606 HW2

Ahsanul Choudhury

September 11, 2016

2.6 Dice rolls. If you roll a pair of fair dice, what is the probability of

(a) getting a sum of 1?

Answer (a)Assuming the pair of dice are fair six sided dice the probablity of getting sum of on is 0 since the minimum possible sum is 2. It is a disjoint probability, we can not roll 2 dice and get sum of 1.

(b) getting a sum of 5?

Answer (b)The number of ways we can get 5 with two dice is 4 (formula (N+R-1)!/((R-1)!*(N)!)), therefore probability of getting sum of 5 is 4/36 = 0.11.

(c) getting a sum of 12?

Answer (c)The number of ways we can get 12 with two dice is 1 (formula (N+R-1)!/((R-1)!*(N)!)), therefore probability of getting sum of 12 is 1/36 =0.028

2.8 Poverty and language. The American Community Survey is an ongoing survey that provides data every year to give communities the current information they need to plan investments and services. The 2010 American Community Survey estimates that 14.6% of Americans live below the poverty line, 20.7% speak a language other than English (foreign language) at home, and 4.2% fall into both categories.

(a) Are living below the poverty line and speaking a foreign language at home disjoint?

Answer(a)No, they are not disjoint, a person can both live under poverty line anc can speak a foreign language at home, 4.2% fall into both catagories.

(b) Draw a Venn diagram summarizing the variables and their associated probabilities.

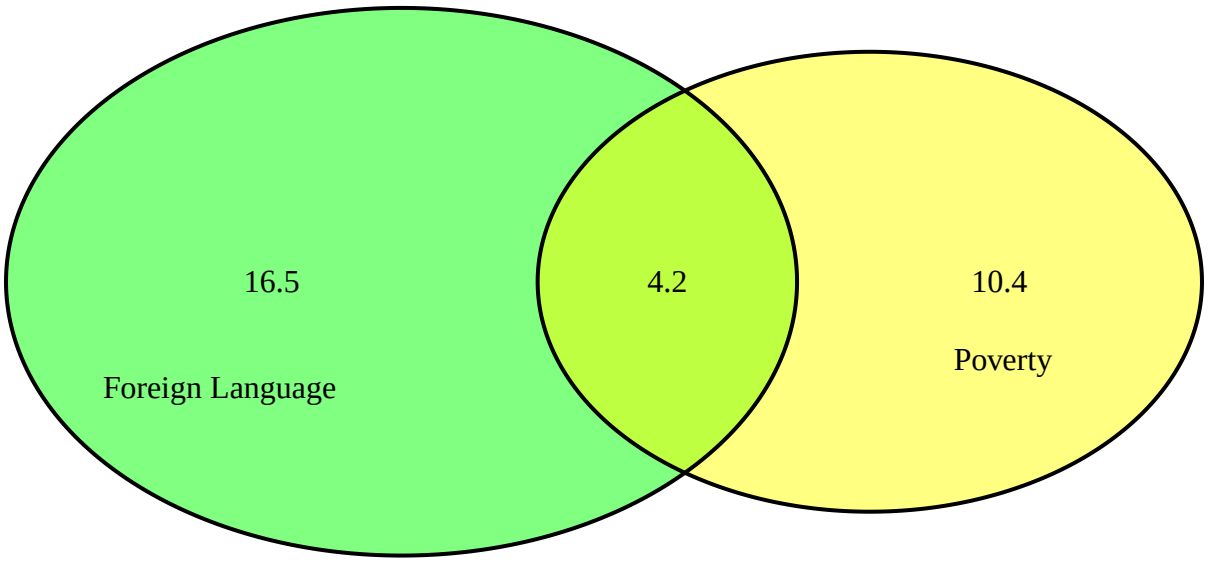
```
library(VennDiagram)
```

```
## Loading required package: grid
```

```
## Loading required package: futile.logger
```

```
pline <- 14.6
fxlang <- 20.7
both <- 4.2
```

```
venn.plot <- draw.pairwise.venn(pline,
                    fxlang,
                    cross.area=both,
                    c("Poverty", "Foreign Language"),
                    fill=c("yellow", "green"),
                    cat.dist=-0.08,
                    ind=FALSE)
grid.draw(venn.plot)
```



(c) What percent of Americans live below the poverty line and only speak English at home?

Answer(c)

```
engonly <- pline - both
engonly
```

```
## [1] 10.4
```

10.4% of Americans live below the peverty line and only speak English.

(d) What percent of Americans live below the poverty line or speak a foreign language at home?

Answer(d)

```
porfx <- fxlang + pline - both
porfx
```

```
## [1] 31.1
```

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

(e) What percent of Americans live above the poverty line and only speak English at home?

Answer(e)

```
apne <- 100 - fxlang - (pline - both)
apne
```

```
## [1] 68.9
```

68.9% of Americans live above poverty line and only speak English.

(f) Is the event that someone lives below the poverty line independent of the event that the person speaks a foreign language at home?

Answer(f)Events are independent if it satisfy the multiplication rule, if, P(A and B) = P(A)*P(B). In our case, P(A and B) = 0.207* 0.146 = 0.030222 is not equal P(A)*P(B) = 0.042. The events are not independent.

2.20 Assortative mating. Assortative mating is a nonrandom mating pattern where individuals with similar genotypes and/or phenotypes mate with one another more frequently than what would be expected under a random mating pattern. Researchers studying this topic collected data on eye colors of 204 Scandinavian men and their female partners. The table below summarizes the results. For simplicity, we only include heterosexual relationships in this exercise.

