**19 ביוני 2021**

עבודת ביניים

**רקע כללי**

1. **בנייה של קובץ דירות** - כהכנה לביצוע הפרויקט בנינו קובץ של דירות, כך שלכל עיר דירה הייתה רשומה אחת.

הקוד המתייחס לבנייה כתוב ב- Part 1 להלן בנוסף צרפנו את הדוח של קובץ הדירות בפורמט של אקסל.

1. **ניתוח אומדנים** - ב- Part 2 ניתחנו את הנתונים הבאים:

א. סך שייכים לכח העבודה עבור כל האוכלוסייה, גברים, ונשים.

ב. אחוז השייכים לכח העבודה מתוך כלל האוכלוסייה עבור כל האוכלוסייה, גברים, ונשים

.ג. אחוז הלא-עובדים מתוך השייכים לכח העבודה עבור כל האוכלוסייה, גברים, ונשים.

ד. ממוצע שעות העבודה בשבוע לאלה שעובדים עבור כל האוכלוסייה. בכל אחד מהסעיפים א-ד חישבנו שלושה אומדנים.

1. **חלוקה לתתי מדגם וחישוב שונויות** –
2. ב- Part 3 ניתחנו את הנתונים הבאים:

כהכנה לחישוב חלקנו את המדגם ל-5 תתי מדגם בהתאם לכללים שלמדו בתחילת הקורס, עבור כל אחד מהאומדנים הדרושים אמדנו את השונות, את טעות הדגימה, ואת טעות הדגימה היחסית.

את השונויות אמדנו בשיטת הקבוצות המקריות ובשיטת jackknife

1. **שפת תכנות ותוצרים** - כפי שניתן לראות הפרויקט נכתב בשפת – R. הקוד והתוצאות והפלטים הרלוונטים מופיעים בקובץ.
2. **לסיכום**: נהננו מאד מהעבודה ולמדנו ליישם כפי שמבוצע בלמ"ס המון נושאים חשובים שנסקרו בשיעור בנושא אמידת שונות.

variance

library(readxl)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(pander)  
library(rio)  
  
#install.packages("rio")

# Part 1

work\_data <- read\_excel("C:/Users/zivpa/OneDrive/שולחן העבודה/variance/var\_data.xlsx")

## New names:  
## \* `` -> ...2  
## \* `` -> ...3  
## \* `` -> ...4  
## \* `` -> ...5  
## \* `` -> ...6  
## \* ...

work\_data <- na.omit(work\_data)  
work\_data <- work\_data[-1,]  
colnames(work\_data) <- c("city", "apt\_num", "room\_num", "people\_num", "gender", "working\_hours", "age", "code", "weight")

work\_data$female <- ifelse(test = work\_data$gender == 2 & work\_data$age > 14 , yes = 1, no = 0)  
work\_data$male <- ifelse(test = work\_data$gender == 1 & work\_data$age > 14 , yes = 1, no = 0)  
work\_data$all <- work\_data$female + work\_data$male  
  
work\_data$belongs <- ifelse(test = work\_data$code < 5 , yes = 1, no = 0)  
work\_data$not\_belongs <- ifelse(test = work\_data$code > 4 , yes = 1, no = 0)  
work\_data$working <- ifelse(test = work\_data$code < 4 , yes = 1, no = 0)  
work\_data$not\_working <- ifelse(test = work\_data$code == 4 , yes = 1, no = 0)  
  
work\_data$belonging\_male <- ifelse(test = work\_data$code < 5 & work\_data$gender == 1, yes = 1, no = 0)  
work\_data$not\_belonging\_male <- ifelse(test = work\_data$code > 4 & work\_data$gender == 1 , yes = 1, no = 0)  
work\_data$working\_male <- ifelse(test = work\_data$code < 4 & work\_data$gender == 1 , yes = 1, no = 0)  
work\_data$not\_working\_male <- ifelse(test = work\_data$code == 4 & work\_data$gender == 1 , yes = 1, no = 0)  
  
