

Scaling

$$\text{Score} = A + B \log(\text{odds})$$

offset factor $\frac{P}{1-P}$, where P : good 확률

여기서, A 와 B 를 구하자! \leftarrow 2개의 조건이 필요하다

i) PDO: points to double odds

(odds를 2배 올리는데 필요한 score)

$$\text{즉, } \text{Score} + \text{PDO} = A + B \log(2 \times \text{odds}) \quad (1)$$

$$\therefore PDO = B \log(2)$$

$$B = PDO / \log(2)$$

$$, \quad \boxed{A = \text{Score} - B \log(\text{odds})}$$

↑ 1개의 조건이 더 필요

예) Score 600점일때, odds를 50이라고 하자!
(50:1)

$$\& \quad PDO = 20$$

$$\Rightarrow \left(B = 20 / \log(2) = 28.8539 \right.$$

$$A = 600 - 28.8539 \times \log(50) = 487.123$$

그러므로, Scaling 하는 공식은

$$\text{Score} = 487.123 + 28.8539 \times \log(\text{odds})$$

Logistic regression이
산출하는 Score

Dummy Variable의 의미

Age : 20 | \Rightarrow 20세
30세
40세
50세 이상
└──────────┘
classing
즉 binning

<u>Age₁</u>	<u>Age₂</u>	<u>Age₃</u>
1	0	0
0	1	0
0	0	1
0	0	0

└──────────┘

dummy variables
* 해당 classing의 level = k
 \Rightarrow dummy var's는 k-1개 필요

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_{A_1} * Age_1 + \beta_{A_2} * Age_2 + \beta_{A_3} * Age_3 + \varepsilon$$

만약, 20대라면 log-odds 값은 $\hat{\beta}_0 + \hat{\beta}_{A_1}$

만약, 40대라면 log-odds 값은 $\hat{\beta}_0 + \hat{\beta}_{A_3}$

만약, 50대 이상인 경우는 $\hat{\beta}_0$

Age와 checking이 사용되는 경우,

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_{A_1} \text{Age}_1 + \beta_{A_2} \text{Age}_2 + \beta_{A_3} \text{Age}_3 \\ + \beta_{c_1} \text{ch}_1 + \beta_{c_2} \text{ch}_2 + \beta_{c_3} \text{ch}_3 + \varepsilon$$

예) 24이름 2024, checking = 1

$$\Rightarrow \hat{\beta}_0 + \hat{\beta}_{A_1} + \hat{\beta}_{c_1}$$