

HOMEWORK 4
BIOSTATISTICS 755
DUE MARCH 16TH, 2023

1. **(50 points)** Respond to the “To address” item from your project.
2. **(50 points)** The Study of Assets and Health Dynamics among the Oldest Old (AHEAD) is a national longitudinal study with initial sample of 7444 respondents aged 70 years and older, and their spouses (if married). Objectives of the study include: (i) to monitor transitions in physical, functional, and cognitive health; (ii) to examine the relationship of late-life changes in physical and cognitive health to patterns of saving and income flows; (iii) to relate changes in health to economic resources and intergenerational transfers; (iv) to examine how the mix and distribution of economic, family, and program resources affect key outcomes, including institutionalization, saving, and health declines. In this problem, we look at data from the first four waves of the study, collected in 1993, 1995, 1998 and 2000. A reduced data set `ahead.xlsx` is on the course website. We explore the relationship of a test of cognitive function to some physical functioning indicators. The hypothesis is that, as physical function declines, movement becomes more difficult, and hence there are fewer stimuli, leading to cognitive decline. Here is a description of the variables contained:

`id`- subject identification number
`year`- study wave
`sex`- 1=male, 2=female
`age`- age in years
`immword`- immediate word recall
`delword`- delayed word recall
`blks`- difficulty walking several blocks without help
`strs`- difficulty climbing flight of stairs without help
`push`- difficulty pulling or pushing a living room chair
`bag`- difficulty lifting a bag of groceries
`dime`- difficulty picking up a dime
`iadlany`- perform any instrumental activities of daily living

The focus of the study is to assess the association of difficulty performing instrumental activities of daily living (IADLs) to memory and cognition problems. IADLs include activities such as shopping, managing money, taking medications, using the telephone, preparing meals. They are supposed to represent the types of integrative activities that one needs to be able to perform in order to live independently in the community.

The response variable here will be total word recall (`totword = immword + delword`) and the main exposure is `iadlany`, an indicator for whether or not the subject reports difficulty with any (at least one) IADL. The goal of this analysis is to determine if there is a relationship between total word recall and the IADL variable using GEE.

Consider the variables age and sex as possible confounders and/or effect modifiers for this analysis.

- (a) **(5 points)** Discuss whether you feel that GEE is the appropriate analysis choice for this project.
- (b) **(45 points)** Complete an analysis of this data using a GEE that results in the best possible answer to the the scientific aims of the study. This should start with exploratory analysis of the data, explore different statistical models based on model fit and scientific hypotheses, and finish with a clear summarization of your findings including the interpretation of the coefficients of interest. Please include only relevant output to how you found the best model. It is encouraged to give the code to replicate your findings.