work\_data$belonging\_female <- ifelse(test = work\_data$code < 5 & work\_data$gender == 2, yes = 1, no = 0)  
work\_data$not\_belonging\_female <- ifelse(test = work\_data$code > 4 & work\_data$gender == 2 , yes = 1, no = 0)  
work\_data$working\_female <- ifelse(test = work\_data$code < 4 & work\_data$gender == 2 , yes = 1, no = 0)  
work\_data$not\_working\_female <- ifelse(test = work\_data$code == 4 & work\_data$gender == 2 , yes = 1, no = 0)

fixed\_data <- aggregate(list(as.numeric(work\_data$working\_hours), work\_data$male, work\_data$female, work\_data$all , work\_data$belongs, work\_data$not\_belongs, work\_data$working, work\_data$not\_working, work\_data$belonging\_male, work\_data$not\_belonging\_male, work\_data$working\_male, work\_data$not\_working\_male, work\_data$belonging\_female, work\_data$not\_belonging\_female, work\_data$working\_female, work\_data$not\_working\_female), by = list(work\_data$city, as.numeric(work\_data$apt\_num), work\_data$room\_num, work\_data$people\_num, round(as.numeric(work\_data$weight), 4)), sum)  
colnames(fixed\_data) <- c("city", "apt\_num", "room\_num", "people\_num","weights", "working\_hours","tot\_males\_over15", "tot\_females\_over15","all", "belonging","not\_belonging", "working", "not\_working", "belonging\_male", "not\_belonging\_male", "working\_male", "not\_working\_male", "belonging\_female", "not\_belonging\_female", "working\_female", "not\_working\_female")  
  
fixed\_data <- fixed\_data %>% arrange(city, apt\_num)

# Part 2:

**Q1**

df1 <- fixed\_data %>% mutate(tot\_est = fixed\_data$belonging \* weights)  
total\_belonging\_est <- round(sum(df1$tot\_est))  
paste("The total estimation of belonging male is:", total\_belonging\_est)

## [1] "The total estimation of belonging male is: 522794"

df1 <- df1 %>% mutate(male\_est = fixed\_data$belonging\_male \* weights)  
male\_belonging\_est <- round(sum(df1$male\_est))  
paste("The total estimation of belonging male is:", male\_belonging\_est)

## [1] "The total estimation of belonging male is: 259487"

df1 <- df1 %>% mutate(female\_est = fixed\_data$belonging\_female \* weights)  
female\_belonging\_est <- round(sum(df1$female\_est))  
paste("The total estimation of belonging male is:", female\_belonging\_est)

## [1] "The total estimation of belonging male is: 263308"

#presentation  
population\_type <- c("Total", "Men", "Women")  
labor\_force\_est <- data.frame(population\_type = population\_type, Y\_hat = c(total\_belonging\_est,male\_belonging\_est, female\_belonging\_est))  
pander(labor\_force\_est)

|  |  |
| --- | --- |
| population\_type | Y\_hat |
| Total | 522794 |
| Men | 259487 |
| Women | 263308 |

**Q2**

df2 <- df1 %>% mutate(tot\_pop = fixed\_data$all \* weights)  
sum\_total <- sum(df2$tot\_pop)  
belong\_tot\_perc <- total\_belonging\_est/sum\_total  
   
  
df2 <- df1 %>% mutate(weight\_pop\_male = fixed\_data$tot\_males\_over15 \* weights)  
sum\_tot\_male <- sum(df2$weight\_pop\_male)  
belong\_tot\_perc\_male <- male\_belonging\_est/sum\_tot\_male  
   
  
df2 <- df1 %>% mutate(weight\_pop\_female = fixed\_data$tot\_females\_over15 \* weights)  
sum\_tot\_female <- sum(df2$weight\_pop\_female)  
belong\_tot\_perc\_female <- female\_belonging\_est/sum\_tot\_female  
  
  
#presentation  
labor\_force <- data.frame(population\_type = population\_type,Y\_hat = c(belong\_tot\_perc, belong\_tot\_perc\_male, belong\_tot\_perc\_female))  
pander(labor\_force)

|  |  |
| --- | --- |
| population\_type | Y\_hat |
| Total | 0.5746 |
| Men | 0.5925 |
| Women | 0.558 |

