) Birim

Herhangi bir tanımlayama ihtiyaç duyulmayan birimdir. Özelliğin ne kadarlık kısmına bir birim deneceği uzmanlar tarafından belirtilmemiştir.

1. Bahçeden 4 ağaç var. Burada uzmanlar tarafından kesin olarak belirlenmiş bir ağaç boyutu yoktur. Yani ağacın boyu 2 metrede olsa 3 metrede olsa bir ağacı ifade eder.
2. Bu kitap 220 sayfadır. Burada da uzmanlar tarafından belirlenmiş bir kitap sayfa sayısı yoktur. Yani kitap 110 sayfa olsa da 350 sayfa olsa da bir kitaptır. Kitabın kaç sayfa olacağı uzmanlar tarafından belirlenmemiştir.
3. Masanın uzunluğu 6 karıştır. Burada uzmanlar tarafından belirlenmiş bir karış boyutu yoktur. 20 cm olsa da 30 cm olsa da bir karıştır.

**Not:** Doğal birimde genellikle fiziksel özellikler vardır.

### b. Yapay (Standart, İzafi) Birim

Tanımlanmış birimlerdir. Özelliğin ne kadarlık kısmına bir birim deneceğine uzmanlar karar vermiştir.



1. Bir günün 24 saat olmasına uzmanlar tarafından karar verilmiştir. Bu sebeple yapay birimdir.
2. Bir yılın 365 gün olacağına uzmanlar karar vermişlerdir. Bu sebeple yapay birimdir.
3. Para, ağırlık birimleri, uzunluk, metre gibi birimler uzmanlar tarafından belirlenmiştir. Bu sebeple yapay birimlerdir.

### **Birimin Sahip Olması Gereken Özellikler**

Bir birimin sahip olması gereken bazı özellikler vardır. Bu özellikler eşitlik, genellik, kullanışlılık olmak üzere üçe ayrılır. Aşağıdaki gibi sıralayabiliriz:

#### **a. Eşitlik**

Mesafenin eşit olmasıdır. Nereden bakılırsa bakılsın mesafe aynıdır.

1. Metre üzerindeki 1. ve 2. cm'ler arasındaki mesafe ile 41. ve 42. cm'ler arasında da aynı olmasıdır.
2. Termometre üzerindeki 15. ve 16. C derece arasındaki mesafe ile 32. ve 33. C derece arasındaki mesafenin aynı olmasıdır.
3. Bir saat üzerinde saatin 2:00 ile 3:00 arasındaki sürenin 13:00 ile 14:00 arasındaki süreyle eşit olmasıdır.

#### **b. Genellik**

Aynı toplumda herkes tarafından aynı birimin kullanılmasıdır. Bu özellik sayesinde gündelik yaşam kolaylaşır.

1. Bir toplum içerisinde herkesin metreyi kullanması genellik ilkesine uygundur.
2. Bir toplum içerisinde herkesin kilogramı kullanması genellik ilkesine uygundur.
3. Bir toplum içerisinde herkesin aynı para birimini kullanması genellik ilkesine uygundur.

**Not:** Bir ülke içinde aynı işlemi görmek üzere birden fazla değişik birim kullanıldığında gündelik hayatta sıkıntılar çıkabilecektir. Ancak burada üzerinde durulması gereken kısım bir toplumun içerisinde herkes tarafından aynı birimin kullanılmasıdır. Ege Bölgesi'nde metre kullanılması, İç Anadolu Bölgesi'nde karış kullanılması genellik ilkesine aykırıdır. Yoksa

Türkiye’de uzunluk birimi olarak metrenin kullanılması, İngiltere’de inç kullanılması genellikle ilkesine aykırı değildir.

### **c. Kullanışlık**

En uygun birimin kullanılmasıdır. Zaman tasarrufu sağlar.

- 1.** Boy ölçmek için metrenin kullanılması uygundur. Eğer ki boy ölçmek için kilometre kullanılsaydı kullanışlı olmazdı.
- 2.** Bir insanın yaşını sorarken “Kaç yaşındasın?” demek kullanışlıdır. “Kaç günlüksün?” diye sormak kullanışlı değildir.
- 3.** Ayrıca altını ölçerken gram kullanmak, kömürü ölçerken kilogram kullanmak daha kullanışlıdır.

**Not:** Ölçmede hata miktarı, bir ölçme işlemine karışan hata miktarı o ölçme işleminde kullanılan birimin yarısı kadardır. Örneğin bir sınavda 10’ar puandan 10 soru varsa buradaki hata miktarı; her bir soru 10’ar puan olduğu için 5 puanlık bir hata var demektir. Yani öğrenci sınavdan 65 almış ise; öğrenci gerçekte o sınavdan 60 ile 70 arasında bir puan almış demektir. Eğer ki toplam 100 puana karşılık gelecek şekilde 4 soru sorulmuş ise her bir soru 25 puan olacağı için ölçme işlemine 12,5 puanlık bir hata karışmış demektir. Bu sınavdan öğrenci 70 almışsa, öğrenci gerçekte 57,5 ile 82,5 arasında bir puan alacaktır denebilir. Sınavda sadece 2 adet soru varsa bu ölçme işlemine 25 puanlık bir hata karışmış demektir. Bu sınavdan öğrenci 50 almışsa, çok yüksek ihtimal bu öğrencinin gerçek alacağı not 25 ile 75 arasında demektir. Bu sebeple ölçmedeki hata miktarını azaltmak için ölçme işlemine öğrencileri sıkmayacak şekilde yeteri kadar soru sorulmalıdır. Örneğin sınavda 20-30 arasında soru sormak daha net sonuç verecektir denebilir.

## ÖLÇMEDE SIFIR

Ölçme işlemlerinde genel olarak *gerçek sıfır*, *itibari sıfır* olmak üzere iki adet sıfır vardır. Bunları aşağıdaki gibi sıralayabiliriz:

### Gerçek (Doğal, Mutlak) Sıfır

Doğal sıfır, ölçmeye konu olan özelliğin gerçek anlamda yokluğunu ifade eder (Yaşar, 2010). Eğer ki bir şeye 0 (sıfır) diyorsak, ortada hiçbir şeyin olmamasıdır. Yani yokluğu ve hiçliği belirtir. Aynı zamanda o özelliğin olmadığını ifade eder.

Aşağıdaki örneklerle bunu daha iyi kavrayabiliriz:

1. Sınıfta 0 (sıfır) öğrenci var demişsek sınıfta öğrencinin olmadığı anlaşılır. Bu gibi durumlarda kullanılan 0 (sıfır) gerçektir.
2. Karpuz 15 kilogram ağırlığındadır. Yani 0 (sıfır) kilogram demiş olsaydık, ortada karpuzun olmadığını söylemiş olacaktık.
3. Tahtanın uzunluğu 220 cm'dir. Yani 0 (sıfır) cm'dir demiş olsaydık, ortada bir tahta yok demektir.
4. Ali'nin sınavda 10 yanlış vardır. Yani 0 (sıfır) yanlış var demiş olsaydık, Ali'nin sınavda yanlış yapmadığını anlardık.
5. Ölçme ve değerlendirme dersimiz 55 dakika sürdü. Yani ölçme ve değerlendirme dersimiz 0 (sıfır) dakika sürdü demiş olsaydık, bu dersin hiç olmadığını anlardık.
6. Nabız dakikada 85 atıyor. Burada gerçek 0 (sıfır) vardır. Çünkü nabız atmasa yaşamsal faaliyet gösteremezdi.

### İzafi (İtibari, nispi, tanımlanmış, göreceli, keyfi) Sıfır

Göreceli (itibari) sıfır ise, gerçek sıfır olmadığı halde ölçeğin üzerinde belli bir noktanın sıfır olarak kabul edildiği durumlar için geçerlidir (Turgut ve Baykul, 1992:23). Göreceliği tanımlanmış 0 (sıfır)'dır. Yokluğu ve hiçliği ifade etmez. Başlangıç noktası teknik sebeplerden dolayı 0 (sıfır) kabul edilir. Yani bir başlangıç noktasına ihtiyaç duyarız.

1. Kenan, Matematik dersinden 0 (sıfır) aldı demişsek, burada Kenan'ın bu ders hakkında hiçbir şey bilmediğini ifade etmez. Sadece sorulan soruları bilmediğini gösterir.

2. Hava sıcaklığı 23 C derece ölçülmüştür. Yani hava 0 (sıfır) C derece olsaydı havada sıcaklık yoktur diyemeyeceğimize göre burada izafi 0 (sıfır) vardır.
3. Ayşe'nin zekâ puanı 145 çıkmıştır. Yani 0 (sıfır) çıktı deseydik, Ayşe'de zekâ yoktur diyemeyeceğimiz için burada izafi 0 (sıfır) vardır.
4. Erciyes Dağı'nın yüksekliği 3917 metredir. Yani biz burada dağın boyunu ölçmek için dağın boyunu düzlükten ya da deniz kıyısından ölçmüş olabiliriz. Ancak dağın boyu daha aşağıdan da başlar.
5. Biz bir ağacın boyunu ölçerken toprağa değdiği yerden ölçeriz ancak ağacın boyu daha da aşağıya gitmektedir.
6. Van'da Richter Ölçeği'ne göre 7.2 şiddetinde deprem olmuştur. Deprem, yer sarsıntısı demektir ve yer her daim bir sarsıntı içerisinde. Yani bizler ölçememiş olsak bile yer sarsılmaya devam etmekte olduğu için burada izafi 0 (sıfır) vardır.
7. Ölçme ve değerlendirme dersi 09:10'da başlıyor gibi ifadeler izafi 0 (sıfır)'dır. Çünkü sürekli devam eden bir olay bulunmaktadır. Saat 00:00 dediğimizde burada zamanın yokluğundan bahsedemeyiz.

**Not:** Gerçek 0 (sıfır) ile izafi 0 (sıfır) arasındaki birkaç farkı aşağıdaki gibi sıralayabiliriz:

1. Gerçek 0 (sıfır) asla negatif değer almaz. İzafi 0 (sıfır) negatif değer alabilir.
2. Gerçek 0 (sıfır)'da katlamalı ve oranlı ilişkiler kurulur. İzafi 0 (sıfır) da ise katlamalı ve oranlı ilişkiler kurulamaz.
3. Hava sıcaklığını C derece ile ölçtüğümüzde izafi 0 (sıfır) varken, hava sıcaklığını Kelvin ile ölçmüşsek burada gerçek 0 (sıfır) vardır. Çünkü Kelvin gerçek yokluğu ifade eder.
4. Verilen bir örnekte sürede varsa gerçek 0 (sıfır), zaman varsa ise izafi 0 (sıfır) vardır.
5. Enlem ve boylamlarda izafi 0 (sıfır) vardır.
6. Gölge boyu izafi sıfır ile ifade edilir çünkü biz gölgemizi göremesek bile gölgemiz ayaklarımızın altındadır.
7. Bir kişi için sıfır bedendir demişsek burada izafi sıfır vardır, çünkü bu kişinin hala bir beden ölçüsü vardır.
8. "Moralim sıfır." diyen birisi içinde izafi sıfır vardır çünkü yine ortada mutlak bir yokluktan söz edemeyiz.

## ÖLÇEK TÜRLERİ

Ölçek kelimesi bazı kaynaklarda birim bazı kaynaklarda ise ölçme aracı olarak tanımlanmaktadır. Ölçekler dört gruba ayrılmaktadır. Bunlar *sınıflama ölçekleri*, *sıralama ölçekleri*, *eşit aralıklı ölçekler* ve *eşit oranlı ölçeklerdir*.

### a. Sınıflama (Nominal, Adlandırma) Ölçekleri

En kaba, en genel ölçeklerdir. Benzer özelliklere sahip olanların aynı sınıfa aynı kategoriye aynı kümeye koymaktır. Matematiksel işlem yapılamaz. Ancak tanıma ve ayırt etme amaçlı sayı kullanılır.

Sınıflama ölçeklerine ile ilgili örnekler aşağıdaki gibi sıralanabilir:

1. Sınıftaki öğrencileri kızlar ve erkekler olarak ikiye ayırmak sınıflama ölçeklerine örnektir.
2. Galatasaraylılar ve Fenerbahçeliler olarak ikiye ayırmak sınıflama ölçeklerine örnektir.
3. Kimlik numarası bir sınıflama ölçeğidir. Çünkü burada adlandırma vardır. Öğrenci numarası, sokak numarası, futbolcunun formasındaki sayı, illerin plakaları da bir adlandırma olduğu için sınıflama ölçeklerinin içerisinde yer alır.
4. Bir yayınevi deposundaki kitapları kişisel gelişim, klasikler, çok satanlar, indirimdekiler şeklinde ayırması sınıflandırma ölçeklerinin içerisinde yer alır.
5. Sınıflama ölçeğinde geçişlilik ( $A=B$  ise  $B=C$  ise  $A=C$ ) ve simetriklik vardır. Yani  $A=B$  ise  $B=A$  gibi. Sınıflama ölçeklerinde *Mod*, *Frekans* ve *Yüzde* bulunabilir.

### b. Sıralama (Derecelendirme, Ordinal) Ölçekleri

Sıralama ölçekleri, incelenen varlık veya olayların, ölçülen özelliğe sahip olma dereceleri bakımından kendi grupları içindeki sıralarını göstermekle birlikte, böyle bir sıranın ardışık iki konumunda bulunan varlık ve olaylar arasındaki farkın miktarını göstermeyen ölçüler verirler (Özçelik, 1992:18) Varlıkları aldığı özelliğin miktarına göre bir sıralamaya tabi tutmaktır. Büyüklük sırası önemlidir ancak sıralamada miktar bilinmez. Başlangıç noktası belli değildir ve sıralar arası mesafe eşit değildir. Özelliğin farkı miktar cinsinden verilmez.

Sıralama ölçekleri ile ilgili örnekler aşağıdaki gibi sıralanabilir:

1. Dünya'nın en yüksek dağı Everest Tepesi'dir (Dağın yüksekliği bizi ilgilendirmez).

2. Okulumuz bilgi yarışmasında birinci olmuştur. (Kaç puan alarak birinci olduğu sınıflama ölçeğinde yer almaz).
3. En kalın kitap ölçme ve değerlendirme dersi kitabıdır (Kitabın kaç sayfa olduğu bilinmemektedir).
4. Sınıfın en uzun boylu öğrencisi Süleyman'dır (Süleyman'ın boyunu bilmemekteyiz).
5. En soğuk ay Ocak ayıdır (Ancak Ocak ayının sıcaklık değerini bilmiyoruz).
6. Öğrencileri boy sırasına göre dizmek sıralama ölçeğinin içerisinde.

Not: Ancak öğrencilerin bir sınavdan aldıkları ham puanlar sıralama ölçeği düzeyindedir.

**Not:**

- Sıralama ölçeğinde birimler eşit değildir. Çünkü iki özellik arasındaki mesafe belli değildir.
- Sıralama ölçeğinde sıfır anlamsızdır. Çünkü örneğin Matematik dersinden bir öğrenci 80 başka bir öğrenci 70 almışsa, aradaki 10 puanlık farkın, Matematik bilgisinde ne kadarlık bir kısma karşılık geldiği belli değildir.
- Medyan, mod, yüzde hesapları ve sıralar arası fark bulunabilir.

### **c. Eşit Aralıklı (Interval) Ölçek**

Başlangıç noktası olan 0 (sıfır) izafidir. Birimler arası fark eşittir. Bu ölçekte katlamalı oranlı bir ilişki yoktur. Negatif değer alabilir. Bu ölçek türünde sadece toplama ve çıkarma işlemi yapılabilir.

Eşit aralıklı ölçeklerle ilgili örnekler aşağıdaki gibi sıralanabilir:

1. İki öğrenciden biri sınavdan 40, diğeri 80 almışsa; “biri diğerinin iki katı biliyor.” diyemeyiz.
2. Hava sıcaklığı dün 17 bugün 34 derece ölçülmüşse; “dünün sıcaklığı bugünün sıcaklığının yarısıdır.” diyemeyiz.
3. Mersin’de sıcaklık 28 C derece, Kayseri’de 14 C derece ise; “iki şehir arasında 14 C derece fark vardır.” diyebiliriz.
4. Saat eşit aralıklı bir ölçektir. Saat 16:00, saat 8:00’in iki katıdır diyemeyiz. 8 saatlik fark var diyebiliriz.

**Not:**

- Eğitimde en sık kullanılan ölçektir. Daha çok dolaylı ölçme türünde kullanılır. Ersin sınavdan 80 puan aldı (Tekilse eşit aralıklı).
- Takvim, enlem/boylam, rihter ölçeği eşit aralıklı ölçek düzeyindedir.
- Öğrencilerin Z ve T puanları (standart puanlar) eşit aralıklı ölçek düzeyindedir.

**d. Eşit Oranlı (Ratio) Ölçek**

Başlangıç noktası olan sıfır, gerçek 0 (sıfır) olan ölçeklerdir. Her türlü matematiksel işlem yapılabilir. Özellik farkını miktar cinsinden verir. Katlamalı oranlı ilişki vardır. Negatif değer yoktur. Kıyaslama yapılabilir. Eğitimin hedefleri arasında yer almaz (Örneğin öğrencilerin boyları, kiloları eğitim için çok fazla gerekli değildir). En çok bilgi veren ölçektir.

Eşit oranlı ölçeklerle ilgili örnekler aşağıdaki gibi sıralanabilir:

1. “Arabalardan biri 160 km hızla diğeri 80 km hızla gidiyor” denmişse, “arabalardan biri diğerrinin iki katı hızla gidiyor” diyebiliriz. Çünkü bu tür ölçeklerde her türlü matematiksel işlem yapılabilmektedir.
2. Ahmet’in 100 TL parası var. Mehmet’in 25 TL parası var. O halde Ahmet’in parası Mehmet’in parasının 4 katıdır. Katlamalı oranlı ilişki kurulabilmesi sebebiyle eşit oranlı ölçek düzeyindedir.
3. 6/A sınıfının uzunluğu 18 metredir. 6/B sınıfının uzunluğu ise 12 metredir. 6/A sınıfı 6/B sınıfının  $3/2$  oranına sahiptir. Katlamalı oranlı ilişki vardır, gerçek sıfır kullanılmıştır o halde eşit oranlı ölçek düzeyindedir.

**Not:** Uzunluk ölçüleri, ağırlık ölçüleri, elektrik ölçüleri, ses şiddeti, basınç birimleri gibi araçlar eşit oranlı ölçekler içinde en çok kullanılan araçlardır.

## DEĞERLENDİRME

Ölçme sonuçlarının belirli bir kıstasa vurularak, bir karar verme sürecidir. Değerlendirmede dört temel öge bulunmaktadır. Bunlar aşağıdaki gibi sıralanmıştır:

1. Ölçme sonucu (ölçüm),
2. Ölçüt (kriter),
3. Değer yargısı,
4. Karar.

Değerlendirme kavramı kendi içinde çeşitli alt dallara ayrılmaktadır. Bir ölçme işleminde en son öge değerlendirmedir. Değerlendirmenin kendi içindeki alt dalları aşağıdaki tabloda verilmiştir.

Değerlendirme						
Ölçüte Göre Değerlendirme		Amacına Göre Değerlendirme			Diğer Değerlendirme Türleri	
Mutlak	Bağıl	Tanıma ve Yerleştirme	Biçimlendirme-Yetiştirme	Sonuç Görme	Öğretimin Değerlendirilmesi	Rehberlik Amaçlı Değerlendirme

### 1. Ölçüte Dayalı Değerlendirme

Ölçüte göre değerlendirme kendi içerisinde *mutlak değerlendirme* ve *bağıl değerlendirme* olarak ikiye ayrılmaktadır.

#### a. Mutlak Değerlendirme (Kriter Dayanaklı/Ölçüt Dayanaklı)

Genellikle ölçütün daha önceden belirlendiği değerlendirmedir. Elimizde bulunan sabit bir sayı ya da ölçüte göre değerlendirmeye varılmasıdır. Bir başka deyişle önceden belirlenmiş mutlak (sabit) ölçüte göre değerlendirmenin yapılmasıdır. Ölçüt; kesin, sabit ve değişmezdir.

1. Sınavdan 50 alan herkes bu dersten başarılı olacaktır.
2. Yüksek Öğretime Giriş Sınavı'ndan 450 ve üzeri alanlar tıp fakültesine girecektir.
3. Final sınavından 50'nin altında alanlar ilgili dersten başarısız sayılacaklardır.



**Not:**

- Geçti/kaldı gibi kararlar verilirken mutlak değerlendirme yapılması daha uygundur.
- Mutlak değerlendirmenin geçerliği ve güvenilirliği daha yüksektir çünkü hedef davranışlar kazanılmadan kimse dersi geçemez.
- Hayati önem taşıyan programlarda, mutlak değerlendirme tercih edilmelidir.
- Programın, dersin, ünitenin, konunun hedeflerine ulaşılması isteniyorsa; öğrencilerin öğrenme eksiklikleri belirlenecekse mutlak değerlendirme kullanılmalıdır.
- Öğretmen kanısı kullanılacaksa mutlak değerlendirme vardır çünkü öğretmen, öğrencinin hedef davranışa ulaşp ulaşmadığı hakkında kanısını ortaya koyacaktır.
- Mutlak değerlendirme sayesinde sınıf içerisinde rekabet oluşmaz.
- Mutlak değerlendirmenin olduğu yerde dersi herkes geçebilir ya da hiç kimse geçemeyebilir.
- Muafiyet sınavları, mezuniyet durumlarında mutlak ölçüt kullanılmalıdır.

