

# Binary logistic Regression $E[y_i | x_{ij}]$

Consider  $n$  independent observations  $y_1, \dots, y_n \sim B(p)$  conditionally on a set of  $p$  categorical or numerical predictors  $x_j, j = 1, \dots, p$ . The model is

$$g(E[y_i | x_{ij}]) = g(\pi_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} = \underline{x}_i^T \underline{\beta}$$

canonical link function:

$$g(E[y_i | x_{ij}]) = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} = \underline{x}_i^T \underline{\beta}$$

It follows naturally that

$$E[y_i | x_{ij}] = \pi_i = \frac{e^{\underline{x}_i^T \underline{\beta}}}{1 + e^{\underline{x}_i^T \underline{\beta}}}$$

Indeed, we have that

$$\begin{aligned} e^{\ln(\pi_i / (1 - \pi_i))} &= e^{\underline{x}_i^T \underline{\beta}} \\ \pi_i / (1 - \pi_i) &= e^{\underline{x}_i^T \underline{\beta}} \\ \pi_i &= e^{\underline{x}_i^T \underline{\beta}} (1 - \pi_i) \\ \pi_i + \pi_i e^{\underline{x}_i^T \underline{\beta}} &= e^{\underline{x}_i^T \underline{\beta}} \end{aligned}$$

$$\Leftrightarrow \pi_i = 1 / (1 + e^{-\underline{x}_i^T \underline{\beta}}) = \frac{e^{\underline{x}_i^T \underline{\beta}}}{1 + e^{\underline{x}_i^T \underline{\beta}}}$$