İSTATİSTİK II
OLASILIK

Dr. Öğretim Üyesi Muhlis ÖZDEMİR

OLASILIK

- Olasılık; bir deneme sonrasında, ilgilenilen olayın, tüm mümkün olaylar içinde ortaya çıkma ihtimali olarak tanımlanabilir.

- Bir olayın gerçekleşme olasılığının hesaplanması, buna ilişkin olarak geliştirilen olasılık kuramına dayanır.

- Bu kuram denemelerin olası sonuçları ile ilgilenmektedir.

OLASILIK

- Araştırmacılar, çoğunlukla evrenin tamamına zaman, maliyet gibi nedenlerden ötürü erişemezler bu nedenle örneklemden hareketle çıkarımda bulunulmaya çalışılır.

- Örneklemden elde edilen bulguların gerçekte evrende olup olmadığının araştırılmasında olasılık kuramından yararlanılabilir.

OLASILIK

- Herhangi bir olayın mümkün tüm sonuçlarına örneklem uzayı denir ve "S" ile ifade edilir. *S* harfi İngilizce *’Space’* kelimesinden gelmektedir.

- Bir deneye ait tüm mümkün sonuçlar S={} şeklinde ifade edilir.

- Örneğin bir para deneyine ait örneklem uzayı S={Yazı, Tura} şeklinde,

ya da bir zar deneyine ait örneklem uzayı, bir zarın yüzeyinde yer alan tüm sayılar olacağından S={1,2,3,4,5,6} şeklinde olacaktır.

Bir olayın olasılığı

- Bir olay için olasılık 0 ve 1 arasında yer alacaktır. Bir başka ifade ile bir olayın gerçekleşmesi %0 ile %100 arasında olabilir. Olayın gerçekleşmesi mümkün değilse %0 olasılık;

$$\frac{1}{100} * 0 \ = 0$$

Gerçekleşmesi kesin ise %100 olasılık;

$\frac{1}{100} * 100 \ = 1$ olacaktır.

Olay

- İstatistikte olaylar bağımlı olaylar ve bağımsız olaylar olarak ikiye ayrılır.

- Bir olayın gerçekleşmesi başka bir olayın gerçekleşmesinden etkileniyorsa bağımlı olay,

- Bir olayın gerçekleşmesi başka bir olayın gerçekleşmesinden etkilenmiyorsa bağımsız olay olarak adlandırılır.

Bağımsız olay

- Bağımsız bir olayın olasılığının hesaplanması için iki temel kural söz konusudur. Bunlar toplama kuralı ve çarpma kuralıdır.
- Toplama Kuralı:

$$P(A\ veya\ B) = P(A) + P(B)$$

Yukarıda yer alan eşitlik toplama kuralını göstermektedir.

10 kişilik bir banka kuyruğunda yer alanların 6'sı erkek, 4'ü kadın ise bu kuyruktan rastgele seçilecek birinin erkek olması olasılığı;

$P(Erkek) = \frac{6}{10} = 0{,}6 = \%60$ olacaktır.

Bağımsız olay

10 kişilik bir banka kuyruğunda yer alanların 6’sı erkek, 4’ü kadın ise bu kuyruktan rastgele seçilecek birinin kadın olması olasılığı;

$P(Kadın) = \frac{4}{10} = 0,4 = \%40$ olacaktır.

Erkek veya kadın olması olasılığı ise;

$$P(Erkek\ veya\ Kadın) = P(Erkek) + P(Kadın)$$
$$P(Erkek\ veya\ Kadın) = 0,6 + 0,4 = 1$$

Bağımsız olay

İki mümkün sonucu bulunan bir olayın(erkek ya da kadın olma) gerçekleşmesi ihtimali P ile ifade edilirse eğer gerçekleşmemesi ise Q ile ifade edilecektir.;

$P + Q = 1 = \%100$ olacaktır.

Birbirini dışta tutan ve tutmayan olaylar:

Banka kuyruğundan çekilecek bir kişinin hem erkek hem de kadın olması mümkün olmayacağından bu iki olay birbirini dışlayan olaylardır.

Birbirini dışta tutmayan olaya örnek olarak ise kişinin hem kadın hem de sırada en önde yer alması verilebilir.

