Project1

Garima Lohani

27 October 2018

# PART I

# Question i

ds<-read.table("hearttransplant.csv", header=TRUE, stringsAsFactors = FALSE,sep=",",na.strings="-999")  
str(ds)

## 'data.frame': 10168 obs. of 12 variables:  
## $ female : int 0 0 1 1 1 1 0 0 0 0 ...  
## $ medicaid : int 0 0 1 0 0 0 0 0 0 0 ...  
## $ ethcat4 : int 2 1 2 1 1 2 1 1 1 1 ...  
## $ diag : int 2 1 2 1 1 1 6 1 1 6 ...  
## $ diab4 : int 1 1 3 1 1 3 1 1 1 3 ...  
## $ age\_list : int 58 62 57 57 54 60 68 52 52 56 ...  
## $ bmi\_list : num 23.6 30.3 26.8 29.5 32.6 ...  
## $ gfr\_list : num 50.4 25.1 47 86.2 39.2 ...  
## $ age\_tx : int 58 NA NA 57 55 60 69 NA 53 56 ...  
## $ bmi\_tx : num 22.7 30.8 26.8 29.5 32.1 ...  
## $ gfr\_tx : num 41.8 NA NA 86.2 46.6 ...  
## $ pstatusdc: int 0 NA NA 0 0 0 0 NA 0 0 ...

dim(ds)

## [1] 10168 12

# Question ii

ds$agecat <-   
ds$agecat [ds$age\_list <=18 & ds$age\_list <=39]<-1  
ds$agecat [ds$age\_list >=40 & ds$age\_list<=59]<-2  
ds$agecat [ds$age\_list >=60 & ds$age\_list <=69]<-3  
  
  
table(ds$agecat)

##   
## 1 2 3   
## 1707 5278 3183

# Question iii

table(ds$diab4)

##   
## 1 2 3   
## 7342 327 2260

ds$diabetes <-NA  
ds$diabetes[ds$diab4==1]<- 0  
ds$diabetes[ds$diab4==2 | ds$diab4==3]<- 1

# Question iv

str(ds$diabetes)

## num [1:10168] 0 0 1 0 0 1 0 0 0 1 ...

ds<- subset(ds, !is.na(ds$diabetes))  
str(ds$diabetes)

## num [1:9929] 0 0 1 0 0 1 0 0 0 1 ...

dim(ds)

## [1] 9929 14

table(ds$pstatusdc)

##   
## 0 1   
## 6020 342

ds$tx\_complete <-  
ds$ tx\_complete[ds$pstatusdc==is.na(ds$pstatusdc)]<-0  
ds$tx\_complete[ds$pstatusdc==0 | ds$pstatusdc==1]<-1  
table(ds$tx\_complete)

##   
## 0 1   
## 3567 6362

# Question vi

ds$gfr\_delta <- ds$gfr\_tx-ds$gfr\_list

# Question vii

ds$bmi\_delta<- ds$bmi\_tx-ds$bmi\_list  
  
ds2<-ds #saving it in a variable as later on its value changes when we label

# Part II

# part a

library(tableone)

## Warning: package 'tableone' was built under R version 3.4.4

ds$ethcat4 <- factor(ds$ethcat4, levels=c(1,2,3), labels=c("White","Black/African American","Hispanic"))  
  
ds$diag <- factor(ds$diag,levels=c(1,2,3,4,5,6,7),labels=c("Cardiomyopathy","Ischemic","Congenital heart disease","Hypertrophic","Restrictive " ,"Valvular","Other"))  
  
ds$female<- factor(ds$female, levels=c(0,1), labels=c("male","females"))  
  
ds$medicaid <- factor(ds$medicaid,levels=c(0,1), labels=c("nonmedicaid insurance","medicaid insurance"))  
  
ds$tx\_complete<- factor(ds$tx\_complete,levels=c(0,1),labels=c("Notransplant","Transplant"))  
   
ds$diabetes[ds$diabetes==0]<- "No diabetes"  
ds$diabetes[ds$diabetes==1]<- "Diabetes"  
  
myCats<-c("age\_list","ethcat4","female","medicaid","diag","bmi\_list","gfr\_list","tx\_complete")  
  
tab1 <-CreateTableOne(vars=myCats, strata="diabetes",data=ds)  
print(tab1)