**b. Bağlı Değerlendirme (Norm Dayanaklı)**

Değerlendirme yapabilmek için daha önceden belirlenmiş net bir ölçüt bulunmamaktadır. Ölçütün (kaç puan olacağı ya da ne kadar olacağı) daha sonra netlik kazanacağı söylenebilir. Ölçütün kaç olduğu bulunan grubun durumuna göre değişiklik gösterecektir.

1. Aritmetik ortalamanın altında alanlar bu dersten kalacaklardır.
2. Kamu Personeli Seçme Sınavı'ndan en yüksek puan alan ilk 100 kişi işe yerleştirilecektir.
3. Yüksek Öğretime Giriş Sınavı'ndan en yüksek puan alan ilk 150 kişi tıp fakültesi tercihi yapabilecektir.

**Not:**

- Bağlı değerlendirme öğrencilerin performansına göre değişir.
- Öğrencilerin yetenekleri söz konusuysa bağlı değerlendirme yapılır çünkü en yetenekliler seçilecektir.
- Ülke çapındaki normlar bağlı değerlendirmedir.
- Z ve T puanları bağlı değerlendirmedir çünkü her iki puanda hesaplanırken sınıfın aritmetik ortalaması hesaplanır.

- Seçme amaçlı sınavlarda bağıl değerlendirme tercih edilmelidir.
- Bağıl değerlendirme sınıf içerisinde rekabet oluşturur. Çünkü öğrencilerin sınıf içi başarısı önemlidir.
- Bağıl değerlendirme de mutlaka ilgili dersten “kalan” olur.
- Bağıl değerlendirmenin geçerlik ve güvenirliği düşüktür çünkü öğrenilmesi gereken davranışlar öğrenilmeden ders geçilebilir.
- Standart sapma, mod, medyan, yüzdelik dilim, çan eğrisi, eriş testleri gibi durumlar bağıl değerlendirmenin içerisinde yer alır.

## **2. Amacına Göre Değerlendirme**

Amacına göre değerlendirme kendi içerisinde *tanıma ve yerleştirme*, *biçimlendirme-yetiştirme* ve *düzey belirlemeye yönelik* olarak üçe ayrılmaktadır. Bu değerlendirme türleri sırasıyla *girdi*, *süreç* ve *çıkış* şeklinde sıralanacaktır.

### **a. Tanıma ve Yerleştirmeye Yönelik Değerlendirme (Diagnostik)**

Bireyin öğrenme sürecine girerken sahip olduğu davranışları ortaya koyabilmek amacıyla yapılır. Bireyin o andaki sahip olduğu mevcut durumunu ortaya koymak öğrenciyi tanımak amacı için yapılan değerlendirmedir. Bireyin sahip olduğu mevcut durumun seviyesi de belirlenir. Bu sebeple iki tür değerlendirme yapılabilir. Bunlar *hazırbulunuşluğu belirleme* ve *ön koşul öğrenmeleri belirleme*dir.

**Not:** Hazırbulunuşluk, bireyin sahip olduğu bilgileri, onun bir sonraki bilgileri öğrenmesinde alt yapı oluşturur. Örneğin bir birey, alfabedeki harfleri öğrenememişse, onun kelime ya da cümle kurmasında sorunlar yaşanabilir. Yani öğrenci ‘Ayşe’ kelimesini okuyabilmek için a, y, ş, e harflerini daha önceden öğrenmiş olması gerekir.

Tanıma ve yerleştirmeye dönük değerlendirme yapılırken yararlanılan bazı testler vardır. Bunlar; hazırbulunuşluk testleri, kelime işlemleri, problem çözme testleri, genel yetenek testleri, kişilik testleri, tutum ölçekleri, okuma-anlama testleri ve ilgi envanterleri olarak sıralanabilir. Muafiyet sınavları, özel kurum sınavları, özel yetenek sınavları, Kamu Personeli Seçme Sınavı, Yükseköğretim Kurumları Sınavı, Liselere Giriş Sınavı örnek olarak verilebilir.

**Not:** Envanter (stok), belirli bir tarihe ilişkin borç, alacak ve varlıkların miktarlarının ve değerlerinin, sayım, kontrol ve düzeltme yaparak saptanmasıdır. Psikolojide kullanılan envanter ise (envanter tipi test), sadece evet ve hayır ile cevaplanacak soruları içerir (wikipedia.org).

#### **b. Biçimlendirme-Yetiştirmeye- Geliştirmeye Yönelik Değerlendirme (Formative)**

Eğitim-öğretim süreci devam ederken bireyin öğrenme eksikliklerini, öğretimin hatalarını ve eksikliklerini bulmak, bunları düzenlemek, dönüt almak amacıyla yapılır. Öğrenme eksikliklerini belirleyebilmek için öğrenciler; anlık değerlendirmeye veya ara değerlendirmeye tabii tutulabilirler. Yanlış ya da eksik öğrenmeler, öğrenme eksiklikleri, öğrenme güçlükleri, kavram yanlışları tespit edildiğinde bunların düzeltilmesini yani biçimlendirilmesi sağlanabilir. Ayrıca eksik öğrenmeleri kapatıcı çalışmalar yapılabilir.

Biçimlendirme-yetiştirmeye dönük değerlendirmelerde *gözlemler, ünite testleri, öğrenme testleri, ara sınavlar, uygulama testleri* bulunmaktadır.

#### **Not:**

- Öğretmen öğrenciden, öğrencide öğretmenden dönüt alır.
- Dönemin içerisinde, her bir ünitenin/konunun sonunda yapılabilir.
- Bu tür sınavlarda not ya da puan yoktur. Motivasyon amaçlı puanlama yapılır ancak bu puanlamanın ders geçme üzerindeki etkisi %10'u geçmez.

#### **c. Sonuç Görmeye, Düzey Belirlemeye, Değer Biçmeye Yönelik Değerlendirme (Summative)**

Sonuç görmeye yönelik değerlendirmede eğitim-öğretim etkinliğinin amacına ulaşp-ulaşmadığına, öğrencinin hedef davranışları kazanıp kazanmadığına bakılır. Öğrencilerin tümel (bilişsel, duyuşsal, psiko-motor, kişilik vb.) olarak değerlendirmeleri yapılır. Hedeflere ulaşılma düzeyi de önemlidir. Genellikle öğrencilere yapılan başarı testleriyle karara varılmaya çalışılır. Öğrencinin eğitim-öğretim etkinliğine girişindeki mevcut düzeyi ile etkinlik sonunda çıktı düzeyi arasında belirli farkın olması istenir. Giriş ile çıkış arasındaki farka eriş denir.

**Not:**

- Erişi, öğrencinin son test puanlarıyla, ön test puanları arasındaki hedeflerle tutarlı farktır.
- Bu testler/sınavlar dönem sonunda yapılır ve bu sınavlar sonucunda öğrenci hakkında geçti/kaldı kararı verilir.
- Belge veya sertifika veriliyorsa bu testler kullanılır.
- Bu testler sayesinde hedef davranışların kazanılıp kazanılmadığına bakılır.
- Vizeler, finaller sonuç görmeye yönelik sınavlardır.
- Kritik davranışların ölçülmesi sonuç görmeye yönelik sınavlardır.

**Örnek Sorular:**

Aşağıdaki maddelerde geçen cümlelerin ölçme, ölçüt, değerlendirme, ölçme kuralı, sınav kuralı olarak incelediğimizde aşağıdaki sonuçlar çıkacaktır.

1. Murat bilim proje yarışmasında 1. oldu. (Ölçme) 1. olmak sıralama ölçeğidir.
2. Serhat yüksek onur alarak sınıfı geçti. (Değerlendirme)
3. Ezgi iyi bir proje hazırladı. (Ölçme)
4. İş tecrübesi 3 yılın üzerinde olanlar işe alınacak. (Ölçüt)
5. Her soru 10 puan değerindedir. (Ölçme kuralı)
6. Sınavda hesap makinesi kullanmak yasaktır. (Sınav kuralı)
7. Yasemin pekiyi aldı. (Ölçme)
8. Melih KPSS'den 91 alarak 1. oldu. (Ölçme) (Bir karar belirtilmediği için ölçmedir).

***Ölçme bir durum tespittir. Değerlendirme ise bir karara varmaktır.***

## ÖLÇMEDE HATA

Hata, ölçme sonuçlarına genellikle istenmeyen değişkenlerin karışmasıdır. Hataların bir kısmı kontrol dışında ölçme sonuçlarına karışabilir. Ölçmede hata, bir özelliğin gerçek değeri ile ölçülen değeri arasındaki farktır. Hataların azaltılması ölçme işleminin güvenilirliğini ve geçerliğini artıracaktır.

Ölçme işlemleri içerisinde hata karışma ihtimali her zaman vardır. Çünkü insanın olduğu her yerde hata vardır. Ölçme işlemleri yapılırken de ölçme sonuçlarına istemli veya istemsiz şekilde hatalar karışır. Bu hata şekilleri *sabit hata*, *sistematik hata* ve *tesadüfi hata* olmak üzere üçe ayrılır. Bunları aşağıdaki gibi sıralayabiliriz:

### a. Sabit Hata

Ölçmeden ölçmeye ölçmeciden ölçmeciye değişmeyen ve bütün ölçme sonuçlarına karışan hatalardır. Bir ölçme işleminde özellikle hatalardan bahsederken “her” kelimesi geçiyorsa, buradaki hata çok yüksek ihtimalle sabit hatadır. Sabit hatada miktar herkes için aynıdır. Hatanın yönü ve miktarı belli olduğu için düzeltilebilir.

1. Cetvelle ölçüm yaparken her seferinde 0 (sıfır) yerine 1 (bir)’den başlamak sabit hatadır.
2. Öğretmenin yaptığı sınav sonuçlarının tamamına (yani her bir öğrencinin sınav sonucuna) +15 puan eklemesi sabit hatadır. Ancak her öğrenciye +15 puan eklendiğinde öğrencilerden birinin toplam puanı 100’ü geçiyorsa burada artık sabit hata değil sistematik bir hata var demektir.
3. Bir tartının her ölçme işleminde 300 gram fazla göstermesi sabit hatadır.
4. Kolumuzdaki saate her bakışımızda saatimizin 10 dakika ileri de göstermesi sabit hatadır.

### Not:

- Bir öğretmenin her öğrenciye +5, +10 veya +15 vermesi bir hatadır. Çünkü ölçme ve değerlendirme sürecinin amacı olanı olduğu gibi ölçmektir.
- Bir ölçme işleminde sabit hata yapılması, sınavın öncelikle geçerliğini düşürür.
- Sabit hata yapılan bir sınavda sınavın, aritmetik ortalaması değişir ancak sınavın standart sapması, ranjı, varyansı değişmez.
- Bir sınavda örneğin zaman yetmemesi sebebiyle tüm öğrenciler son 5 soruyu işaretleyememişse burada sabit hata vardır.

## **Sistemantik Hata**

Ölçmeden ölçmeye belli oranda artış ya da azalış gösteren hatalardır. Öğretmenin puanlama da taraf tutması ile ortaya çıkar. Ölçme sonuçlarında oluşan hatanın miktarı değişiklik gösterecektir. Hatanın yönü, miktarı veya oranı belli olduğu için gerektiğinde düzeltme yapılabilir.

1. Öğretmenin, sınav sonucunda her bir öğrenciye aldıkları puanın %10 fazlasını vermesi sistemantik hatadır.
2. Öğretmenin yazısı güzel olanlara +15 puan vermesi sistemantik hatadır.
3. Öğretmenin hemşerilerine +10 puan vermesi sistemantik hatadır.

### **Not:**

- Bir ölçme işleminde sistemantik hata yapılması sınavın öncelikle geçerliğini düşürür.
- Sistemantik hata yapılan bir sınavda sınavın aritmetik ortalaması, standart sapması, ranjı, varyansı değişir.
- Öğretmenlerin kullandıkları kanaat notları sistemantik hatadır çünkü burada yanlılık vardır.
- 30 cm'lik bir cetvelin ucu bir santimetrelik kırılmışsa, 30 cm'den daha kısa ölçümlerde sabit hata verir ancak yapılan ölçümler 30 cm'den büyükse ve tekrarlı bir ölçüm yapılıyorsa bu durumda sistemantik hata var demektir.

### **Sabit hata ile sistemantik hata arasındaki farklar:**

- Kefeli bir terazide kaç kilogramlık ölçüm yaparsak yapalım her seferinde +100 gr fazla çıkıyorsa sabit hata; kefeli terazide ölçüm yaparken 1000 gr ölçtüğümüzde sonuç 1100 gr çıkıyorsa, 2000 gr ölçtüğümüzde sonuç 2200 gr çıkıyorsa ve her seferinde bu şekilde bir sistemantikle yükseliyorsa burada sistemantik hata vardır.
- Bir öğretmen hazırladığı bir sınavda toplam 41 soru sormuş ve “Her soru 2,5 puan değerindedir” demiştir. Burada tüm sorular doğru cevaplandığında toplam 102,5 puan alınabilecektir. Burada dikkat edilmesi gereken nokta, eğer öğrenci 41. soruyu doğru cevaplırsa puan alacağı için burada sistemantik bir hata vardır. Bu durumda burada sabit hatadan söz edemeyiz.

## **Tesadüfî (Rastlantısal/Random) Hata**

Ölçme sonuçlarına nasıl karıştığı bilinmeyen, yönü, miktarı ve nedeni başka bir deyişle kaynağı belli olmayan hatalardır.

1. Kaygı ve stres nedeniyle öğrencinin sınavdan düşük not alması tesadüfî hatadır.
2. Bir öğretmenin bazı öğrenciye +5 puan bazısına +3 puan verilmesi tesadüfî hatadır.
3. Öğretmenin sınav kâğıdındaki puanları toplarken yanlış toplaması tesadüfî hatadır.

## **HATA KAYNAKLARI**

Ölçme sonuçlarına karışan hataların çeşitli sebepleri vardır. Hataların ortadan kaldırılması ölçme işlemini daha doğru yapmamızı sağlayacaktır. Çünkü amacımız en doğru ölçümü yapıp, gerçek puanı bulmaktır. Hatanın kaynaklarını aşağıdaki gibi sıralayabiliriz:

### **1. Öğrenciden Kaynaklanan Hatalar**

Öğrencinin sınava geç gelmesi, sınav esnasında hasta olması, sınav sebebiyle kaygılı olması sınavda kaydırma yapması gibi durumlar öğrenciden kaynaklı hatalardır. Bu hataların ortadan kaldırılması için sınav saatinin herkesin rahatça gelebileceği bir saatte olması ayrıca ailenin, öğretmenlerin ya da sınav yapan kurumun öğrencileri sınav hakkında bilgilendirmesi ve öğrencilerin kaygısını azaltması önerilebilir.

### **2. Ölçme Aracından Kaynaklanan Hatalar**

Soruların öğrenciler tarafından yeterince anlaşılamaması ve yazıların öğrenciler tarafından okunmasında sıkıntılar yaşanması, sınav içerisindeki soruların yanlış olması, birden fazla seçeneğin doğru veya yanlış olması, soruların işlenen konuları iyi temsil etmemesi gibi nedenler ölçme aracından kaynaklanan hatalardır.

### **3. Ölçmeciden (Öğretmenden) Kaynaklanan Hatalar**

Sınavın olması gerektiğinden daha zor ya da kolay olması (sınavlar, bilenle bilmeyeni birbirinden ayırt edebilecek düzeyde olmalıdır), sınav kâğıtlarının okunurken tarafsızlığın korunamaması, yorgunluk, dikkatini toplayamamak gibi nedenlerle puanlamanın yanlış yapılması, sınavda uygulanacak test türünün öğrencilere söylenenden farklı yapılması gibi nedenler öğretmen kaynaklı hatalardır.

#### 4. Ortamdan Kaynaklanan Hatalar

Sınavın yapıldığı mekânın sınav yapma koşullarını yeterince taşınamaması sebebiyle ölçme işlemi tam olarak gerçekleştirilemeyebilir. Örneğin sınıfın oldukça sıcak ya da soğuk olması, gürültülü olması, kötü kokması, yeterince ışık alamaması gibi nedenler sınav ortamından kaynaklanan hatalar arasındadır.

#### 5. Ölçme Yönteminden Kaynaklanan Hatalar

Ölçme işlemi gerçekleştirilirken ölçülecek davranışın derecesini belirleyebilmek için ölçme yöntemini de uygun seçmek gerekmektedir. Matematik sınavının sözlü olarak yapılması, aynı test türünün tüm sınavda uygulanması, iki seçenekli sınav yapılması gibi durumlar olanı olduğu gibi ölçmeyeceği için ölçme yönteminden kaynaklanan hatalar arasındadır.

**Not:** Bir sınavda birden fazla test türünün kullanılması daha doğrudur. Yani sınavın içerisinde yazılı yoklama, çoktan seçmeli testler, kısa cevaplı testler, doğru-yanlış, eşleştirmeli sorular gibi her türden sınav çeşidi olmalıdır.

### KORELASYON

#### Korelasyon (ilişki, bağlantı, bağıntı)

İki değişken arasındaki ilişkinin yönünü ve miktarını verir. Korelasyon genel olarak üçe ayrılır. Bunlar *pozitif korelasyon*, *negatif korelasyon* ve *nötr korelasyon* olarak üçe ayrılmaktadır.

##### A. Pozitif Korelasyon

İki değişken arasında doğru orantı varsa, aynı yönlü bir ilişki varsa, benzer özellikleri ölçüyorsa burada pozitif korelasyon vardır. Pozitif korelasyon 0.00 ile +1.00 arasından değer alır. Yani iki değişkenden biri artarken diğeri de artıyorsa veya biri azalırken diğeri de azalıyorsa burada pozitif korelasyon vardır.

1. Hava sıcaklığının artması ile klima satışlarının artması arasında pozitif bir korelasyon vardır.
2. Ders çalışma saatindeki artış ile sınavdan alınan puanlardaki artış arasında pozitif bir korelasyon vardır.



3. Televizyonda bir markanın reklamının azalması ile o markanın ürünlerinin satışının azalması pozitif bir korelasyon vardır.

## **B. Negatif Korelasyon**

İki değişken arasında ters orantı varsa, zıt yönlü bir ilişki varsa, farklı özellikleri ölçüyorsa burada negatif korelasyon vardır. Negatif korelasyon 0.00 ile -1.00 arasından değer alır. Yani iki değişkenden biri artarken diğeri de azalıyorsa veya biri azalırken diğeri de artıyorsa burada negatif korelasyon vardır.

1. Enflasyon arttıkça insanların alım gücü arasında negatif bir korelasyon vardır (Biri artarken diğeri azalmaktadır).
2. Uyku arttıkça insanın dikkatini toplaması arasında negatif bir korelasyon vardır.
3. Şehirlerin nüfusunun artması ile köylerin nüfusunun azalması arasında negatif bir korelasyon vardır.

**Not:** Tekrar ile hatırlama arasında pozitif korelasyon, tekrar ile unutma arasında negatif korelasyon vardır.

## **C. Nötr Korelasyon**

İki değişken arasında herhangi bir ilişkinin, bağlantının olmadığı durumları ifade eder. Korelasyon değeri sıfıra (0.00)'a yaklaştıkça iki değişken arasında ilişkinin azaldığı görülür.

1. Türkiye'deki kitap okuma oranının artması ile kahve tüketiminin artması arasında nötr bir korelasyon vardır.
2. Matematikte çok başarılı olanlarla 1994 yılında doğanlar arasında nötr bir korelasyon vardır.
3. Çok yemek yiyenlerin cesur olması arasında nötr bir korelasyon vardır.

**Not:** İki değişken arasındaki değerler +1 ve -1'e yaklaştıkça, iki değer arasında ilişkinin arttığından söz edilebilir. - ve + değerler ilişkinin yönünü belirlerken, - ve + 'dan sonraki sayılar ise ilişkinin miktarını verir. - ve + 'dan sonraki değer, 0 (Sıfır)'a ne kadar yaklaşırsa iki değişken arasındaki bağlantıdan o kadar az olduğundan söz edilebilir.

**Uyarı:** Korelasyon değeri yüzde hesabı olarak kullanılamaz. Korelasyon değeri başarı ya da başarısızlığı yorumlama da kullanılamaz.

# İYİ BİR ÖLÇME ARACININ SAHİP OLMASI GEREKEN ÖZELLİKLER

İyi bir ölçme aracının sahip olması gereken özellikler *güvenirlilik*, *geçerlik* ve *kullanışlıktır*.

## 1. GÜVENİRLİK

Güvenilir bir ölçme aracı, aynı özellikle ilgili olarak ardarda yapılan ölçmelerde yaklaşık olarak benzer sayısal sonuçları vermelidir. Bir testin aynı gruba iki ya da üç kez uygulandığında, gruptaki her bir kişinin bütün uygulamalarda yaklaşık olarak aynı puanı alması testin güvenilir olduğunu gösterir. Bir testten elde edilen puanların güvenirliliğine birçok etken etki etmektedir. Bu etkenler, testin uzunluğu, testi alan öğrenci, testin uygulanma koşulları ile uygulandığı grup ve testin puanlamasıdır (Tekin, 2004).

