Project 1 FYS-STK3155

Jonatan Hoffmann Hanssen   Eric Emanuel Reber

September 7, 2022

Abstract

a

We have defined $\boldsymbol{y}$ as a function of a matrix multiplication $X\boldsymbol{\beta}$ plus an error vector $\boldsymbol{\epsilon}$. This means that each element of the vector $\boldsymbol{y}$ can be expressed as follows:

$$y_i = \sum_j x_{ij}\beta_j + \epsilon_i$$

If we take the expectation value of this expression we get the following:

$$E[y_i] = E\left[\sum_j x_{ij}\beta_j + \epsilon_i\right]$$

$$E[y_i] = E\left[\sum_j x_{ij}\beta_j\right] + E[\epsilon_i]$$

However, the elements of X are not stochastic, and neither are the elements of β the first expectation value is simply the sum itself. Furthermore, ϵ is explicitly defined as a normal distribution $N(0, \sigma^2)$, and will by definition have the expectation value 0. Therefore, we end up with the final expression:

$$E[y_i] = \sum_j x_{ij}\beta_j = \boldsymbol{X}_{i,*}\boldsymbol{\beta}$$

We can use expectation values to calculate the variance as well:

$$var[y_i] = E\left[(y_i - E[y_i])^2\right]$$