

PS7_{schmidt}

akeaneschmidt1998

March 2024

1 Introduction

Q: At what rate are log wages missing? Do you think the logwage variable is most likely to be MCAR, MAR, or MNAR?

A: The missing rate is 0.24933, so approximately one instance in four is missing this data. I think it's most likely MNAR- I hypothesize that people who are uncomfortable with the amount they make (either being bashful about it being too high, or ashamed of it being low) will not report it. So I think both the left and right tail are under reported.

Q: Tell me about the progress you've made on your project. What data are you using? What kinds of modeling approaches do you think you're going to take?

A: I decided to go with the data on Oklahoma crime rates. I think I'm making progress, especially in another class- I'm taking an EIPT class on hierarchical linear models, so I'm thinking it might be interesting to examine different facets that feed into various crime rates- for instance, maybe population only APPEARS to correlate with crime rate, but in reality it's a minor correlation once you control for poverty rate, which itself correlates with each.

2 Summary Table of the Cleaned Data Frame and Multiple Imputation Table

	Unique	Missing Pct.	Mean	SD	Min	Median	Max
logwage	669	0	1.6	0.4	0.0	1.7	2.3
hgc	16	0	12.6	2.3	0.0	12.0	18.0
tenure	233	0	5.2	5.1	0.0	3.3	24.8
age	12	0	39.2	3.1	34.0	39.0	45.0

	Complete Cases	Mean Imputation	Predicted Imputation	Multiple Imputation
(Intercept)	0.534 (0.146)	0.708 (0.116)	0.534 (0.112)	0.667 (0.130)
hgc	0.062 (0.005)	0.050 (0.004)	0.062 (0.004)	0.058 (0.005)
collegenot college grad	0.145 (0.034)	0.169 (0.026)	0.145 (0.025)	0.105 (0.030)
tenure	0.050 (0.005)	0.038 (0.004)	0.050 (0.004)	0.044 (0.005)
I(tenure^2)	-0.002 (0.000)	-0.001 (0.000)	-0.002 (0.000)	-0.001 (0.000)
age	0.000 (0.003)	0.000 (0.002)	0.000 (0.002)	-0.001 (0.003)
marriedsingle	-0.022 (0.018)	-0.027 (0.014)	-0.022 (0.013)	-0.021 (0.016)
Num.Obs.	1669	2229	2229	2246
Num.Imp.				5
R2	0.208	0.146	0.277	0.226
R2 Adj.	0.206	0.144	0.275	0.224
AIC	1179.9	1093.8	925.5	
BIC	1223.2	1139.5	971.1	
Log.Lik.	-581.936	-538.912	-454.737	
RMSE	0.34	0.31	0.30	