Koşullar ve şartlar değişmeden, birbiri ardına yapılan denemelerde aynı sonucun elde edilmesidir. Güvenirlilik 0 (Sıfır) ile +1 (bir) arasında bir değer alır. Eğitim bilimlerinde bir testin güvenilir olarak kabul edilebilmesi için güvenirlilik katsayısının 0.70 ve üzeri olması gerekmektedir. Ayrıca güvenirlilik, ölçme sonuçlarının tesadüfi hatalardan arınık olmasıdır. Bir ölçme aracının güvenirlilik katsayısı 0.70'in üzerindeyse; ölçme aracının tesadüfi hatalardan arınık olduğu, ölçme ortamının uygun olduğu, soruların açık ve anlaşılır olduğu, öğrencilerin başarı sıralamalarının benzer olduğu, yapılan ölçme sonuçlarının benzer olduğu, sınav uygulanan grubun heterojen olduğu, sınavı oluşturan maddelerin aynı özellikleri ölçtüğü, testin tek boyutlu olduğu gibi yorumlar yapılabilir.

**Güvenirlilik kavramlarını *duyarlılık*, *kararlılık*, *objektiflik*, *iç tutarlılık* ve *tutarlılık* olarak sıralayabiliriz.**

**1. Duyarlılık (Hassaslık):** Birim sayısının çok olmasıdır. Örneğin bir sınavda 40 adet sorunun olması ile başka bir sınavda sadece 4 adet sorunun sorulması karşılaştırıldığında, 40 soruluk testin daha duyarlı olduğu söylenecektir. Çünkü daha ayrıntılı ve daha net sonuç verecektir. Buna benzer şekilde bir kişinin ağırlığını kantarla ölçüp 74 kg ile aynı kişiyi elektronik terazi ile ölçtüğümüzde 74,6 kg olarak gördüğümüzde elektronik terazinin daha duyarlı ve daha net sonuç verdiği görülecektir. Soru sayısının çok olması daha duyarlı sonuç

vereceği gibi sırf daha fazla duyarlı olsun diye, öğrencilere çok fazla soru sormak onların sıkılmalarına sebep olabilir.

**2. Kararlılık:** Aynı ölçme aracının aynı öğrenci grubuna belirli aralıklarla uygulanması sonucunda benzer sonuçların elde edilmesidir. Uygulanan testin veya sınavın her seferinde benzer sonuçlar vermesidir. Örneğin öğrenciler bir zekâ testine tabi tutulmuşsa bu zekâ testi ne zaman uygulanırsa uygulansın sonuçların birbirine yakın olmasıdır. Bununla birlikte aynı ölçme aracının uygulanması önemlidir. Bankadan para çektiğimizde parayı iki kez kendimizin sayması kararlılığını verecektir.

**3. Objektiflik (Yansızlık):** Puanlamanın yansızlığıdır. Uygulanan sınava kim girerse girsin, kim uygularsa uygulasin, sınav kim tarafından okunursa okunsun alınan sonuçlara hiçbir yanlılığın katılmamasıdır. Sınav kâğıdının üzerinde yazan ismin ölçme sonuçları açısından hiçbir öneminin olmamasıdır. Örneğin bir öğrenci, sınav sonucu olarak 76 alacaksa sınav kâğıdını kim okursa okusun sonucun değişmemesidir. Ayrıca standart hatanın düşük olması güvenilirliğin göstergesidir.

**4. İç Tutarlılık:** Uygulanan sınavın veya testin içerisindeki sorular aynı veya benzer özellikleri ölçmesidir. Örneğin 45 soruluk Coğrafya Testini öğrencilere uyguluyorsak, 45 sorunun tamamının Coğrafya ile ilgili olmasıdır. Ayrıca öğrenciler belli bir dersin belirli bir konusu için sınava tabi tutulmuşlarsa, soruların tamamının o konuyla ilgili olması gerekmektedir. Testin homojen olması yani testin tek boyutlu olmasıdır.

**5. Tutarlılık:** Birbirine benzer özelliklere sahip iki testin aynı gruba uygulanması sonucunda benzer sonuçların elde edilmesidir. Yani farklı ölçme araçları ile özelliğin ölçülmesidir. Bankadan para çekince önce bizim saymamız sonrasında da arkadaşımıza parayı saydırmamız örnek olarak verilebilir.

### **Güvenirliği Hesaplama Yöntemleri:**

**a) Birden Fazla Uygulama Yapılırsa:** Uygulanan bir testin güvenilirliğini sağlayabilmek için testin birden fazla kez uygulanmasıdır. Aynı veya benzer sonuçlar alınması hedeflenir. Birden fazla uygulama yöntemi kısaca ikiye ayrılır. Bunlar aşağıdaki gibi sıralanır.

- i. Test-Tekrar Test:** Öğrencilere uygulanan testin güvenilirliğini sağlayabilmek için aynı testin aynı öğrencilere belirli bir zaman sonrasında tekrar yapılmasıdır. Bu yöntemde aynı test öğrencilere hemen uygulanmamalı ve aynı zamanda çokta geciktirilmemelidir. Aynı test 3-4 hafta içerisinde tekrar yapılmalıdır. Öğrencilerin ilk uygulama ile ikinci uygulama arasındaki sonuçlar birbirine benziyorsa, testin

güvenirliğinden söz edebiliriz. Örneğin Melike ilk uygulamadan 58, ikinci uygulamadan da 60 almışsa testin güvenilirliğinden söz edebiliriz. Aynı durum sınıf ortalaması temel alınarak da yapılabilir.

Başarı testlerinde tercih edilmemelidir. Burada hata kaynağı zamandır çünkü iki test arasında beklemek gerekmektedir. Test- tekrar test yönteminde elde edilen güvenilirlik katsayısı **kararlılık katsayısı** olarak adlandırılır.

- ii. **Paralel Test (Eşdeğer Form):** Uygulanan sınavın paralelinde, eşdeğerinde/muadilinde başka bir sınavın öğrencilere uygulanmasıdır. Öğrencileri İngilizce dersinden final sınavı yaptığımızda alınan sonuçlar ile final sınavının muadili olan başka bir sınavı da öğrencilere uyguladığımızda iki sınavın sonuçları arasındaki ilişkinin karşılaştırılmasıdır.

Burada hata kaynağı benzer sınavın hazırlanmasının zorluğudur. Çünkü benzer kapsamda, benzer soru türünde, benzer soru sayısında, benzer soru güçlüğünde, benzer davranışları ölçen soru hazırlamak oldukça zordur. Her iki testte öğretmen tarafından hazırlanmaktadır. Uygulanan iki testten elde edilen güvenilirlik katsayısı **tutarlılık katsayısı** olarak adlandırılır.

**b) Tek Uygulamayla Güvenirlik Belirleme:** Burada öğrencilere uygulanan tek bir sınav ile sınavın güvenilirliğinin sağlanması vardır. Tek uygulamayla güvenilirlik belirleme yöntemi kısaca ikiye ayrılır. Bunlar aşağıdaki gibi sıralanır.

- i. **Paralel (Eşdeğer) Yarılar:** Öğrencilere uygulanan testin ilk bölümü ile ikinci bölümünün birbiriyle karşılaştırılmasıdır. Ancak burada dikkat edilmesi gereken bir durum vardır. Örneğin öğrencilere 100 soru sorulmuş ise ilk 50 soru ile ikinci 50 sorunun birbiriyle karşılaştırılması anlaşılmamalıdır. Çünkü öğrenci sınavın son sorularına doğru dikkatini toparlayamamış, rastgele işaretlemiş, uykusu gelmiş olabilir. Bu sebeple öğrencinin sınav sorularından tek numaralı sorular ile çift numaralı sorulara verdiği cevapların karşılaştırılması yapılmalıdır. Burada Spearman&Brown formülü kullanılmalıdır.

**Spearman-Brown Formülü**

$$r_x = \frac{2r_1}{1 + r_1}$$

Burada sorular kolaydan zora doğru hazırlanmalıdır. Soru sayısı mutlaka çift olmalıdır. Ayrıca testin tüm güvenilirliği testin yarı güvenilirliğinden fazladır. Çünkü testin tümündeki soru sayısı testin yarısındaki soru sayısından fazla olmasıdır.

## ii. Maddelerin Tutarlılığına Dayalı Yöntemler:

**KR-20 ve KR-21 (İç tutarlılık güvenilirlik yöntemi):** Bir testte yer alan tüm maddelerin birbirleriyle ne derece tutarlı olduğunu belirlemek için kullanılır. İç tutarlılık güvenilirliğini hesaplamak için Kuder ve Richardson tarafından geliştirilmiş olan **KR-20** ve **KR-21** formüllerinden yararlanılır. Bu formülle elde edilen güvenilirlik katsayısı yüksekse, test maddeleri birbirleriyle tutarlıdır. KR-20 ve KR-21 ile elde edilen sonuçlar ne derece yüksek ise güvenirlikte o oranda yüksek demektir. Hem KR-20 hem de KR-21; doğruların 1 yanlışların ve boş bırakılan soruların ise 0 kabul edildiği testlerde kullanılırlar. Test bir kere uygulanır ve testte yer alan tüm sorular arasında teker teker korelasyona bakılır. Elde edilen KR değeri 0,70'ten büyükse test güvenilirirdir.

**KR-20:** Soruların güçlük düzeylerinin farklı olduğu, madde istatistiklerinin bilindiği veya hesaplanabildiği, madde güçlüğüne/madde varyansının/madde ayırt ediciliğinin bilindiği durumlarda karşımıza çıkar. KR-21'e göre daha duyarlıdır çünkü madde istatistikleri farklıdır.

**KR-21:** Soruların güçlük değerlerinin aynı, test istatistiklerinin (aritmetik ortalama, standart sapma) bilindiği durumlarda karşımıza çıkar. **KR-21** ile testin sapması, dağılımın homojenliği ve iç tutarlılık katsayısı hesaplanır.

**Not:** KR-20 madde güçlüklerinin farklı olması sebebiyle KR-21'den daha büyüktür. Ancak madde güçlüklerinin aynı olduğu durumlarda  $KR-20 = KR-21$  olarak çıkar.

**Cronbach Alpha:** Doğruların 1 yanlışların 0 olarak kabul edildiği testlerde kullanılabildiği gibi aynı zamanda Likert Tipi ölçeklerinde güvenilirliğini hesaplamak için kullanılır. Tüm güvenilirlik hesaplamalarında olduğu gibi burada da testin güvenilirliğinin 0,70 ve üzeri çıkması gerekmektedir. KR-20 ve KR-21'den likert tipi ölçeklerde de kullanılmasıyla ayrılır.

### Güvenirlik Belirleme Yöntemlerinin Karşılaştırılması

Güvenirlik Belirleme Yöntemi	Anlamı	Form Sayısı	Uygulama Sayısı	Hata Kaynağı
Test-Tekrar Test	Kararlılık	Bir	İki	Zaman
Paralel Formlar	Tutarlılık	İki	İki	Kapsam
Eş Değer Yarılar (Testi Yarılama)	İç Tutarlılık	Bir	Bir	Kapsam
KR-20 KR-21	İç Tutarlılık	Bir	Bir	Kapsam ve heterojenliği
Cronbach Alpha	İç Tutarlılık	Bir	Bir	Kapsam ve heterojenliği

**Puanlama Güvenirliği:** Ölçme işlemini gerçekleştiren kişinin sınav kâğıdını puanladıktan belirli bir süre sonra tekrar puanlamasıdır. Örneğin Ölçme ve Değerlendirme dersi ders öğretmeninin sınav kâğıdını puanlaması ve aradan geçen üç sonrasında yeniden puanlamadır. Aynı işlemi aynı kişinin iki kez yapması kararlılıkla ilgilidir ancak buradaki amaç **objektifliği** sağlamaktır.

**Puanlayıcı Güvenirliği:** Yapılan bir ölçme işleminde puanlamanın en az iki farklı kişi tarafından puanlanmasıdır. Örneğin yapılan sınavı Ölçme ve Değerlendirme ders öğretmeni puanladıktan sonra aynı sınav kâğıdını başka bir kişi tarafından puanlanmasıdır. Burada aynı ölçme puanlama işleminin iki farklı kişi tarafından yapılması tutarlılıkla ilgilidir ancak buradaki amaç **objektifliği** sağlamaktır.

#### Testteki Hatayı Puan Cinsinden İfade Etme ve Ölçmenin Standart Hatası:

Güvenirlik katsayısı ölçmede yapılan hataların miktarını belirtmez. Ölçmeye karışan hatayı puan cinsinden ifade etmek için kullanılır. Ölçmede yapılan hataların standart değerini hesaplamak için şu yöntem kullanılır:

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

$S_e$ : Ölçmenin standart hatası

$S_x$ : Test puanlarının standart sapması

$r_x$ : Testin güvenilirliği

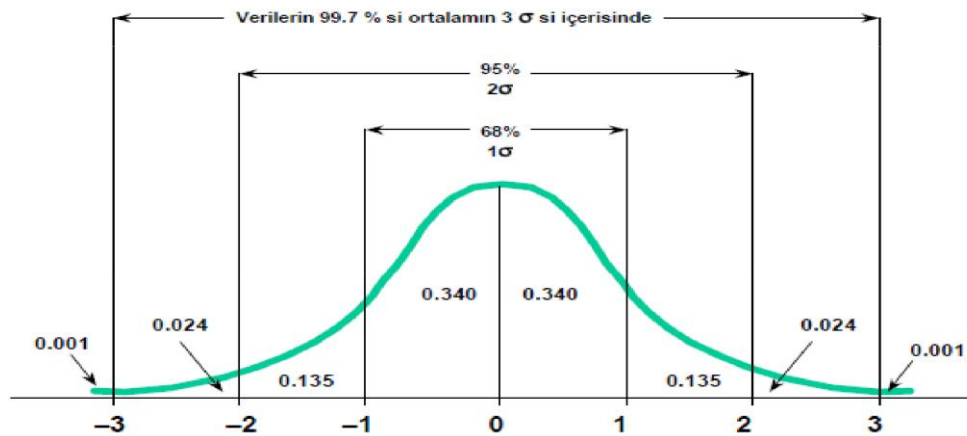
Standart sapma yükseldikçe standart hata yükselir, standart sapma düştükçe standart hata düşer. Bununla birlikte güvenilirlik arttıkça standart hata azalır, güvenilirlik azaldıkça standart hata artar.

Standart hata, ölçme işlemine karışan hataları standart hale getirmek amacıyla hesaplanmaktadır. Ortaya çıkan sonuç ölçme işleminin sonucuna ekleme veya çıkarma yapılarak en doğru sonuca ulaşmaya çalışılır.

Bir ölçme sonucunun doğruluğu için üç farklı hesaplama (%68, %95, %99) karşımıza çıkar. Bunu bulabilmek için öncelikle ölçmenin standart hatası bulunmalıdır. Bulunan standart hata, ölçülen bireyin ölçümüne eklenir veya çıkarılır ve bu açıklık ölçülen bireyin ölçüm aralığını gösterir.

Örneğin standart hata 5 ve ölçülen bireyin sınav sonucu 50 ise; %68 ihtimalle (%32 yanılma payı) öğrencinin puan aralığı 45-55, aynı bireyin %95 ihtimalle (%5 yanılma payı) puan aralığı 40-60, %99 ihtimalle (%1 yanılma payı) puan aralığı 35-65 arasındadır. Bu üç ihtimalden başka bir ihtimal olamaz. Çünkü oluşturulan bu değerler normal dağılım eğrisi kullanılarak hesaplanmıştır.

Normal dağılım eğrisi aşağıda gösterilmiştir:



Bir testte veya bir sınavda güvenilirliği artırmak için;

- i. Soru sayısı artırılmalıdır.
- ii. Sorular açık, net ve anlaşılır olmalıdır.

- iii.** Sınavın yönergesi olmalıdır.
- iv.** Madde güçlük indeksinin 0.50 civarında çıkması gerekir.
- v.** Sınav koşullarının tüm öğrenciler için eşit olması gerekmektedir.
- vi.** Sınav evrakındaki yazıların puntolarının ve yazı karakterlerinin öğrencilerin rahatlıkla okuyacağı şekilde olması gerekmektedir.
- vii.** Şans başarısını azaltmak amacıyla çoktan seçmeli sorularda seçenek sayısı artırılmalıdır.
- viii.** Sınavın sorularının homojen, sınavın uygulandığı grubun heterojen olması güvenilirliği olumlu yönde etkileyecektir.

## **2. GEÇERLİK**

Ölçme aracının amacına hizmet etme derecesidir. (-1) ile (+1) arasında değişir. Bir ölçme aracı +0,30 üzeri ise geçerlidir ancak geçerlikte +0.40 ve üzeri ideal olanıdır. 20 soru ile bir Ölçme Ve Değerlendirme dersi bilgisi ölçmek için eğer aynı test içerisinde Öğretim İlke Ve Yöntemleri dersinin bilgisini ölçmeye kalkarsak sıkıntı olur.

Örneğin: Kendi üniversitesinde okuyacak olan öğrencilere iş garantisi vadeden özel bir üniversite, birinci sınıfın derslerinden olan İngilizce dersinin final sınavında öğrencilerine Fizik dersi sorularını sormuştur. Düşük not alan öğrenciler, mahkemeye başvurmuş ve mahkeme yapılan final sınavının “amaca hizmet etmediği” gerekçesiyle iptal etmiştir. Çünkü sınavın geçerliği yok hükmündedir.

**Not:**“Bir ölçme aracı tamamen geçerli ya da tamamen geçersiz” diyemeyiz.

Geçerliğin çeşitleri bulunmaktadır. Bunları aşağıdakiler gibi sıralayabiliriz:

- 1. Kapsam Geçerliği**
- 2. Bir Ölçüte Dayalı Geçerlik**
  - a.** Yordama Geçerliği
  - b.** Uygunluk Geçerliği
- 3. Yapı Geçerliği**
- 4. Görünüş Geçerliği**
- 5. Sonuçsal Geçerlik**



### a) Kapsam Geçerliği

Bir yıl ya da ya da bir dönem içerisinde aktarılan bilgilerin tamamını ölçecek şekilde sınav sorusunun hazırlanmasıdır. Kısaca, soruların konuları örnekleyebilmesidir. Örneğin öğrenciler, kitabın ilk 100 sayfasından sorumlu tutulmuşsa; ölçme aracımız yani sınavımız 100 sayfayı kapsayacak şekilde soru barındırmalıdır. Anlatılan içeriğin (kapsamın) tamamını içerisinde barındıracak şekilde soruların hazırlanması sayesinde ilgili konulara hangi öğrenci ne kadar hazırlanmışsa sınavdan da o kadar not alması sağlanacaktır.

*Kapsam geçerliği sağlamak için daha önceden geçerliği ve güvenirliği yapılmış sınavlar öğrencilere uygulanabileceği gibi bir öğretmen bir **uzmana başvurarak da** sınav hazırlayabilir.* Ancak bir uzmana danışma imkânının her zaman olmaması ve öğretmenin soracağı sorularında geçerliği ve güvenirliği yapılmamışsa bir öğretmen; *Belirtke Tablosu* hazırlayarak kapsam geçerliğini sağlayabilir.

### Belirtke Tablosu

	Hedef Alanları					
Konular	Bilgi	Kavrama	Uygulama	Analiz	Sentez	Değerlendirme
1. Ünite	1	1	1	1	1	1
2. Ünite	1	1	1	1	1	1
3. Ünite	1	1	1	1	1	1
4. Ünite	1	1	1	1	1	1
5. Ünite	1	1	1	1	1	1

Sınav soruları hazırlanırken her ünitenin her hedef alanından soru hazırlanması sayesinde kapsam geçerliği sağlanmış olacaktır. Ancak öğretmen örneğin 3. Üniteden daha fazla soru sorabileceği gibi 1. Üniteden daha az soru sorabilir. Kapsam geçerliği açısından her üniteden eşit şekilde soru sorulması daha uygundur. Ayrıca yukarıdaki tabloya uygun şekilde soru hazırlandığında her hedef alanının her basamağında soru hazırlanmış ve öğrencilerin bilgilerinin ne düzeyde olduğu da tespit edilmiş olacaktır. Belirtke tablosu soruların niteliği, soru sayısı, soruların konusu ve üniteleri, testin türü gibi konularda bilgi verir.

### Not:

- Başarı/izleme testlerinde öncelikli olarak kapsam geçerliğine bakılır.

- Kapsam geçerliği için mutlak ölçüt kullanılmalıdır.

## **b) Bir Ölçüte Dayalı Geçerlik**

Seçme sınavlarında aranan özelliktir (Örneğin, Lisans Yerleştirme Sınavı, Kamu Personeli Seçme Sınavı). Öğrencilerin tanıma ve yerleştirmeye yöneliktir. Öğrencilerin ilgili bölümde başarılı olup olamayacağı tespit edilmeye çalışılır. Bir ölçüte dayalı geçerlik; yordama geçerliği ve uygunluk geçerliği olmak üzere ikiye ayrılır.

