**HAP 823 | Assignment 06 Stratified Covariate Balancing Pt 1| March 8, 2020 | Chelsea Zabowski**

**Summary**

Question 4: My answer matches the Teach One Answer.

**Question 4**

**4.1.** The following data have been taken from nurses rounding in a facility.  The time they spent with patients has been recorded.  In addition, several characteristics of the patients have also been recorded and standardized.  Using stratified covariate balancing indicate if any of the nurses have a significant impact on overall satisfaction in the unit?

**Note:** The R packaged “StratifiedBalancing” written by Professor Alemi and Amr ElRafey is no longer maintained on CRAN.

Many functions have been deleted from the R package since the Teach One video was posted, therefore answers may differ.

> print(packageVersion("StratifiedBalancing"))

[1] ‘0.3.0’

Only the stratify() function is in the package

> # Print functions within the package:

> lsf.str("package:StratifiedBalancing")

stratify : function (Treatment, Outcome, Matrix, Discretize = TRUE, Synthetic = TRUE, Ordered = TRUE, Markov = FALSE)

# function to install packages if missing by user

fun.usePackage<-function(p){

 if (!is.element(p, installed.packages()[,1])){

 print(paste("Package:",p,"Not found, Installing Now..."))

 install.packages(p, dep = TRUE)}

 print(paste("Loading Package :",p))

 require(p, character.only = TRUE)

}

fun.usePackage('tidyverse') # data manipulation

fun.usePackage('writexl') # read and write excel files

fun.usePackage('openxlsx') # read and write excel files

fun.usePackage('readxl') # read and write excel files

fun.usePackage('StratifiedBalancing') # stratified balancing package

# do some memory cleanup.

gc(verbose = TRUE, reset = TRUE)

format(memory.size(), units = "MB")

rm(list = ls(all.names = TRUE))

#### This follows Polly Shih's Teach One video: https://www.youtube.com/watch?v=3Hy\_f9\_nzpc&feature=youtu.be

# Import Data

Q4\_Data <- read\_excel("HAP823/Wk6\_Stratified\_Covariate\_Balancing\_Pt1/Q4/StratCovariateBalancing\_Q4Data\_statisfaction\_w\_rounding\_time.xlsx")

# Convert Nurse 1 to Binary

Q4\_Data\_Binary <- Q4\_Data %>%

 mutate(Nurse1 = if\_else(Nurse1 > mean(Nurse1),1,0))

# Convert Satisfaction to Binary

Q4\_Data\_Binary <- Q4\_Data\_Binary %>%

 mutate(Satisfaction = if\_else(Satisfaction > mean(Satisfaction),1,0))

# Convert Age to Binary

Q4\_Data\_Binary <- Q4\_Data\_Binary %>%

 mutate(Age = if\_else(Age > mean(Age),1,0))

# Drop unnecessary columns

subdata = Q4\_Data\_Binary[-c(3,4,12)]

# Balance Data using stratadisc function

# Use covariates MI, CHF, Diabetes, Injuries, Lung Cancer, Age, Under Staffed to see effect of Nurse 1 on Satisfaction

# Note: The repository on CRAN is owned by Professor Alemi but not longer maintained\*\*\*\*

**# The Stratadisc function no longer exist**

# Convert from dataframe to a matrix

subdata <- data.matrix(subdata, rownames.force = NA)

# Balance the data using the stratify function (stratdisc no longer exist)

balanced = stratify(2,1,subdata) # 2 = Nurse1 (Treatment) , 1 = Satisfaction (The Outcome variable)



> balanced = stratify(2,1,subdata)

[1] "Before usning synthetic mathcing, percent of cases mathced was: 91.7335473515249"

[1] "After usning synthetic mathcing, percent of cases mathced became: 100"

 Results

Odds Ratio Of Impact Of Treatment on Outcome 4.649

Mantel Haenszel P-Value 0.334

No. Of Cases Matched 1246.000

95% C.I Upper Bound 5.340

95% C.I Lower Bound 4.535

**Summary:** Because the p-value > 0.05 and the Odds Ratio is 4.649, it means that Nurse 1 does not have a significant impact on satisfaction in the unit because we used all available covariates. If more covariates are dropped, the impact of Nurse 1 on satisfaction may change.