## Stratified by diabetes  
## Diabetes No diabetes p   
## n 2587 7342   
## age\_list (mean (sd)) 55.86 (8.71) 50.04 (12.96) <0.001  
## ethcat4 (%) 0.002  
## White 1810 (70.0) 5235 (71.3)   
## Black/African American 538 (20.8) 1588 (21.6)   
## Hispanic 239 ( 9.2) 519 ( 7.1)   
## female = females (%) 519 (20.1) 2004 (27.3) <0.001  
## medicaid = medicaid insurance (%) 286 (11.1) 1106 (15.1) <0.001  
## diag (%) <0.001  
## Cardiomyopathy 1011 (39.1) 3716 (50.6)   
## Ischemic 1305 (50.4) 2245 (30.6)   
## Congenital heart disease 17 ( 0.7) 287 ( 3.9)   
## Hypertrophic 18 ( 0.7) 189 ( 2.6)   
## Restrictive 35 ( 1.4) 223 ( 3.0)   
## Valvular 36 ( 1.4) 168 ( 2.3)   
## Other 165 ( 6.4) 514 ( 7.0)   
## bmi\_list (mean (sd)) 29.45 (5.43) 26.87 (5.06) <0.001  
## gfr\_list (mean (sd)) 61.69 (26.04) 68.65 (37.24) <0.001  
## tx\_complete = Transplant (%) 1555 (60.1) 4807 (65.5) <0.001  
## Stratified by diabetes  
## test  
## n   
## age\_list (mean (sd))   
## ethcat4 (%)   
## White   
## Black/African American   
## Hispanic   
## female = females (%)   
## medicaid = medicaid insurance (%)   
## diag (%)   
## Cardiomyopathy   
## Ischemic   
## Congenital heart disease   
## Hypertrophic   
## Restrictive   
## Valvular   
## Other   
## bmi\_list (mean (sd))   
## gfr\_list (mean (sd))   
## tx\_complete = Transplant (%)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **No Diabetes**  **(N=xxx)** | **Diabetes**  **(N=xxx)** | **p-value** |
| **Age at Listing**, years | 50.04 (12.96) | 55.86 (8.71) | <0.001 |
| **Female Sex** | 2004 (27.3) | 519 (20.1) | 0.001 |
| **Race/Ethnicity Group** |  |  | 0.002 |
| White | 5235 (71.3) | 1810 (70.0) |  |
| Black/African-American | 1588 (21.6) | 538 (20.8) |
| Hispanic | 519 ( 7.1) | 239 ( 9.2) |
| **Medicaid Insurance** | 1106 (15.1) | 286 (11.1) |  |
| **Diagnosis** |  |  |  |
| Cardiomyopathy | 3716 (50.6) | 1011 (39.1) |  |
| Ischemic | 2245 (30.6) | 1305 (50.4) |
| Congenital heart disease | 287 ( 3.9) | 17 ( 0.7) |
| Hypertrophic | 189 ( 2.6) | 18 ( 0.7) |
| Restrictive | 223 ( 3.0) | 35 ( 1.4) |
| Valvular | 168 ( 2.3) | 36 ( 1.4) |
| Other | 514 ( 7.0) | 165 ( 6.4) |
| **Body mass index at Listing, kg/m2** | 26.87 (5.06) | 29.45 (5.43) | 0.001 |
| **Glomular filtration rate at Listing, mL/min/1.73m2** | 68.65 (37.24) | 61.69 (26.04) | 0.001 |
| **Underwent Transplant** | 4807 (65.5) | 1555 (60.1) | 0.001 |

# Part b

ds1<-subset(ds,!tx\_complete==0) #patient who received transplant  
  
dim(ds1)

## [1] 9929 17

library(epitools)

## Warning: package 'epitools' was built under R version 3.4.2

tb1 <- table(ds1$diabetes,ds1$pstatusdc)  
  
  
  
prop.table(tb1,2)

##   
## 0 1  
## Diabetes 0.2443522 0.2456140  
## No diabetes 0.7556478 0.7543860

res1<-chisq.test(tb1,correct=FALSE)  
  
res1$expected

##   
## 0 1  
## Diabetes 1471.408 83.59164  
## No diabetes 4548.592 258.40836

res1

##   
## Pearson's Chi-squared test  
##   
## data: tb1  
## X-squared = 0.0027903, df = 1, p-value = 0.9579

oddsratio(t(tb1),method="wald")

## $data  
##   
## Diabetes No diabetes Total  
## 0 1471 4549 6020  
## 1 84 258 342  
## Total 1555 4807 6362  
##   
## $measure  
## odds ratio with 95% C.I.  
## estimate lower upper  
## 0 1.000000 NA NA  
## 1 0.993201 0.7710852 1.279299  
##   
## $p.value  
## two-sided  
## midp.exact fisher.exact chi.square  
## 0 NA NA NA  
## 1 0.9501963 0.9484856 0.957873  
##   
## $correction  
## [1] FALSE  
##   
## attr(,"method")  
## [1] "Unconditional MLE & normal approximation (Wald) CI"