### **i. Yordama Geçerliği**

Yordamak, eldeki bilgilerden yararlanıp geleceğe yönelik tahminde bulunmaktır. Bilinenden tahminle bilinmeyenleri ortaya çıkarmaktır. Örneğin Yüksek Öğretime Giriş sınavı ile öğrencinin akademik başarısı arasındaki ilişki. Bununla birlikte bir öğrencinin 2012 yılında birincilikle girdiği bir bölümü 2016 yılında birincilikle bitirmesi gibi tahminlerde bulunulmasıdır. **Seçme sınavında aranan bir özelliktir.**

Yordama geçerliğinde ölçüt daha sonra ortaya çıkmaktadır. Bir işe alım sınavında, işe alım sınavından yüksek not alan bir kişinin, aradan geçen belli bir süre sonunda iş performansı yüksekse ilgili sınavın yordama geçerliği yüksektir.

### **ii. Uygunluk Geçerliği (Eş zamanlılık, zamandaş, mevcut hal, uyum)**

Uygunluk geçerliğine iki farklı şekilde bakılması daha doğru olacaktır.

**1-** Uygunluk geçerliğinde ölçüt sınavın öncesinde bellidir. Bir işe alım sınavında, işe alım sınavından yüksek not alan bir kişinin, sınava girmeden önce de başarılı bir performansa sahipse ilgili sınavın uygunluk geçerliği yüksektir.

**2-** Bir öğretmenin hazırladığı bir ölçme aracını daha önce uzmanlarca hazırlanmış bir ölçme aracıyla karşılaştırmasıdır. Örneğin öğretmen tarafından hazırlanan bir sınavdan öğrencinin 45 alması ancak aynı öğrencinin bu sınava eşdeğer olacak ve uzmanlarca hazırlanmış başka bir sınav yapıldığında da 45 puanına yakın puan alması gerekir. Her iki sınavda da birbirine yakın puanlar alınıyorsa sınavımız *uygundur* denebilir. Ancak uygulanan iki sınavdan (içlerinden birinin geçerlik ve güvenirliği yapılmış ise) biri; uzmanlar tarafından hazırlanmışsa öğretmenin kendi hazırladığı sınavın uzmanların hazırladığı sınava yakın sonuçlar vermesi gerekir.

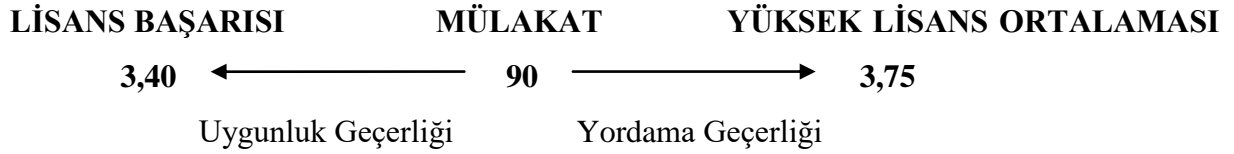
Bir öğrencinin girdiği sınavlar ve bu sınavlardan aldığı notlar ilgili tabloda verilmiştir.

Sınavlar ve Alınan Notlar		
Öğretmen Yapımı Sınav	ÖSYM Yapımı Sınav	MEB Yapımı Sınav
45	62	60

Tabloya bakıldığında Öğrenci Seçme ve Yerleştirme Merkezi ve Milli Eğitim Bakanlığı tarafından yapılan sınavlardan alınan notlar birbirine yakın sonuçlar vermiştir. Böyle bir durumda öğretmen yapımı sınavın biraz daha kolaylaştırılması gerektiği söylenebilir. Buradaki durum tam tersi şekilde de olabilirdi.

#### Not:

- Ölçüt puan, uygulanan testten önce elde edilmişse uygunluk geçerliği vardır.



#### c) Yapı Geçerliği

İnsanın zihninde olan direkt ölçülemeyen psikolojik özelliklerin ortaya çıkarılmasıdır. Bir ölçme aracının ölçmek istediği teorik yapıyı ölçüp ölçmediğiyle ilgilidir. Amaç olanı olduğu gibi çıkarmaktır. Örneğin bireyin çoklu zekâsı ölçülmek isteniyorsa bu durumda mevcut sınavla çoklu zekâ hangi kademelerden oluşuyorsa elimizdeki sınavda o kademelerden oluşmalıdır ve bu dağılımı sergilemesi gerekir. Bununla birlikte bir ölçme aracının ölçmek istediği özelliğe sahip olanla olmayanı ayırt edebilmesidir. Bir ölçme aracının ölçmek istediği özelliği başka bir özelliklerle karıştırmadan ölçmesidir. **Psikolojik testlerde özellikle aranan geçerliktir.** Örneğin zekâ testleri, kişilik testleri, ilgi testleri, kaygı testleri, stres testleri, yaratıcı düşünme testleri psikolojik testlere örnek olarak verilebilir.

Başka bir örneği ise şu şekilde verebiliriz. Bir bireyin hangi mesleğe daha yatkın olduğu ölçülmek istenmektedir. Ancak bu ölçme işlemindeki bulunan maddelerden 29. soru meslek seçimi ile ilgili değilse, 29. soru ölçme aracının yapı geçerliğini düşürmektedir. Genel

olarak bir testteki yapı geçerliği, test maddeleri arasındaki ilişkinin boyutunu gösterir. Her test maddesi, testi oluşturan diğer bütün maddelerle ilişkili olmalıdır.

**Not:**

- Yapı geçerliği olan testler, grup farklılıklarını ortaya çıkarabilmelidir.
- Ön test ve son test arasında farkın olabilmesi yapı geçerliğiyle ilgilidir.
- Öğretmenin sınav yaptığı grubu, soruları nasıl cevapladıklarıyla ilgili dinlemesi yapı geçerliğiyle ilgilidir.
- Yapılan sınavın cevap anahtarının ipucu oluşturmaması, sınavın yapı geçerliğini düşürür.

	A	B	C	D
1	X			
2		X		
3			X	
4				X
5			X	
6		X		
7	X			
8		X		
9			X	
10				X

**d) Görünüş Geçerliği**

Bir ölçme aracının ölçmek istediğini ne derecede ölçebildiğiyle ilgilidir. Ayrıca testin kapağında yazanla içeriğinin yani içindekilerin benzer olmasıdır. Karşımızdaki muhatap olduğumuz kişi, testi eline aldığı anda “Türkçe Testi” yazıyorsa kitapçığın içerisindekilerin Türkçe dersi ile ilgili olduğunu anlamasıdır. Burada muhatapın ilk tesiri önemlidir. Ancak psikolojik testlerde görünüş geçerliği olmayabilir.

**e) Sonuçsal Geçerlik**

Çağdaş ölçme değerlendirme yaklaşımlarının değerlendirilmesidir. Alternatif ölçme araçlarında aranan bir özelliktir. Yapılandırmacılığa dayalı olarak ortaya çıkmıştır.

**Geçerliğı artırmak için:**

- i. Sınav öncesi sorular öğrencilere verilmemelidir.
- ii. Sınav esnasında öğrencilere ipucu verilmemelidir.
- iii. Her yıl aynı sorular sorulmamalıdır.
- iv. Kopyaya imkân verilmemelidir.
- v. Sorunun cevabı sorunun bünyesinde veya testin içinde olmamalıdır. Örneğin testteki 7.Sorunun cevabının testteki 21. soruda verilmesi gibi.
- vi. Jestler ve mimikler ile öğrencilerin doğru cevabı bulması sağlanmamalıdır.
- vii. Sabit ve sistematik hatadan kaçınılmalıdır.
- viii. Test planı hazırlanmalıdır.

**Not:**

- Bir testin güvenirlik katsayısının karekökü, o testin geçerliğinin varabileceğı üst sınırıdır.

$$\sqrt{0,81} = 0,90$$

$$\sqrt{0,64} = 0,80$$

**GEÇERLİK İLE GÜVENİRLİK ARASINDAKİ İLİŞKİ**

- Bir ölçme aracı geçerliyse aynı zamanda güvenilirirdir.
- Bir ölçme aracı güvenilir olsa bile geçerli olmayabilir ancak geçerli olması için güvenilir olması ön koşuldur.
- Güvenirlik sadece tesadüfî hatalardan etkilenir.
- Geçerlik; sabit ve sistematik hatadan doğrudan etkilenirken tesadüfî hatadan dolaylı olarak etkilenir. Çünkü güvenirlik düştükçe geçerlik sınırlanır.

**3. KULLANIŞLIK**

Bir ölçme aracının zaman, emek ve materyal yönünden ekonomik olmasıdır. Yani hazırlanacak olan bir ölçme aracı hazırlama, çoğaltma, uygulama, puanlama yönünden ekonomik olmalıdır. Ancak ölçme aracımız sırf kullanışlı olsun diye geçerlik ve güvenirlik tehlikeye atılmaz.

## İNSAN DAVRANIŞLARININ AŞAMALI SINIFLANDIRILMASI

### BLOOM'UN TAKSONOMİSİ

	Bilişsel Alan	Duyuşsal Alan	Psiko-Motor (Devinişsel) Alan
1	Bilgi	Alma	Algılama
2	Kavrama	Tepkide Bulunma	Kurulum
3	Uygulama	Değer Verme	Kılavuz Denetiminde Yapma
4	Analiz	Örgütlenme	Mekanikleşme
5	Sentez	Kişilik Haline Getirme	Beceri Haline Getirme
6	Değerlendirme		Uyarlama/Uyum
7			Yaratma

Bilişsel alan ve güncellenen bilişsel alan hakkında yer alan bilgilendirmeler aşağıdadır.

	Bilişsel Alan	Bilişsel Alanın Güncel Hali	
1	Bilgi	Hatırlama	Tanımlar, düzenler, açıklar, toplar, tarif eder, yineler, sayar.
2	Kavrama	Anlama	İlişkilendirir, değiştirir, netleştirir, sınıflar, düzenler, karşılaştırır.
3	Uygulama	Uygulamak	Uygular, hesaplar, değiştirir, seçer, düzenler, gösterir, geliştirir.
4	Analiz	Analiz etmek	Analiz eder, değer biçer, düzenler, kategorize eder, sınıflar.
5	Sentez	Değerlendirme	Aslını öğrenir, tartışır, değerlendirir, ilişitir, sonuçlandırır.
6	Değerlendirme	Yaratma	Yaratır, tasarlar, geliştirir, planlar, yerleştirir, formüle eder.

### Test Çeşitleri

Ölçme sonuçlarındaki hata miktarının en aza indirilebilmesi için uygulanacak olan testlerin yani sınavların amaca hizmet etmesi gerekmektedir. Hangi test türünün hangi derslerde uygulanmasının daha uygun olacağını bilmek hem öğretmenin işini kolaylaştıracaktır hem de ölçme sonuçlarına karışan hataların oranını azaltacaktır. Bu sebeple bir test hazırlanırken hangi aşamalardan geçeceği aşağıdaki gibi anlatılmıştır.

## Test Nedir?

Test, bireyin davranışlarını standart altında gözlemlemek ve tanımlamak amacına yönelik sistematik bir yöntemdir.

## Test Hazırlama Süreci

Öğrencilere uygulanacak olan bir testin sorunsuz ve yanlışsız olabilmesi için aşağıdaki süreci işlenmesi oldukça yerinde olacaktır. Bu sayede öğrencilerden elde edilecek ölçme sonuçlarının daha kesin ve hatasız olması sağlanabilir.

- 1. Testin Amacının Belirlenmesi:** Bir testin hangi amaçla hazırlandığının belirlenmesi gerekmektedir. Örneğin: Uygulanacak bu sınavın amacı öğrencilerin Coğrafya bilgilerini ölçmektir.
- 2. Kapsam ve Ölçülecek Davranışın Belirlenmesi:** Yapılacak olan sınavın kapsamının belirlenmesi gerekmektedir. Örneğin kitabın ilk 100 sayfasından ya da kitabın ilk 4 ünitesinden sınavın yapılacağına öğrencilere bildirilmesi gibi. Aynı zamanda öğrencilerde görülmek istenen davranışların da belirlendiği kısımdır. Bilgi, kavrama veya analiz seviyesinde davranışların belirlenmesi bu aşamada yapılmalıdır.
- 3. Belirtke Tablosu Hazırlama:** Sınavda sorulacak olan soruların belirli bir üniteye yada sayfalara sıkışmasının engellenmesi ve öğrencilere her alandan (bilgi, kavrama, uygulama, analiz, sentez, değerlendirme) soru sorabilmek için hazırlanmalıdır. Bu sayede hem ilgili kapsamın tamamından soru sorulabildiği gibi hem de öğrencinin bilgiye sahip olma, kullanma basamaklarından sorular sorulmuş olacaktır.
- 4. Kullanılacak Soru Türünü Belirleme:** Sınavda kullanılacak soru türünün belirlendiği kısımdır. Çoktan seçmeli, kısa cevaplı, boşluk doldurma, yazılı yoklama türünden sorulacak olan soruların türünün belirlenmesidir. Ancak burada özellikle belirtilmelidir ki, bir ölçme aracıda yani sınavda mümkün olduğu kadar fazla soru türü kullanılmalıdır.
- 5. Madde (Soru) Yazma:** Sınavda kullanılmak üzere soruların oluşturulduğu yani yazıldığı kısımdır. Bol bol soru yazılması/hazırlanması gerekmektedir. Bazı kaynaklarda testte sorulacak soru sayısının üç katı kadar soru yazılmasını tavsiye eder.
- 6. Soruların Gözden Geçirilip Seçilmesi:** Sınav konularını ilgilendiren soruların yazılmasından sonra bu soruların içerisinden seçilecek olan soruların gözden geçirilmesi ve sınav formatına getirildiği aşamadır. Soruların dikkatlice gözden geçirilmesine, sorularda

yanlışlıklar olmamasına, bir sorunun cevabının başka bir soru da olmamasına dikkat edilmelidir.

- 7. Ayrıntılı Puan Anahtarının Hazırlanması:** Her ölçme aracının mutlaka ayrıntılanmış bir cevap anahtarının olması gerekmektedir. Cevap anahtarı sayesinde daha objektif bir ölçme işlemi gerçekleştirilebilir. Sınav sonrasında ilan edilen bir cevap anahtarı sayesinde sınava giren öğrenciler, sınav sonuçları hakkında daha sağlıklı tahminlerde bulunabilecekleri gibi kendi yanlışlarının da aslında nasıl doğru olabileceğini görmüş olacaklardır. Cevap anahtarı sayesinde kişisel yargılardan kurtulmuş oluruz. Sistemik hatanın önüne geçilebilir.
- 8. Soruların Çoğaltılmaya Hazırlanması:** Sınav evraklarının sınavdan birkaç gün öncesinde hazırlanmış olması gerekmektedir. Çünkü sınav kâğıdının çoğaltılmasının son dakikaya bırakılması sebebiyle sınavın başlaması gereken zamanı erteletmeye sebebiyet verilebilir. Ayrıca daha önceden çoğaltılması sayesinde sınav evraklarında görülen bir hatanın fark edilmesini sağlayabilir.
- 9. Uygulama:** Sınavın uygulanması aşamasıdır. Sınavın uygun ortamda (sıcaklık, ışık, sessizlik vb.) yapılması sağlanmalıdır. Uygulama esnasında sınavın akışını bozacak durumlara müsaade edilmemelidir.
- 10. Puanlama:** Sınavın uygulanmasından sonra sınav sorularına verilen cevapların okunduğu aşamadır. Puanlama mümkün olduğu kadar hızlı şekilde yapılmalı ve ilan edilmelidir. Geç okunan ve ilan edilen bir sınavın sonucu sebebiyle öğrenciler bir sonraki sınava hazırlık yapmalarında gecikmeler olabilecektir.

Test hazırlama sürecine etkin bir şekilde uyulması sayesinde, ölçme sonuçlarındaki hataların da azalmasını sağlayacaktır. Eğitim-öğretim faaliyetlerinin en önemli öğelerinden olan ölçme ve değerlendirme işlemleri sayesinde, programların güçlü ve zayıf yönleri de fark edilmiş olacaktır.

### Testin Hazırlanış Amaç ve Şekilleri

Puanlanış Bakımından	Öğrenci Sayısı Bakımından	Testi Hazırlayan Bakımından	Süre Bakımından	Ölçtüğü Özellik Bakımından
Objektif Puanlanan	Grup	Uzman Yapımı (Standart)	Hız	Maksimum Performans
Subjektif Puanlanan	Bireysel	Öğretmen Yapımı	Güç	Tipik Performans
			Yeterlilik	



Ölçme ve değerlendirme işlemi için daha test hazırlanmadan öncesinde testin amacı belirlenmelidir. Testin amacının belirlenmesi, ölçme işleminde hangi test türünün kullanılmasını daha uygun hale getirecektir. Bir testin hazırlanış amaç ve şekilleri aşağıda sırasıyla açıklanmıştır.

### 1. Puanlanış Bakımından

- a. Objektif Puanlanan:** Kim olursa olsun sınavdan alınacak not üzerinde herhangi bir değişiklik yapılmamasını ifade eder. Her şey şeffaftır. Çoktan seçmeli testler objektif olması yönünden oldukça kullanışlıdır.
- b. Subjektif Puanlanan:** Kişisel görüşler etkilidir. Puanlamaya, çeşitli etkenler karışabilir. Yanlılık olmasıdır. Sözlü sınavların en çok eleştirilmesinin sebebi subjektif puanlamaya yatkın olmasıdır.

### 2. Öğrenci Sayısı Bakımından

- a. Grup:** Kısa zamanda çok kişinin sınav yapılmasıdır. Örneğin Ölçme Ve Değerlendirme Dersinin sınavında aynı anda 250 kişinin sınav yapılmasıdır.
- b. Bireysel:** Grup testine göre daha fazla emek ve maliyet gerektirmektedir. Ancak özel eğitime muhtaç bireyler sınava yapılırken kullanılması uygundur.

### 3. Testi Hazırlayan Bakımından

- a. Uzman Yapımı (Standart):** Gerekli geçerlilik ve güvenirlik çalışmalarının uzmanlar tarafından yapıldığı testlerdir. Örneğin ÖSYM'nin yaptığı sınavlar uzman yapımı olarak değerlendirilir.
- b. Öğretmen Yapımı:** Öğretmenler tarafından hazırlanan testlerdir. Daha sık tercih edilmelidir. Çünkü dersin akışı içerisinde hangi konular üzerinde daha çok durulmuşsa o konular üzerinde öğrenciler yoğunlaşacaktır. Ayrıca bir dersin işlenen kısımlarının önemli yerleri öğretmenler tarafından daha iyi bilinmektedir.

### 4. Süre Bakımından

- a. Hız:** Sınav sorularının genellikle kolay ancak zamanı daha iyi kullananların yüksek not alacağı sınavlardır. Örneğin ALES (Akademik Personel ve Lisansüstü Eğitim Sınavı) bir hız testidir. Yeteri kadar zaman olduğunda birçok öğrenci soruların daha fazlasını yapabilecektir.

**b. Güç:** Soruların genel itibariyle çok zor olduğu ancak zaman konusunda belirli bir sınırlamanın olmadığı sınavlardır. Örneğin İngilizce çeviri sınavına girildiğinde sınav süresinin net olarak söylenmeyip, sınav ne zaman bitirilirse o zaman teslim edilmesidir.

**c. Yeterlilik:** Güçlük testlerine göre soruların biraz daha kolay ancak hız testlerine göre de zamanın daha yeterli olduğu sınavlardır. Örneğin KPSS (Kamu Personeli Seçme Sınavı) bir yeterlilik sınavıdır. Bu sınav ilk yapıldığı zamanlarda ki ismi yeterlilik sınavıydı.

## 5. Ölçtüğü Özellik Bakımından

**a. Maksimum Performans:** Sahip olunan bilginin her yönüyle ölçülmesidir. Örneğin İngilizce bilgisi için yapılan IELTS(Uluslararası İngilizce Dil Yeterlilik Sınavı) bir maksimum performans sınavıdır. Çünkü sınavın içeriğinde konuşma, okuma, dinleme, yazma davranışlarının hepsi vardır. Ayrıca bireylerin belirlenen bir performansı ne derece yapabildiğini belirleme de kullanılır.

**b. Tipik Performans:** Genellikle psikolojik özellikleri ölçen testlerdir. Örneğin bir doktorun bir hastasının hangi psikolojik hastalığı olduğu teşhis edebilmesi için hastasına bir takım sorular yönlendirmesidir.

## Test Nasıl Hazırlanmalı?

Test hazırlanırken takip edilmesi gereken aşamaları izledikten sonra testin daha işlevsel ve hatasız olabilmesi için uyulması gereken birtakım kriterleri aşağıdaki gibi sıralayabiliriz:

1. Test hazırlanırken sınav sorularının Times New Roman 10-12 punto olmasına dikkat edilmelidir. Daha küçük puntuyla yazılması okumayı güçleştireceği gibi gereğinden fazla büyük harflerle yazılması da daha fazla kâğıt harcanmasına sebebiyet verebilecektir. Öğrencilerin yaşları ve gelişimleri dikkate alınarak yapılması daha uygundur.
2. Soru maddesi yani soru kökü belirgin olmalıdır. İki soru arasında yeteri kadar boşluk bırakılmalıdır. Soruların birbirinin içerisine girmiş şekilde hazırlanan bir ölçme aracı düzensiz duracağı gibi öğrencilerinde kafasının karışmasına neden olabilecektir.
3. Çoktan seçmeli testlerde seçeneklerde çeldiriciler kuvvetli olmalıdır. Sorunun cevabının bulunmasında çeldiriciler kuvvetli değilse sınava giren öğrencilerin bilenle bilmeyenin birbirinden ayrılması güçleşecektir.

