Final Proposal

Chloe Wong, Hanwen Yang, Junhao Lai

2024-03-01

1 Introduction

Our research looks at how the quality of living arrangements affects how satisfied older adults are with their lives, especially during the COVID-19 economic downturn. We're using data from a study focusing on people aged 55 and up during the early stages of the pandemic, "https://www.openicpsr.org/openicpsr/project/131022/version/V3/view?path=/openicpsr/131022/fcr:versions/V3/C19CS-Data-Labeled.xlsx&type=file" The main question we want to find out is whether there's a direct link between living conditions and life satisfaction, or if we can predict satisfaction based on how good someone's living situation is. We're focusing on how people rate their own health as our main measure (Y), and one thing we're looking at is how likely it is that they'll run out of money because of COVID-19 (X). We're using different statistical methods like simple logit regression, lasso regression, and ridge regression to figure out these relationships and understand how living conditions, health, and happiness are connected during tough economic times.

```
# setwd('/Users/laix1/OneDrive/Desktop/Final_Proposal')
knitr::opts_knit$set(root.dir = "C:/Users/laix1/OneDrive/Desktop/Final_Proposal/ECON124")

library(haven)

covid_data <- data.frame(read_dta("C19CS Data.dta"))


dim(covid_data)
```

```
## [1] 6938 1867
```

```
colnames(covid_data)[c(204, 451)]
```

```
## [1] "t0_slfhlth" "t1_money"
```

```
# names(covid_data)
```

1.1 Part (b)

2 Data

3 Methodology

4 Main Results

we first did a raw correlation on the smoking status and COVID-19 positivity

```
SM_data <- na.omit(covid_data[, c(201, 851), ])  #
head(SM_data)
```

```
##   t0_smoke t5_cov_self_t
## 1        0            -2
## 2        0            -2
## 3        0            -2
## 4        0            -2
## 5        0            -2
## 7        2            -2
```

```
naive_SM_OLS <- glm(t5_cov_self_t ~ t0_smoke, data = SM_data)
summary(naive_SM_OLS)
```

```
##
## Call:
## glm(formula = t5_cov_self_t ~ t0_smoke, data = SM_data)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.72000    0.01853 -92.840   <2e-16 ***
## t0_smoke     0.02444    0.02612   0.936     0.35
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.5401954)
##
##     Null deviance: 1329.4  on 2461  degrees of freedom
## Residual deviance: 1328.9  on 2460  degrees of freedom
## AIC: 5474.7
##
## Number of Fisher Scoring iterations: 2
```

A person what has an active smoking status is associated with a 2.4% increased in the chance of tested positive for COVID-19 past month.