Homework 3

Alexander Brandt
STAT 215A

November 15, 2017

1 Classification

7.B.1

a

True: since from the definition of the CDF/PDF relationship:

$$\frac{\partial \Phi(x)}{\partial x} = \phi(x)$$

b

True: again, this is a basic CDF/PDF relationship:

$$\int_{-\infty}^{x} \phi(z)dz = \Phi(x)$$

c

False: this is a basic calculus definition:

$$\int_{x}^{x} \phi(z)dz = 0$$

d

True: From above, and part e:

$$\int_{x}^{x} \phi(z)dz + \int_{-\infty}^{x} \phi(z)dz = 0 + \Phi(x)$$

e

True: Basic CDF/PDF relationship:

$$\int_{-\infty}^{x} \phi(z)dz = \Phi(x)$$

f

True: Given CDF/PDF relationship:

$$\begin{aligned} P(x < Z < x+h) &= \lim_{x \to h} \int_x^{x+h} \phi(z) dz \\ &= \Phi(x+h) - \Phi(x) \end{aligned}$$

And basic calculus:

$$\lim_{x \to h} \frac{\Phi(x+h) - \Phi(x)}{h} = \phi(x)$$

Thus:

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

7.B.2

a

X_i is a column vector of 1 (the intercept), the educational level in years of school the individual has attended, the income level of the individual in dollars, and the gender of the individual (1 for male, 0 for female). β_i is a column vector of the weights associated with for the probit model.

b

Random and Latent

c

U_i is i.i.d. $N(0, 1)$, and independent of the prediction variables.

d

The log likihood function is a sum with one term for each subject.

1.1

False.

$$\begin{aligned} \text{Difference} &= \Phi(X'_{Harry}\beta) - \Phi(X'_{George}\beta) \\ &= \Phi(.29) - \Phi(.19) \\ &= 3.87\% \end{aligned}$$

7.B.3

I think I will be using x_i in the opposite orientation that the authors of the paper do.

Proof of 5.1

Given $(X_i'X_i)^{-1} = (X'X - x_i x_i')^{-1}$, and the Sherman Morrison formula:

$$\begin{aligned} (X_i'X_i)^{-1} &= (X'X)^{-1} + \frac{(X'X)^{-1} x_i x_i' (X'X)^{-1}}{1 - x_i' X^{-1} x_i} \\ &= (X'X)^{-1} + \frac{(X'X)^{-1} x_i x_i' (X'X)^{-1}}{1 - h_i} \end{aligned}$$

Proof of 5.5

Continuing from above:

$$\begin{aligned}
\hat{\beta}_i &= (X_i'X_i)^{-1}(X'Y - x_iy_i) \\
&= \left[(X'X)^{-1} + \frac{(X'X)^{-1}x_ix_i'(X'X)^{-1}}{1-h_i}\right](X'Y - x_iy_i) \\
&= (X'X)^{-1}X'Y - (X'X)^{-1}x_iy_i + \frac{(X'X)^{-1}x_ix_i'X'Y}{1-h_i} - \frac{(X'X)^{-1}x_ix_i'x_iy_i}{1-h_i} \\
&= \hat{\beta} - \frac{(X'X)^{-1}x_i}{1-h_i}\left[y_i(1-h_i) - x_i\hat{\beta} + h_iy_i\right] \\
&= \hat{\beta} - \frac{(X'X)^{-1}x_i}{1-h_i}\left[y_i - x_i\hat{\beta}\right] \\
&= \hat{\beta} - \frac{(X'X)^{-1}x_ir_i}{1-h_i} \\
\hat{\beta} - \hat{\beta}_i &= \frac{(X'X)^{-1}x_ir_i}{1-h_i}
\end{aligned}$$

I found this proof to be tricky, so I spent a lot of time reading through linear algebra/linear regression texts. In particular Linear Regresion Analysis by Seber and Lee (2003), especially Chapter 10, we really helpful in helping me complete these proofs.