4. Özellikle soru kökündeki olumsuz ifadeler (değildir, değinilmemiştir, olamaz vb.) **kalin**, altı çizili ya da *italik* şekilde yazılmalıdır. Yanlış okumadan kaynaklanacak durumlar nedeniyle ölçme işlemine hata karışabilecektir.
5. Puanlamalar belirtilmelidir. Yani öğrenci verdiği her doğru cevap için kaç alacağını bilmelidir. Örneğin birinci soru yazıldıktan hemen sonrasında parantez içerisinde kaç puan olduğu (5 Puan gibi) yazılabilir. Bu şekilde olabileceği gibi testin sonuna “Her doğru cevap 5 puandır” şeklinde de yazılabilir.
6. Öğrencilerin sorulara cevap verebilecekleri kadar zaman tanınmalıdır. Bu zamanın çok kısa olması öğrencilerin sona bıraktıkları soruları okuyamamalarına sebep olabileceği gibi gereğinden fazla zaman verilmesi, öğrencilerin sınavlarını bitirdikten sonra kopya çekmelerine sebebiyet verebilmektedir. Bir sınav için uygun zaman bulunurken bu sınavın daha önceden uygulanmış olması daha net sonuç verecektir. Pilot uygulamada sınav yapılırken sınavını bitiren öğrencilerin kâğıtlarını teslim ederken kaç dakikada sınavı bitirdikleri bir yere not edilir. Daha sonra bunların aritmetik ortalaması alınarak ideal süre bulunabilir.
7. Sınavdaki sorulardan bazılarının yine aynı sınavdaki başka sorulara cevap olacak şekilde olmamasına dikkat edilmelidir. Örneğin sınavdaki 5. Sorunun cevabı aynı sınavdaki 17. soruda olmamasına dikkat edilmelidir.
8. Çoktan seçmeli sorularda üst üste aynı seçeneklerin olmaması gerektiği gibi test boyunca aynı seçeneğe sıkça yer verilmemelidir. Her seçenekten eşit sayıda doğru cevap bulunmalıdır. Örneğin 40 soruluk bir testte mümkünse 8 adet A seçeneği, 8 adet B seçeneği, 8 adet C seçeneği, 8 adet D seçeneği, 8 adet E seçeneği olmalıdır. Aynı seçenekten çok fazla olursa öğrenciler bilemedikleri sorular için hep aynı seçeneği doğru olarak işaretleyebilirler. Aynı seçeneğe sıkça yer verilmesi yapı geçerliliğini riske atar.
9. Hangi türden soru sorulursa sorulsun, sorularda imla hatalarının olmaması gerekmektedir. Sınav kâğıdı defalarda incelenmeli hatta gerekirse dil açısından bir kişiye de incelenmelidir. Sınav esnasında bir virgül bile sorunun yanlış anlaşılmasına ve yanlış cevap verilmesine neden olabilir.
10. Noktalama işaretlerinin hemen sonrasında bir karakter boşluk bırakılmalıdır. Genellikle dikkat edilmeyen bir husustur ancak ölçme aracının daha düzenli görünebilmesi için bu tür şekilselliklere dikkat edilmelidir.
11. Sınav için mutlaka bir ayrıntılı bir cevap anahtarı hazırlanmalıdır. Öğrencilerin görebileceği bir yerde ilan edilmelidir. Bu sayede öğrencilerde gerçekleşmiş olan yanlış öğrenmelerin önüne geçilmiş olacaktır.

12. Sınav sonuçları mümkün olan en hızlı şekilde ilan edilmelidir. Genel olarak yasal süre 15 takvim günü olsa da, sınav sonucunun erken ilan edilmesi hem öğrencilerin bir sonraki sınava hazırlanmaları için daha erken plan yapmalarını hem de sınavın puanlamasında yapılmış olan bir hata varsa gelen itirazlara göre değerlendirilip gerekirse yeniden puanlanması sağlanabilir. Amaç, her şeyi olduğu gibi ölçmektir. Ölçme işlemine hata karışmasının önüne geçilmelidir.
13. Bir ölçme aracı hazırlanırken çok soru hazırlanması önemli olduğu kadar, ölçme aracı içerisinde sadece aynı tip soru maddelerinin olmamasına önem verilmelidir. Sınav sadece çoktan seçmeli olmamalı ya da sadece yazılı yoklama türünden olmamalıdır. Sınavın içerisindeki soru çeşitlerinde yazılı yoklama türünden sorular, çoktan seçmeli sorular, boşluk doldurma türünde sorular, eşleştirmeli sorular, doğru-yanlış türünde sorular olmalıdır.

### GELENEKSEL (KLASİK) VE ÇAĞDAŞ (ALTERNATİF) TEST ÇEŞİTLERİ

	Klasik Ölçme Değerlendirme Araçları	Alternatif-Otantik (Yaşantılara dayalı) Ölçme Değerlendirme Araçları
1.	Kısa cevaplı testler	Portfolyo
2.	Sözlü sınavlar	Tanılayıcı dallanmış ağaç
3.	Yazılı yoklamalar	Kontrol listesi
4.	Doğru-yanlış testleri	Dereceleme ölçeği
5.	Eşleştirmeli testler	Kelime ilişkilendirme testi
6.	Çoktan seçmeli testler	Yapılandırılmış grid
7.	Ödevler	Öz değerlendirme
8.		Akran değerlendirmesi
9.		Rubrik

#### 1. YAZILI YOKLAMALAR

Soruların öğrencilere yazılı olarak verildiği, öğrencilerinde soruların cevabını düşünüp yazarak cevapladığı sınav türüdür. Hazırlanması kolay puanlanması güçtür. Soru sayısının azlığı nedeniyle geçerlik ve güvenilirliği (objektiflik, duyarlılık, kararlılık, tutarlılık, ayırt

edicilik yönünden) düşüktür. Üst düzey bilişsel (analiz, sentez, değerlendirme) ürünlerin ölçülmesine elverişlidir.

### **Özellikleri:**

- Kapsam geçerliği düşüktür.
- Soru sayısı azdır, bu sebeple güvenilirliği düşüktür.
- Şans başarısı yoktur.
- Soruların hazırlanması kolaydır ancak puanlaması zaman alır.
- Üst düzey davranışları ölçmekte kullanılır.
- Öğrenciler şişirme cevaplar verebilirler.
- Öğrenci sayısı azsa kullanılabilir.
- Kopya ihtimali düşüktür.

### **Yazılı Yoklamaların Puanlama Yöntemleri**

- 1. Sınıflama Yoluyla Puanlama Yöntemi:** Sınav kâğıtlarına genel olarak bakılır ve kâğıtlar kendi içerisinde *iyi-orta-kötü* şeklinde sıralanır. Her biri kendi içerisinde değerlendirilir. Örneğin iyi olanlar kâğıtlara 80-100; orta olan kâğıtlara 60-80; kötü olan kâğıtlara da 40-60 arasında bir puanın, ölçmeci tarafından kâğıdın durumuna göre takdir edilmesidir.
- 2. Sıralama Yoluyla Puanlama Yöntemi:** Kâğıtlar sıralanır. Bulunduğu sıraya göre puan verilir. Örneğin öğretmen sınıftaki öğrencilerin kâğıtlarına genel bir göz gezdirir ve kâğıtları sıraya koyar. Birinci olan öğrenciye 95, ikinciye 90, üçüncüye 85 gibi puanlar verir. Yanlılık vardır. Öğrencilerin sınıf içerisindeki durumları sınav kâğıdına olumlu ya da olumsuz olarak yansıtacaktır.
- 3. Genel İzlenime Göre Puanlama Yöntemi:** En az kullanılan puanlama yöntemidir. Kâğıdın hepsini okuyup kâğıtla ilgili genel izlenimle öğrenciye puan verilir. Öğrenci için olumlu ya da olumsuz düşünce sınav sonuçlarını etkileyecektir. Ancak öğrencilerin yazdığı bir kompozisyon, yazdıkları bir makale ya da deneme gibi çalışmalarda öğretmenin, çalışmanın tamamını okuyup genel bir puan vermesi şeklinde yorumlanabilir.
- 4. Cevap Anahtarı Kullanılarak Puanlama Yöntemi:** En sağlıklı puanlama yöntemidir. Öğrencinin ismi kapatılabilir. Öncelikle tüm öğrencilerin birinci sorusu sonra ikinci sorusu okunur, bu şekilde devam edilir. Gerekirse birden fazla kişinin (gözlemcinin) okuması sağlanabilir.

## Yazılı Yoklama Sorularının Taşınması Gereken Özellikler

1. Sorular açık, net ve anlaşılır olmalıdır.
2. Geçerliliği sağlayabilmesi için olanı olduğu gibi ölçebilmelidir.
3. Soru sayısı artırılmalıdır. Bu sayede sınavın duyarlılığı artırılmış olacaktır. Soruların cevaplarının kısa ama çok fazla soru olmasına dikkat edilmelidir. Sorular, kapsamın (içeriğin) tamamını karşılayacak nitelikte olmalıdır.

## 2. SÖZLÜ SINAVLAR

Soruların sözlü olarak verildiği ve cevaplarında sözlü olarak istendiği sınavlardır.

- Soru sayısı azdır ve puanlama güvenirliği düşüktür.
- Kişinin ölçmecide bıraktığı izlenim, sözlüye kalkış sırası, kişinin psikolojik durumu vb. pek çok durum geçerlik ve güvenirliği olumsuz yönde etkiler.
- Kapsam geçerliği düşüktür.
- Sorulara verilen cevaplardan çok, öğrencilerin konuşma becerisi ve beden dili etkilidir.
- Okuma-yazma bilmeyen gruplarda ve özel eğitime muhtaç öğrencilerde etkilidir.
- Öğrenciler sözlüye kalkarken cesaretlendirilmelidirler.
- Üst düzey davranışları ölçmede kullanılabilir.
- Sınıfta yoklama sırasına göre değil öğrenciler rastgele seçilerek sözlü yapılmalıdır.
- Sürekli eleştirilen bir sınav şeklidir.

## 3. KISA CEVAPLI TESTLER (BOŞLUK DOLDURMA)

Cevabı öğrenciler tarafından düşünülüp bulunulan ve tasarlanıp yazılan test çeşididir. Cevap bir kelime, bir sözcük grubu, bir tarih, bir sayı olabilir. Kısa zamanda hazırlanır ve uygulanır. Ancak öğrencileri ezbere yönlendirir. Öğrencileri ezbere yönlendirmemek için abartmadan kullanılabilir (bir sınavda 20 soru varsa 2-3 âdetinin kısa cevaplı test olması yeterlidir). Ayrıca mümkün olduğunca aynı tarz soru tiplerinden sorulmamalıdır. İfadeler kaynaklardan aynen alınmamalıdır. Bilgi ve hatırlama düzeyinde kullanılır. Daha üst basamaklarda (analiz, sentez, değerlendirme) kısa cevaplı testler genelde kullanılmaz. Çoktan

seçmeli testler kadar güvenilir değildir ancak yazılı yoklama ve sözlü sınavlara göre yanlılığı azaltacağı için daha güvenilirdir diyebiliriz.

#### **Özellikleri:**

- Hazırlaması kolaydır.
- Puanlaması kolaydır.
- Puanlama objektifliği yüksektir.
- Kısa cevaplı testlerde şans faktörü yoktur.
- Çok sayıda soru sorulabilir.
- Alt düzey basamakları (bilgi, kavrama) ölçer.
- Öğrencileri ezbere iter.

#### **Notlar:**

- Boşluk doldurma sınavları, şans faktörünün başladığı sınavlar olarak kabul edilirler.
- Boşluk doldurma sınavlarında boşluktan sonra ekler kullanılmamalıdır.
- Boşluklar belirli bir uzunlukta ve standart olmalıdır. Biri uzun biri kısa olmamalıdır.
- Boşluklar cümlelerin başında veya sonunda olmamalıdır.

#### **4. DOĞRU – YANLIŞ TESTLERİ**

Cevaplayıcıların (öğrencilerin) test maddelerini Doğru (D) veya Yanlış (Y) olarak sınıflamasına dayanan bir ölçme aracıdır. Yazma, işaretleme, doldurma, evet-hayır, (-) (+) gibi farklı alternatifleri vardır.

#### **Özellikleri:**

- Doğru-yanlış testlerinde şans başarısı oldukça yüksektir.
- Çok fazla soru sormaya imkân tanır.
- Öğrencilerin bilgi ve hatırlama basamaklarındaki becerileri ölçülür.
- Puanlaması kolaydır.
- Soruların cevaplarının tamamı **sadece doğru** ya da **sadece yanlış** olmamalıdır. Öğrencilerin sadece doğru ya da sadece yanlış seçeneğini işaretleyerek tam puan

almasının önüne geçebilmek için doğru cevaplar seçeneklere eşit miktarda dağıtılmalıdır.

- Kopya çekmeye elverişlidir.
- İlke ve genellemelerin yokladığı testtir.
- Üst düzey becerileri ölçemez.
- Tüm derslerde kullanılabilir.
- Düzeltme formülü kullanılırsa şans başarısı ortadan kalkar.
- Maddeler olumsuz ifadeler şeklinde sorulmamalıdır.
- Maddeler çift olumsuzluk şeklinde olmamalıdır.
- Maddelerde daima, her zaman, sık sık kelimeleri kullanılmamalıdır.

### **Örnekler:**

1. Geçerli olan her şey her defasında güvenilirdir. (D) (Y)
2. Bir ağacın yüksekliğini ölçmek mutlak sıfırla ifade edilir. (D) (Y)
3. Hava sıcaklığı 303 Kelvin'dir. Burada bağıl sıfır vardır. (D) (Y)
4. Pozitif korelasyonda doğru orantı vardır. (D) (Y)

## **5. EŞLEŞTİRMELİ TESTLER**

İki grup halinde verilen ve birbirleriyle ilgili olan bilgi öğelerinin, belli bir açıklamaya göre eşleştirilmesi esasına dayanan madde türüdür. Seçmenin tahmine dayanmaması için ve sona kalanların birebir eşlenmemesi için iki sütundan biri diğerinden fazla olacak şekilde hazırlanır. Örneğin bilgi verilip karşılığında tarih istenebilir. Eşleştirmeli testlerin sınav evrakının aynı yüzünde olması gerekmektedir. Bir kısmı sayfanın ön yüzünde kalanlarında sayfanın arka yüzünde olmaması gerekmektedir. Mümkün olduğu kadar ok işaretleriyle eşleştirme yapılmamalıdır. Çünkü öğrenciler yanlış yaptıklarını düşünüp silmeye çalıştıklarında temiz bir şekilde silemezlerse yanlış puanlamaya sebebiyet verebilir. Belirli haneler açılmalı ve oralara sayıların ya da harflerin yazılması istenmelidir. Çok sık şekilde kullanılmamalıdır.

### **Özellikleri:**

- Seçenek sayısı öncül sayısından en az iki adet fazla olmalıdır.



- Öncül sayısı en az 6, en fazla 15 olmalıdır.
- Öncüller sol tarafta, seçeneklerde sağ tarafta olmalıdır.
- Seçenekler belirli bir düzende verilmelidir.
- Öncüllerin ve seçeneklerin tamamı aynı sayfada olmalıdır.
- Bir öncül birden fazla seçenekle eşleşebilir.
- Öğrencileri ezbere itebilir.
- Kapsam geçerliği yüksektir.
- Objektiftir.

**Not:**

- Öğretim amaçlı eşleştirmeli test yapılıyorsa öncül ile seçenek sayısı eşit olabilir.
- Eşleştirmeli test kullanılan bir sınavda soru öncesi bir yönerge yani bilgilendirme olmalıdır.

**Örnek:** Aşağıdaki tabloda ilgili ifadelerin karşısına gelecek olan doğru cevapları eşleştiriniz. Eşleştirme yaparken ifadelerin önündeki rakamları kullanınız.

	Öncül		Seçenek
1	İçinde matematiksel işlemler bulunan ölçme türüne denir.		Sistematik hata
2	Bir öğretmen öğrencilerin sınav kâğıtlarını puanlarken öncelikle bütün öğrencilere +5 puan vermesiyle oluşan hatadır.		Türetilmiş ölçme
3	Öğretmenin yazısı güzel olan öğrencilere +5 puan daha vermesiyle oluşan hatadır.		Tesadüfî hata
4	Öğretmenin sınav sonuçlarını öğrenci bilgi sistemine girerken 55 alan Ahmet adında ki öğrencinin notunu yanlışlıkla 65 olarak yazmasıyla oluşan hatadır.		Sabit hata
5			Doğrudan ölçme

## 6. ÇOKTAN SEÇMELİ TESTLER

Akla ilk gelen sınav tekniğidir. Soruların öğrenciye yöneltip, doğru cevabı hatalı cevapların arasından seçmesinin beklendiği türdeki testlerdir.

**Özellikleri:**

- Hazırlanması zordur (Geleneksel ölçme araçları içerisinde hazırlanması en zor olan ölçme aracıdır). Bu sebeple uzmanlık gerektiren bir ölçme aracıdır.
- Çoktan seçmeli testlerde öğrenciler, zamanlarının büyük bir kısmını soruları okuyarak harcarlar.
- Puanlaması kolaydır.
- Kapsam geçerliği yüksektir.
- Güvenirliği ve geçerliği yüksektir.
- Objektiftir.
- Sentez basamağı hariç her basamağı ölçebilir.
- Şans başarısı vardır.
- Kopya çekme ihtimali vardır.
- Öğrenciler hakkında önemli kararlar alınırken kullanılması uygundur.

#### Notlar:

- Sorular açık, net ve anlaşılır olmalıdır.
- Her soru sadece bir hedef davranışı ölçmelidir. Bir soru birden fazla hedef davranışı yoklama amacıyla sorulmamalıdır.
- Soru köklerinde iki farklı olumsuz ifade kullanılmamalıdır. “Aşağıdakilerden hangisi belirtilmeyenlerden değildir?” gibi soru sorulmamalıdır.
- Seçeneklerin tamamı birbirleriyle eşit uzunlukta olmalıdır.
- Seçenekler soruyla ilgili olmalıdır ve seçenekler aynı konudan olmalıdır.
- Seçenekler bilenle bilmeyeni ayırmaya yönelik olmalıdır. Öğrencilerin dikkatini ölçmek amacıyla **yazılmamalıdır**.
- Çeldiriciler, doğru seçeneğe benzedikçe soru zorlaşacaktır.

#### Örnekler:

Ankara'nın başkent olduğu yıl aşağıdaki seçenlerin hangisinde doğru olarak verilmiştir? (Madde Kökü)

- |         |   |             |   |         |
|---------|---|-------------|---|---------|
| A. 1923 | → | Doğru cevap | → | Seçenek |
| B. 1924 | → | Çeldirici   | → | Seçenek |

- C. 1925 → Çeldirici → Seçenek  
D. 1926 → Çeldirici → Seçenek  
E. 1927 → Çeldirici → Seçenek

Aşağıdakilerden hangisi ölçmenin öğelerinden biri **değildir**?

- A. Ölçülecek nitelik  
B. Birim  
C. Ölçme aracı  
D. Ölçüt  
E. Ölçen kişi ya da makine

**Not:** Çoktan seçmeli testlerin kapsam geçerliliği yüksektir. Çoktan seçmeli test tekniğinin artıları ve eksikleri vardır. Seçeneğe göre şans başarısı artar ya da azalır. Tahminle sonuca gidilebilir. Çünkü cevap sorunun içerisindeydir. Çoktan seçmeli testlerin hazırlanması zor ancak puanlaması kolaydır. Ayrıca test ve madde istatistikleri yapılabilir. İstatistik sonucunda testin güçlü ve zayıf yanları bulunabilir.

#### **Düzeltilme Formülü**

<b>Düzeltilmiş puan = Doğru cevap – Yanlış sayısı / Seçenek sayısı-1</b>
--

Örneğin;

5 seçenekli bir sınavda 4 yanlış bir doğruyu;

4 seçenekli bir sınavda 3 yanlış bir doğruyu;

3 seçenekli bir sınavda 2 yanlış 1 doğruyu;

2 seçenekli bir sınavda 1 yanlış 1 doğruyu götürür.

Düzeltilme formülü sonunda “Kaç netin var?” sorusunun cevabı düzeltme formülünden elde edilen puandır.

## 7. ÖDEVLER

Öğrencilerin bilgi ve becerilerini ölçmek amacıyla kullanılan ve bir ölçme aracı niteliği taşıyan ödevler, aynı zamanda öğrencilerin öğrendikleri konuları pekiştirmeleri amacıyla da kullanılmaktadır.

### Ödevler Hangi Amaçlar Verilmez?

- Aileler istediği için ödev verilmez.
- Öğrencilere ceza amacıyla ödev verilmez.
- Okul idarecileri istediği için ödev verilmez.
- Öğrenciler sınıfta sessizce durması amacıyla ödev verilmez.
- Öğrenciler boş zamanlarını değerlendirsin amacıyla ödev verilmez.

