İstatistik: Olasılık Dağılımları

Atil Samancioglu

1 Olasılık Kütle Fonksiyonu (PMF)

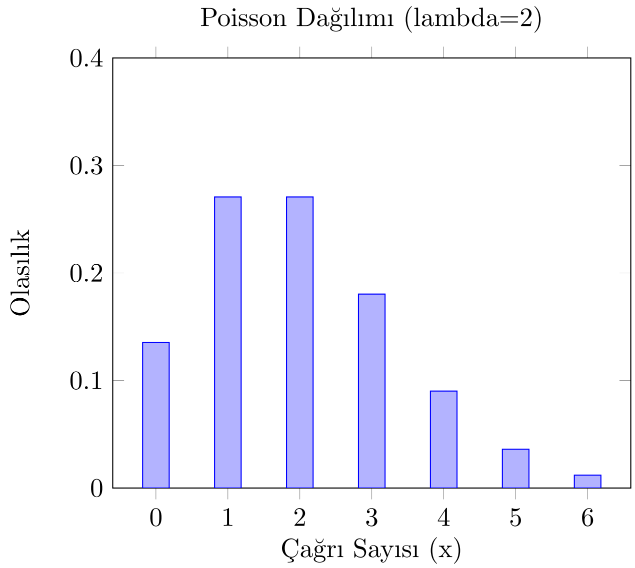
PMF (Probability Mass Function - Olasılık Kütle Fonksiyonu), **kesikli** (discrete) rastgele değişkenlerin aldığı belirli bir değerin olasılığını gösteren bir fonksiyondur.

Örnek: Bir çağrı merkezinde belirli bir dakikada kaç çağrı geleceğini düşünelim. Bir dakikada 0, 1, 2, 3, ... çağrı gelebilir ama **2.5 çağrı gibi kesirli bir sayı olamaz**. Bu yüzden kesikli bir değişkenden bahsediyoruz.
Matematiksel olarak PMF şu şekilde ifade edilir:

$$P(X = x) = f(x)$$

Burada X rastgele değişken, $f(x)$ ise o değişkenin belirli bir değeri alma olasılığıdır.

Örnek: Poisson Dağılımı ile PMF
Poisson dağılımı, belirli bir zaman aralığında nadiren gerçekleşen olayların olasılığını gösterir. Örneğin, **bir çağrı merkezine dakikada kaç çağrı geleceğini** modellemek için kullanılır.



Ne Görüyoruz?
- En olası çağrı sayısı 1 veya 2.
- Daha fazla çağrı gelme olasılığı gittikçe azalıyor.
—

2 Olasılık Yoğunluk Fonksiyonu (PDF)

PDF (Probability Density Function - Olasılık Yoğunluk Fonksiyonu), sürekli (continuous) değişkenlerin belirli bir aralıkta bulunma olasılığını tanımlar. **Örnek:** Bir insanın boy uzunluğunu düşünelim. Boy, kesikli değil, sürekli bir değişkendir çünkü bir kişinin boyu 170.3 cm veya 180.5 cm olabilir.

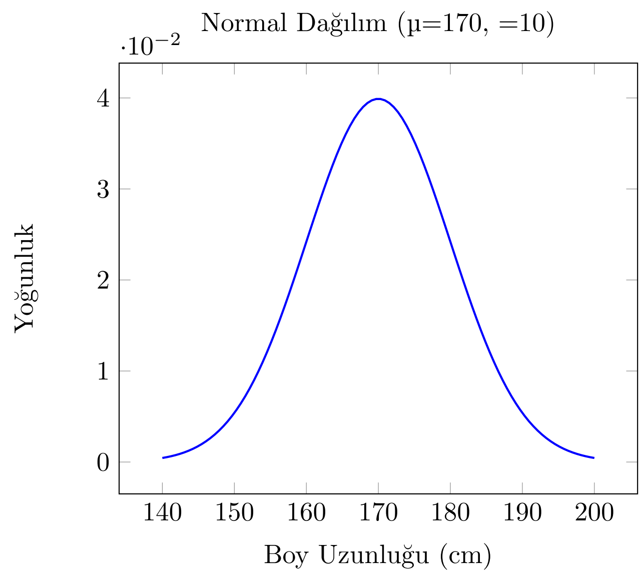
Matematiksel olarak PDF şu şekilde tanımlanır:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Burada, $f(x) \geq 0$ olmalı ve toplam alan 1 olmalıdır.

Örnek: Normal Dağılım ile PDF

İnsan boy uzunluklarının normal dağılıma uyduğunu düşünelim. Örneğin, ortalama boy $170cm$, standart sapma $10cm$ olsun.



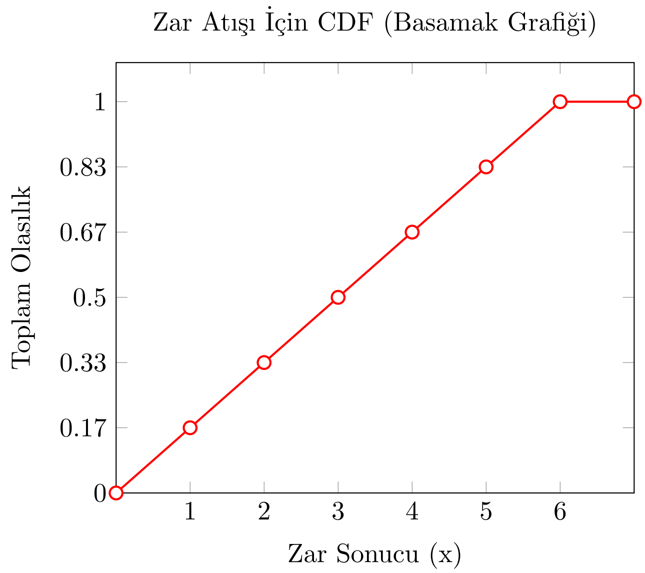
Ne Görüyoruz?
- Ortalama 170cm etrafında en yüksek yoğunluk var.
- Aşırı kısa ve uzun boylar daha az görülüyor.
—

3 Kümülatif Dağılım Fonksiyonu (CDF)

CDF (Cumulative Distribution Function - Kümülatif Dağılım Fonksiyonu), rastgele değişkenin belirli bir değere kadar olan toplam olasılığını gösterir. Örneğin, bir zar atışında şu olasılıkları hesaplayabiliriz:

- $P(X \leq 1) = 1/6$
- $P(X \leq 2) = 2/6$
- $P(X \leq 3) = 3/6$
- $P(X \leq 4) = 4/6$
- $P(X \leq 5) = 5/6$
- $P(X \leq 6) = 1$

Şimdi bu dağılımı basamaklı (staircase) bir histogram olarak çizelim:



Ne Görüyoruz?
- Her basamak, belirli bir değere kadar olan toplam olasılığı gösteriyor.
- **x=6'da toplam olasılık 1 oluyor**, çünkü zar sonucu en fazla 6 olabilir.
- Basamaklı görünüm, kesikli değişkenleri daha iyi temsil eder.

—