Homework 4

Classification

① a) $\boxed{\text{True}}$ defn of CDF, PDF: $\frac{d}{dx}\Phi(x) = \phi(x)$

b) $\boxed{\text{True}}$ defn of CDF, PDF: $\int_{-\infty}^{x} \phi(z)\,dz = \Phi(x)$

c) $\boxed{\text{False}}$ property of any continuous dist: $P(Z=x) = \int_x^x \phi(z)\,dz = 0$

d) $\boxed{\text{True}}$ $P(Z<x) = \int_{-\infty}^{x} \phi(z)\,dz = \Phi(x)$ by defn

e) $\boxed{\text{True}}$ $P(Z \le x) = P(Z<x) + P(Z=x) = \Phi(x) + 0 = \Phi(x)$ [from part c, d]

f) $\boxed{\text{True}}$ From defn and calculus rules,

$$P(x<Z<x+h) = \int_x^{x+h} \phi(z)\,dz = \Phi(x+h) - \Phi(x)$$

For small, positive h: $\lim_{x \to h} \frac{\Phi(x+h) - \Phi(x)}{h} = \phi(x)$

$\Rightarrow P(x<Z<x+h) = h\,\phi(x)$

② a) X_i : row vector $[1, ED_i, INC_i, MAN_i]$, where ED_i (years of schooling), INC_i (annual income), MAN_i (1 for man, else 0) for subject i

β: col vector of weights associated with probit model

b) random, latent

c) U_i is iid $N(0,1)$ and independent from prediction vars

d) Log likelihood is a <u>sum</u>, with one <u>term</u> for each <u>subject</u>. Likelihood is a product (after multiplying probabilities for each subject. Taking log makes it a sum.

③ False: $X_{harry}\beta = X_{george}\beta + 0.1$

$\Rightarrow$ Diff in probabilities $= \Phi(X_{harry}\beta) - \Phi(X_{george}\beta) \approx \boxed{0.03}$

The Hat Matrix

(5.1) We know $(X_{(i)}^T X_{(i)})^{-1} = (X^T X - x_i x_i^T)^{-1}$

Using Sherman Morrison:

$$(X_{(i)}^T X_{(i)})^{-1} = (X^T X)^{-1} + \frac{(X^T X)^{-1} x_i x_i^T (X^T X)^{-1}}{1 - x_i^T X^{-1} x_i}$$

$$= (X^T X)^{-1} + \frac{(X^T X)^{-1} x_i x_i^T (X^T X)^{-1}}{1 - h_i} \qquad \square$$

(5.5)

$$\hat{\beta}_i = (X_{(i)}^T X_{(i)})^{-1} (X^T Y - x_i y_i)$$

$$= \left[(X^T X)^{-1} + \frac{(X^T X)^{-1} x_i x_i^T (X^T X)^{-1}}{1 - h_i} \right] (X^T Y - x_i y_i)$$

$$= (X^T X)^{-1} X^T Y - (X^T X)^{-1} x_i y_i + \frac{(X^T X)^{-1} x_i x_i^T X^T Y}{1 - h_i} - \frac{(X^T X)^{-1} x_i x_i^T x_i y_i}{1 - h_i}$$

$$= \hat{\beta} - \frac{(X^T X)^{-1} x_i}{1 - h_i} \left(y_i (1 - h_i) - x_i^T \hat{\beta} + \cancel{h_i} y_i \right)$$

$$= \hat{\beta} - \frac{(X^T X)^{-1} x_i r_i}{1 - h_i} \Rightarrow \hat{\beta} - \hat{\beta}_i = \frac{(X^T X)^{-1} x_i r_i}{1 - h_i} \qquad \square$$