# PS7 Cook

natcook97

March 2021

1. 25 percent of logwage values are missing. I would guess that the observations are most likely MNAR.

2. Each of the estimates from the models are less than the true value of 0.093. Model 1 with the list-wise deletion gives the largest and closest estimate followed by model 3 which used the predicted value from the complete cases regression. I can't get Model 4 to run but it is probably the most legit method. Model 1 is not ideal because the observations are almost certainly not MCAR. Model 2 will cause the variance to drop which we see in the relatively lower F stat. Models 3 and 4 are probably the best for imputing the missing values.

3. Progress on the project is not so bueno. I haven't landed on a data set yet but I'd like to find one/a topic that's fairly straightforward and just gives me a good opportunity to apply the techniques we've been learning. I don't feel like I need to cure cancer necessarily with this research - meaning if I just confirm returns to education or something like that I will be happy.

| | Unique (#) | Missing (%) | Mean | SD | Min | Median | Max |
|---|---|---|---|---|---|---|---|
| logwage | 670 | 25 | 1.6 | 0.4 | 0.0 | 1.7 | 2.3 |
| hgc | 16 | 0 | 13.1 | 2.5 | 0.0 | 12.0 | 18.0 |
| tenure | 259 | 0 | 6.0 | 5.5 | 0.0 | 3.8 | 25.9 |
| age | 13 | 0 | 39.2 | 3.1 | 34.0 | 39.0 | 46.0 |

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| (Intercept) | 0.534 | 0.708 | 0.563 |
| | (0.146) | (0.116) | (0.112) |
| hgc | 0.062 | 0.050 | 0.059 |
| | (0.005) | (0.004) | (0.004) |
| as.factor(college)not college grad | 0.145 | 0.168 | 0.177 |
| | (0.034) | (0.026) | (0.025) |
| poly(tenure, 2, raw = T)1 | 0.050 | 0.038 | 0.047 |
| | (0.005) | (0.004) | (0.004) |
| poly(tenure, 2, raw = T)2 | -0.002 | -0.001 | -0.002 |
| | (0.000) | (0.000) | (0.000) |
| age | 0.000 | 0.000 | 0.000 |
| | (0.003) | (0.002) | (0.002) |
| as.factor(married)single | -0.022 | -0.027 | -0.028 |
| | (0.018) | (0.014) | (0.013) |
| Num.Obs. | 1669 | 2229 | 2229 |
| R2 | 0.208 | 0.147 | 0.223 |
| R2 Adj. | 0.206 | 0.145 | 0.221 |
| AIC | 1179.9 | 1091.2 | 956.8 |
| BIC | 1223.2 | 1136.8 | 1002.4 |
| Log.Lik. | -581.936 | -537.580 | -470.382 |
| F | 72.917 | 63.973 | 106.573 |