Homework\_2

Priya Shaji

2/15/2019

## Probability

### 2.6

1. getting a sum of 1?

We assume a fair dice with values 1-6

Zero, minimum sum that can be obtained from a pair of dice is 2(1+1).

1. getting a sum of 5?

Possible combinations for argetting the sum of 5 are (1,4), (2,3), (3,2), (4,1).

So probability of getting a sum of 5 is

4/36 = 1/9

1. getting a sum of 12?

Possible combinations for argetting the sum of 12 are (6,6)

So probability of getting a sum of 12 is =1/36

### 2.8

1. Are living below the poverty line and speaking a foreign language at home disjoint?

No. As according to the statistics 4.2% fall into both living below the proverty line and speaking a foreign language.

1. Draw a Venn diagram summarizing the variables and their associated probabilities.

The following venn diagram summarizze the variables and their associated probabilities:

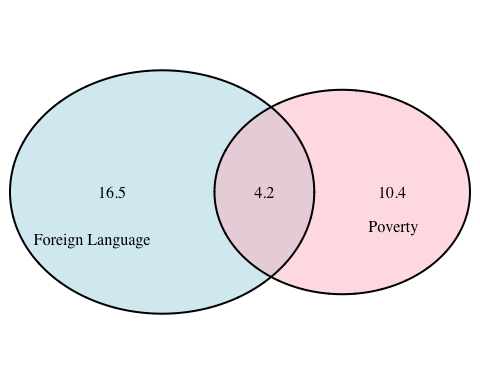
library(VennDiagram)

## Loading required package: grid

## Loading required package: futile.logger

poverty <- 14.6  
foreign\_Language <- 20.7  
both <- 4.2  
povertyOnly <- poverty - both  
foreign\_Language\_Only <- foreign\_Language - both  
venn.plot <- draw.pairwise.venn(poverty,   
 foreign\_Language,  
 cross.area=both,   
 c("Poverty", "Foreign Language"),   
 fill=c("pink", "lightblue"),  
 cat.dist=-0.08,  
 ind=FALSE)

grid.draw(venn.plot)



1. What percent of Americans live below the poverty line and only speak English at home?

14.6-4.2

## [1] 10.4

1. What percent of Americans live below the poverty line or speak a foreign language at home?

14.6+20.7-4.2

## [1] 31.1

1. What percent of Americans live above the poverty line and only speak English at home?

100-(14.6+20.7-4.2)

## [1] 68.9

1. Is the event that someone lives below the poverty line independent of the event that the person speaks a foreign language at home?

The event that someone lives below the poverty line is not independent of the event that the person speaks a foreign language at home.

### 2.20

1. What is the probability that a randomly chosen male respondent or his partner has blue eyes?

(114+19+11)/204

## [1] 0.7058824

1. What is the probability that a randomly chosen male respondent with blue eyes has a partner with blue eyes?

(78/114)

## [1] 0.6842105

1. Whatistheprobabilitythatarandomlychosenmalerespondentwithbrowneyeshasapartner with blue eyes? What about the probability of a randomly chosen male respondent with green eyes having a partner with blue eyes?

(19/54)

## [1] 0.3518519

1. Does it appear that the eye colors of male respondents and their partners are independent? Explain your reasoning.

The eye colors of male respondents and their partners are NOT independent.

P(blue male | blue female) = P(blue male) \* P(blue female) (78/204) = (114/204) \* (108/204).The values are not equal Therefore, eye colors of male respondents and their partners are NOT independent.

### 2.30

1. Find the probability of drawing a hardcover book first then a paperback fiction book second when drawing without replacement.

(28/95) \* (59/94)

## [1] 0.1849944

1. Determine the probability of drawing a fiction book first and then a hardcover book second, when drawing without replacement.

(72/95) \* (28/94)

## [1] 0.2257559

1. Calculate the probability of the scenario in part (b), except this time complete the calculations under the scenario where the first book is placed back on the bookcase before randomly drawing the second book.

