Project 1 FYS-STK3155

Jonatan Hoffmann Hanssen   Eric Emanuel Reber

September 8, 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:

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

Distributing the outer expectation value function:

$$var[y_i] = E[y_i^2] - 2E\left[E[y_i]\right]E[y_i] + E\left[E[y]^2\right]$$
$$= E[y_i^2] - 2E\left[E[y_i]\right]E[y_i] + E\left[E[y]^2\right]$$

$$= E[y_i^2] - 2E\left[E[y_i]\right]E[y_i] + E\left[E[y]\right]E\left[E[y]\right]$$

The result of calculating the expectation value is non-stochastic. This means that $E[E[X]] = E[X]$. From this it follows that

$$\begin{aligned}
&= E[y_i^2] - 2E[y_i]E[y_i] + E[y_i]E[y_i] \\
&= E[y_i^2] - E[y_i]E[y_i] \\
&= E[y_i^2] - E[y_i]^2
\end{aligned}$$

The expectation value in the second summand has been proven to be equal to $\boldsymbol{X}_{i,*}\boldsymbol{\beta}$ above. We therefore have

$$\begin{aligned}
var[y_i] &= E[y_i^2] - (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2 \\
&= E\left[(X_{i,*}\boldsymbol{\beta})^2 + X_{i,*}\boldsymbol{\beta}\epsilon_i + \epsilon_i^2\right] - (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2 \\
&= E\left[(X_{i,*}\boldsymbol{\beta})^2\right] + E[X_{i,*}\boldsymbol{\beta}\epsilon_i] + E[\epsilon_i^2] - (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2
\end{aligned}$$

$X_{i,*}\boldsymbol{\beta}$ and ϵ_i are both scalars. Therefore the expectation value can be written as $E[X_{i,*}\boldsymbol{\beta}]E[\epsilon_i]$. However, $E[\epsilon_i]$ is by definition 0, because ϵ_i is defined as a normal distribution of mean 0 and variance σ^2. Furthermore, the expectation value of the non-stochastic $(X_{i,*}\boldsymbol{\beta})^2$ is simply the expression itself. We therefore have

$$\begin{aligned}
var[y_i] &= (X_{i,*}\boldsymbol{\beta})^2 + E[X_{i,*}\boldsymbol{\beta}] \cdot 0 + E[\epsilon_i^2] - (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2 \\
&= E[\epsilon_i^2]
\end{aligned}$$

We can prove that this is equal to the variance of ϵ_i:

$$E[\epsilon_i^2] = \frac{1}{n}\sum_i \epsilon_i^2$$

$$var[\epsilon_i] = \frac{1}{n}\sum_i (\epsilon_i - \bar{\epsilon}_i)^2$$

$$var[\epsilon_i] = \frac{1}{n}\sum_i (\epsilon_i - 0)^2$$

$$var[\epsilon_i] = \frac{1}{n}\sum_i \epsilon_i^2 = E[\epsilon_i^2]$$

And of course, we know that the variance of ϵ_i by definition is σ^2.

$$var[y_i] = E[\epsilon_i^2] = var[\epsilon_i] = \sigma^2$$

Proof of $var(\beta) = \sigma^2(X^TX)^{-1}$

$$var(\beta) = var((X^TX)^{-1}X^Ty)$$

Because we are assuming $(X^TX)^{-1}X^T$ to be deterministic and y to be stochastic, we can rewrite the $var(\beta)$ as following:

$$var(\beta) = (X^TX)^{-1}X^T var(y)((X^TX)^{-1}X^T)^T$$

$$= (X^TX)^{-1}X^T var(y)(X^T)^T((X^TX)^{-1})^T$$

Since $(X^T)^T = X$ and $(X^TX)^{-1})^T = (X^TX)^T)^{-1} = (X^TX)^{-1}$

$$= (X^TX)^{-1}X^T var(y)X(X^TX)^{-1}$$

We know var(y) = σ^2, and since it is a scalar it is commutative. Thus we may move it freely

$$= \sigma^2(X^TX)^{-1}X^TX(X^TX)^{-1}$$

$$= \sigma^2(X^TX)^{-1}$$

■