Bağımsız olay

- Bağımsız bir olayın olasılığının hesaplanması için iki temel kural söz konusudur. Bunlar toplama kuralı ve çarpma kuralıdır.
- Çarpma Kuralı:

$$P(A\ ve\ B) = P(A) * P(B)$$

Yukarıda yer alan eşitlik çarpma kuralını göstermektedir.

Birbirini dışta tutan, yani bağımsız olan olayların birlikte ortaya çıkış olasılığı, olayların tek tek ortaya çıkış olasılıklarının çarpımına eşittir.

10 kişilik bir banka kuyruğunda yer alanların 6'sı erkek, 4'ü kadın ise bu kuyruktan rastgele seçilecek iki kişinin de erkek olması olasılığı;

$P(Erkek\ ve\ Erkek) = \frac{6}{10} * \frac{6}{10} = 0{,}6 * 0{,}6 = \%60 * \%60 = \%36$ olacaktır.

Bağımsız olay

10 kişilik bir banka kuyruğunda yer alanların 6’sı erkek, 4’ü kadın ise bu kuyruktan rastgele seçilecek ilk kişinin erkek ikinci kişinin ise kadın olması olasılığı;

$P(Erkek\ ve\ Kadın) = \frac{6}{10} * \frac{4}{10} = 0{,}6 * 0{,}4 = \%60 * \%40 = \%24$ olacaktır.

Bağımsız olay

İstatistik dersi vize sınavının test sorularından oluştuğunu göre A isimli öğrenci, şansına güvenerek soruların tamamını sallamaya karar vermiştir. 10 sorudan oluşan sınavdaki her bir soru için 5 şık söz konusu olduğuna göre bu öğrencinin sınavdan 100 alma ihtimalini hesaplayınız.

$$P(1,2,3,4,5,6,7,8,9,10) = \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} * \frac{1}{5} =$$

$$0{,}2 * 0{,}2 * 0{,}2 * \ 0{,}2 * \ 0{,}2 * \ 0{,}2 * \ 0{,}2 * \ 0{,}2 * \ 0{,}2 * \ 0{,}2 = 0{,}0000001024$$

olacaktır.(%0,000001024 ile %0’a oldukça yakın bir değerdir.)

Bir olayın olasılığını etkileyen faktörler

Tesadüfi olayların tümü aynı değildir.

Olayların bağımsız veya bağımlı olması: Bir olayın ortaya çıkması diğer olayın ortaya çıkmasından etkilenmiyorsa **bağımsız**, etkileniyorsa **bağımlı** olaylardır. Örneğin araba yıkatmak ile yağmur yağması olayı **bağımsız olay** iken ders çalışmak ile sınavdan alınan not **bağımlı olay**a örnektir.

Bir olayın olasılığını etkileyen faktörler

Bir olayın olasılığı örneklem tipinden etkilenir. İadeli örneklem ve iadesiz örneklem olmak üzere ikiye ayrılır. Bir oyun destesinden iki kart seçilirken ilk örneklemeden sonra kart desteye yeniden yerleştirilirse iadeli yerleştirilmezse iadesiz örnekleme olur.

NOT: istatistikte genellikle olayların bağımsız olduğu ve iadeli örnekleme yapıldığı varsayılır.

Standart Normal Dağılım

Evrende gözlenen değişkenlerin büyük çoğunluğunun çan eğrisine benzer bir dağılım gösterdikleri kabul edilmektedir. Değişkenlere ilişkin verilerin oluşturduğu çan eğrisine benzer olan bu eğriye normal dağılım eğrisi, bu eğrinin yatay eksene göre gösterdiği dağılıma da normal dağılım denilmektedir.