# HO: Among the patient who received transplant the true probabilties of the event are same in diabetes status and survival to discharge  
  
#H1:Among the patient who received transplant the true probabilties of the event are different in diabetes status and survival to discharge  
  
# The test statistic is 0.002 with degree of freedom 1, p value is 0.95. Since the p value is greater than 0.05, we do not have enough evidence to reject null hypothesis. Thus, we conclude that the true probablities of the event are same in diabetes status and survival to discharge.

# part c

sd<-subset(ds, !is.na(ds$bmi\_list))  
sd<-subset(ds, !is.na(ds$bmi\_tx))  
sd<-subset(ds, !is.na(ds$gfr\_tx))  
sd<-subset(ds, !is.na(ds$gfr\_list))  
sd<-subset(ds, !is.na(ds$bmi\_delta))  
sd<-subset(ds, !is.na(ds$gfr\_delta))  
dim(sd)

## [1] 6279 17

dim(ds)

## [1] 9929 17

sd$diabetes[sd$diabetes==0]<- "No diabetes"  
sd$diabetes[sd$diabetes==1]<- "Diabetes"  
  
myCats<-c("bmi\_list","bmi\_tx","bmi\_delta","gfr\_list","gfr\_tx","gfr\_delta")  
  
tab1 <-CreateTableOne(vars=myCats, strata="diabetes",data=sd)  
print(tab1)

## Stratified by diabetes  
## Diabetes No diabetes p test  
## n 1537 4742   
## bmi\_list (mean (sd)) 28.85 (5.54) 26.41 (4.74) <0.001   
## bmi\_tx (mean (sd)) 28.70 (4.83) 26.43 (5.44) <0.001   
## bmi\_delta (mean (sd)) -0.15 (4.39) 0.02 (3.33) 0.114   
## gfr\_list (mean (sd)) 63.41 (26.03) 69.08 (34.32) <0.001   
## gfr\_tx (mean (sd)) 63.17 (25.43) 70.70 (39.18) <0.001   
## gfr\_delta (mean (sd)) -0.24 (24.78) 1.62 (29.64) 0.027

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **No Diabetes** | | | | | **Diabetes** | | | | |
| **N=4742** | **Mean (SD)** | | | **p-value**  comparing listing  to transplant | **N**  **=1537** | **Mean (SD)** | | | **p-value**  comparing listing  to transplant |
| Listing | Transplant | Change | Listing | Transplant | Change |
| Body mass index, kg/m2 |  | 26.41 (4.74 | 26.43 (5.44 | 0.02 (3.33) | 0.114 |  | 28.85 (5.54) | 28.70 (4.83) | -0.15 (4.39) | 0.114 |
| Glomular Filtration Rate, mL/min/1.73m2 |  | 69.08 (34.32) | 70.70 (39.18) | 1.62 (29.64) | 0.027 |  | 63.41 (26.03) | 63.17 (25.43 | -0.24 (24.78) | 0.027 |

# 

# part d

dim(ds2)

## [1] 9929 17

dia<- subset(ds2,!(ds2$diabetes==0))  
dim(dia)

## [1] 2587 17

t.test(dia$bmi\_delta, mu=0)

##   
## One Sample t-test  
##   
## data: dia$bmi\_delta  
## t = -1.2024, df = 2585, p-value = 0.2293  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -0.21192699 0.05081586  
## sample estimates:  
## mean of x   
## -0.08055557

t.test(dia$gfr\_delta, mu=0)

##   
## One Sample t-test  
##   
## data: dia$gfr\_delta  
## t = -0.37765, df = 1536, p-value = 0.7057  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -1.478247 1.000932  
## sample estimates:  
## mean of x   
## -0.2386577

#HO: There is a no change in BMI between listing and transplant  
#H1: There is change in BMI between listing and transplant  
  
#The test statistic is -1.2024, df=2585 with p value= 0.2293. Since the p value is larger than 0.05 we cannot reject null hyposthesis and thus conclude there is no significant change in BMI between listing and transplant.  
  
  
  
  
  
#HO: There is a no change in GFR between listing and transplant  
#H1: There is change in GFR between listing and transplant  
  
#The test statistic is -0.37765, df=1536 with p value= 0.7057. Since the p value is larger than 0.05 we cannot reject null hyposthesis and thus conclude there is no significant change in GFR between listing and transplant in case of diabetes.