**Q3**

df3 <- df2 %>% mutate(tot\_not\_working = fixed\_data$not\_working \* weights)  
sum\_tot\_not\_working <- sum(df3$tot\_not\_working)  
not\_working\_tot\_perc <- sum\_tot\_not\_working/total\_belonging\_est  
   
  
df3 <- df2 %>% mutate(tot\_not\_working\_male = fixed\_data$not\_working\_male \* weights)  
sum\_not\_working\_male <- sum(df3$tot\_not\_working\_male)  
not\_working\_male\_perc <- sum\_not\_working\_male/male\_belonging\_est  
   
  
df3 <- df2 %>% mutate(tot\_not\_working\_female = fixed\_data$not\_working\_female \* weights)  
sum\_not\_working\_female <- sum(df3$tot\_not\_working\_female)  
not\_working\_female\_perc <- sum\_not\_working\_female/female\_belonging\_est  
  
  
#presentation  
labor\_force <- data.frame(population\_type = population\_type,Y\_hat = c(not\_working\_tot\_perc, not\_working\_male\_perc, not\_working\_female\_perc))  
pander(labor\_force)

|  |  |
| --- | --- |
| population\_type | Y\_hat |
| Total | 0.0464 |
| Men | 0.05354 |
| Women | 0.03936 |

**Q4**

df4 <- df3 %>% mutate(tot\_working\_hours = fixed\_data$working\_hours \* weights)  
df4 <- df4 %>% mutate(tot\_workers = fixed\_data$working \* weights)  
  
sum\_working\_hours <- sum(df4$tot\_working\_hours)  
sum\_workers <- sum(df4$tot\_workers)  
  
  
   
paste("The average estimation of woking hours per weak is:", round(sum\_working\_hours/sum\_workers, 2))

## [1] "The average estimation of woking hours per weak is: 37.2"

export(fixed\_data, "variance\_estimation.xlsx")

# Part 3:

*חלוקה לתתי מדגם לפי ערים*

jerusalem\_df <- fixed\_data %>% filter(city == 3000)  
tlv\_df <- fixed\_data %>% filter(city == 5000)  
haifa\_df <- fixed\_data %>% filter(city == 4000)  
haifa\_df <- haifa\_df[-1,]  
  
c1 <- rep(c(1,2,3,4,5), length(jerusalem\_df$city)/5)  
c2 <- rep(c(1,2,3,4,5), length(tlv\_df$city)/5)  
c3 <- rep(c(1,2,3,4,5), length(haifa\_df$city)/5)  
  
jerusalem\_df$sub <- c1  
tlv\_df$sub <- c2  
haifa\_df$sub <- c3  
  
full\_data <- full\_join(jerusalem\_df, tlv\_df)

## Joining, by = c("city", "apt\_num", "room\_num", "people\_num", "weights", "working\_hours", "tot\_males\_over15", "tot\_females\_over15", "all", "belonging", "not\_belonging", "working", "not\_working", "belonging\_male", "not\_belonging\_male", "working\_male", "not\_working\_male", "belonging\_female", "not\_belonging\_female", "working\_female", "not\_working\_female", "sub")

full\_data <- full\_join(full\_data, haifa\_df)

## Joining, by = c("city", "apt\_num", "room\_num", "people\_num", "weights", "working\_hours", "tot\_males\_over15", "tot\_females\_over15", "all", "belonging", "not\_belonging", "working", "not\_working", "belonging\_male", "not\_belonging\_male", "working\_male", "not\_working\_male", "belonging\_female", "not\_belonging\_female", "working\_female", "not\_working\_female", "sub")

group\_1 <- full\_data %>% filter(sub == 1)  
group\_2 <- full\_data %>% filter(sub == 2)  
group\_3 <- full\_data %>% filter(sub == 3)  
group\_4 <- full\_data %>% filter(sub == 4)  
group\_5 <- full\_data %>% filter(sub == 5)  
  