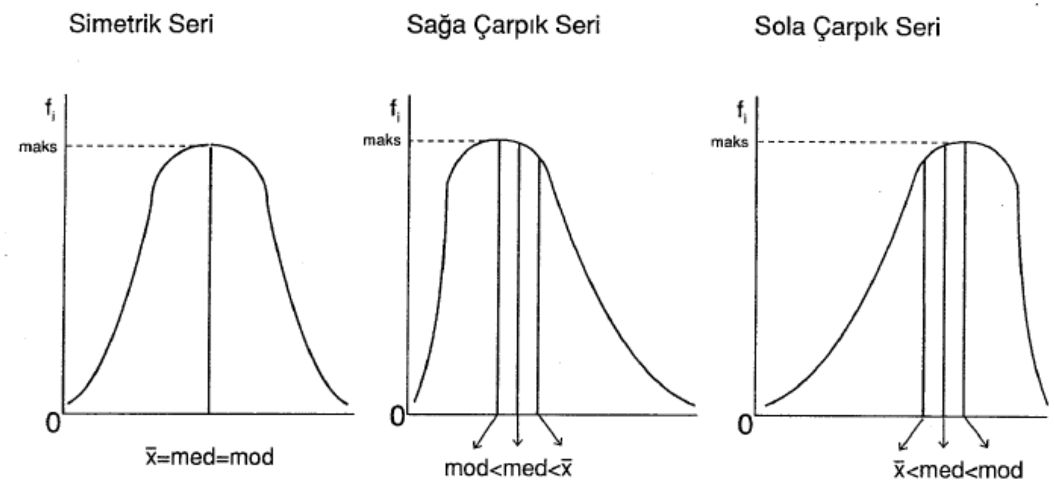
Normal dağılım, gerçekte, her biri aritmetik ortalama ve standart sapma değerleri ile tanımlanabilen dağılımların bir kümesidir. Bu dağılımların bazıları daha geniş ve basık, bazıları da daha dar ve sivridir.

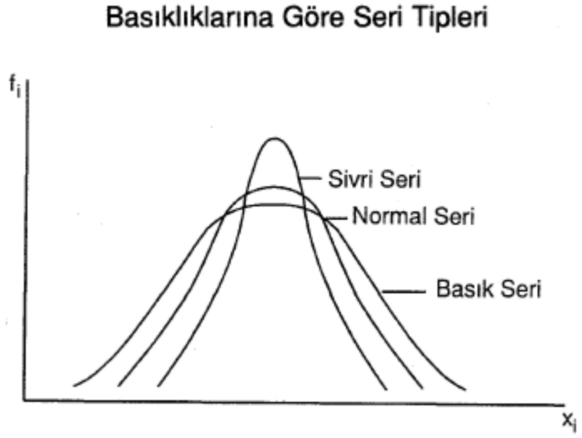
Standart Normal Dağılım

Normal dağılım aşağıda yer alan 4 özelliği taşır;

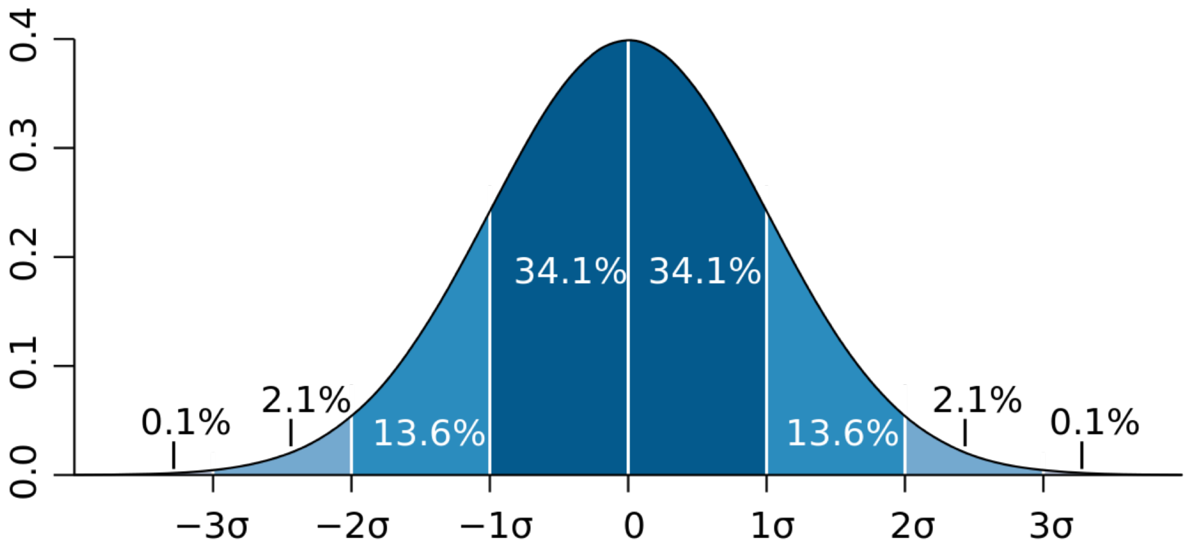
- Eğri dikey eksene göre simetriktir. Puanların yarısı eksenin sağında, diğer yarısı ise eksenin sol tarafındadır.
- Puanlar merkez(aritmetik ortalama ya da dikey eksen) etrafında kümelenme eğilimi gösterir.
- Mod, medyan ve aritmetik ortalama birbirine eşittir.
- Dağılımın her iki ucu giderek yatay eksene yaklaşır, ancak hiçbir zaman bu eksene değmez. Normal dağılım altında kalan alan sınırsızdır.

