* IRLS

$$g(\mu) = \log\frac{\pi}{1-\pi} = \log\frac{\mu}{1-\mu}$$

$$\dot{g}(\mu) = \frac{1}{\mu(1-\mu)} = \frac{(1+e^{\theta})^2}{e^{\theta}}$$

$$\ddot{b}(\theta) = \frac{e^{\theta}}{(1+e^{\theta})^2}$$

$$w_i = \left\{(\dot{g}(\mu))^2 \cdot \ddot{b}(\theta)\right\}^{-1} = \frac{e^{\theta}}{(1+e^{\theta})^2} = \frac{e^{x_i\beta}}{(1+e^{x_i\beta})^2}$$

$$e_i = \dot{g}(\mu_i)(y_i - \mu_i) = \frac{(1+e^{\theta})^2}{e^{\theta}}\left(y_i - \frac{e^{\theta}}{1+e^{\theta}}\right)$$

$$z_i = x_i\beta + e_i = x_i\beta + \frac{(1+e^{x_i\beta})^2}{e^{x_i\beta}}\left(y_i - \frac{e^{x_i\beta}}{1+e^{x_i\beta}}\right)$$

$$\therefore \beta^{(t+1)} = (X^T W^{(t)} X)^{-1} X^T W^{(t)} z^{(t)}.$$

$$W^{(t)} = \mathrm{diag}(w_i^{(t)}) \quad \text{and} \quad z^{(t)} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

$$(X^T W^{(t)} X)^{-1} = \left\{ (X_1^T \cdots X_n^T) \begin{pmatrix} w_1 X_1 \\ \vdots \\ w_n X_n \end{pmatrix} \right\}^{-1}, \qquad W^{(t)} z^{(t)} = \begin{pmatrix} w_1 z_1 \\ \vdots \\ w_n z_n \end{pmatrix}$$

$$\therefore \beta^{(t+1)} = \left(X^T \begin{pmatrix} w_1 x_1 \\ \vdots \\ w_n x_n \end{pmatrix} \right)^{-1} X^T \begin{pmatrix} w_1 z_1 \\ \vdots \\ w_n z_n \end{pmatrix}$$

1st derivative : $\dot{\ell}(\beta) = X'(Y - \pi) = X'\left\{Y - \left(1 - \frac{1}{1+e^{\theta_i}}\right)\right\}$