JK\_data1 <- full\_data %>% filter(sub != 1)  
JK\_data2 <- full\_data %>% filter(sub != 2)  
JK\_data3 <- full\_data %>% filter(sub != 3)  
JK\_data4 <- full\_data %>% filter(sub != 4)  
JK\_data5 <- full\_data %>% filter(sub != 5)

#Variance estimation

**Q1** #שיטת קבוצות מקריות

#total  
random\_tot\_estimation <- rbind(  
 round(sum(group\_1$belonging \* group\_1$weights \* 5)),  
 round(sum(group\_2$belonging \* group\_2$weights \* 5)),  
 round(sum(group\_3$belonging \* group\_3$weights \* 5)),  
 round(sum(group\_4$belonging \* group\_4$weights \* 5)),  
 round(sum(group\_5$belonging \* group\_5$weights \* 5)))  
  
Var\_random\_tot1 <- sum((random\_tot\_estimation - mean(random\_tot\_estimation))^2)/20  
sd\_random\_tot1 <- sqrt(Var\_random\_tot1)  
se\_random\_tot1 <- sd\_random\_tot1/total\_belonging\_est  
  
  
#male  
random\_male\_estimation <- rbind(  
 round(sum(group\_1$belonging\_male \* group\_1$weights \* 5)),  
 round(sum(group\_2$belonging\_male \* group\_2$weights \* 5)),  
 round(sum(group\_3$belonging\_male \* group\_3$weights \* 5)),  
 round(sum(group\_4$belonging\_male \* group\_4$weights \* 5)),  
 round(sum(group\_5$belonging\_male \* group\_5$weights \* 5)))  
  
Var\_random\_male1 <- sum((random\_male\_estimation - mean(random\_male\_estimation))^2)/20  
sd\_random\_male1 <- sqrt(Var\_random\_male1)  
se\_random\_male1 <- sd\_random\_male1/male\_belonging\_est  
  
  
#female  
random\_female\_estimation <- rbind(  
 round(sum(group\_1$belonging\_female \* group\_1$weights \* 5)),  
 round(sum(group\_2$belonging\_female \* group\_2$weights \* 5)),  
 round(sum(group\_3$belonging\_female \* group\_3$weights \* 5)),  
 round(sum(group\_4$belonging\_female \* group\_4$weights \* 5)),  
 round(sum(group\_5$belonging\_female \* group\_5$weights \* 5)))  
  
Var\_random\_female1 <- sum((random\_female\_estimation - mean(random\_female\_estimation))^2)/20  
sd\_random\_female1 <- sqrt(Var\_random\_female1)  
se\_random\_female1 <- sd\_random\_female1/female\_belonging\_est  
  
  
  
  
names <- c("Variance", "SD", "SE")  
random\_est1 <- data.frame(Random = names , total = c(Var\_random\_tot1, sd\_random\_tot1,se\_random\_tot1), male = c(Var\_random\_male1, sd\_random\_male1,se\_random\_male1), female = c(Var\_random\_female1, sd\_random\_female1,se\_random\_female1))  
  
pander(random\_est1)

|  |  |  |  |
| --- | --- | --- | --- |
| Random | total | male | female |
| Variance | 36715401 | 52336070 | 94578935 |
| SD | 6059 | 7234 | 9725 |
| SE | 0.01159 | 0.02788 | 0.03693 |

#Jacknife

#total  
jk\_tot\_estimation <- rbind(  
 round(sum(JK\_data1$belonging \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$belonging \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$belonging \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$belonging \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$belonging \* JK\_data5$weights \* 5/4)))  
  
Var\_jk\_tot1 <- sum((jk\_tot\_estimation - mean(jk\_tot\_estimation))^2)\*4/5  
sd\_jk\_tot1 <- sqrt(Var\_jk\_tot1)  
se\_jk\_tot1 <- sd\_jk\_tot1/total\_belonging\_est  
  
  
#male  
jk\_male\_estimation <- rbind(  
 round(sum(JK\_data1$belonging\_male \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$belonging\_male \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$belonging\_male \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$belonging\_male \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$belonging\_male \* JK\_data5$weights \* 5/4)))  
  