Standart Normal Dağılım





Standart Normal Dağılım



Standart Normal Dağılım

- Standart normal dağılım altındaki puanların ortalaması 0 ve standart sapması 1'dir. Ortalamanın sol tarafındaki birimler negatif(-1,-2 gibi), sağ tarafındaki birimler pozitiftir(+1, +1 gibi). İki standart sapma değeri arasındaki uzaklıklar birbirine eşittir.

- Standart normal dağılım tablolarından yararlanılarak birimlerin %68,26 , %95,5 , %99 ve %99,73'ünün aritmetik ortalama etrafında hangi değerler arasında değer aldığı ortaya çıkarılabilir.

Standart Normal Dağılım

Birimlerin;

- **%68,26'sı :** $\overline{x} \pm \boldsymbol{\sigma}$
- **%95'i :** $\overline{x} \pm \mathbf{1,96}\boldsymbol{\sigma}$
- **%95,5',:** $\overline{x} \pm \mathbf{2}\boldsymbol{\sigma}$
- **%99'u :** $\overline{x} \pm \mathbf{2,58}\boldsymbol{\sigma}$
- **%99,73'ü :** $\overline{x} \pm \mathbf{3}\boldsymbol{\sigma}$

sınırları arasında değer almaktadır.

Standart Normal Dağılım

- Örneğin ortalaması 50, standart sapması 10 olan bir sınav için normal dağılımda birim sapmalar arasındaki alanlar kullanılarak aşağıdaki yorumlar yapılabilir:

- Puanların %68,26’sı ortalama ile $\bar{\boldsymbol{x}} \pm \boldsymbol{\sigma}$ arasındadır. $\mathbf{50} \pm \mathbf{10}$ **= 40 – 60** arasındadır.
- Puanların %99,73’ü ortalama ile $\bar{\boldsymbol{x}} \pm \mathbf{3}\boldsymbol{\sigma}$ arasındadır. $\mathbf{50} \pm \mathbf{3} *$**10 = 20 – 80** arasındadır.

Standart Puanlar

Karşılaştırma yapabilmek adına standardizasyon işlemine gerek duyulur.

İstatistikte en yaygın kullanılan standart puanları Z ve t standart puanlarıdır.

Z-Puanı

Karşılaştırma yapabilmek adına ortalaması sıfır($\bar{X} = 0$) ve standart sapması($\sigma = 1$) olan ve normal dağılım gösteren standart bir puana dönüşüm gerçekleştirilir.

Z puanı, verilen bir puanın ortalamanın ne kadar altında ya da üstünde olduğunu anlamamıza yardımcı olur.

Z değeri ölçümlerin ortalamadan uzaklıklarının standart sapmaya oranını gösterir. Buna göre $1\sigma = 1Z, 2\sigma = 2Z, 3\sigma = 3Z$'dir. z puanları, normal dağılım eğrisi üzerinde karşılaştırılabilir. Çünkü puanlar aynı birimlere dönüştürülmüştür.

- $Z = \frac{x - \bar{x}}{\sigma}$ formülü ile hesaplanır.

T-Puanı

İstatistikte en çok kullanılan puan türünden bir diğeri ise T-Puanıdır.

Ham puanlardan elde edilen Z puanı, T puanlarına dönüştürmek için;

$T = 10(z) + 50$

veya

$T = 10\left(\frac{X-\bar{X}}{\sigma}\right) + 50$

yazılabilir.

Z Puanlarına Karşılık Gelen T Puanları