INSH5301 Intro Computational Statistics

Ali Banijamali

01/26/2020

1.a. What's the chance of getting a sequential pair on two rolls of a die (eg, a 3 then a 4 counts, but a 4 then a 3 does not). (Hint: you can calculate this manually if you like, by counting up the sample space and finding the fraction of that sample space that consists of ordered pairs.)

ANS.
The sample space would be: {(1,2), (2,3), (3,4), (4,5), (5,6)} -> 5

Number of all possible outcomes are: 6 * 6 = 36

Therefore the answer is: P = $\frac{5}{36}$

1.b. Given a dartboard with a inner circle that is 2/3 of the total area, and a bulls-eye that is 5% of the total area (and entirely within the inner circle): if you are throwing a random dart (that is guaranteed to hit somewhere on the board, but everywhere inside is equally likely), what is the chance of hitting the bulls-eye conditional on knowing your dart is somewhere inside the innner circle?

ANS.
This problem must be solved using Bayesian probabilities:
A: Inner Circle
B: Bulls-Eye
P(A) = $\frac{2}{3}$
P(B) = $\frac{5}{100}$
P(A|B) = 1
P(B|A) = $\frac{P(A)*P(A|B)}{P(B)} = \frac{0.05}{\frac{2}{3}} = 0.033$

1.c. You take a test for a scary disease, and get a positive result. The disease is quite rare − 1 in 1000 in the general population. The test has a sensitivity of 95%, and a false positive rate of only 5%. What is the chance you have the disease?

ANS.
D: Disease
H: Healthy
P: Positive Result
N: Negative Result
We are looking for the probability of P(D|P):
P(D|P) = $\frac{P(D)*P(P|D)}{P(P)}$

P(D) = 0.001
P(P|D) = 0.95
P(P) = P(D)P(P|D) + P(H)P(P|H)
P(H) = 1-0.001 = 0.999
P(P|H) = 0.05
=> P(P) = (0.001 * 0.95) + (0.999 * 0.05) = 0.050

=> $P(D|P) = \frac{0.001*0.95}{0.050} = 0.019$

1.d. What is the chance you have the disease if everything remains the same, but the disease is even rarer, 1 in 10,000?

ANS.
$P(D) = 0.0001$
$P(P|D) = 0.95$
$P(P) = P(D)P(P|D) + P(H)P(P|H)$
$P(H) = 1\text{-}0.0001 = 0.9999$
$P(P|H) = 0.05$
=> $P(P) = (0.0001 * 0.95) + (0.9999 * 0.05) = 0.050$

=> $P(D|P) = \frac{0.001*0.95}{0.050} = 0.019$

$P(D|P) = \frac{0.0001*0.95}{0.051} = 0.0019$

1.e. What does this tell you about the dangers of tests for rare diseases?

ANS.
It shows that when the disease is very rare, We don'y get a reliable probability for the test. The probabilities barely changed.

2.a. You have a 20-side die. Using sample, roll it 1000 times and count the number of rolls that are 10 or less.

```
die <- c(1:20) # 20 side die
rolls <- sample(die, 1000, replace = T) # Roll 1000 times, with replacement
length(rolls[rolls<=10])
```