## ÇAĞDAŞ (ALTERNATİF-TAMAMLAYICI) TEST ÇEŞİTLERİ

Şu ana kadar işlenilen konular gelenekseldir, klasiktir. Bundan sonrakiler ise otantiktir. Otantik, yaşantılara dayalı olan demektir.

Alternatif ölçme ve değerlendirme, geleneksel ölçme ve değerlendirmeye göre daha fazla gerçek hayatla ilişkili (otantik) ve öğrenci merkezlidir. Alternatif değerlendirmeler ürün kadar sürecin de değerlendirilmesini dikkate alır. Öğrencilerin yüksek düzeydeki düşüncelerini, problem çözme ve yaratıcılıklarını ön plana çıkarır. Çoklu cevaplar, stratejiler ve oluşturulan süreç öğretmen tarafından değerlendirilir, onaylanır ve ödüllendirilir. Değerlendirmedeki bu değişiklikler birçok açıdan öğrencilerin öğrenme konusunda sorumluluk sahibi olmasını ve öğrendikleriyle gurur duymasını sağlayabilir (Bahar ve diğerleri, 2006).

Alternatif ölçme ve değerlendirme, öğrenciyi merkez kabul ederek onların aktif olarak derse katılmalarını sağlar, düşündüklerini açıkça söyleyebilme fırsatı verir ve kendilerini rahatlıkla ifade etmelerine yardımcı olur (Yıldız ve Uyanık, 2004).

## 1. PORTFOLYO

Literatürde geçen diğer isimleri *bireysel gelişim dosyası, ürün seçki dosyası, tümel değerlendirme, alternatif değerlendirme, tamamlayıcı değerlendirme* ve *portföy*dür.

Eğitim süreci boyunca geliştirilen öğrenci çalışmalarının sistematik şekilde toplanmasıdır. Öğrencinin merkezde bulunduğu aktif bir süreçtir. Öğrencinin kendi kendine öğrenmesi ile ilgili sorumluluk almasına yardımcı olan alternatif bir değerlendirmedir (Elimizde böyle bir dosya varsa öğrenci hakkında çok şey öğrenebiliriz.)

### Portfolyo hangi bölümlerden oluşur?

- 1. Özgeçmiş:** Bu kısım öğrencinin ortaya koyduğu çalışmaların hangi aşamalardan geçtiğini yansıtır.
- 2. Ürünler:** Portfolyo içinde öğrenci tarafından konulan tüm çalışmaları kapsar.
- 3. Yansıtma:** Öğrenci yaptığı çalışmaların bir ölçüde muhakemesini yapmak amacı ile kendisi ile ilgili görüşlerini yansıtır.

Günlük, tamamlanmış yazılar, öğretmen-öğrenci değerlendirmeleri portfolyo içinde vardır. Kalabalık sınıflarda bu şekilde bir dosyanın oluşturulması güçtür. Uzun zaman ister. Özellikle ilköğretim sınıf öğretmenleri tarafından sıklıkla başvurulur. Değerlendirme sübjektiftir. Rubrik gerekir.

### Portfolyonun Özellikleri:

- **Özgünlük:** Her portfolyo diğerinden farklı özelliklere sahiptir, yani özgündür.
- **Etkileşimlilik:** Öğretmeni, öğretmen, aile iletişimi gerçekleştirir.
- **Yansıtıcılık:** Öğrenci kendini en iyi yansıtacak olanları belirleyebilir.
- **Seçicilik:** Öğrenci içerisinden istediğini seçebilir, isterse çıkarıp değiştirebilir.
- **Bireysellik:** Her portfolyo bireysel özellikleri barındırır.
- **Gelişimcilik:** Öğrencinin nereden başlayıp nereye ilerlediğini gösterir.

**Portfolyo Örneği:**

**Ek 1**

**ÖLÇME VE DEĞERLENDİRME  
DERSİ  
ÜRÜN DOSYASI**

Adı : Aytekin  
Soyadı : Aslan  
Numarası : 147  
Sınıfı : 4/A

**Ek 2**

**İÇİNDEKİLER**

1. Birinci Sınav Kâğıdı
2. İkinci Sınav Kâğıdı
3. Devamsızlık Durumu
4. Ek Ödevler Çetelesi

**ÖĞRENCİ ÜRÜN DOSYASI  
DEĞERLENDİRME FORMU**

**Yönerge:** Aşağıdaki her bir ölçütün ne düzeyde yeterli olduğunu göz önüne alarak dosyayı değerlendiriniz.

ÖLÇÜTLER	Dereceler				
	1	2	3	4	5
Çalışmaların içeriğinin tam olması					
Çalışmalardaki çeşitlilik					
Çalışmaların amaçları karşılaması					
Çalışmaların doğruluğu					
Dosyanın düzenliliği					
Harcanan çabaları gösterme					
Yaratıcılığı gösterme					
Öğrencinin gelişimini gösterme					
Kendini değerlendirme					

## 2. KONTROL LİSTESİ

Psiko-motor davranışları ölçmede daha sık kullanılan kontrol listesinde bir davranışın yapılıp yapılmadığına bakılır. Eğer davranış kişi tarafından yapılmışsa “Evet” ya da “Hayır” , “Var” ya da “Yok” , “+” ya da “-” şeklinde işaretlenir.

### Kontrol Listesinin Özellikleri:

- Psiko-motor davranışları ölçmede önemlidir.
- Kriterler yapılış sırasına göre hazırlanmalıdır.
- Duyarlı bir ölçme aracı değildir.
- Davranışların ne kadarına sahip olup olunmadığı hakkını bilgi verir.
- Hayati tehlikeli olan durumlarda önemlidir.
- Sınıflama ölçekleri düzeyindedir.

### Kontrol Listesi Örneği

	Davranışlar	Evet	Hayır
1	Takla atmak için gerekli hazırlıklarını yaptı.	✓	
2	Takla atmak için işaret verilmesini bekledi.	✓	
3	Takla atma davranışını kendisine verilen yerde yaptı.	✓	
4	Takla attıktan sonra minderden doğru şekilde kalktı.	✓	
5	Takla attıktan sonra ayakta düzgün şekilde durdu.		✓
6	Takla attıktan sonra hakemi selamladı.		✓

Tabloda da görüleceği gibi öğrencinin takla atma becerisi ölçülürken altı davranış belirlenmiş ve öğrenci bu davranışlardan dördünü yapmıştır. Ancak öğrencinin yaptığı davranışları ne derece doğru yaptığı hakkında herhangi bir bilgi vermemektedir.

## 3. DERECELEME ÖLÇEĞİ

Kontrol listesinin bir eksikliği olan bir davranışın ne düzeyde yapıldığını belirtmek amacıyla yapılır. Bir davranışın yapılış derecesi önemlidir. Örneğin yapılan bir davranışa 1-10 arasında bir puan vermek dereceleme ölçeği düzeyindedir.

### Dereceleme Ölçeğinin Özellikleri:

- Sıralama ölçeği düzeyindedir.
- Kontrol listesine göre daha duyarlıdır.
- %100 objektiftir denemez.
- Sahip olunan bir davranışın derecesini ortaya koyar.
- Dereceleme ölçeği, ölçme aracının adıdır.

### Dereceleme Ölçeği Örneği

	Davranışlar	Çok İyi (4)	İyi (3)	Orta (2)	Zayıf (1)
1	Takla atmak için gerekli hazırlıklarını yaptı.		3		
2	Takla atmak için işaret verilmesini bekledi.		3		
3	Takla atma davranışını kendisine verilen yerde yaptı.	4			
4	Takla attıktan sonra minderden doğru şekilde kalktı.		3		
5	Takla attıktan ayakta düzgün şekilde durdu.				1
6	Takla attıktan sonra hakemi selamladı.				1

Tabloda da görüleceği gibi öğrencinin takla atma becerisi ölçülürken altı davranış belirlenmiş ve öğrencinin bu davranışlardan hangisini ne derece de yaptığı ortaya konmuştur. Öğrencinin yaptığı davranışları ne derece doğru yaptığını ortaya koyduğu için kontrol listelerine göre daha duyarlıdır ve daha objektiftir.

### 4. TANILAYICI DALLANMIŞ AĞAÇ

Klasik doğru-yanlış tipindeki sorulardan oluşan fakat doğru-yanlış testlerinden farklı olarak soruların birbirinden bağımsız cevaplanması yerine her bir soruya verilen kararın bir sonraki soruyu etkilemesiyle oluşturulan alternatif ölçme aracıdır. Öğrencilerin bilişsel yapılarındaki yanlış öğrenmelerin ortaya çıkarılarak öğrenme-öğretme süreci tekrar düzenlenebilir.

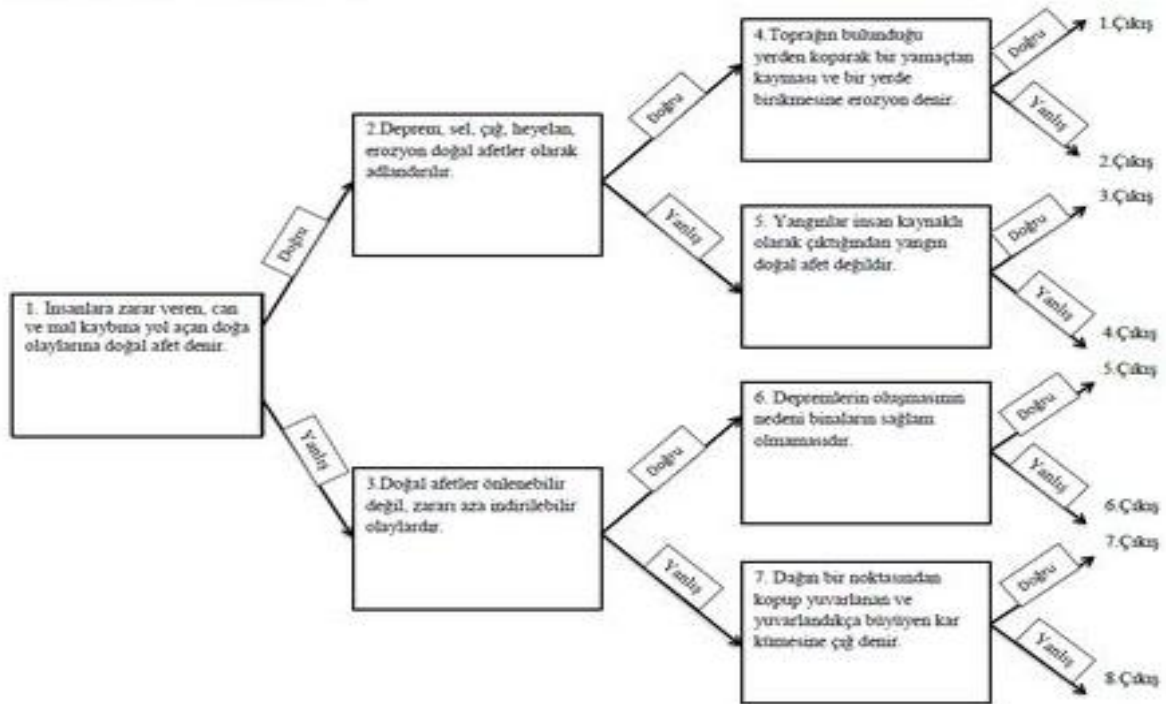
### Tanılayıcı Dallanmış Ağacın Özellikleri:

- Genelden özele doğru bir sıra izlenir.
- Amaç not vermek değil, öğrenme eksikliklerini ortaya çıkarmaktır.



- Temelden ayrıntıya gitmek vardır.
- Üst düzey öğrenme becerilerini ölçmekte yeterli değildir.
- Belirli bir konuda nelerin öğrenildiğini nelerin öğrenilmediğini tespit eder.
- Yanlış öğrenmelerin tam olarak nerede olduğunu gösteren tek ölçme aracıdır.
- Kapsam geçerliği doğru-yanlış testlerine göre daha azdır.
- İç tutarlığı yüksektir.

### Tanılayıcı Dallanmış Ağaç Örneği:



## 5. YAPILANDIRILMIŞ GRİD

Öğrencilerin bilgileri ne şekilde ilişkilendireceğini yani öğrencilerin anlamlı öğrenmelerini ölçmeyi amaçlayan alternatif bir tekniktir. Öğrencilerin kavram yanlışlarını, bilişsel yapılarındaki eksikliklerini ve aksaklıklarını ortaya çıkarır. Bir soruda “kutucuklar” gibi bir ifade geçiyorsa sorunun doğru cevabının yapılandırılmış grid olma ihtimali yüksektir.

### Yapılandırılmış Grid’in Özellikleri:

- Öğrenme eksikliklerini belirlemede kullanılırlar.

- Kavram yanlışlarını ortaya çıkarmada etkilidir.
- Alt düzey basamakları (bilgi, kavrama) ölçmekte yeterlidir.
- Şans başarısı yok denecek kadar azdır.
- Genellikle 9, 12 ya da 16 kutucuk kullanılır.
- Analitik düşünme becerisi kazandırır.
- Bütün parça ilişkisi kurmaya yardımcı olur.

### Yapılandırılmış Grid Örneği:

1	Ölçme	5	Ölçüt	9	Objektiflik	13	Sistematiik
2	Geçerlik	6	Güvenirlik	10	Kararlılık	14	Rastgele
3	Kullanışlılık	7	Ayırt edicilik	11	Ölçme Kuralı	15	Doğrudan Ölçme
4	Değerlendirme	8	Sabit	12	Duyarlılık	16	Dolaylı Ölçme

1. Kutucuklardan hangileri bir ölçme aracında olması gereken özelliklerdir?
2. Kutucuklardan hangileri ölçmede hata çeşitleri içerisinde sayılabilir?
3. Kutucuklardan hangileri ölçme çeşididir?
4. Kutucuklardan hangileri güvenilirlik kavramlarıyla ilgilidir?

Yapılandırılmış gride öğrencilerin sorulara verdikleri cevaplardan aldıkları puan hesaplanırken, aşağıdaki formül kullanılır.

## DEĞERLENDİRME

Seçilen doğru kutucuk sayısı (C1)  
Toplam doğru kutucuk sayısı (C2)  
Seçilen yanlış kutucuk sayısı (C3)  
Toplam yanlış kutucuk sayısı (C4)

$$\left[ \left( \frac{C1}{C2} - \frac{C3}{C4} \right) + 1 \right] \times 5$$



Soru kaç puanlıysa o puanın yarısı yazılır. Soru 10 puanlık

## 6. ÇETELEME ARACI (GÖZLEM ARACI)

Belli davranışların kaç kez tekrar edildiğini ölçmek için kullanılan ölçme aracıdır. Gözlenecek davranışlar ölçme aracının içerisine sırasıyla yazılırlar ve davranış ortaya çıktıkça yanına işaret konulur.

### Çeteleme Aracı Örneği:

Davranışlar	Çeteleme (Davranışın Kaç Kez Yapıldığı)
Tahtayı Kullanma	XXX
Öğrencilerle Göz Teması Kurma	XXXXXX
Öğrencilere Söz Hakkı Verme	/
Öğrencilere Soru Sorma	XXXXXXXX
Öğrencilerden Cevap Alma	XX/

## 7. RUBRİKLER (DERCELİ PUANLAMA ANAHTARI)

Bir öğrenciye performans görevi (makale, kompozisyon, etkinlik) verildiğinde yanında verilen ve öğrenciye ne yaparsa ne kadar puan alacağını gösteren ölçektir. Öğrenci neyi nasıl yapacağını bilir ve ona göre hareket eder. Yol göstericidir.

Rubrikler *analitik rubrik* ve *holistik rubrik* olmak üzere ikiye ayrılır.

### a) Analitik Rubrik:

Ayrıntılı (süreç) değerlendirilir yani süreç odaklıdır Daha objektiftir. Velilere kolaylık sağlar. Açık bir kriter tanımlar. Değerlendirme işini basitleştirir. Öğrenci kendi öğrenme sorumluluğunu alır.

### Analitik Rubrik Örneği:

	3 Puan	2 Puan	1 Puan
<b>Makalenin Başlığı</b>	Makalenin başlığı içeriği iyi şekilde yansıtmaktadır. İlgi çekici bir başlıktır. Yaratıcı bir başlıktır.	Makalenin başlığı içeriği iyi şekilde yansıtmaktadır. İlgi çekici bir başlıktır. Yaratıcı bir başlık değildir.	Makalenin başlığı içeriği iyi şekilde yansıtmaktadır. Ancak ilgi çekici ya da yaratıcı bir başlık değildir.
<b>İçerik</b>	Bilimseldir. Gerekli atıflar yapılmıştır. APA kurallarına uyulmuştur.	Bilimseldir. Gerekli atıflar yapılmıştır. APA kurallarına uyulmamıştır.	Bilimseldir. Gerekli atıflar yapılmamıştır. APA kurallarına uyulmuştur.
<b>Yöntem</b>	Doğru yöntem seçilmiştir. Yöntem iyi açıklanmıştır. Veriler doğru toplanmıştır.	Doğru yöntem seçilmiştir. Yöntem iyi açıklanmıştır. Veriler doğru toplanmamıştır.	Doğru yöntem seçilmiştir. Yöntem iyi açıklanmamıştır. Veriler doğru toplanmamıştır.
<b>Sonuçlar</b>	Mantıksal çıkarımlar yapılmıştır. Bulgular doğru yorumlanmıştır.	Mantıksal çıkarımlar yapılmıştır.	Mantıksal çıkarımlar yapılmıştır. Bulgular doğru

	Bilimsel olarak desteklenmiştir.	Bulgular doğru yorumlanmıştır. Bilimsel olarak desteklenmemiştir.	yorumlanmamıştır. Bilimsel olarak desteklenmemiştir.
<b>Kaynakça</b>	APA kurallarına uyulmuştur. Yeterlidir. Alfabetik şekilde düzenlenmiştir.	APA kurallarına uyulmuştur. Yeterlidir. Alfabetik şekilde düzenlenmemiştir.	APA kurallarına uyulmuştur. Yeterli değildir. Alfabetik şekilde düzenlenmemiştir.

**b) Holistik rubrik:**

Sonuç değerlendirilir. Bütüncüdür. Rehberdir. Her alan için yenisi hazırlanmalıdır. Hazırlama esnasında öğretim programı dikkate alınmalıdır. Holistik rubrikler ürün odaklı çalışır.

**Holistik Rubrik Örneği:**

	<b>11-15 Puan</b>	<b>6-10 Puan</b>	<b>0-5 Puan</b>
<b>Makale</b>	Makalenin başlığı, içeriği, yöntemi, sonuçları ve kaynakçası doğru şekilde oluşturulmuştur. Başlık ilgi çekici ve yaratıcı olduğu gibi içerik de yeteri kadar bilimsel olarak desteklenmiştir. Anlamli çıkarımlar yapılmıştır. Bir makale için gerekli olan her şey yapılmıştır.	Makalenin başlığı, içeriği, yöntemi, sonuçları doğru şekilde oluşturulmuştur. Ancak kaynakçası sorunludur. Başlık ilgi çekici ve yaratıcı olduğu gibi içerik de yeteri kadar bilimsel olarak desteklenmiştir. Anlamli çıkarımlar yapılmıştır. Bir makale için gerekli olan her şey tam olarak yapılmamıştır.	Makalenin başlığı ve içeriği doğru şekilde oluşturulmuştur. Ancak yöntem doğru seçilmemiştir. Kaynakçası sorunludur. Başlık ilgi çekici ve yaratıcı olduğu gibi içerik de yeteri kadar bilimsel olarak desteklenmiştir ama geliştirilmesi gerekir. Anlamli çıkarımlar yapılmamıştır. Bir makale için gerekli olan çoğu şey eksiktir.

## 8. ÖZ DEĞERLENDİRME

Bireyin kendini değerlendirmesidir. Biliş bilgisidir (Farklı bakış açılarıyla bireylerin kendilerine bakmalarıdır). Kendi güçlü zayıf yanlarını öğrenir. Bireyde farkındalık duygusu gelişir. Yeni öğrenme hedefleri belirlenir. Birey kendi yeteneklerini ortaya çıkarmayı öğrenir. Subjektif bir değerlendirmedir.

- Bu çalışmada neler yaptım?
- Bu çalışmada neler öğrendim?
- Bu çalışmada başarılı olduğum bölümler nelerdir?
- Bu çalışmada zorlandığım bölümler hangileridir?
- Bu çalışmada kendime kaç puan veririm?

## 9. AKRAN DEĞERLENDİRME

Bireyin eleştirel düşüncesi gelişir. Diyalog ve içsel konuşma becerisi gelişir. Öğretmen dışında kişinin de değerlendirmesi sağlanır. Subjektif bir değerlendirmedir.

- Çalışmalara gönüllü olarak katılır.
- Bildiklerini arkadaşlarıyla paylaşır.
- Arkadaşlarının görüşlerine saygılıdır.
- Aldığı görevleri zamanında yerine getirir.

