

İstatistik: Merkezi Limit Teoremi (CLT)

Atıl Samancıoğlu

1 Merkezi Limit Teoremi (CLT)

Merkezi Limit Teoremi (CLT - Central Limit Theorem), büyük örneklemlemlerle çalışıldığında, bağımsız rastgele değişkenlerin toplamı veya ortalamasının yaklaşık olarak normal dağıldığını söyler.

- Örneklem büyüklüğü arttıkça, dağılımın şekli normal dağılıma yaklaşır.
- Başlangıçtaki dağılımın şekli ne olursa olsun, yeterince büyük bir örneklemde dağılım normalleşir.
- İstatistiksel testler ve güven aralıklarının hesaplanmasında kritik bir rol oynar.

Matematiksel olarak, örneklem ortalamasının dağılımı şu şekilde ifade edilir:

$$\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$$

Burada:

- $\bar{X} \rightarrow$ Örneklem ortalaması
- $\mu \rightarrow$ Ana kütlenin ortalaması
- $\sigma^2 \rightarrow$ Ana kütlenin varyansı
- $n \rightarrow$ Örneklem büyüklüğü

Önemli nokta: n büyüdükçe, dağılım normalleşir.

1.1 Örnek: Zar Atışları

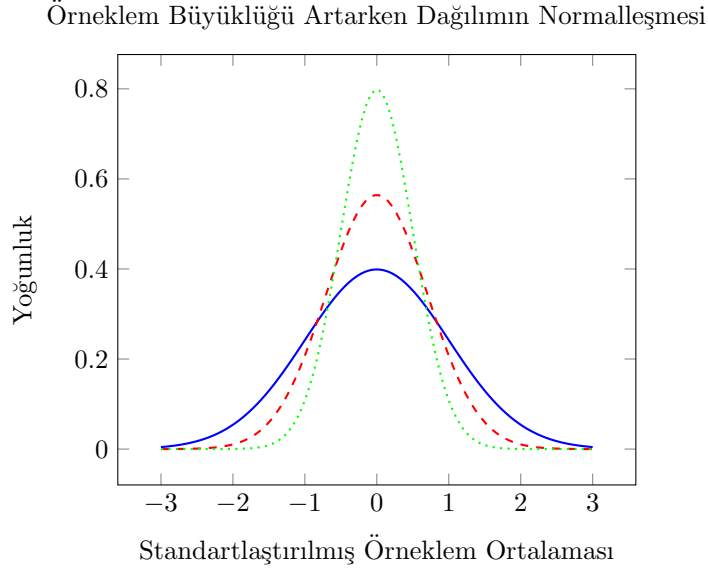
Normalde bir zarı attığımızda 1, 2, 3, 4, 5 veya 6 gelme olasılığı eşittir. Yani tek bir zar atışının dağılımı düzgün dağılım gösterir (bütün sayılar eşit olasılıkla gelir).

Şimdi, tek bir atışı değil, her seferinde 10 zar atıp bunların ortalamasını alalım.

- İlk deney: 10 zar attık, gelen sayıların ortalaması 3.7
- İkinci deney: 10 zar attık, gelen sayıların ortalaması 3.9
- Üçüncü deney: 10 zar attık, gelen sayıların ortalaması 3.5

Bunu binlerce kez tekrar ettiğimizi düşünelim. Sonuç olarak elde edilen ortalamaların histogramını çizdiğimizde, bu dağılımın normal dağılıma benzediğini görürüz.

1.2 Merkezi Limit Teoreminin Görselleştirilmesi



Grafikten ne anlıyoruz?

- Mavi eğri: $n = 10$ örneklem büyüklüğü için dağılım.
- Kırmızı kesikli eğri: $n = 30$ örneklem büyüklüğü için dağılım (daha dar ve yüksek).
- Yeşil noktalı eğri: $n = 100$ için dağılım neredeyse mükemmel normal dağılım.

2 Merkezi Limit Teoreminin Sonuçları

Merkezi Limit Teoremi'nin en önemli sonuçları şunlardır:

- Çoğu istatistiksel yöntem, örneklem dağılımının normal olmasını varsayar.
- Büyük örneklem kullanıldığında, gerçek dağılım ne olursa olsun normal kabul edilebilir.
- Güven aralıkları ve hipotez testleri Merkezi Limit Teoremi'ne dayanır.

- Gerçek hayatta birçok şeyin tahmin edilmesini kolaylaştırır. Örneğin, büyük bir fabrikada üretilen ampullerin ömrü değişkenlik gösterebilir. Ama ampullerden alınan rastgele örneklerin ortalaması, normal dağılım gösterdiği için kalite kontrol yapmak kolaylaşır.