(72/95) \* (28/95)

## [1] 0.2233795

1. The final answers to parts (b) and (c) are very similar. Explain why this is the case.

The final answers to parts (b) and (c) are very similar because the possible events are considerable large so the outcome will not be affected much.

### 2.38

1. Build a probability model, compute the average revenue per passenger, and compute the corresponding standard deviation.

bag\_piece <- c(0, 1, 2)  
bag0 <- 0  
bag1 <- 25  
bag2 <- bag1 + 10  
bag\_fee <- c(bag0, bag1, bag2)  
bag\_percent <- c(0.54, 0.34, 0.12)  
bag\_revenue\_per\_person <- sum(bag\_fee \* bag\_percent)  
bag\_revenue\_per\_person

## [1] 12.7

variance <- (25^2\*.34 + 35^2\*.12) - 12.7^2  
variance

## [1] 198.21

standart\_deviation <- sqrt(variance)  
standart\_deviation

## [1] 14.07871

Revenue per person is $12.7, Variance is 198.21 and the standard deviation is 14.07871

1. About how much revenue should the airline expect for a flight of 120 passengers? With what standard deviation? Note any assumptions you make and if you think they are justified.

revenue\_120 <- 12.7 \* 120  
revenue\_120

## [1] 1524

standart\_deviation\_120 <- sqrt(variance \* 120)  
standart\_deviation\_120

## [1] 154.2245

The expected revenue for the 120 passengers is $1524. The standard deviation will be $154.2245

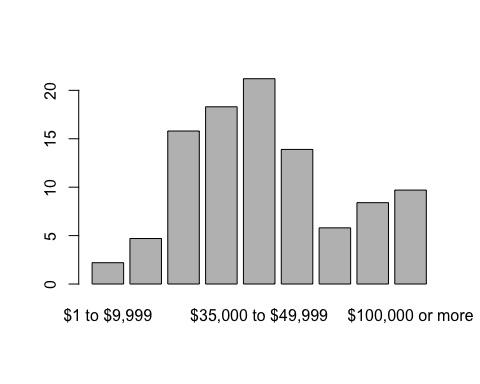
### 2.44

1. Describe the distribution of total personal income.

income <- c("$1 to $9,999","$10,000 to $14,999","$15,000 to $24,999","$25,000 to $34,999","$35,000 to $49,999","$50,000 to $64,999","$65,000 to $74,999","$75,000 to $99,999","$100,000 or more")  
  
total <- c(2.2,4.7,15.8,18.3,21.2,13.9,5.8,8.4,9.7)  
  
distribution <- data.frame(income, total)  
distribution

## income total  
## 1 $1 to $9,999 2.2  
## 2 $10,000 to $14,999 4.7  
## 3 $15,000 to $24,999 15.8  
## 4 $25,000 to $34,999 18.3  
## 5 $35,000 to $49,999 21.2  
## 6 $50,000 to $64,999 13.9  
## 7 $65,000 to $74,999 5.8  
## 8 $75,000 to $99,999 8.4  
## 9 $100,000 or more 9.7

barplot(distribution$total, names.arg=income)



It seems normalcontinuous distribution.

1. What is the probability that a randomly chosen US resident makes less than $50,000 per year?

sum(distribution[1:5,2]/100)

## [1] 0.622

1. What is the probability that a randomly chosen US resident makes less than $50,000 per year and is female? Note any assumptions you make.

0.622 \* 0.41

## [1] 0.25502

Sample consists of 41% females. The probability that a randomly chosen US resident makes less than $50,000 per year and is female is 0.25502

Assumption: The probability of an income of less than $50,000 and being female are independent events.

1. The same data source indicates that 71.8% of females make less than $50,000 per year. Use this value to determine whether or not the assumption you made in part (c) is valid. Income Total

If 71.8% of females make less than $50,000 per year the equation will be 0.718 = 0.622 \* 0.41, which is not correct, so we can conclude that the assumption we made for step (c) is wrong. with the equation it looks like events are not independent.