Homework Assignment #1

Arun Krishnaraj, eid:ak37738

2020-09-10

Problems 1-5

```
792/11
## [1] 72
sqrt(pi)
## [1] 1.772454
exp(2)
## [1] 7.389056
256^(1/4)
## [1] 4
log(2)
## [1] 0.6931472
```

Problem 6

```
a <- 3
b <- 7
a*b
## [1] 21
```

Problem 7

```
v <- -2:4
w <- seq(from=1,length.out = 7)
v*w
## [1] -2 -2  0  4 10 18 28
```

`v*w` returns a vector composed of products of each element of `v` multiplied by the corresponding element in `w`.

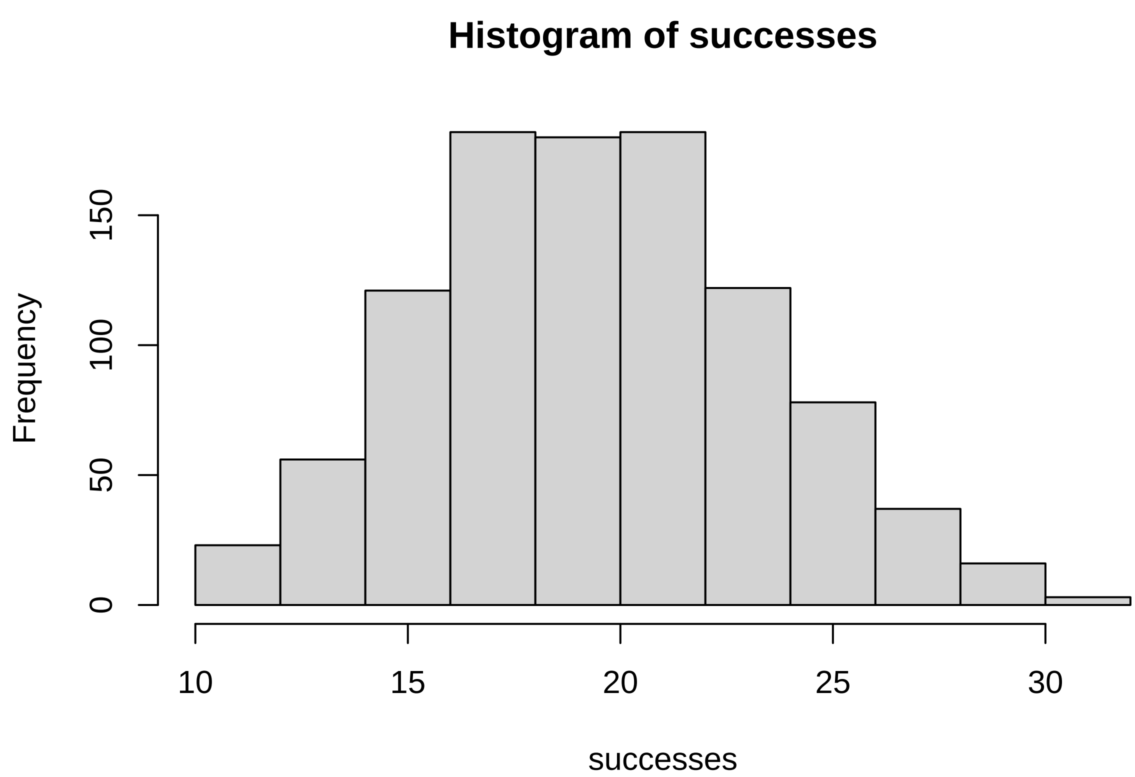
Problem 8

```
successes <- rbinom(100,1,.2)
successes
##   [1] 0 0 0 0 0 0 0 0 0 0 1 0 1 1 0 1 1 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0
##  [38] 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0
##  [75] 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 1 0 1 0 0 0
sum(successes)/100
## [1] 0.2
```

The simulated batch of 100 trials has a proportion of success of .2.

Problem 9

```r
successes <- rbinom(1000,100,.2)
hist(successes)
```



Problem 10

For fair sided dice, the probability of any of the numbers 1 to 6 being rolled is $\frac{1}{6}$. We can consider the two dice to be independent, so the the probability of rolling a 6 on both is $\frac{1}{6}^2$; since the probability of rolling any number on a fair die is the same, the probability of rolling a 3 on both dice is also $\frac{1}{6}^2$. The second roll of dice is independent of the first, making the probability of rolling pairs of 6's twice in a row $\frac{1}{6}^4$; this is true for rolling pairs of 3's twice in a row as well, so both observed roll patterns have $\frac{1}{1296}$ probability, and are just as probable.