Var\_jk\_male1 <- sum((jk\_male\_estimation - mean(jk\_male\_estimation))^2)\*4/5  
sd\_jk\_male1 <- sqrt(Var\_jk\_male1)  
se\_jk\_male1 <- sd\_jk\_male1/male\_belonging\_est  
  
  
#female  
jk\_female\_estimation <- rbind(  
 round(sum(JK\_data1$belonging\_female \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$belonging\_female \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$belonging\_female \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$belonging\_female \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$belonging\_female \* JK\_data5$weights \* 5/4)))  
  
Var\_jk\_female1 <- sum((jk\_female\_estimation - mean(jk\_female\_estimation))^2)\*4/5  
sd\_jk\_female1 <- sqrt(Var\_jk\_female1)  
se\_jk\_female1 <- sd\_jk\_female1/female\_belonging\_est  
  
  
  
  
JK\_est1 <- data.frame(Jacknife = names , total = c(Var\_jk\_tot1, sd\_jk\_tot1, se\_jk\_tot1), male = c(Var\_jk\_male1, sd\_jk\_male1, se\_jk\_male1), female = c(Var\_jk\_female1, sd\_jk\_female1,se\_jk\_female1))  
  
pander(JK\_est1)

|  |  |  |  |
| --- | --- | --- | --- |
| Jacknife | total | male | female |
| Variance | 36719296 | 52338633 | 94578729 |
| SD | 6060 | 7235 | 9725 |
| SE | 0.01159 | 0.02788 | 0.03693 |

**Q2**

#שיטת קבוצות מקריות

#total  
random\_belong\_tot\_perc <- random\_tot\_estimation / rbind(  
 round(sum(group\_1$all \* group\_1$weights \* 5)),  
 round(sum(group\_2$all \* group\_2$weights \* 5)),  
 round(sum(group\_3$all \* group\_3$weights \* 5)),  
 round(sum(group\_4$all \* group\_4$weights \* 5)),  
 round(sum(group\_5$all \* group\_5$weights \* 5)))  
  
  
Var\_random\_tot2 <- sum((random\_belong\_tot\_perc - mean(random\_belong\_tot\_perc))^2)/20  
sd\_random\_tot2 <- sqrt(Var\_random\_tot2)  
se\_random\_tot2 <- sd\_random\_tot2/belong\_tot\_perc  
  
  
#male  
random\_belong\_tot\_perc\_male <- random\_male\_estimation / rbind(  
 round(sum(group\_1$tot\_males\_over15 \* group\_1$weights \* 5)),  
 round(sum(group\_2$tot\_males\_over15 \* group\_2$weights \* 5)),  
 round(sum(group\_3$tot\_males\_over15 \* group\_3$weights \* 5)),  
 round(sum(group\_4$tot\_males\_over15 \* group\_4$weights \* 5)),  
 round(sum(group\_5$tot\_males\_over15 \* group\_5$weights \* 5)))  
  
Var\_random\_male2 <- sum((random\_belong\_tot\_perc\_male - mean(random\_belong\_tot\_perc\_male))^2)/20  
sd\_random\_male2 <- sqrt(Var\_random\_male2)  
se\_random\_male2 <- sd\_random\_male2/belong\_tot\_perc\_male  
  
  
  
#female  
random\_belong\_tot\_perc\_female <- random\_female\_estimation / rbind(  
 round(sum(group\_1$tot\_females\_over15 \* group\_1$weights \* 5)),  
 round(sum(group\_2$tot\_females\_over15 \* group\_2$weights \* 5)),  
 round(sum(group\_3$tot\_females\_over15 \* group\_3$weights \* 5)),  
 round(sum(group\_4$tot\_females\_over15 \* group\_4$weights \* 5)),  
 round(sum(group\_5$tot\_females\_over15 \* group\_5$weights \* 5)))  
  
