SUPPORT2 DATASET: HOSPITAL OUTCOMES OF CRITICALL ILL PATIENTS

Anthony Barrios, Elif Yildirim, &
Max Kucera

ABOUT THE DATA

- Dataset: SUPPORT2
- Multi-Center Clinical Study
- Goal: Understand medical decisions, prognoses, and outcomes of seriously ill, hospitalized patients
- ~9,000 critically ill patients
- age, sex, race, dzgroup, dzclass, hospital death, length of stay, and various clinic test results

Our Variables of Interest:

OUR RESEARCH QUESTIONS

1. Linear Regression final guiding question here
2. Logistic Regression final guiding question here

Visuals of what data looks like?

DATA PREPROCESSING

1. Cleaning
2. Factoring
3. Etc
4. Anthony’s cleaning
5. Dropping variables vif etc

EXPLORATORY ANALYSIS

Scatterplots here? key trends? what tests to include here?

CATEGORICAL V CATEGORICAL

sex v dnr? maybe don’t include because not as interesting

disease class v dnr, stronger relationship

Chi-Squared details, etc

BASELINE LINEAR MODEL

What are we predicting and explain variables of interest,

findings and changes, include initial RSE and R squared, etc, maybe VIF here instead

FINAL LINEAR MODEL

findings and changes, include final RSE and R squared, etc, maybe incremental f-test?

LOGISTIC BASELINE MODEL

What are we predicting and explain variables of interest,

Initial findings and accuracy metrics, preview the changes that are to come

LOGISTIC FINAL MODEL

How changes affected model, findings, etc

SUMMARY OF FINDINGS

Insight A, How does x1,x2,x3 predict hosp death 0/1
Insight B How does x1,x2,x3 contribute to slos
Insight C

Talk about if our findings were useful
If enhancements to the model worked?