Stats 215A, Fall 2020
Homework 9

Classification

1. (a) By definition $\Phi(x) = \int_{-\infty}^{x} \phi(y)dy$, taking the derivative

$$\frac{d}{dx}\Phi(x) = \frac{d}{dx}\int\limits_{-\infty}^{x} \phi(y)d = \phi(x)\frac{d}{dx}x - \phi(-\infty)\frac{d}{dx} - \infty = \phi(x)$$

   Hence the answer is True.

   (b) True, since by definition $\Phi(x) = \int_{-\infty}^{x} \phi(y)dy$

   (c) False, $\mathbb{P}(Z = x) = 0$ while $\phi(x) > 0\ \forall x$

   (d) True, $\mathbb{P}(Z < x) = \int_{-\infty}^{x} \phi(y)dy = \Phi(x)$

   (e) True, $\mathbb{P}(Z \leq x) = \mathbb{P}(Z = x) + \mathbb{P}(Z < x) = \Phi(x)$

   (f) False, $\mathbb{P}(x < Z < x + h) = \Phi(x + h) - \Phi(x)$, by definition

$$\frac{d}{dx}\Phi(x) = \lim_{h \to 0} \frac{\Phi(x + h) - \Phi(x)}{h}$$

   which is different from $\Phi(x + h) - \Phi(x)$

2. (a) X_i- features for observation i, $\beta-$ vector of model parameters

   (b) random, latent

   (c)
   - Has a normal distribution
   - Independent of X

   (d) sum, term , subject. Since we assume independence between subjects the likelihood is the multiplication of the marginals for each subject.Taking log this becomes a sum.

3. George's probability is 0.58, while Harry's is 0.61

4.
   - Simple linear algebra calculations show that

$$X^T X = X_{(i)}^T X_{(i)} + (x_i)\, x_i^T$$

     Therefore

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

     Denote $A = X^T X, u = -x_i, v = x_i$, we simply apply the formula and get

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

     We conclude by noting that $x_i^T (X^T X)^{-1} x_i = h_i$

   - First by definition

$$\hat{\beta} - \hat{\beta}_{(i)} = \left(X^T X\right)^{-1} X^T y - \left(X_{(i)}^T X_{(i)}\right)^{-1} X_{(i)}^T y_{(i)}$$

Noting that $X_{(i)}^T y_{(i)} = X^T y - X_i^T y_i$ and applying the inverse formula

$$
\begin{aligned}
\hat{\beta} - \hat{\beta}_{(i)} &= \left(X^T X\right)^{-1} X^T y - \left((X^T X)^{-1} + \frac{(X^T X)^{-1} x_i^T x_i (X^T X)^{-1}}{1 - h_i} \right) \left(X^T y - X_i^T y_i\right) \\
&= (X^T X)^{-1} X_i^T y_i - \frac{(X^T X)^{-1} x_i^T \hat{y}_i}{1 - h_i} + \frac{(X^T X)^{-1} x_i^T x_i (X^T X)^{-1} x_i^T y_i}{1 - h_i} \\
&= \frac{(X^T X)^{-1} x_i^T}{1 - h_i} \left((1 - h_i) y_i + h_i y_i - \hat{y}_i\right) \\
&= \frac{(X^T X)^{-1} x_i^T}{1 - h_i} (y_i - \hat{y}_i) \\
&= \frac{(X^T X)^{-1} x_i^T r_i}{1 - h_i}
\end{aligned}
$$