Lineer Regresyon Nedir?

- Sayısal verilerin modellenmesi için kullanılan tekniktir.
- İki veya daha fazla değişkenin arasındaki ilişkiyi kullanarak bir değişkenin değerine göre diğer değişken hakkında bilgi edinebiliriz.
- Regresyon, tahmin, değerlendirme, hipotez testi ve nedensel ilişkileri modellemek için kullanılabilir.

$$Y = X1 + X2 + X3$$

Bu denkleme göre;
X_1, X_2,X_3 değişkenleri tahmin edici, açıklayıcı, bağımsız değişkendir.
Y değişkeni ise tahmin edilen çıktı, bağımlı, yanıt değişkenidir.

Neden Lineer Regresyon Kullanırız?

- Olasılıksal bir çerçeveden doğrusal regresyonun temel kavramlarını geliştirmek,
- Doğrusal modellerle parametrelerin tahmin edilmesi ve hipotez testi
- R'da doğrusal regresyon

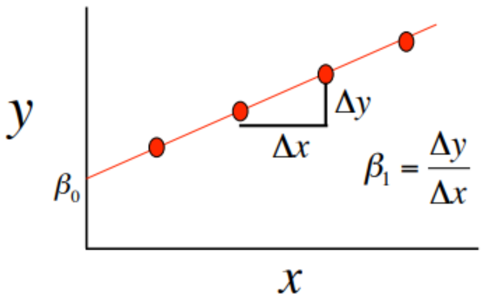
Bağımlı değişken Y'yi üç belirleyici X_1, X_2, X_3açısından modellemek istediğimizi var sayalım.
Y = f (X_1, X_2, X_3)

Genellikle denemek ve doğrudan tahmin etmek için yeterli veriye sahip olmayacağız.
Bu nedenle, genelde bazı sınırlı formlara sahip olduğunu varsaymak zorundayız.
Örneğin, Y=X_1+ X_2+ X_3

Doğrusal Regresyon Olasılıksal Bir Modeldir

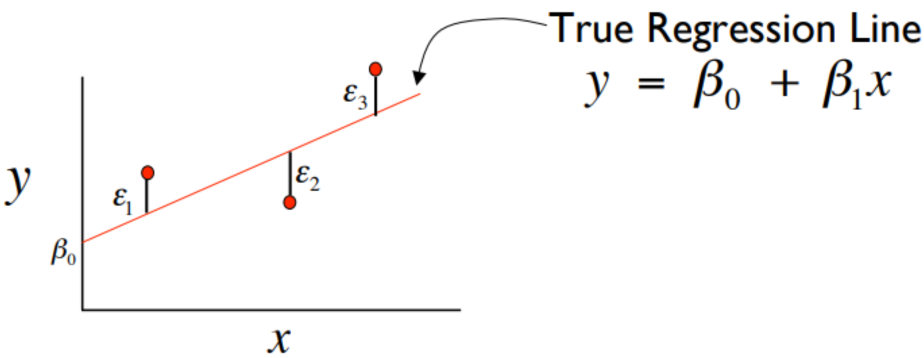
Matematik çoğunlukla deterministik olarak birbiriyle ilişkili değişkenleri incelemeye adanmıştır.
Ancak deterministik olmayan bir şekilde ilişkili değişkenler arasındaki ilişkiyi anlamakla ilgileniyoruz.

$$y = \beta_0 + \beta_1 x$$



- Doğrusal Olasılık Modeli

β_0,$\beta_1$ve bağımsız değişken X in herhangi bir sabit değeri için bağımlı değişkenin model denklemi aracılığıyla x ile ilişkili olduğu parametreler vardır
• ε, N (0,σ^2) olarak kabul edilen bir rasgele değişkendir.



- Y'nin beklenen değeri X in doğrusal bir fonksiyonudur, ancak sabit x için, Y değişkeni beklenen değerinden rastgele bir miktar ile farklıdır

- X * ' in bağımsız değişken X in belirli bir değerini göstermesine izin verin, o zaman doğrusal olasılık modelimiz şöyle diyor:

E (Y | x*) = $\mu_{Y|x*}$ = x, x* iken Y’nin ortalama değeri
V (Y | x*) = $\sigma^2{}_{Y|x*}$ x, x* iken Y’nin varyans değeri