## 10. KELİME İLİŞKİLENDİRME TESTİ (KİT)

Herhangi bir kavramın kısa bir sürede öğrencinin zihninde anlamsal çağrışım yaratan anahtar sözcüklerin listelenmesidir. 30-45 saniye arasında uygulanır. Bir sınavda doğru cevabın kelime ilişkilendirme testi olması için gerekli olan kelime “**anahtar kelime**”dir. Bu sayede öğrencilerin kavram yanılgıları ortaya çıkar. Öğrencilerin yanlış öğrenmeleri ve bilişsel düzeylerindeki eksiklikler görülebilecektir.

### Kelime İlişkilendirme Testi Örneği:

Öğrencilerin geçerlik kelimesi ile ilişkili olduğunu düşündüklerini yazmalarıdır.

Geçerlik .....

Geçerlik .....

Geçerlik .....

Geçerlik .....

Geçerlik .....

Geçerlik .....

Geçerlik .....

Öğrencilerin yazdığı kelimelerde yanlışlıklar varsa kelime ilişkilendirme testi sayesinde görülmüş olacaktır. Öğrencilerin bilişsel alanlarındaki eksiklikler görülür ve düzeltilir.

# TEST İSTATİSTİKLERİ

## Merkezi Eğilim (Yığılım, Vasat) Ölçüleri

Veri grubunu temsil edebilecek, merkez noktasına en yakın değerlerdir. Merkezi eğilim ölçüleri *aritmetik ortalama*, *mod (tepe değer)* ve *medyan (ortanca)* dır.

### 1. ARİTMETİK ORTALAMA

Bir veri grubundaki sayıların toplamının, veri sayısına bölünmesiyle elde edilen ortalamadır. Aritmetik ortalama duyarlı bir ortalamadır. Veri grubuna veri eklenmesi veya çıkarılması, veri grubundaki sayıların büyüklükleri ortalamayı etkiler. Bir veri grubundaki sayılar birbirine yakın sayılardan oluşuyorsa aritmetik ortalama, veri grubu hakkında sağlıklı bir değerlendirme sunar. Aritmetik ortalama, bir sayı dizisindeki en küçük sayıdan daha küçük olamaz. Aynı zamanda bir aritmetik ortalama, bir sayı dizisindeki en büyük sayıdan daha büyük olamaz.

$$\text{Aritmetik Ortalama Formülü} = \text{Verilerin Toplamı} / \text{Veri Sayısı}$$

### 2. MEDYAN (ORTANCA)

Küçükten büyüğe veya büyükten küçüğe doğru sıralanmış bir sayı dizisinde tam ortadaki sayı medyan (ortanca)'dır. Veri grubunda terim sayısı tekse tam ortadaki sayı, veri grubunda terim sayısı çift ise ortadaki iki terimin aritmetik ortalaması medyana verir. Çarpık dağılımlarda en iyi ölçü birimi medyandır.

**Not 1:** Veri grubundaki sayıların büyüklüklerinden etkilenmediği için duyarlı olmayan bir ortalamadır. Veri grubunda birbirine uzak değerlerin bulunması durumunda medyan, aritmetik ortalamaya göre daha sağlıklı bir değerlendirme sunar.

**Not 2:** Aritmetik ortalama ile medyan arasındaki farkı anlamak için şöyle bir örnek verilebilir. Bir sınıfta 19 kişi 100, bir kişi sıfır almışsa aritmetik ortalama 95 olacaktır ve her öğrenci 95

almış gibi bir durum ortaya çıkacaktır. Ancak medyan sayesinde sınıfın genelinin 100 aldığı ortaya çıkar ve bu durum daha sağlıklı bir yorum yapmamızı sağlar.

**Uyarı:** Medyan, veri grubunun elemanı olmak zorunda değildir.

**Uyarı:** Medyan uç değerlerden etkilenmez. Medyan ölçüm sayısından etkilenir. Sayı dizisine yeni sayıların eklenmesi veya çıkarılması medyanı etkiler.

Puan	Frekans	Yığılmalı Frekans
100	9	51
90	6	42
80	11	36
70	5	25
60	9	20
50	6	11
40	5	5

Toplam 51 kişi bulunmaktadır. Medyan değerini bulmak için yığılmalı frekansa 1 eklenir.  $51+1=52$ .  $52/2=26$ . 26. Kişi bize medyanı verecektir. Bu sebeple medyan 80'dir.

### 3. MOD (TEPE DEĞER):

Bir sayı dizisinde en çok tekrar eden sayı moddur. Ayrıca frekansı en çok yüksek olan değer moddur. Mod, dağılımın en belirgin özelliğinin ortaya çıkmasını sağlayan bir değerdir. Bir dağılımda birden fazla mod olabilir. Eğer sayı dizisinde iki mod varsa çift modlu dağılım, ikiden fazla mod varsa çok modlu dağılım ismini alır.

**Uyarı:** Frekansları eşit olan dağılımlarda mod yoktur. Bunun sebebi ise dağılımdaki hiçbir sayının diğerinden daha fazla tekrar etmemesidir.

Puanlar	Frekans
100	8
90	11
80	13
70	10
60	9

50	3
----	---

Bu sayı dizisine bakıldığında en çok tekrar eden sayının 80 olduğu görülecektir. Çünkü sayı dizisinde frekansı en büyük olan sayı bize modu verecektir. Ancak aynı sayı dizisinde frekansların benzer olduğu alttaki tablo incelendiğinde;

Puanlar	Frekans
100	8
90	11
80	13
70	10
60	13
50	3

Tablomuzdaki sayı dizisinde frekansı en çok tekrar eden iki sayı görülmektedir. Bu sebeple bu tabloda iki adet mod bulunmaktadır. Modlar, 60 ve 80 sayılarıdır. Ancak frekansları en yüksek olan iki sayı değeri arasında başka sayının olmadığı aşağıdaki tablo incelendiğinde;

Puanlar	Frekans
100	8
90	11
80	13
70	13
60	9
50	3

Frekansları en çok tekrar eden 70 ve 80 sayıları görülecektir. İki mod değeri arasında başka bir sayı değeri yoksa bu iki modun aritmetik ortalaması bizlere mod sayısını verecektir. Yani  $70+80=150$ .  $150/2=75$  mod değerimiz olarak ortaya çıkacaktır.

Ayrıca belirli bir puan aralığı verilir de mod değeri istenirse frekansı en yüksek olan puan aralığının aritmetik ortalaması bizlere mod değerini verecektir. Örneğin;

Puan Aralığı	Frekans
76- 80	8



71-75	11
66-70	13
61-65	10
56-60	9
51-55	3

Frekansı en çok tekrar eden puan aralığı 66-70'dir. Bu puan aralığının aritmetik ortalaması 68'dir. Yani mod 68'dir.

#### Mod değerinin bulunamadığı durumlar;

- Sayı dizisindeki bütün sayıların frekansı aynıysa (1, 2, 3, 4, 5),
- Sayı dizisindeki bütün sayılar eşit sayıda tekrar etmişse (1, 1, 2, 2, 3, 3, 4, 4, 5, 5),
- Sayı dizisindeki bütün sayılar aynıysa (1, 1, 1, 1, 1) şeklinde kısaca yazılabilir.

#### 4. AĞIRLIKLIL ORTALAMA

Ağırlıklı ortalama, puanların ortalamaya katkılarının farklı olduğu durumlarda kullanılır. Farklı ağırlıklar verilerek hesaplanan ortalamaya “ağırlıklı ortalama” denir.

- Üniversitede geçme notu hesaplanırken vize sınavının %40'ının alınması ve final sınavının %60'ının alınması ağırlıklı ortalamaya örnektir.

Geçme notu 50 olan bir puan sisteminde vize puanının %40'ı, final sınavının %60'ı değerlendirmeye katılmaktadır. Vize sınavından 70, final sınavından 90 alan bir öğrenci kaç ortalamayla dersi geçmiştir?

$$70 \times 0,40 = 28$$

$$90 \times 0,60 = 54$$

$$28 + 54 = \mathbf{82} \text{ ortalamayla dersi geçmiştir.}$$

## MERKEZİ DAĞILIM ÖLÇÜLERİ

Veri grubunu temsil edebilecek, merkezden uzaklaşan değerlerdir. Merkezi dağılım ölçüleri *standart sapma*, *varyans*, *ranj* (açıklık, dizi genişliği, çeyrekler açıklığı ve bağıl değişkenlik katsayısı) 'dır.

### 1. STANDART SAPMA

Bir veri grubundaki her bir verinin aritmetik ortalamadan ne kadar uzaklaştığını gösteren bir merkezi yayılım (dağılım) ölçüsüdür. Bir puan dağılımındaki ölçme sonuçlarının aritmetik ortalamadan ne kadar uzaklaştığını bulmak amacıyla yapılır.

Formülü:

Standart Sapma

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

$X_i$  : i'inci öğrencinin ham başarı puanı.

$\bar{X}$  : Sınıfın aritmetik ham başarı ortalaması

$S$  : Standart sapma

$n$  : Bağıl değerlendirmeye katma sınırı dikkate alınarak sınıftaki devam koşulunu yerine getirmiş öğrenci sayısı

**Standart sapma bulunurken;**

1. Veri grubunun aritmetik ortalaması bulunur.
2. Her bir verinin aritmetik ortalama ile farkının kareleri toplamı bulunur.
3. Bulunan toplam, veri sayısının bir eksiğine bölünür.
4. Bölümün karekökü alınır. Elde edilen sayı standart sapmadır.

Standart sapma değeri büyüdükçe veya küçüldükçe bize, o dağılıma ait farklı bilgiler verir.

### Standart sapma büyük olduğunda;

1. Öğrenciler arasındaki farklılaşmanın büyük olduğu,
2. Uygulanan testin ayırt ediciliğinin yüksek olduğu,
3. Testin uygulandığı grubun heterojen olduğu gibi yorumlar yapılabilir.

Standart sapma büyük olduğunda puanlar geniş bir alana yayılacak ve basık bir görünüme kavuşacaktır. Bu yüzden bu şekildeki dağılımlara “**basık dağılımlar**” denir.

### Standart sapma küçük olduğunda ise;

1. Öğrenciler arası farklılaşmanın az olduğu,
2. Uygulanan testin ayırt ediciliğinin düşük olduğu,
3. Testin uygulandığı grubun homojen olduğu gibi yorumlar yapılabilir.

Standart sapma küçük olduğunda puanlar dar bir alanda kalacak ve sivri bir görünüme kavuşacaktır. Bu yüzden bu şekildeki dağılımlara “**sivri dağılımlar**” denir.

	Standart Sapma	Büyük	Küçük
1	Grup	Heterojen	Homojen
2	Öğrencilerin yokladığı davranışları bakımından	Farklı	Benzer
3	Alınan puanlar birbirine	Uzak	Yakın
4	Öğrenciler arası farklılaşma	Fazla	Az
5	Öğrencilerin öğrenme düzeyleri	Farklı	Benzer
6	Dağılım	Basık	Sivri

## 2. VARYANS

Varyans, standart sapma değerinin karesinin alınmasıyla hesaplanır. Aynı zamanda varyans, standart sapma formülündeki karekökün işleme dâhil edilmemesi suretiyle de elde edilebilir. Standart sapma için yapılmış olan tüm yorumlar varyans için de geçerlidir. Zaten varyansın karekökü, standart sapma olarak adlandırılır.

Varyans, bir puan dağılımına ait her bir değerin, dağılımın ortalamasından ne kadar uzak olduğuyla ilgilidir. Varyans söz konusu sapmaların ortalama değerini ölçmektedir. Özellikle çok fazla kişinin katıldığı (2000, 3000 kişi vb.) ölçmelerde daha duyarlı ölçümler yapar.

### 3. RANJ (AÇIKLIK, DİZİ GENİŞLİĞİ)

Puan dağılımında en yüksek ölçme sonucuyla en düşük ölçme sonucu arasındaki farktır. Ranj büyüdükçe grup ölçülen özellik bakımından heterojenleşir. Küçüldükçe homojen hale gelir.

**Örnek:** 10, 12, 15, 15, 20, 30, 30 sayı dizisinin açıklığı kaçtır?

**Çözüm:** Sayı dizisindeki en büyük sayı ile en küçük sayı arasındaki farka *açıklık* denir. Buradaki en büyük sayı 30, en küçük sayı 10 olduğu için  $30-10=20$  işlemi sonucunda elde edilen sayımız bize açıklık değerini verecektir.

**Örnek:** 23, 26, 45, 89, 87, 99, 63, 12, 74, 71, 58, 60 sayı dizisinin açıklığı kaçtır?

**Çözüm:** Sayı dizisindeki en büyük sayı ile en küçük sayı arasındaki farka *açıklık* denir. Buradaki en büyük sayı 99, en küçük sayı 12 olduğu için  $99-12=87$  işlemi sonucunda elde edilen sayı, bize açıklık değerini verecektir.

**Örnek:** Bir sınıftaki öğrencilerin aldıkları notların aralığı ve bu notları alan öğrencilerin sayısı aşağıdaki tabloda verilmiştir. Bu tabloya göre alınan notların açıklığı kaçtır?

Alınan Puan Aralığı	Frekans
80-100	5
60-80	4
40-60	9
20-40	8
0-20	5

**Çözüm:** Sayı dizisindeki en büyük sayı ile en küçük sayı arasındaki farka *açıklık* denir. Buradaki en büyük sayı 80-100 arasında, en küçük sayı 0-20 arasında olduğu için en büyük

sayıyı bulmak için  $(80+100)/2$ ; en küçük sayıyı bulmak için  $(0+20)/2$  işlemi sonucunda elde edilen  $90-10=80$  sayımız bize açıklık değerini verecektir.

#### **4. ÇEYREKLER AÇIKLIĞI**

Bir veri grubundaki terimler küçükten büyüğe doğru sıralandığında tam ortadaki sayı medyandır. Medyan sayı dizisini ortadan ikiye böler. Medyan bulunduktan sonra hem medyandan küçük olan sayıların medyanı bulunur hem de medyandan büyük olan sayıların medyanı bulunur.

Medyandan (ortancadan) küçük olan sayıların meydana alt çeyrek (Q1), medyandan büyük olan sayıların medyanın üst çeyrek (Q3) denir.

$$\text{Çeyrekler açıklığı} = \text{Üst çeyrek (Q3)} - \text{Alt çeyrek (Q1)}$$

Çeyrekler açıklığı, bir sayı dizisindeki aşırı uç değerlerin sayı dizisinin yorumlanmasında oluşabilecek yanlışlıkların önüne geçer. Bir sayı dizisinde çoğunluk yüksek not almasına rağmen sadece bir kişinin 0 (sıfır) alması, sayı dizisinin heterojen olduğu yorumuna neden olur. Ancak çeyrekler açıklığı ile sayı dizisinin aşırı uç değerleri arındırılır ve daha sağlıklı yorumlar yapılmasını sağlar.

12, 16, 17 19, 21, 25, 26, 30, 35, 40, 49 bu sayı dizisinin meydana 25'tir. Sayı dizisi medyanla ortadan ikiye bölünmüştür. Sayı dizisinin üst grubunda yer alan sayı dizisinin medyanı 35'tir. Sayı dizisinin alt grubunda yer alan sayı dizisinin medyanı 17'dir. Çeyrekler açıklığını bulmak için  $35-17=18$  işlemi yapılır. Çeyrekler açıklığı 18'dir. Çeyrekler açıklığı aşırı uç noktalardan etkilenmemiştir. Ancak bu işlem yapıldıktan sonra ortaya çıkan sonuç ne kadar yüksekse puan dağılımı heterojen; ortaya çıkan sonuç ne kadar düşükse puan dağılımı homojendir yorumu yapılabilir.

#### **5. BAĞIL DEĞİŞKENLİK KATSAYISI**

Standart sapmayı aritmetik ortalamaya bölüp, daha sonra bu bölümü 100 (yüz) ile çarparak hesapladığımız bir değişkenlik ölçüsüdür. Aritmetik ortalama ve standart sapmayı içerisinde bağladığı için dağılım hakkında daha çok bilgi verdiğini söyleyebiliriz. Bağlı değişkenlik katsayısı "V" sembolü ile ifade edilir.

$$\text{Formülü} = v = \frac{s}{\bar{X}} \times 100$$

V: Bağıl değişkenlik katsayısı

S: Standart sapma

X: Aritmetik ortalama

Bağıl değişkenlik katsayısı 25 üzerinde bir sonuç çıkarsa heterojen dağılım vardır ve basık bir grafik çıkar. 20-25 arasında bir sonuç çıkarsa simetrik bir dağılım vardır ve normal grafik çıkar. 20'den aşağı bir sonuç çıkarsa homojen bir dağılım vardır ve sivri bir grafik çıkar.

Farklı puanlama sistemlerinden elde edilen standart sapmaları doğrudan karşılaştırmak yanlıştır. İşte bu durumda bağıl değişkenlik katsayısını kullanmak gerekir. Bağıl değişkenlik katsayısı büyük olan grup, diğer gruba göre daha heterojen olacaktır.

Bağıl değişkenlik katsayısı, birbirinden farklı grupları, grup farklılaşmaları açısından kıyaslarken ya da iki farklı puanlama sisteminde (birimlerin farklı olduğu iki ölçme sisteminde) olan grupların farklılaşmalarının karşılaştırılmasında kullanılır.

Bağıl değişkenlik katsayısının büyük olduğu grubun diğerine göre daha heterojen, bağıl değişkenlik katsayısının küçük olduğu grubun diğerine göre homojen olduğunu söylenebilir.

**Not:** Bağıl değişkenlik katsayısı sadece eşit oranlı ölçeklerden elde edilen verilerle hesaplanır. Eşit aralıklı, sınıflama ve sıralama ölçeklerinden elde edilen verilerle kullanılamaz.

## STANDART PUANLAR

### Standart Puanlar:

Bir veri grubundaki bir puanın, standart sapmaya göre ortalamadan ne kadar uzakta olduğunu tespit edebilmek ya da farklı veri gruplarına ait verileri birbiriyle karşılaştırabilmek için bu puanın standartlaştırılması yani aynı standart puana dönüştürülmesi için Z ve T puanı kullanılır. Her iki puanda öğrencinin sınıf içerisindeki başarı sırasını belirlemek amacıyla kullanılır. Z ve T değeri büyüdükçe öğrencinin sınıf içerisindeki başarısı artar. Z ve T değeri küçüldükçe öğrencinin sınıf içerisindeki başarısı azalır. Standart puanlar; Z ve T puanlar olarak ikiye ayrılır:

### Z Puanı:

Z puanı, yapılan ölçümlerden elde edilen puanların aritmetik ortalamasının 0 (Sıfır), standart sapmanın 1 (Bir) kabul edildiği puanlardır.

$$Z \text{ Puanı} = \frac{\text{Öğrenci Puanı} - \text{Aritmetik Ortalama}}{\text{Standart Sapma}}$$

Z Puanının 0 (Sıfır) çıkması dönüştürülen puanın ortalamaya eşit, pozitif çıkması ortalamadan büyük, negatif çıkması ortalamadan küçük olduğunu gösterir.

Z Puanı değeri yükseldikçe öğrencinin sınıf içerisindeki başarısı artar, Z Puanı değeri azaldıkça öğrencinin sınıf içerisindeki başarısı azalır. Z Puanı değeri genel olarak -3Z ile +3Z arasında değişir. Yani öğrencinin Z puanı -3'e yaklaştıkça sınıf içi başarısı azalır, Z puanı +3'e yaklaştıkça sınıf içi başarısı artar.

### T Puanı:

Z puanının yorumlanmasının zor olduğu durumlarda Z puanının tam sayıya dönüştürülmesi T puanı yardımıyla olur. T puanı, verilen puanların ortalaması 50 (Elli), standart sapması 10 (On) olan puanlara dönüştürülür.

Z puanından T puanına geçiş;

$T = 10Z + 50$  formülüyle yapılır.

## STANDART PUANLAR

**Z puanı:**

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

$X_i$  = Bireyin puanı

$\bar{X}$  = Grup ortalaması

$S_x$  = Grup standart sapması

**T puanı:**

$$T = 10 \cdot Z + 50$$

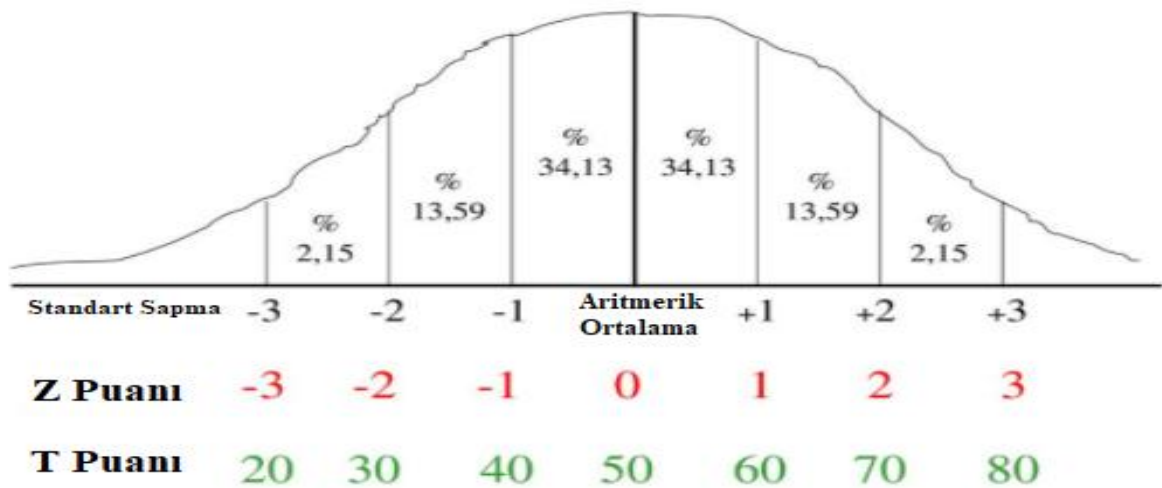
### Notlar:

Z ve T değeri büyüdükçe sınıf içinde başarı sırası artar.

