# W271, Unit 2 Question 2

Artem Lebedev

May 27, 2023

**Variance of Linear Probability model**

Bias term is inconsequential to this derivation. Omitting the bias we can express LPM as the following:

$$y_i = \beta x_i + \epsilon_i \text{ where}$$
$$\begin{cases} y = 1 : p(x_i) \\ y = 0 : 1 - p(x_i) \end{cases}$$

Note that expectation of $\epsilon_i$ is 0, because of the properties of OLS. Therefore:

$$\begin{aligned} Var(\epsilon_i) &= E[(\epsilon_i - E(\epsilon_i))^2] \\ &= E[\epsilon_i^2] \end{aligned}$$

Given the fact that y can only take two values, $\epsilon_i$ also can only take two values:

$$\begin{cases} 1 - \beta x_i : p(x_i) \\ 0 - \beta x_i : 1 - p(x_i) \end{cases}$$

Using this fact and definition of the expectation:

$$\begin{aligned} Var(\epsilon_i) &= E[\epsilon_i^2] \\ &= (1 - \beta x_i)^2 p(x_i) + (0 - \beta x_i)^2 (1 - p(x_i)) \end{aligned}$$

Now, using the fact that $E[y|x_i] = p(x_i)$, derived in the lecture, we can write:

09 gh 0p,nnnn999 qdsssssskjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjjj

$\sqcup$

$999999hj$

$doublebarwedg,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,dfs-$

000000000000000000$p(x_i) = E[y_j|x_i]$
$= E[\beta jjjj 6bbbbbbbbbbbbbbbbbbbbbbb$