Verzani Problem Set

Next are considered the problems from Verzani's book on page 47.

Problem 6.1

Generate 10 random numbers from a uniform distribution on [0,10].

$$X \in Unif(0, 10)$$

```
> x <- runif(10, min = 0, max = 10); x
 [1] 5.1961457 4.8886519 4.5406558 3.4014172 9.8342093
5.1126519 4.8233336
 [8] 0.1848366 9.7662502 5.3869546
```

Use R to find the maximum and minimum values.

```
> min(x)
[1] 0.1848366
> max(x)
[1] 9.834209
```

Problem 6.2

Generate 10 random normal numbers with mean 5 and standard deviation 5 /normal(5,5)/.

$$X \in N(5, 5^2)$$

```
> x <- rnorm(10, mean = 5, sd = 5); x
 [1] 11.3526648  2.5924390 -0.1936728  5.1189532
-0.2718476  5.4082892
 [7]  8.1593152  8.7738163 -0.9402545  4.6674340
```

How many are less than 0?

```
> sum(x < 0)
[1] 3
```

Estimate the probability to observe value less than 0. Compute it theoretically.

```
> sum(x < 0) / length(x)
[1] 0.3
> pnorm(0, mean = 5, sd = 5)
[1] 0.1586553
```

Problem 6.3

Generate 100 random normal numbers with mean 100 and standard deviation 10.

$$X \in N(100, 10^2)$$

```
> x <- rnorm(100, 100, 10)
```

How many are more than 2 standard deviations from the mean (smaller than 80 or bigger than 120)?

```
> sum(x < 80 | x > 120)
[1] 2
```

Calculate the mean and the standard deviation from the sample and use them to calculate how many are more than 2 standard deviations from the mean.

```
> mean(x)
[1] 100.3669
> sd(x)
[1] 9.094089
> sum(abs(x - mean(x)) / sd(x) > 2)
[1] 6
```

Problem 6.4

Toss a fair coin 50 times (using R). How many heads do you have?

```
> coin <- sample(c("Head","Tail"), size = 50, replace =
TRUE)
> sum(coin == "Head")
```

```
[1] 26
```

Or use the binomial distribution

$Bi(50, 0.5)$

directly.

```
> rbinom(1, 50, 0.5)
[1] 27
```

Problem 6.5

Roll a "die" 100 times. How many 6's did you see?

```
> die <- sample(1:6, size = 100, replace = TRUE)
> sum(die == 6)
[1] 21
```

Or use the binomial distribution $Bi\left(100, \frac{1}{6}\right)$ directly.

```
> rbinom(1, 100, 1/6)
[1] 16
```

Problem 6.6

Select 6 numbers from a lottery containing 49 balls. What is the largest number? What is the smallest?

```
> lottery <- sample(1:49, 6, replace = FALSE); lottery
[1] 18 21 45 48 38 34
> min(lottery)
[1] 18
> max(lottery)
[1] 48
```

How many even numbers do you have? Estimate the probability to have an even number.

```
> sum(lottery %% 2 == 0)
[1] 4
> sum(lottery %% 2 == 0) / length(lottery)
```

```
[1] 0.6666667
```

Or use the hypergeometric distribution $HG(24,25,6)$ directly.

```
> rhyper(1, 24, 25, 6)
[1] 0
```

Compute the theoretical probability to have 2 even numbers?

```
> dhyper(2, 24, 25, 6)
[1] 0.2496743
```

Problem 6.7

For $Z \in N(0,1^2)$, find a number z^* solving $\mathbb{P}(Z \leq z^*) = 0.05$ (use `qnorm`).

```
> qnorm(p = 0.05, mean = 0, sd = 1)
[1] -1.644854
```

Problem 6.8

For $Z \in N(0,1^2)$, find a number z^* solving $\mathbb{P}(-z^* \leq Z \leq z^*) = 0.05$ (use `qnorm` and symmetry).

```
> qnorm(p = 0.05 / 2, mean = 0, sd = 1)
[1] -1.959964
> qnorm(p = 1 - 0.05 / 2, mean = 0, sd = 1)
[1] 1.959964
```