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

Answer(a)In the study there are 204 male with female partner, 114 male with blue eye, 108 female with blue eyes and in 78 case both male and female have blue eyes, therefore the probability of male with blue eyes or partner blue eyes is:

114/204 + 108/204 - 78/204 = 0.706

(b) What is the probability that a randomly chosen male respondent with blue eyes has a partner with blue eyes?

Answer(a)We are lookin for blue eyed male with blue eyed female partner,

78/204 = 0.382

(c) What is the probability that a randomly chosen male respondent with brown eyes has a partner with blue eyes? What about the probability of a randomly chosen male respondent with green eyes having a partner with blue eyes?

Answer(c)Porbability brown eyed male with blue eyed partner = 19/54 = 0.352

Porbability green eyed male with blue eyed partner = 11/36 = 0.306

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

From the table it does not appear that the eye colors of male respondents and their partners are independent, the prportion of male with same color eyes as their partner's appears to be higher than male with different eye color then theie partner's.

2.30 Books on a bookshelf. The table below shows the distribution of books on a bookcase based on whether they are nonfiction or fiction and hardcover or paperback.

(a) Find the probability of drawing a hardcover book first then a paperback

```r
chkb <- c(0, 1, 2)
prob <- c(0.54, 0.34, 0.12)
fees <- c(0, 25, 25+35)
df <- data.frame(chkb, prob, fees)
df$weightedrev <- df$prob * df$fees
df
```

```
##   chkb prob fees weightedrev
## 1    0 0.54    0         0.0
## 2    1 0.34   25         8.5
## 3    2 0.12   60         7.2
```

```r
##Average revenue per passenger
avgrevpp <- sum(df$weightedrev)
avgrevpp
```

```
## [1] 15.7
```

```r
##Standard deviation
sdiv1 <- sd(df$weightedrev)
sdiv1
```

```
## [1] 4.578573
```

(b) 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.

Answer(b)

```r
##Expected revenue for a flight of 120 passengers
pas <- 120
rev <- pas * avgrevpp
rev
```

```
## [1] 1884
```

```r
##Standard deviation
sdiv2 <- pas * sdiv1
sdiv2
```

```
## [1] 549.4288
```

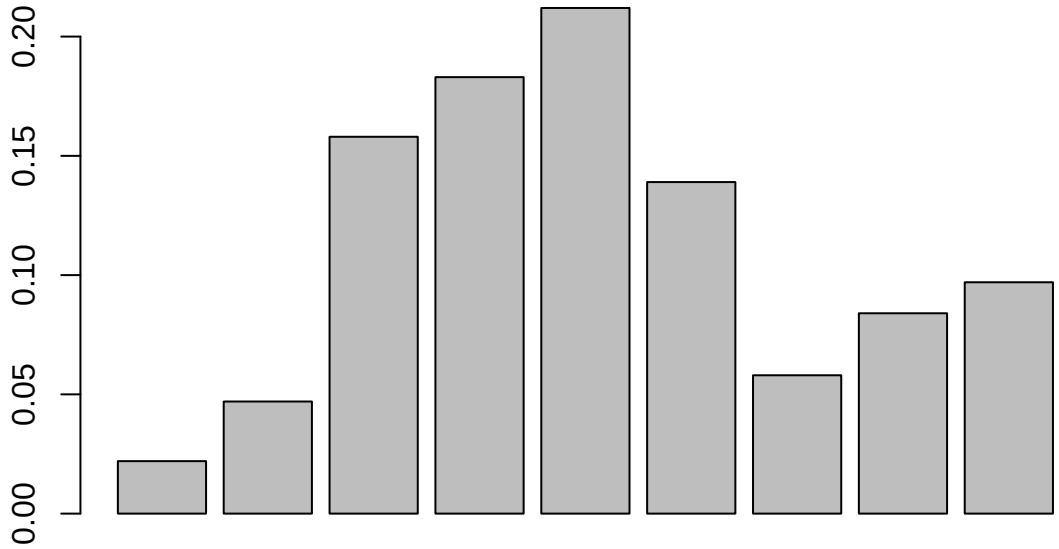
Assumption: There is no other random variable

2.44 Income and gender. The relative frequency table below displays the distribution of annual total personal income (in 2009 inflation-adjusted dollars) for a representative sample of 96,420,486 Americans. These data come from the American Community Survey for 2005-2009. This sample is comprised of 59% males and 41% females.

(a) Describe the distribution of total personal income.

Answer(a)

```
prob1 <- c(0.022, 0.047, 0.158, 0.183, 0.212, 0.139, 0.058, 0.084, 0.097)
barplot(prob1)
```



Distribution is bimodal peaks at $35,000 to $49,000 and again at $100,000 or more, also shows a slight skewness to the right.

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

Probability that a randomly chosen US resident makes less than $50,000 per year:

(0.022 + 0.047 + 0.158 + 0.183 + 0.212) = 0.622

Answer(b)

(c) What is the probability that a randomly chosen USresident makes less than $50,000 per year and is female? Note any assumptions you make.

Answer(c)Probability that a randomly chosen USresident makes less than $50,000 per year and is female:

0.622 * 0.41 = 0.25502

Assumption: Variable are independent

(d) 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.

Answer(d)0.718 * 0.41 = 0.29438, which is close not equal to 0.25502, refering to the rule of multiplication we can assume being female and making less then $50,000 are not independent event.