Var\_random\_female2 <- sum((random\_belong\_tot\_perc\_female - mean(random\_belong\_tot\_perc\_female))^2)/20  
sd\_random\_female2 <- sqrt(Var\_random\_female2)  
se\_random\_female2 <- sd\_random\_female2/belong\_tot\_perc\_female  
  
  
random\_est2 <- data.frame(Random = names , Total = c(Var\_random\_tot2, sd\_random\_tot2,se\_random\_tot2), Male = c(Var\_random\_male2, sd\_random\_male2,se\_random\_male2), Female = c(Var\_random\_female2, sd\_random\_female2, se\_random\_female2))  
  
pander(random\_est2)

|  |  |  |  |
| --- | --- | --- | --- |
| Random | Total | Male | Female |
| Variance | 5.161e-05 | 3.378e-05 | 0.000242 |
| SD | 0.007184 | 0.005812 | 0.01556 |
| SE | 0.0125 | 0.009809 | 0.02788 |

#Jacknife

#total  
jk\_belong\_tot\_perc <- jk\_tot\_estimation / rbind(  
 round(sum(JK\_data1$all \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$all \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$all \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$all \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$all \* JK\_data5$weights \* 5/4)))  
  
  
Var\_jk\_tot2 <- sum((jk\_belong\_tot\_perc - mean(jk\_belong\_tot\_perc))^2)\*4/5  
sd\_jk\_tot2 <- sqrt(Var\_jk\_tot2)  
se\_jk\_tot2 <- sd\_jk\_tot2/belong\_tot\_perc  
  
  
#male  
jk\_belong\_male\_perc <- jk\_male\_estimation / rbind(  
 round(sum(JK\_data1$tot\_males\_over15 \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$tot\_males\_over15 \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$tot\_males\_over15 \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$tot\_males\_over15 \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$tot\_males\_over15 \* JK\_data5$weights \* 5/4)))  
  
  
Var\_jk\_male2 <- sum((jk\_belong\_male\_perc - mean(jk\_belong\_male\_perc))^2)\*4/5  
sd\_jk\_male2 <- sqrt(Var\_jk\_male2)  
se\_jk\_male2 <- sd\_jk\_male2/belong\_tot\_perc\_male  
  
  
#female  
jk\_belong\_female\_perc <- jk\_female\_estimation / rbind(  
 round(sum(JK\_data1$tot\_females\_over15 \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$tot\_females\_over15 \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$tot\_females\_over15 \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$tot\_females\_over15 \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$tot\_females\_over15 \* JK\_data5$weights \* 5/4)))  
  
  
Var\_jk\_female2 <- sum((jk\_belong\_female\_perc - mean(jk\_belong\_female\_perc))^2)\*4/5  
sd\_jk\_female2 <- sqrt(Var\_jk\_female2)  
se\_jk\_female2 <- sd\_jk\_female2/belong\_tot\_perc\_female  
  
JK\_est2 <- data.frame(Jacknife = names , total = c(Var\_jk\_tot2, sd\_jk\_tot2, se\_jk\_tot2), male = c(Var\_jk\_male2, sd\_jk\_male2, se\_jk\_male2), female = c(Var\_jk\_female2, sd\_jk\_female2,se\_jk\_female2))  
  
pander(JK\_est2)

|  |  |  |  |
| --- | --- | --- | --- |
| Jacknife | total | male | female |
| Variance | 5.181e-05 | 3.307e-05 | 0.0002535 |
| SD | 0.007198 | 0.005751 | 0.01592 |
| SE | 0.01253 | 0.009706 | 0.02853 |

**Q3**

#שיטת קבוצות מקריות

#total  
random\_not\_working\_tot\_perc <- rbind(  
 round(sum(group\_1$not\_working \* group\_1$weights \* 5)),  
 round(sum(group\_2$not\_working \* group\_2$weights \* 5)),  
 round(sum(group\_3$not\_working \* group\_3$weights \* 5)),  
 round(sum(group\_4$not\_working \* group\_4$weights \* 5)),  
 round(sum(group\_5$not\_working \* group\_5$weights \* 5))) / random\_tot\_estimation   
  
  
Var\_random\_tot3 <- sum((random\_not\_working\_tot\_perc - mean(random\_not\_working\_tot\_perc))^2)/20  
sd\_random\_tot3 <- sqrt(Var\_random\_tot3)  
se\_random\_tot3 <- sd\_random\_tot3/not\_working\_tot\_perc  
  
