Annand\_assignment5

Joseph Annand

2023-06-11

# Libraries —————————————

library("VennDiagram")

## Loading required package: grid

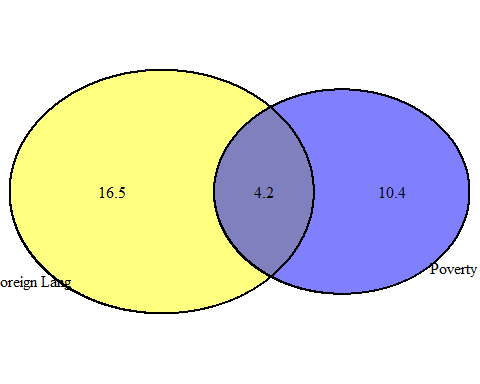
## Loading required package: futile.logger

# Question 3.8 ————————————

a.) Living under the poverty line and speaking a foreign language at home are not disjoint, as 4.2% of Americans surveyed fall into both categories.

b.)

# https://www.geeksforgeeks.org/how-to-create-a-venn-diagram-in-r/  
  
# move to new plotting page  
grid.newpage()  
  
# create pairwise Venn diagram  
draw.pairwise.venn(area1=14.6, area2=20.7, cross.area=4.2,  
 category=c("Poverty", "Foreign Lang"), fill=c("Blue", "Yellow"))



## (polygon[GRID.polygon.1], polygon[GRID.polygon.2], polygon[GRID.polygon.3], polygon[GRID.polygon.4], text[GRID.text.5], text[GRID.text.6], text[GRID.text.7], text[GRID.text.8], text[GRID.text.9])

c.)

# percent below poverty line minus percent below poverty line and speaks foreign language  
# gives percent of families below the poverty line that only speak English at home  
english\_and\_below <- 14.6 - 4.2  
english\_and\_below

## [1] 10.4

10.4% of Americans live below the poverty line and only speak English at home.

d.)

#General Addition rule to determine probability of families that speak a foreign language  
# at home or fall below the poverty line  
foreign\_or\_below <- 14.6 + 20.7 - 4.2  
foreign\_or\_below

## [1] 31.1

31.1% of Americans live below the poverty line or speak a foreign language at home.

e.)

# Families that only speak English is total families minus families that speak foreign lang at home  
english\_only <- 100 - 20.7  
# subtract percentage of English only speaking families below poverty line from  
# total English only speaking families  
english\_only - english\_and\_below

## [1] 68.9

68.9% of Americans are above the poverty line and only speak English at home.

f.)

below\_poverty <- 14.6 / 100  
foreign\_at\_home <- 20.7 / 100  
  
foreign\_and\_below <- below\_poverty \* foreign\_at\_home \* 100  
foreign\_and\_below

## [1] 3.0222

foreign\_and\_below == 4.2

## [1] FALSE

Falling below the poverty and speaking a foreign language are not independent processes because the equation for the multiplication rule for independent processes does not hold.

# Question 3.12 ———————————–

a.)

no\_absences <- 100 - 25 - 15 - 28  
no\_absences

## [1] 32

There is a 32% chance that a student picked at random did not miss a single day fo school.

b.)

no\_more\_one <- 100 - 15 - 28  
no\_more\_one

## [1] 57

There is a 57% chance that a student chosen at random misses no more than one day of school.

c.)

least\_one\_absence <- 25 + 15 + 28  
least\_one\_absence

## [1] 68

There is a 68% chance a student chosen at random misses at least one day of school.

d.)

0.32 \* 0.32

## [1] 0.1024

Assuming that the attendance of one child does not affect the attendance of the second, the attendance of each child is independent of the other. The probability that both children do not miss a day of school due to sickness is 10.24%.

e.)

0.68 \* 0.68

## [1] 0.4624

Assuming that the attendance of one child does not affect the attendance of the second, the attendance of each child is independent of the other. The probability that both children miss at least one day of school due to sickness is 46.2%.

f.) The assumption that the two children missing school due to sickness are independent events does not seem reasonable in actuality. When one child misses school due to sickness, the probability of the other child missing school due to sickness is higher than that of other elementary school students because the two children live together.

# Question 3.16 ———————————–

a.) Being in excellent health and having health coverage are not mutually exclusive. These events may happen at the same time, as their are participants surveyed who do not have health coverage and are in excellent health.

b.) The probability of a randomly selected indivisual having excellent health is 23.29%.

c.)

# probability of excellent health and coverage  
exc\_coverage <- 0.2099  
  
# probability of coverage  
coverage <- 0.8738  
  
# P(health is excellent and coverage is yes | coverage is yes)  
p <- exc\_coverage / coverage  
p

## [1] 0.2402152

The probability that a randomly chosen individual has excellent health given that he has health coverage is 24.02%.

d.)

# probability of excellent health and no coverage  
exc\_no\_coverage <- 0.0230  
  
# probability of no coverage  
no\_coverage <- 0.1262  
  
# P(health is excellent and coverage is no | coverage is no)  
p2 <- exc\_no\_coverage / no\_coverage  
p2

## [1] 0.1822504

The probability that a randomly chosen individual has excellent health given that they have no health coverage is 18.23%.

e.)

# probability having excellent health  
excellent\_health <- 0.2329  
  
   
# determine value of P(A and B) if they are independent processes  
# P(A and B) = P(excellent health) x P(having coverage)  
p\_independent <- excellent\_health \* coverage  
p\_independent

## [1] 0.203508

# compare to P(A and B) using general multiplication rule  
# P(A and B) = P(health is excellent given coverage is yes) x P(coverage is yes)  
p

## [1] 0.2402152

p == p\_independent

## [1] FALSE

Having excellent health and having health coverage do not appear to be independent.

# Question 3.22 ———————————–

data <- matrix(c(37, 44, 81, 16, 3, 19, 53, 47, 100), ncol=3)  
  
rownames(data) <- c("In Favor", "Against", "Total")  
colnames(data) <- c("College Degree", "No College Degree", "Total")  
  
data <- as.table(data)  
  
data

## College Degree No College Degree Total  
## In Favor 37 16 53  
## Against 44 3 47  
## Total 81 19 100

p4 <- 37 / 81  
p4

## [1] 0.4567901

P(in favor of Scott Walker given college degree is yes) is 45.68%.