Problem 11

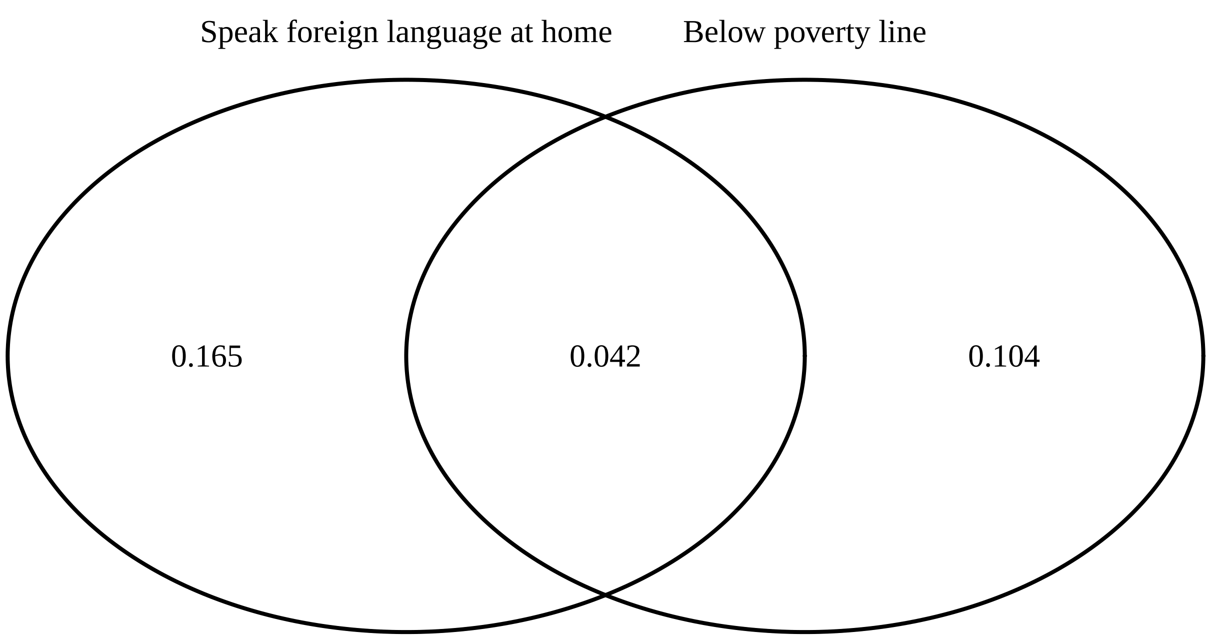
a. 0, since the lowest outcome on each die is a roll of 1; the sum of two dice can never be 1.
b. A sum of 5 can be composed of the following ordered roll pairs $(1,4),(2,3),(3,2),(4,1)$. Each of these pairs occur with probability $\frac{1}{36}$, making the total probability of observing a sum of 5 $\frac{4}{36} = \frac{1}{9}$.
c. A sum of 12 can only occur if both dice result in 6; this probability is thus $\frac{1}{6}^2 = \frac{1}{36}$.

Problem 12

a. No, since there are individuals who are in both categories; this makes the categories nondisjoint by definition.

b.

```r
library(VennDiagram)
## Loading required package: grid
## Loading required package: futile.logger
grid.newpage()
venn.plot <- draw.pairwise.venn(area1      = .146,
                                area2      = .207,
                                cross.area = .042,
                                category   = c("Below poverty line", "Speak foreign language at home"),
                                cat.pos = 0,
                                cat.dist = .035,
                                scaled     = FALSE)
```



c. 10.4% of Americans live below the poverty line and only speak English at home.
d. 31.1% of Americans live below the poverty line or speak a foreign language at home.
e. 68.9% of Americans live above the poverty line and only speak English at home.
f. No; the probability that a person lives below the poverty line is 14.6% as provided. Given that a person speaks a foreign language at home, they have a 0.042/0.165 = 25.45% probability of living below the poverty line. Since P[Living below poverty line] $\neq$ P[Living Below Poverty Line | Speaking Foreign language at home], the two events are not independent.

Problem 13

We can apply Bayes' rule to solve.

$$\text{P[Condition}|+] = \frac{\text{P}[+|\text{Condition}] * \text{P[Condition]}}{\text{P}[+|\text{Condition}] * \text{P[Condition]} + \text{P}[+|\text{No Condition}] * \text{P[No Condition]}}$$

Given the information on proportion of people who have the condition and test accuracy for people with and without the condition, we obtain

$$\text{P[Condition}|+] = \frac{(.99) * (.03)}{(.99) * (.03) + (1 - .98) * (1 - .03)} = \frac{(.0297)}{(.0297) + (.0194)} = 0.605$$

So given that a randomly selected person tests positive for the predisposition, they have a 60.48% probability of actually having the predisposition.

Problem 14

We can again apply Bayes' rule to solve.

$$\text{P[In Favor}|\text{Degree]} = \frac{\text{P[Degree}|\text{In Favor]} * \text{P[In Favor]}}{\text{P[Degree}|\text{In Favor]} * \text{P[In Favor]} + \text{P[Degree}|\text{Against]} * \text{P[Against]}}$$

Using the provided information,

$$\text{P[In Favor}|\text{Degree]} = \frac{(.37) * (.53)}{(.37) * (.53) + (.44) * (1 - .53)} = \frac{(.1961)}{(.1961) + (.2068)} = .487$$

So given that a random exit poll participant had a college degree, there is a 48.67% probability that they voted in favor of Scott Walker.