Homework 4

Aniket Kesari
UC Berkeley Law
25948127

November 16, 2017

1 Freedman Problems

2 1 Let Z be N(0,1) with density function

2.1 a.

True. The definition of the cumulative density function is that it is the antiderivative of the probability distribution evaluated at a point.

2.2 b.

True. Integrating from 0 to x will give the area under the curve up to that point.

2.3 c.

False, the probability of equaling exactly x in a continuous function is infiniteseimally small.

2.4 d.

True. The area represents the sum of the probability that lies below x.

2.5 e.

True. See above.

2.6 f.

True. Incrementally by a small h gives a close approximation of the point estimate for the probability that Z = x.

3 2.

3.1 a.

X_i is the matrix of predictor of variables, and β is the vector of slope coefficients

3.2 b.

Random and Latent variable

3.3 c.

U_i should be distributed iid $N(0,1)$.

3.4 d.

Sum; Term; Subject

4 3.

False, .1 is the difference between the predicted probit output, but we need to apply the inverse Gaussian transformation to get probabilities (R code in .rmd file). In actuality, the difference in probabilities is $\phi(.29) - \phi(.19) = .614 - .575 = .039$

```
## [1] 0.03874645
```

5 Hoaglin and Welsch

First start with the Sherman Morrison Formula:

$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1+v^T A^{-1}u}$

The equation assumes that A is an invertible square matrix, and that u and v are $n*1$ vectors.

5.1 Prove Equation 5.1:

Equation to prove: $(X_{(i)}^T X_{(i)})^{-1} = (X^T X)^{-1} + \frac{(X^T X)^{-1}x_i^T x_i (X^T X)^{-1}}{1-h_i}$

Step 1: Assume we start with a full-rank covariance matrix:

$cov(b) = (X^T X)^{-1}$

Then, imagine we remove row i from the matrix. Denote this as:

$(X_{(i)}^T X_{(i)})^{-1}$

Step 2: This be rewritten as:

$(X^T X - x_i x_i^T)^{-1}$ where x_i denotes the deleted row vector that has been subtracted out.

Step 3: Reverse the signs on Sherman Formula and plug in:

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

We can take this step because the Sherman formula requires that "A" be a invertible square matrix and uv^T be n vectors. $(X^T X)^{-1}$ produces an invertible square matrix, and x_i is definitionally a $1*n$ vector because it is one deleted row. The denominator in the last equation can be rewritten as:

$1 - x_i^T (X^T X)^{-1}x_i = 1 - H_i$

Step 4: Therefore the two expressions are equivalent.

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

5.2 Prove Equation 5.5

Equation to Prove: $\hat{\beta} - \hat{\beta}_i = \frac{(X^T X)^{-1} x_i^T r_i}{1 - h_i}$

Step 1: Definition of β $\hat{\beta} = (X^T X)^{-1} X^T y$

Step 2: Define $\beta_{(i)}$ $\hat{\beta}_{(i)} = (X_{(i)}^T X_{(i)})^{-1} X_{(i)}^T y_{(i)}$

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

Step 3: Subtract $\beta_{(i)}$ from the left side $\hat{\beta} = (X^T X)^{-1} + \frac{(X^T X)^{-1} x_i^T x_i (X^T X)^{-1}}{1 - x_i^T (X^T X)^{-1} x_i} X_{(i)}^T y_{(i)}$

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

Step 4: Recall that: $\hat{\beta} = (X^T X)^{-1} X^T y$

So we can simplify:

$(X^T X)^{-1} + \frac{(X^T X)^{-1} x_i^T x_i \hat{\beta}}{1 - x_i^T (X^T X)^{-1} x_i}$

Step 5: $x_i \hat{\beta}$ is definition of r_i (the vector of residuals), so:

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