Z ve T değerleri için aynı yorumlar yapılabilir.

Z ve T değeri arasında pozitif yüksek korelasyon vardır.

Z ve T puanı bilinirse öğrencinin sınıf içindeki başarısı bilinir. Sınıfın yüzde kaçından daha başarılı olduğu ya da olmadığı ortaya konulabilir. Z değerinin T değerine dönüştürülmesinin amacı öğrenci puanlarını eksi değerlerden kurtarmaktır.



Z değeri genel olarak  $-3z$  ile  $+3z$  arasında bir değer alır. T değeri genel olarak 20 ile 80 arasında bir değer alır. Grafikte görülen bilgiler, yüzdeler, Z ve T değerlerinin aldığı değerler asla değişmez.



## Test Maddelerinin Kapasitesi (Madde İstatistikleri)

### Madde Güçlük İndeksi

Yapılan bir sınavdaki her bir sorunun zorluk derecesini belirlemek amacıyla yapılan istatistik, madde güçlük indeksidir. Madde güçlük indeksinin hesaplanabilmesi için;

Bir sınavda aldıkları notlara göre öğrenciler, en başarılı olandan en başarısız olana doğru sıralanırlar. Daha sonra en başarılı olan %27'lik üst grup ile %27'lik alt grubun soruya verdikleri cevaplara göre karşılaştırmalar sağlanır.

Madde güçlük indeksinin hesaplanması aşağıdaki gibidir:

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

$p_j$ : 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

İşlem yapıldıktan sonra artık yorumlanması kalmıştır. İşlem sonucunda ortaya çıkacak sayı 0 ile 1 arasında bir değerdir. Başka bir sonuç çıkamaz. Çıkan sonucun yorumlanması aşağıdadır.

0.00 – 0.19	Çok zor
0.20 – 0.39	Zor
0.40 – 0.59	Orta güçlükte
0.60 – 0.79	Kolay
0.80 – 1.00	Çok kolay

Seçme yapılan bir sınavda çok zor ve zor, başarı sınavlarında orta güçlükte, izleme testlerinde kolay ve çok kolay sorulardan oluşan bir testin yapılması tercih edilmelidir.

## Madde Ayırt Edicilik İndeksi

Madde ayırt edicilik indeksinde de ilgili sorunun bilenle bilmeyen öğrenciyi ayırıp ayırmadığına bakılmasıdır. Bir sınavda aldıkları notlara göre öğrenciler, en başarılı olandan en başarısız olana doğru sıralanırlar. Daha sonra en başarılı olan %27'lik üst grup ile %27'lik alt grubun soruya verdikleri cevaplara göre karşılaştırmalar sağlanır.

Madde ayırt edicilik indeksi hesaplanma formülü aşağıdaki gibidir:

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

$r_{px}$ : Madde ayırtıcılık gücü indeksi

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

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

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

İşlem yapıldıktan sonra ortaya çıkan sayısal ifadenin yorumlanması önemlidir. Madde ayırt ediciliği -1 ile +1 arasında bir değer alır. Ortaya çıkan sonuçla ilgili yorumlar aşağıdaki gibi yapılabilir.

( -1.00) – (+ 0.19)	Testten atılmalıdır.
(+ 0.20) – (+ 0.29)	Düzeltilmeli.
(+ 0.30) – (+ 0.39)	İyi bir madde (ayırt edici)
(+ 0.40) – (+ 1.00)	Mükemmel bir madde

Bir sınavdaki soruların her birinin madde ayırt ediciliklerinin bilinmesi önemlidir. Ayırt edici olmayan sorular; bilenle bilmeyeni ayırt edemeyeceği gibi sınavın amacına ulaşmasını da engelleyecektir. Sınav sonunda yapılacak olan madde istatistikleri sayesinde testin kalitesi artırılabilir, gerekli düzeltmeler yapılabilir, soruların güçlük ve ayırt edicilikleri geliştirilebilir. Ayırt edici özelliğe sahip bir soru yıllarca kullanılabilecektir.

## TESTİN ORTALAMA GÜÇLÜĞÜ

Elimizde var olan testin bütününün ne kadar zor ve kolay olduğunu gösterir. Ayrıca öğrencilerin başarılarının ne düzeyde olduğunu, testin ayırt ediciliğinin ne düzeyde olduğunu, öğrencilerin hedeflenen davranışları ne düzeyde kazandıklarını gösterir. Testin ortalama

güçlüğü bilirse; öğrencilerin ilgili testteki soruları ne düzeyde bildikleri, testin zorluk düzeyi, programın amaçlarına ulaşip ulaşmadığı hakkında tahminler yürütülebilir.

Testin ortalama güçlüğü, testin aritmetik ortalamasının testteki soru sayısına bölünerek hesaplanır.

$$\frac{\text{Testin Aritmetik Ortalaması}}{\text{Testteki Soru Sayısı}}$$

Test	Aritmetik Ortalama	Testteki Soru Sayısı
1.Test	52	80

İşlem yapıldıktan sonra ortaya çıkan sonuç 0 ile +1 arasındadır. Başka bir sonuç çıkamaz. Çıkan sonuç 0'a (sıfır) yaklaştıkça; testin zor olduğu, öğretimin amacına ulaşmadığı, öğrencilerin konuya ilişkin bilgilerinin az olduğu vb. yorumlar yapılabilir. Çıkan sonuç +1'e (bir) yaklaştıkça testin kolay olduğu, öğretimin amacına ulaştığı, öğrencilerin konuya ilişkin bilgilerinin yeterli düzeyde olduğu vb. yorumlar yapılabilir.

Testler	Aritmetik Ortalama	Testteki Soru Sayısı
1.Test	22	30
2.Test	25	40
3.Test	38	60
4.Test	56	80
5.Test	68	100

**Tabloya göre öğrencilere en kolay ve en zor gelen testler hangileridir?**

**Çözüm:** Soru sayılarının farklı olduğu testlerde, testlerin zorluk derecelerine bakılması için testlerin ortalama güçlüklerinin hesaplanması gerekir. Testin ortalama güçlüğü; testin aritmetik ortalamasının testteki soru sayısına bölünmesiyle hesaplanır. Yukarıdaki tablodaki testlerin ortalama güçlüklerine bakıldığında aşağıdaki gibi sonuç çıkmıştır.

1. Test:  $22/30 = 0.73$
2. Test:  $25/40 = 0.62$
3. Test:  $38/60 = 0.63$
4. Test:  $56/80 = 0.70$
5. Test:  $68/100 = 0.68$

Çıkan sonuçlara bakıldığında en kolay 1. Test, en zor 2. Test olarak görülmektedir.

## MADDE VARYANSI

Madde varyansı, bir maddeyi doğru cevaplayanların oranıyla yani madde güçlük indeksi ile o maddeyi doğru cevaplayamayanların (yanlış yapan ya da boş bırakan) oranının çarpımına eşittir. Madde varyansını, o maddeyle ölçmek istediğimiz özellik açısından öğrenciler arasındaki farklılıkları belirlemek için kullanırız. Yani madde varyansı, öğrencilerin bir soruya verdiği tepkiler bakımından ne kadar çeşit gösterdiklerini gösterir. Eğer öğrenciler bir soruda ne kadar çeşitli olursa madde varyansı da o kadar yüksektir.

$$\text{Madde varyansı} = \text{Madde güçlüğü} \times (1 - \text{Madde güçlüğü})$$

$$\text{Madde güçlüğü} \times (1 - \text{Madde güçlüğü}) = \text{Madde Varyansı}$$

$$1.00 \times 0.00 = 0.00$$

$$0.90 \times 0.10 = 0.09$$

$$0.80 \times 0.20 = 0.16$$

$$0.70 \times 0.30 = 0.21$$

$$0.60 \times 0.40 = 0.24$$

$$0.50 \times 0.50 = \mathbf{0.25}$$
 Madde varyansının varabileceği en yüksek değerdir.

$$0.40 \times 0.60 = 0.24$$

$$0.30 \times 0.70 = 0.21$$

$$0.20 \times 0.80 = 0.16$$

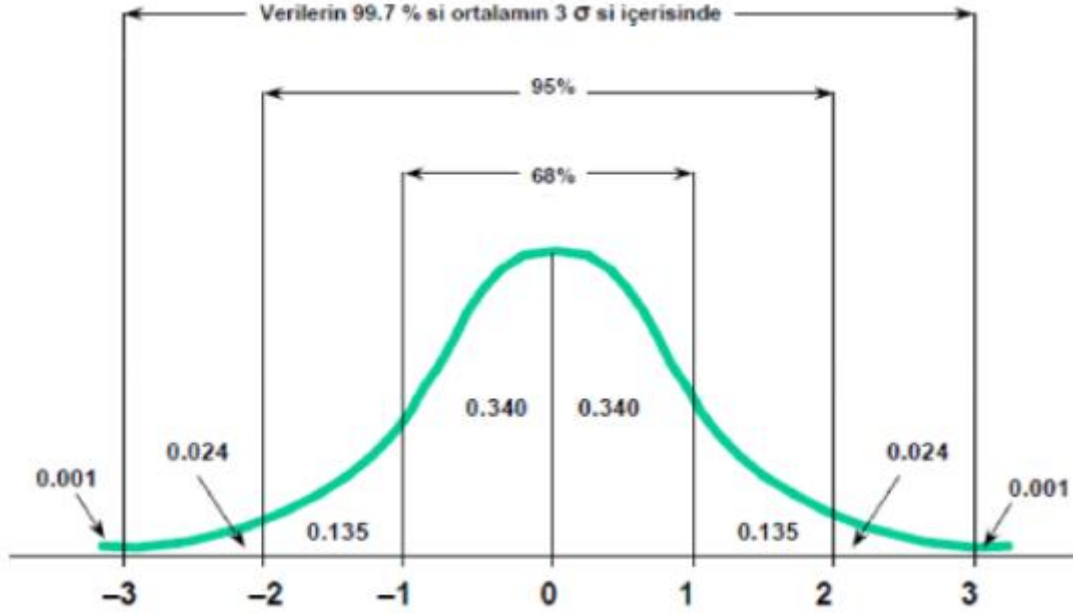
$$0.10 \times 0.90 = 0.09$$

$$0.00 \times 1.00 = 0.00$$

**Not:** Madde varyans değeri küçüldükçe öğrencilerin ilgili soruya verdikleri cevabın benzer olduğu sonucuna varılacaktır. Madde varyansının ulaşabileceği en yüksek değer 0.25'tir. Bu sonucun yorumu, ilgili soruya öğrencilerin yarısı doğru yarısı yanlış cevap vermiş demektir. İlgili soruya öğrencilerin çoğunluğu doğru cevap verse de veya öğrencilerin çoğunluğu yanlış cevap verse de madde varyansı değeri küçük çıkacaktır. Çünkü öğrencilerin verdikleri cevaplar benzerlik göstermektedir.

## DAĞILIM GRAFİKLERİ

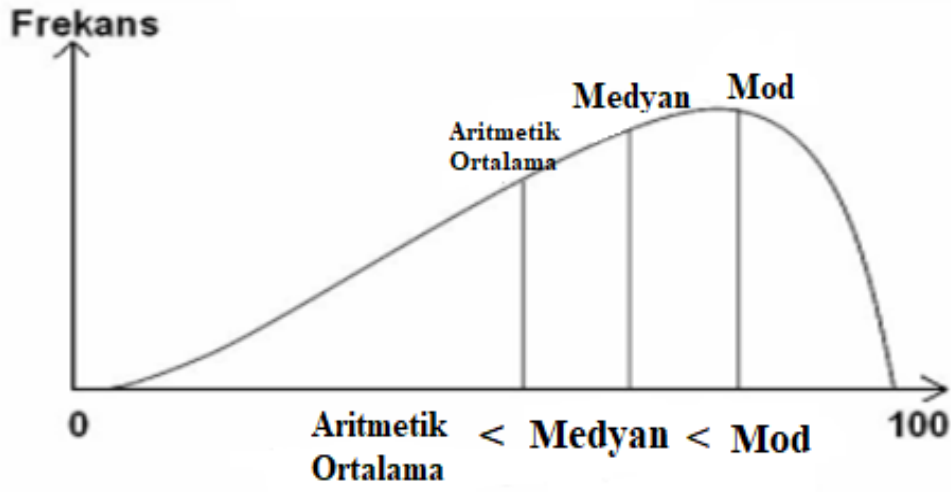
### NORMAL DAĞILIM EĞRİSİ:



### Normal dağılım eğrisinde;

- Mod, medyan ve aritmetik ortalama birbirinin aynısı ya da benzeridir.
- -1 ile +1 standart sapma bölgesinde yer alan notlar, sınava giren öğrencilerin % 68'inin notlarını gösterir.
- -2 ile +2 standart sapma bölgesinde yer alan notlar, sınava giren öğrencilerin % 95'inin notlarını gösterir.
- -3 ile +3 standart sapma bölgesinde yer alan notlar, sınava giren öğrencilerin % 99'unun notlarını gösterir.
- Bilenle bilmeyen birbirinden ayrılmıştır.
- Başarı testlerinde aranan bir grafikdir.

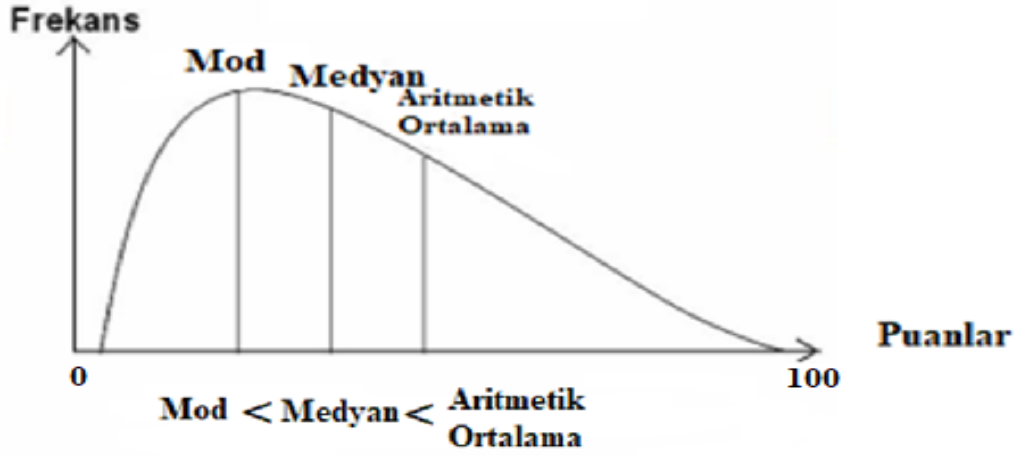
## SOLA ÇARPIK (SAĞA YIĞILIMLI) GRAFİKLER



Sola çarpık (sağa yığılmalı) dağılım eğrisinde;

- Aritmetik Ortalama < Medyan < Mod şeklinde sıralanır.
- Aritmetik ortalamanın küçük olmasının nedeni bazı öğrencilerin çok düşük notlar alarak ortalamayı düşürmeleridir.
- Öğrenci grubu başarılıdır.
- Öğretim amacına ulaşmıştır.
- İzleme testlerinde istenen bir grafikdir.
- Uygulanan sınav öğrencilere kolay gelmiştir.
- Tam öğrenme modeline uygun sonuçlar gözlenmiştir.
- Öğrenciler homojendir.
- Standart sapma küçüktür.

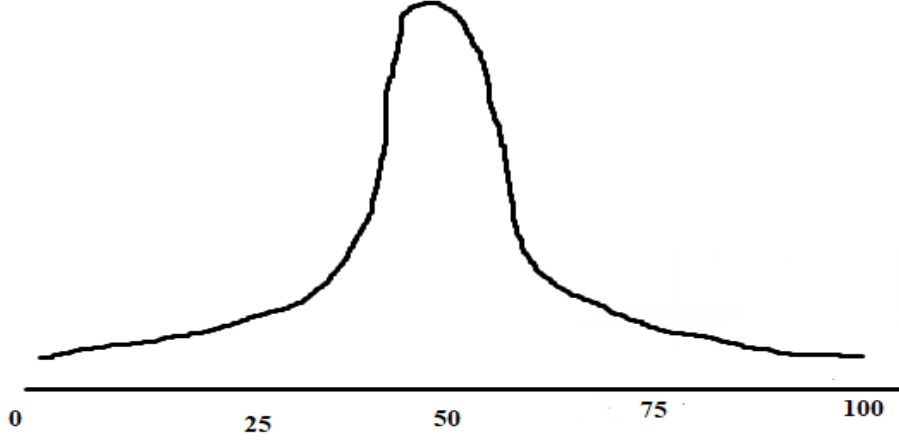
## SAĞA ÇARPIK (SOLA YIĞILIMLI) GRAFİKLER



Sağa çarpık (sola yığılmalı) dağılım eğrisinde;

- $Mod < Medyan < Aritmetik Ortalama$  şeklinde sıralanır.
- Aritmetik ortalamanın büyük olmasının nedeni bazı öğrencilerin çok yüksek notlar alarak ortalamayı yükseltmeleridir.
- Öğrenci grubu başarısızdır.
- Öğretim amacına ulaşmamıştır.
- Seçme sınavlarında istenen bir grafikdir.
- Uygulanan sınav öğrencilere zor gelmiştir.
- Öğrenciler homojendir.
- Standart sapma küçüktür.

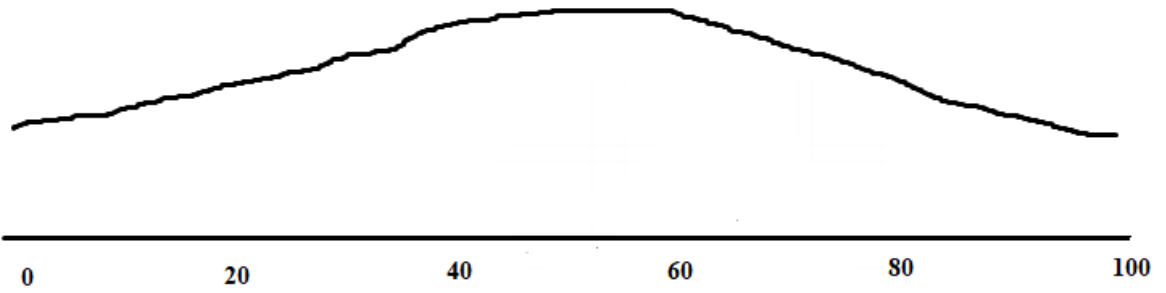
## SİVRİ DAĞILIM GRAFİĞİ



Sivri dağılım grafiğinde;

- Öğrenci puanları belirli bir puan aralığına yığılmıştır.
- Öğrenciler homojendir.
- Öğrenci grubu başarısı hakkında net bir yorum yapılamaz.
- Sınav ayırıcı özelliklere sahip değildir.
- Standart sapma küçüktür.

## BASIK DAĞILIM GRAFİĞİ

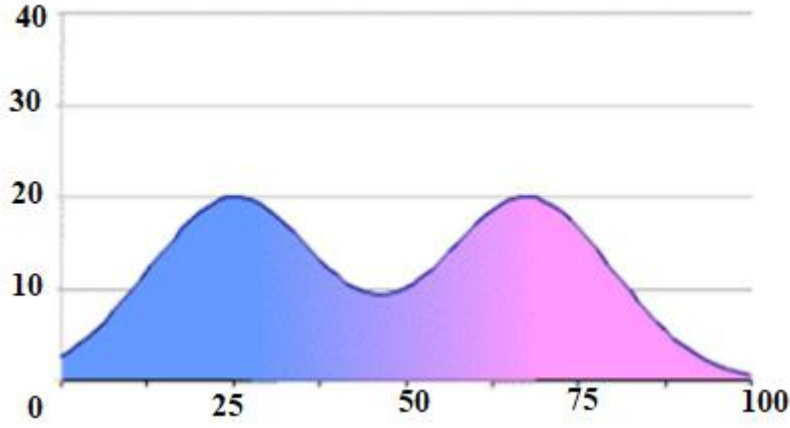


Basık dağılım grafiğinde;

- Öğrenci puanları grafiği geneline yayılmıştır.
- Öğrenciler heterojendir.
- Öğrenci grubu başarısı hakkında net bir yorum yapılamaz.
- Standart sapma büyüktür.



## ÇİFT MODLU (BİMOD, İKİ MODLU) DAĞILIM



Çift modlu dağılım grafiğinde;

- Öğrenci puanları grafiği geneline yayılmıştır.
- Öğrenciler heterojendir.
- Standart sapma büyüktür.
- Başarı düzeyleri çok farklılık gösteren bir sınıfın dağılımı bu şekildedir.