  
#male  
random\_not\_working\_male\_perc <- rbind(  
 round(sum(group\_1$not\_working\_male \* group\_1$weights \* 5)),  
 round(sum(group\_2$not\_working\_male \* group\_2$weights \* 5)),  
 round(sum(group\_3$not\_working\_male \* group\_3$weights \* 5)),  
 round(sum(group\_4$not\_working\_male \* group\_4$weights \* 5)),  
 round(sum(group\_5$not\_working\_male \* group\_5$weights \* 5))) / random\_male\_estimation   
  
  
Var\_random\_male3 <- sum((random\_not\_working\_male\_perc - mean(random\_not\_working\_male\_perc))^2)/20  
sd\_random\_male3 <- sqrt(Var\_random\_male3)  
se\_random\_male3 <- sd\_random\_male3/not\_working\_male\_perc  
  
  
#female  
random\_not\_working\_female\_perc <- rbind(  
 round(sum(group\_1$not\_working\_female \* group\_1$weights \* 5)),  
 round(sum(group\_2$not\_working\_female \* group\_2$weights \* 5)),  
 round(sum(group\_3$not\_working\_female \* group\_3$weights \* 5)),  
 round(sum(group\_4$not\_working\_female \* group\_4$weights \* 5)),  
 round(sum(group\_5$not\_working\_female \* group\_5$weights \* 5))) / random\_female\_estimation   
  
  
Var\_random\_female3 <- sum((random\_not\_working\_female\_perc - mean(random\_not\_working\_female\_perc))^2)/20  
sd\_random\_female3 <- sqrt(Var\_random\_female3)  
se\_random\_female3 <- sd\_random\_female3/not\_working\_female\_perc  
  
random\_est3 <- data.frame(Random = names , Total = c(Var\_random\_tot3, sd\_random\_tot3,se\_random\_tot3), Male = c(Var\_random\_male3, sd\_random\_male3,se\_random\_male3), Female = c(Var\_random\_female3, sd\_random\_female3, se\_random\_female3))  
  
pander(random\_est3)

|  |  |  |  |
| --- | --- | --- | --- |
| Random | Total | Male | Female |
| Variance | 5.641e-06 | 1.573e-05 | 1.278e-05 |
| SD | 0.002375 | 0.003966 | 0.003574 |
| SE | 0.05119 | 0.07407 | 0.09082 |

#Jacknife

#total  
jk\_not\_working\_tot\_perc <- rbind(  
 round(sum(JK\_data1$not\_working \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$not\_working \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$not\_working \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$not\_working \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$not\_working \* JK\_data5$weights \* 5/4))) / jk\_tot\_estimation   
  
  
Var\_jk\_tot3 <- sum((jk\_not\_working\_tot\_perc - mean(jk\_not\_working\_tot\_perc))^2)\*4/5  
sd\_jk\_tot3 <- sqrt(Var\_jk\_tot3)  
se\_jk\_tot3 <- sd\_jk\_tot3/not\_working\_tot\_perc  
  
  
#male  
jk\_not\_working\_male\_perc <- rbind(  
 round(sum(JK\_data1$not\_working\_male \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$not\_working\_male \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$not\_working\_male \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$not\_working\_male \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$not\_working\_male \* JK\_data5$weights \* 5/4))) / jk\_male\_estimation   
  
  
Var\_jk\_male3 <- sum((jk\_not\_working\_male\_perc - mean(jk\_not\_working\_male\_perc))^2)\*4/5  
sd\_jk\_male3 <- sqrt(Var\_jk\_male3)  
se\_jk\_male3 <- sd\_jk\_male3/not\_working\_male\_perc  
  
  
#female  
jk\_not\_working\_female\_perc <- rbind(  
 round(sum(JK\_data1$not\_working\_female \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$not\_working\_female \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$not\_working\_female \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$not\_working\_female \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$not\_working\_female \* JK\_data5$weights \* 5/4))) / jk\_female\_estimation   
  