# part e

dim(ds2)

## [1] 9929 17

ndia<- subset(ds2,!(ds2$diabetes==1))  
dim(ndia)

## [1] 7342 17

t.test(ndia$bmi\_delta, mu=0)

##   
## One Sample t-test  
##   
## data: ndia$bmi\_delta  
## t = 0.52006, df = 7335, p-value = 0.603  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -0.05579099 0.09608241  
## sample estimates:  
## mean of x   
## 0.02014571

t.test(ndia$gfr\_delta, mu=0)

##   
## One Sample t-test  
##   
## data: ndia$gfr\_delta  
## t = 3.7614, df = 4741, p-value = 0.000171  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 0.7751839 2.4628698  
## sample estimates:  
## mean of x   
## 1.619027

#HO: There is a no change in BMI between listing and transplant  
#H1: There is change in BMI between listing and transplant  
  
#The test statistic is 0.52006, df=7335 with p value= 0.603. Since the p value is larger than 0.05 we cannot reject null hyposthesis and thus conclude there is no significant change in BMI between listing and transplant in case of patient with no diabetes.  
  
  
#HO: There is a no change in GFR between listing and transplant  
#H1: There is change in GFR between listing and transplant  
  
#The test statistic is 3.76, df=4741 with p value= 0.000171. Since the p value is smaller than 0.05 we have significant evidence to reject null hyposthesis and thus conclude there is significant change in GFR between listing and transplant in case of patient with no diabetes.

# Part III

#Method:  
#Descriptive statistics for listing were summarized using means and p value.We performed chi square test to find if there is any association between diabetes status and survival to discharge. To know if there is a change in BMI or GFR between listing and transplant both in case of diabetes and not diabetes we applied Chi square test.P-values presented are considered significant at the 0.05 level. All statistical analyses were carried out using R version 1.0.  
  
  
#Results:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **No Diabetes**  **(N=xxx)** | **Diabetes**  **(N=xxx)** | **p-value** |
| **Age at Listing**, years | 50.04 (12.96) | 55.86 (8.71) | <0.001 |
| **Female Sex** | 2004 (27.3) | 519 (20.1) | 0.001 |
| **Race/Ethnicity Group** |  |  | 0.002 |
| White | 5235 (71.3) | 1810 (70.0) |  |
| Black/African-American | 1588 (21.6) | 538 (20.8) |
| Hispanic | 519 ( 7.1) | 239 ( 9.2) |
| **Medicaid Insurance** | 1106 (15.1) | 286 (11.1) |  |
| **Diagnosis** |  |  |  |
| Cardiomyopathy | 3716 (50.6) | 1011 (39.1) |  |
| Ischemic | 2245 (30.6) | 1305 (50.4) |
| Congenital heart disease | 287 ( 3.9) | 17 ( 0.7) |
| Hypertrophic | 189 ( 2.6) | 18 ( 0.7) |
| Restrictive | 223 ( 3.0) | 35 ( 1.4) |
| Valvular | 168 ( 2.3) | 36 ( 1.4) |
| Other | 514 ( 7.0) | 165 ( 6.4) |
| **Body mass index at Listing, kg/m2** | 26.87 (5.06) | 29.45 (5.43) | 0.001 |
| **Glomular filtration rate at Listing, mL/min/1.73m2** | 68.65 (37.24) | 61.69 (26.04) | 0.001 |
| **Underwent Transplant** | 4807 (65.5) | 1555 (60.1) | 0.001 |

The above table show the difference in listing values for diabetes and non diabetes.

To know whether diabetes is associated with survival to hospital discharge we performed Chi square test.

The true probabilities of the event are same in diabetes status and survival to discharge. The test statistic is 0.002 with degree of freedom 1, p value is 0.95. Since the p value is greater than 0.05, we do not have enough evidence to reject null hypothesis.

To find any association between diabetes and change in clinical characteristics between listing and transplant we performed paired t test

There is significant change in GFR between listing and transplant in case of patient with no diabetes. The test statistic is 3.76, df=4741 with p value= 0.000171. Since the p value is smaller than 0.05 we have significant evidence to reject null hypothesis .

#Conclusion:  
#In conclusion, we observed association between the diabetes status and survival to discharge. Similarly, no significant change in BMI or GFR in case of patient with diabetes. In case of patient with no diabetes,there was no significant change in BMI but in case of GFR significant changes were observed.