IntroComp_WeixuanChen_HW3

Weixuan Chen

1/29/2023

1

a

Count manually: The total sample space has 36 outcomes, because for each dice we have 6 choices and we have 6*6 = 36 combinations. For the sequential pairs, we have (1,2), (2,3), (3,4), (4,5), (5,6). So we have 5/36 which is the probability of getting a sequential pair on two rolls.

Probabilistic: Since roll 6 does not have corresponding sequential pair, we have only 1 to 5 on the first roll resulting the probability 5/6. For the second roll, the probability of getting the sequential number of the first roll is 1/6 because only 1 number in 1 to 6 satisfies. Therefore, the total probability is 5/6 * 1/6 = 5/36.

b

By the bayes rules, p(hitting the bull-eyes| dart inside the inner circle) = p(hitting the bull-eyes, dart inside the inner circle)/p(dart inside the inner circle). The probability of intersection of the events hitting the bull-eyes and dart inside the inner circle is p(hitting the bull-eyes) = 0.05 because it is entirely within the inner circle. p(dart inside the inner circle) = 2/3. So we have p(hitting the bull-eyes| dart inside the inner circle) = 1/20 * 3/2 = 3/40.

c

By bayes theorem, p(have disease|test positive) = p(test positive|have disease)*p(have disease) / p(test positive) = 0.95* 1/1000 / p(test positive).By the law of total probability, p(test positive) = p(test positive| have disease)*p(have disease) + p(test positive | do not have disease)p(do not have disease) = 0.95* 1/1000 + 0.05 * 999/1000 = 0.0509. Combining all the result, p(have disease|test positive) = 0.95 * 1/1000 / 0.0509 = 0.01866.

d

(0.95 * 1/10000) / (0.95 * 1/10000 + 0.05 * 999/10000) = 0.00189. This shows that the probability gets smaller.

e

It is obvious that for some rare disease, the probability of actually getting the disease is low even you test positive. However, if doctors or patients have medications based the test result, it would cost unnecessary resources.

2

a

```
my_dice <- sample(1:20,1000,replace=T)
sum(my_dice <= 10)
```

```
## [1] 495
```

```
library(tidyverse)
```

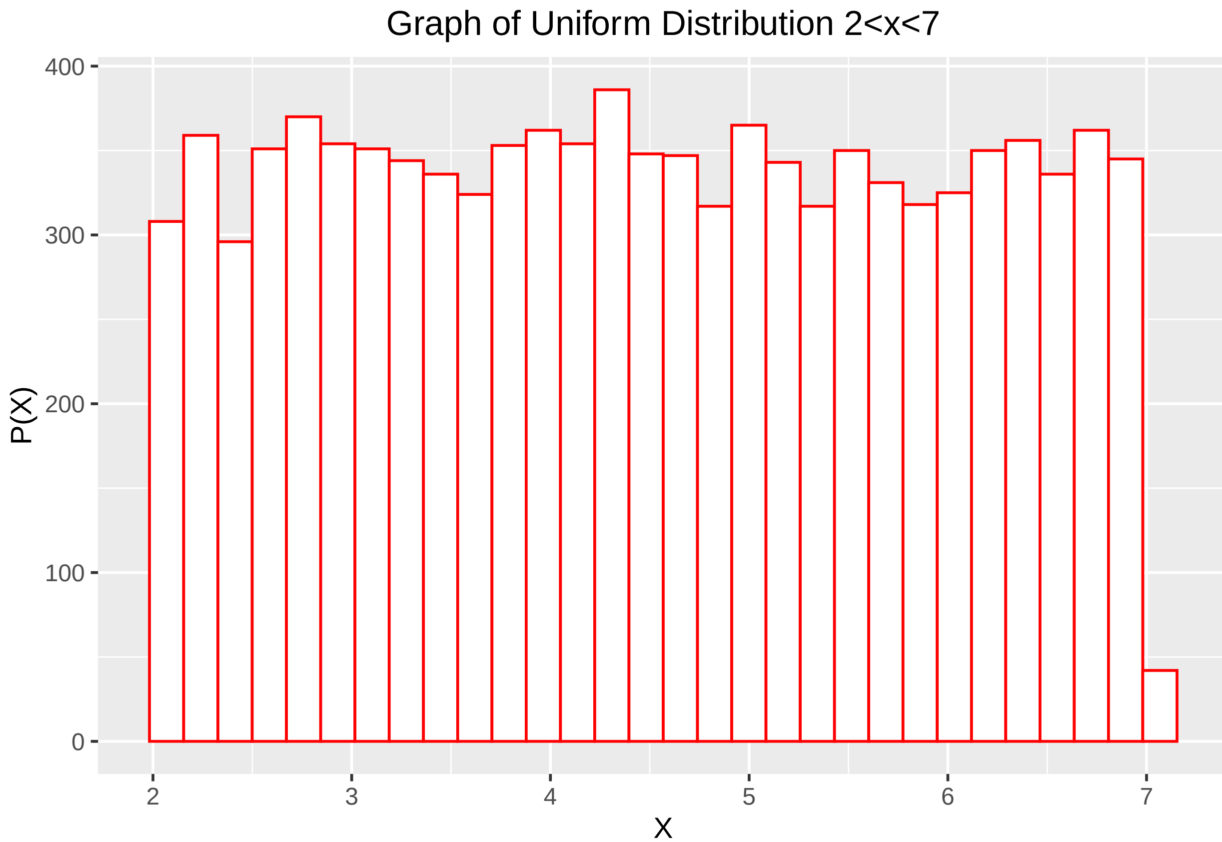
```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   1.0.1
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.3.0      v stringr 1.5.0
## v readr   2.1.3      v forcats 1.0.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

b

```
unif.hist <- as.data.frame(runif(10000,min=2,max=7))
ggplot(data=unif.hist,mapping = aes(x=unif.hist$`runif(10000, min = 2, max = 7)`))+
  geom_histogram(fill='white',color='red')+
  xlab("X")+
  ylab("P(X)")+
  ggtitle("Graph of Uniform Distribution 2<x<7")+
  theme(plot.title = element_text(hjust = 0.5))
```

```
## Warning: Use of `` unif.hist$`runif(10000, min = 2, max = 7)` `` is discouraged.
## i Use `runif(10000, min = 2, max = 7)` instead.
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



c

$$f(x) = \begin{cases} 1/b - a & \text{w.p. } 2 \leqslant x \leqslant 7 \\ 0 & \text{w.p. } x < 2 or x > 7 \end{cases} \tag{1}$$

d

```
sum(unif.hist>1.5&unif.hist<3.2)/10000
```

```
## [1] 0.2412
```

```
punif(3.2,min=2,max=7)-punif(1.5,min=2,max=7)
```

```
## [1] 0.24
```

3

a

```r
pbinom(500,size=10000,prob=(1/20),lower.tail=T)
```

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

```r
sum(my_dice == 20)
```

```
## [1] 45
```

42 rolls are actually 20 for 1000 rolls.

b

```r
hdie <- rbinom(1:100,100,1/100)
sum(hdie == 7)
```

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

c

```r
ppois(1,1,lower.tail=F)
```

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

```r
1-ppois(1,1)
```

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

d

```r
# an 85 or above
pnorm(85,mean=70,sd=10,lower.tail = F)
```

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

```r
# between 50 & 60
pnorm(60,mean=70,sd=10,lower.tail = T)-pnorm(50,mean=70,sd=10,lower.tail = T)
```

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