HW(3)

4.8.1. $p(X) = \dfrac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$ —— (4.2)

$\dfrac{p(X)}{1-p(X)} = e^{\beta_0+\beta_1 X}$ —— (4.3)

From (4.2) $1-p(X) = 1 - \dfrac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$

$1-p(X) = \dfrac{1+\cancel{e^{\beta_0+\beta_1 X}}-\cancel{e^{\beta_0+\beta_1 X}}}{1+e^{\beta_0+\beta_1 X}} = \dfrac{1}{1+e^{\beta_0+\beta_1 X}}$ —— (3)

$\dfrac{(4.3)}{(3)} \Rightarrow \dfrac{p(X)}{1-p(X)} = \dfrac{e^{\beta_0+\beta_1 X}}{\frac{1}{1+e^{\beta_0+\beta_1 X}}}$

$1-p(X) = 1 - \dfrac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}} = \dfrac{1+\cancel{e^{\beta_0+\beta_1 X}}-\cancel{e^{\beta_0+\beta_1 X}}}{1+e^{\beta_0+\beta_1 X}} = \dfrac{1}{1+e^{\beta_0+\beta_1 X}}$

$\dfrac{p(X)}{1-p(X)} = \dfrac{e^{\beta_0+\beta_1 X}}{\cancel{1+e^{\beta_0+\beta_1 X}}} \times \dfrac{\cancel{1+e^{\beta_0+\beta_1 X}}}{1} = e^{\beta_0+\beta_1 X}$

4.8.5 (a) QDA performs better on train test and LDA on the test set

(b) QDA performs better on test set

(c) QDA performs better as n increases due to bias variance tradeoff. A larger training set has lower variance. The training set size increase, decreases variance

(d) False as it will overfit and perform worse on training set

4.8.6 $p(X) = \dfrac{\exp(-6+0.05X_1+X_2)}{1+\exp(-6+0.05X_1+X_2)}$

(a) $p(X_1=40, X_2=3.5) = \dfrac{\exp(-6+0.05(40)+3.5)}{1+\exp(-6+0.05(40)+3.5)} = 0.3775$

b) $\ln\left(\frac{0.5}{1-0.5}\right) = 0 = -6 + 0.05X_1 + 3.5$

$\Rightarrow X_1 = 50$

CamScanner