Problem 6.9

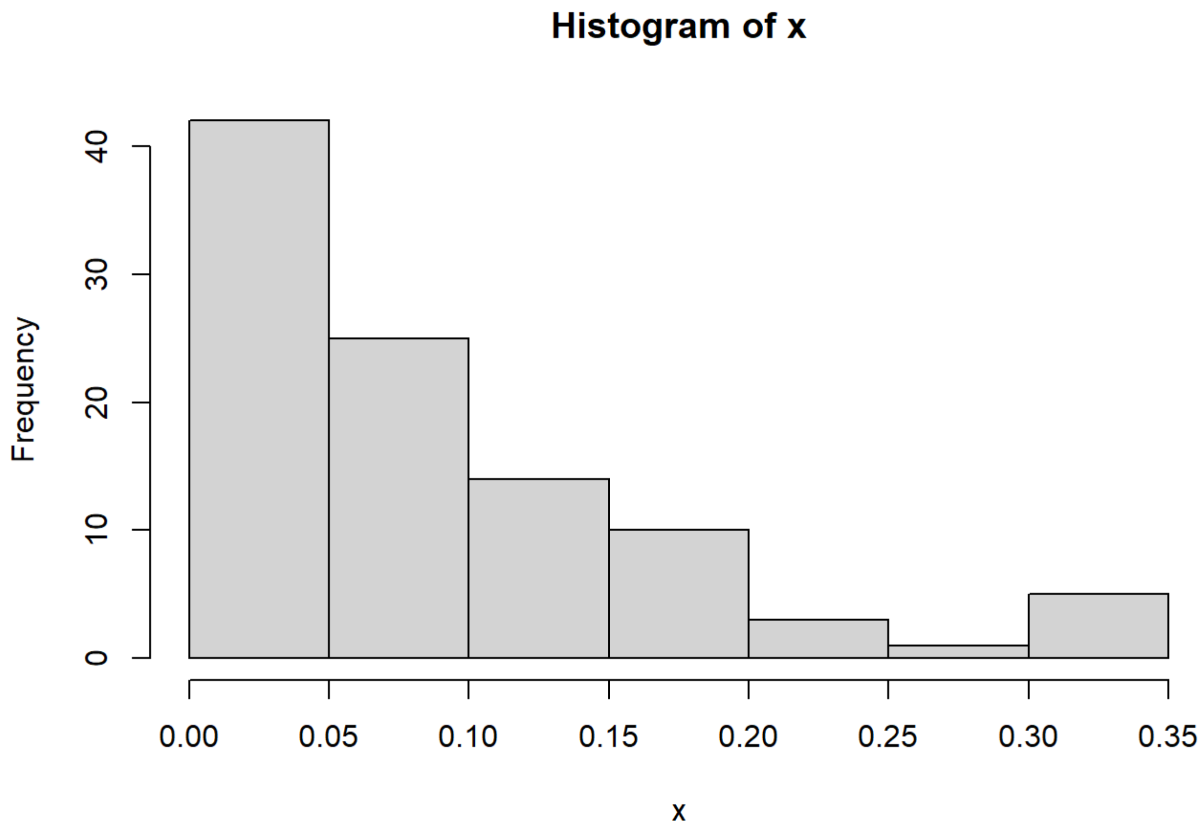
How much area (probability) is to the right of 1.5 for a $N(0,2^2)$?

```
> pnorm(q = 1.5, mean = 0, sd = 2, lower.tail = FALSE)
[1] 0.2266274
```

Problem 6.10

Make a histogram of 100 exponential numbers with mean 10 – $Exp(10)$.

```
> x <- rexp(n = 100, rate = 10)
> hist(x)
```



Estimate the median. Is it more or less than the mean?

```
> median(x)
[1] 0.05937308
> mean(x)
[1] 0.08562395
```

The median is less than the mean.

Is the skewness positive or negative? Is it left-skewed or right-skewed?

```
> library(EnvStats)
Warning: package 'EnvStats' was built under R version
4.0.3

Attaching package: 'EnvStats'
The following objects are masked from 'package:stats':
```

```
    predict, predict.lm
The following object is masked from 'package:base':

    print.default
> skewness(x)
[1] 1.406305
```

We have positive skewness or right-skewed data.

Problem 6.11

Can you figure out what this R command does?

```
> rnorm(5, mean = 0, sd = 1:5)
[1] -0.6512268  0.2426692  6.0827779 -1.3940240
2.2494134
```

We have generated an observation from any one of the distributions

$N(0,1)$, $N(0,2)$, $N(0,3)$, $N(0,4)$, $N(0,5)$.

Problem 6.12

Use R to pick 5 cards from a deck of 52.

```
> cards <- paste(rep(c("Ace",2:10,"Jack","Queen","King"),
4),
+                c("Heart","Diamond","Spade","Club"))
> x <- sample(x = cards, size = 5); x
[1] "7 Heart"    "4 Club"     "3 Spade"    "Ace Heart"  "King
Club"
```

Did you get a pair or better? We say that a pair of cards is two cards with one and the same number(letter).