  
Var\_jk\_female3 <- sum((jk\_not\_working\_female\_perc - mean(jk\_not\_working\_female\_perc))^2)\*4/5  
sd\_jk\_female3 <- sqrt(Var\_jk\_female3)  
se\_jk\_female3 <- sd\_jk\_female3/not\_working\_female\_perc  
  
jk\_est3 <- data.frame(Jacknife = names , Total = c(Var\_jk\_tot3, sd\_jk\_tot3, se\_jk\_tot3), Male = c(Var\_jk\_male3, sd\_jk\_male3,se\_jk\_male3), Female = c(Var\_jk\_female3, sd\_jk\_female3, se\_jk\_female3))  
  
pander(jk\_est3)

|  |  |  |  |
| --- | --- | --- | --- |
| Jacknife | Total | Male | Female |
| Variance | 5.512e-06 | 1.683e-05 | 1.568e-05 |
| SD | 0.002348 | 0.004102 | 0.00396 |
| SE | 0.0506 | 0.07661 | 0.1006 |

**Q4**

#שיטת קבוצות מקריות

random\_working\_avg <- rbind(  
 round(sum(group\_1$working\_hours \* group\_1$weights \* 5)/sum(group\_1$working \* group\_1$weights \* 5)),  
 round(sum(group\_2$working\_hours \* group\_2$weights \* 5)/sum(group\_2$working \* group\_2$weights \* 5)),  
 round(sum(group\_3$working\_hours \* group\_3$weights \* 5)/sum(group\_3$working \* group\_3$weights \* 5)),  
 round(sum(group\_4$working\_hours \* group\_4$weights \* 5)/sum(group\_4$working \* group\_4$weights \* 5)),  
 round(sum(group\_5$working\_hours \* group\_5$weights \* 5)/sum(group\_5$working \* group\_5$weights \* 5)))   
  
  
Var\_random4 <- sum((random\_working\_avg - mean(random\_working\_avg))^2)/20  
sd\_random4 <- sqrt(Var\_random4)  
se\_random4 <- sd\_random4/round(sum\_working\_hours/sum\_workers, 2)  
  
random\_est4 <- data.frame(Random = names , Y\_hat = c(Var\_random4, sd\_random4,se\_random4))  
  
pander(random\_est4)

|  |  |
| --- | --- |
| Random | Y\_hat |
| Variance | 0.34 |
| SD | 0.5831 |
| SE | 0.01567 |

#Jacknife

jk\_working\_avg <- rbind(  
 round(sum(JK\_data1$working\_hours \* JK\_data1$weights \* 5/4)/sum(JK\_data1$working \* JK\_data1$weights \* 5/4)),  
 round(sum(JK\_data2$working\_hours \* JK\_data2$weights \* 5/4)/sum(JK\_data2$working \* JK\_data2$weights \* 5/4)),  
 round(sum(JK\_data3$working\_hours \* JK\_data3$weights \* 5/4)/sum(JK\_data3$working \* JK\_data3$weights \* 5/4)),  
 round(sum(JK\_data4$working\_hours \* JK\_data4$weights \* 5/4)/sum(JK\_data4$working \* JK\_data4$weights \* 5/4)),  
 round(sum(JK\_data5$working\_hours \* JK\_data5$weights \* 5/4)/sum(JK\_data5$working \* JK\_data5$weights \* 5/4)))   
  
  
Var\_JK4 <- sum((jk\_working\_avg - mean(jk\_working\_avg))^2)\*4/5  
sd\_jk4 <- sqrt(Var\_JK4)  
se\_jk4 <- sd\_jk4/round(sum\_working\_hours/sum\_workers, 2)  
  
jk\_est4 <- data.frame(Jacknife = names , Y\_hat = c(Var\_JK4, sd\_jk4, se\_jk4))  
  
pander(jk\_est4)

|  |  |
| --- | --- |
| Jacknife | Y\_hat |
| Variance | 0.64 |
| SD | 0.8 |
| SE | 0.02151 |

pander(random\_est4)

|  |  |
| --- | --- |
| Random | Y\_hat |
| Variance | 0.34 |
| SD | 0.5831 |
| SE | 0.01567 |