magine you have $x \in \mathcal{X} = \{red, green\}$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \downarrow \qquad \downarrow$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad 0 \qquad 1$

irect consequence of $\mathcal{X}$.

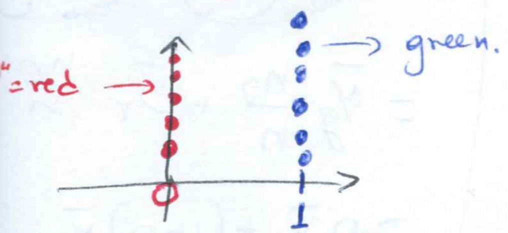
$$\tilde{x} \in \{0, 1\}$$

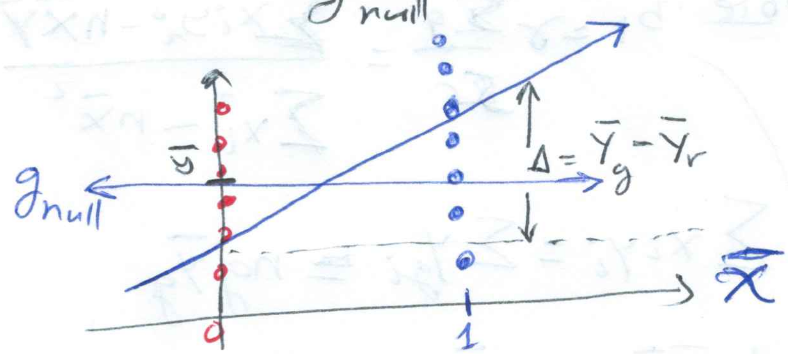
$\mathcal{H} = \{ w_0 + w_1 \tilde{x} : w_0 \in \mathbb{R}, w \in \mathbb{R} \}$ You multiply a number by an index.

$= \{ w_0 + w_1 \mathbb{1}_{x = gree}, w_0 \in \mathbb{R}, w_1 \in \mathbb{R} \}$

$\Rightarrow g(x) = b_0 + b_1 \mathbb{1}_{x = green}$. 1 → give green, 0 → give red.

hat is null model again? Ans: $g_{null} = \bar{Y}$





hat do you think best

The two values will be

$$g(x) = \begin{cases} \bar{y}_r = 0 & \text{if } \tilde{x} = 0 \\ \bar{y}_g = 1 & \text{if } \tilde{x} = 1 \end{cases}$$

$$= \bar{y}_r + (\underbrace{\bar{y}_g}_{\to green} - \underbrace{\bar{y}_r}_{\to red}) \mathbb{1}_{x = g}$$

te: from $g(x) = b_0 + b_1 \mathbb{1}_{x = green}$

$b_0 = \bar{Y}_r$

$b_1 = \bar{y}_g - \bar{y}_r$

why R^2 is bad?

Ans:

Case R^2 is good? yes all pts closed to the line.

How to prove $b_0 = \bar{y}_r$ and $b_1 = \bar{y}_g$?

$b_0 = \bar{y} - b_0 \bar{x}$

$b_1 = \cancel{8} \frac{S_y}{S_r}$

Let's assume $b_0 = \bar{y}_g - \bar{y}_r$ and

$b_0 = (p\bar{Y}_g + (1-p)\bar{Y}_r) - (\bar{y}_g - \bar{y}_i)p$

lest $p = \frac{n_g}{n}$

$= \bar{Y}_r \cdot ((1-p)+p) = \bar{Y}_r$ ✓

Note: $b_1 = r \frac{\Sigma g}{\Sigma r} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$\sum x_i y_i = \sum Y_{gi} = n_g \bar{Y}_g$

$n\bar{x}\bar{y} = np\bar{y}$

$\sum x_i^2 = n_g$

$n\bar{x}^2 = np^2$

$\frac{\sum x_i y_i - n\bar{x}\bar{y} \cdot \frac{1}{n}}{\sum x_i^2 - n\bar{x}^2 \cdot \frac{1}{n}} = \frac{p\bar{y}_g - p\bar{y}}{p - p^2}$

$= \frac{\bar{y}_g - \bar{y}}{1-p} = \frac{\bar{y}_g - (p\bar{y}_g + (1-p)\bar{y}_r)}{1-p} = \frac{\bar{y}_g}{1-p} - \frac{p\bar{y}_g}{1-p} - \bar{y}_r = \bar{y}_g - \bar{y}_r$ ✓

$\bar{Y} = \frac{Y_1 + Y_2 + \cdots + Y_n}{n}$

$= \frac{y_{g_1} + \cdots + y_{g_{n_g}} + y_{n} + \cdots}{n}$

$= \frac{\sum Y_g}{n} + \frac{\sum Y_n}{n}$

$= \frac{\sum y_g}{n} \cdot \frac{n_g}{n_g} + \frac{\sum y_n}{n} \cdot \frac{n_r}{n_r}$

$= \bar{Y}_g \frac{n_g}{n} + \bar{y}_r \frac{n_r}{n}$

$= p\bar{Y}_g + (1-p)\bar{y}$

$\bar{x} = \frac{x_1 + \cdots + x_n}{n}$

$= \frac{x_{g_1} + \cdots x_{g_{n_g}} + \cdots}{n}$

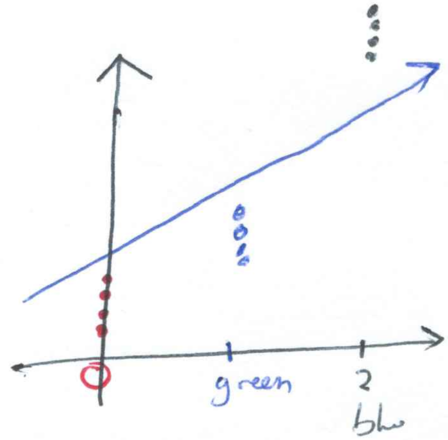
$= \frac{n_g}{n}$

$= p$

Exple: $x \in \mathcal{X} = \{$red, green, blue$\}$

$X = \begin{cases} 1 & \text{non USA} \\ 0 & \underbrace{\text{USA}}_{\text{Called reference category.}} \end{cases}$

"R^2 is how good your model is"



If you go beyond 2 you can't get good model b/c can not chrink.

END - Miterm 1