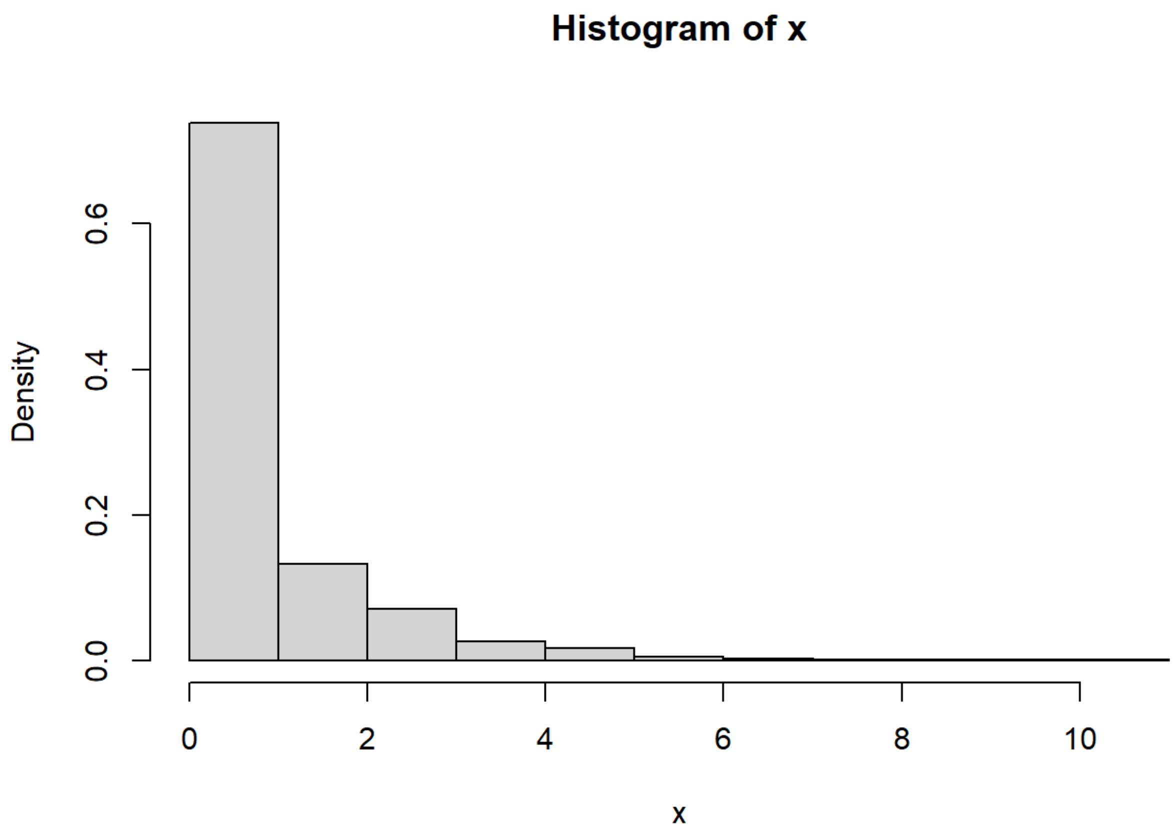
```
> y <- sub(pattern = " .*$",replacement = "" , x); y
[1] "7"     "4"     "3"     "Ace"   "King"
> length(y) - length(unique(y))
[1] 0
```

Repeat until you do. How long did it take?

```
> z <- function(){
+   r <- 0
+   repeat {
+     cards <-
paste(rep(c("Ace",2:10,"Jack","Queen","King"), 4),
+                     c("Heart","Diamond","Spade","Club"))
+     x <- sample(x = cards, size = 5); x
+     y <- sub(pattern = " .*$",replacement = "" , x); y
+     if (length(y) - length(unique(y)) != 0){
+       break
+     } else {
+       r = r + 1
+     }
+   }
+   r
+ }
> z()
[1] 0
```

Repeat the trail 1000 times, plot the histogram and estimate the probability to have the first pair or better on the second trail.

```
> x <- replicate(1000, z())
> hist(x, probability = TRUE)
```




```
> sum(x == 1) / length(x)
[1] 0.248
```

By using the classical definition of probability compute the theoretical probability of the event A to have a pair or better

$$\mathbb{P}(\overline{A}) = \frac{52 \times 48 \times 44 \times 40 \times 36}{52 \times 51 \times 50 \times 49 \times 48} \approx 0.5071$$
$$\mathbb{P}(A) = 1 - \mathbb{P}(\overline{A}) = 1 - 0.5071 = 0.4929$$

Generate the number of the trails before the first pair or better by using the geometric distribution $Geom(0.4929)$.

```
> rgeom(1, prob = 0.4929)
[1] 0
```

Compute the probability to have the first pair or better on the second trail

```
> dgeom(1, prob = 0.4929)
[1] 0.2499496
```

What is the difference if we have replacement?

$$\mathbb{P}(\overline{A}) = \frac{52 \times 48 \times 44 \times 40 \times 36}{52 \times 52 \times 52 \times 52 \times 52} \approx 0.4160$$

$$\mathbb{P}(A) = 1 - \mathbb{P}(\overline{A}) = 1 - 0.4160 = 0.5840$$