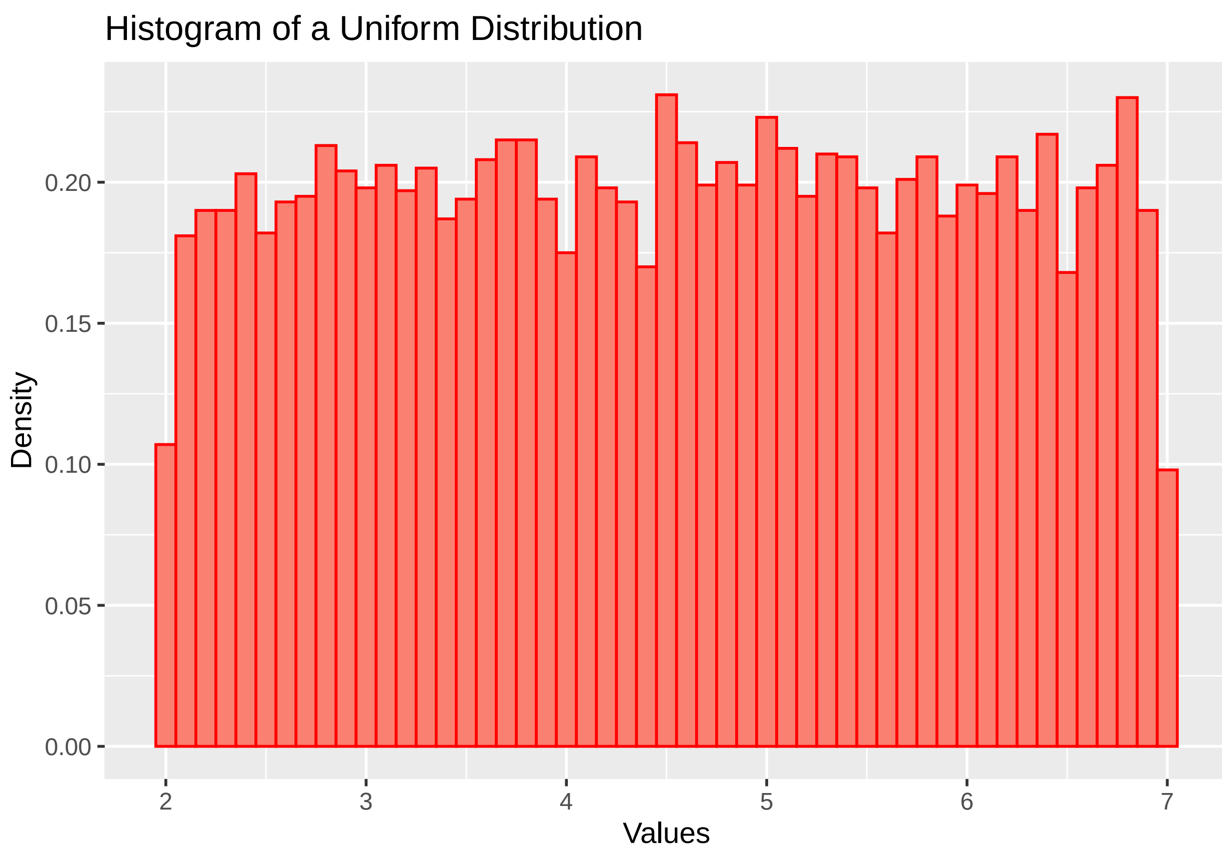
```
## [1] 482
```

2.b. Generate a histogram using ggplot of 10,000 draws from a uniform distribution between 2 & 7.

```
dist <- runif(10000, min=2, max=7)

# Plot:
library(ggplot2)

ggplot(data=as.data.frame(dist), aes(x=dist)) +
  geom_histogram(binwidth=0.1, aes(y=..density..), color="red", fill="salmon") +
  xlab("Values") + ylab("Density") + ggtitle('Histogram of a Uniform Distribution')
```



2.c. Try to write down the equation for this probability density function.

ANS.
The unifrom distribution is the probability distribution of a random number selection from the continuous interval betwee a and b. Its density function is defined by:

f(x) =
$\frac{1}{b-a}$; Where a $\leq$ x $\leq$ b
0 ; Everywhere Else

2.d. What is the probability that a draw from this distribution will be between 1.5 and 3.2?

```
length(dist[dist>1.5 & dist<3.2])/10000
```

```
## [1] 0.2364
```

3.a. Using R's cdf for the binomial, what is the probability of getting 500 or fewer "20"s when rolling your 20-sided die 10,000 times. Looking back at 2a, what proportion of your rolls were actually 20s?

```
p.less.than.500 <- pbinom(500, 10000, 0.05)
p.less.than.500
```

```
## [1] 0.511895
```

```r
# For question 2, we rolled the die 1000 times:
length(rolls[rolls==20])/1000
```

```
## [1] 0.059
```

3.b. Using rbinom, roll a 100-sided die 100 times and report the total number of 7s you get.

```r
roll.100sided <- rbinom(1, 100, 1/100)
# 1/100 Probability of getting a 7 from a 100 sided die
number.of.sevens <- sum(roll.100sided)
number.of.sevens
```

```
## [1] 0
```

3.c. You are a klutz, and the average number of times you drop your pencil in a day is 1. Using the poisson functions in R, what's the chance of dropping your pencil two or more times in a day? (Hint: calculate the chance of dropping it one or fewer times, and then take 1 minus that.)

```r
p.dropping.two.or.more <- 1-ppois(1,1)
p.dropping.two.or.more
```

```
## [1] 0.2642411
```

3.d. Because he is lazy, your teacher has assigned grades for an exam at random, and to help hide his deception he has given the fake grades a normal distribution with a mean of 70 and a standard deviation of 10. What is the chance your exam got a score of 85 or above? What is the chance you got a score between 50 and 60?

```r
# pnorm gives the probability of -inf to q:
p.above.85 <- 1 - pnorm(q=85, mean=70, sd=10)
p.above.85
```

```
## [1] 0.0668072
```

```r
p.50.to.60 <- pnorm(60, mean=70, sd=10) - pnorm(50, mean=70, sd=10)
p.50.to.60
```

```
## [1] 0.1359051
```