University Of Cyprus, M.Sc. In Data Science
DSC 530: Probabilities and Statistics
Project 1

Nikolas Petrou
petrou.p.nikolas@ucy.ac.cy

Problem 1

It is given that grades $\{2, 3, 4, 5, 6\}$ have the same number of students. Assuming that the number of students in grade k is denoted by $G(k)$ where $k \in \{1, ..., 6\}$, and since the number of students in grade 1 is twice the number of the students of each individual class then,

$$G(k) = \begin{cases} x, & k \in \{2, 3, 4, 5, 6\} \\ 2x, & k = 1 \end{cases}$$

Therefore, the total number of the students will be $n = 7x$, which is the sum of the number of students in each grade.

Let the event of a randomly selected student s_i being in the third grade, to be denoted as $A = \{s_i : s_i \text{ belongs in third grade}\}$

In order to calculate the probability that the randomly selected student will be from the third grade, the sample space is denoted by Ω, where in this case $\Omega = \{s_i : i \in \{1, ..., n\}\}$.

Therefore, since the number of students in third grade is x,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{x}{7x} = \frac{1}{7} \tag{1}$$

Following, in order to check-cross the result of (1), a simulation was performed. First, the number of students x was selected to be 10000, since it is considerably a large number. Following, the vector `students` which contains students of the different classes was constructed. The vector contains the relevant frequency of each grade. Next, a sample was drawn by randomly choosing a student from the `students` vector. The experiment was simulated for $M = 100000$ times, and the $P(A)$ was estimated by diving the total number of students which belonged to the third grade by M. dw

```
set.seed(1234) # Seed for reproducibility

# The amount of x students
x <- 10000
students <- rep(c("grade1", "grade2", "grade3",
                  "grade4", "grade5", "grade6"),
                times = c(2*x, x, x, x, x, x))

# Simulating for M times
M <- 100000 # number of simulations
# Pick a student at random
events <- replicate(M, sample(x=students, size=1))
tab <- table(events)
# The students of each grade
tab

## events
## grade1 grade2 grade3 grade4 grade5 grade6
##  28738  14301  14128  14233  14126  14474

# The probability for
# the student to be in grade 3
prob <- tab[3]/M
prob

##  grade3
## 0.14128
```

As it is shown, by comparing the actual value of $\frac{1}{7}$ and the outcome of the simulation 0.14128, the simulated results are pretty close to the actual value.

Finally, it was observed, that by increasing the number M, the estimated results had better accuracy.

Problem 2

Properties of Student's-t and χ^2 Distributions:

Problem 2.1: Student's t-Distribution

Student's t-Distribution is a symmetric and continuous distribution which appears when estimating the mean of a normally distributed population with a short sample size and unknown variance.

Probability Density and Cumulative Distribution Functions:

Student's t-Distribution's $\mathcal{T}(\nu)$ Probability Density Function (p.d.f) with $\nu > 0$ degrees of freedom is [1]:

$$f(x) = \frac{\frac{\Gamma(\nu+1)}{2}}{\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \qquad x \in R$$

while the Cumulative Distribution Function (c.d.f) of the Student's t-Distribution $\mathcal{T}(\nu)$ with $\nu > 0$ degrees of freedom is [1]:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} \frac{\left(1 + \frac{z^2}{\nu}\right)^{-(\nu+1)/2} \frac{\Gamma(\nu+1)}{2}}{\sqrt{\nu\pi}\Gamma(\nu/2)} dz \\ &= \frac{1}{2} + \frac{x\Gamma(\frac{v+1}{2}) {}_2F_1(\frac{1}{2}, \frac{v+1}{2}; \frac{3}{2}; -\frac{x^2}{v})}{\sqrt{\pi v}\Gamma(\frac{v}{2})}, \qquad x \in R \end{aligned}$$

where ${}_2F_1$ is the generalized hypergeometric function of order 2.1 [5].

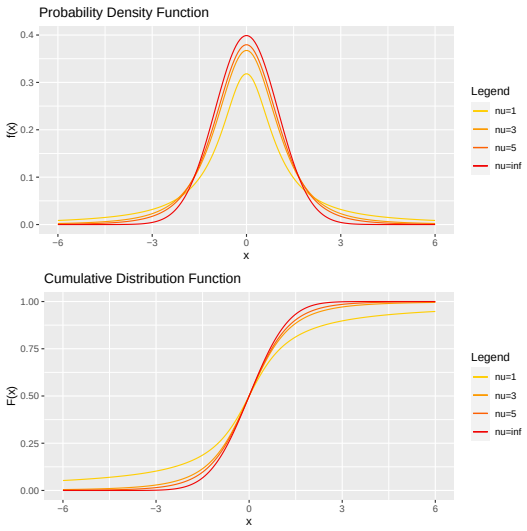
Properties regarding the degrees of freedom:

Moreover, the Student's t-Distribution has a bell shape, and it is similar to the Normal Distribution, but it has heavier tails. Its heaviness is determined by a parameter called degrees of freedom ν. Smaller values of ν give heavier tails, while higher values make the distribution more dense.

The above properties as well as the symmetric property can be easily distinguished from the following plot in R, which shows the Student's t p.d.f and c.d.f. for different values of ν:

```
# Different pdf and cdf plots
library(ggplot2)
library(ggpubr) # Package for ggarrange

x <- seq(-6, 6, len = 200)
pdfs <- cbind(dt(x, df=1), dt(x, df=3), dt(x, df=5),
              dt(x, df=2 ^ .Machine$double.digits))

cdfs <- cbind(pt(x, df=1), pt(x, df=3), pt(x, df=5),
              pt(x, df=2 ^ .Machine$double.digits))

colors <- c("nu=1" = "#FFCE03","nu=3" = "#FD9A01",
```

```
            "nu=5" = "#FD6104", "nu=inf" = "#F00505")

plt.pdfs <- ggplot(data.frame(pdfs), aes(x=x)) +
  geom_line(aes(y = X1, color="nu=1")) +
  geom_line(aes(y = X2, color="nu=3"))  +
  geom_line(aes(y = X3, color="nu=5")) +
  geom_line(aes(y = X4, color="nu=inf")) +
  labs(x = "x", y = "f(x)", color = "Legend")  +
  scale_color_manual(values = colors) +
  ggtitle("Probability Density Function")

plt.cdfs <- ggplot(data.frame(cdfs), aes(x=x)) +
  geom_line(aes(y = X1, color="nu=1")) +
  geom_line(aes(y = X2, color="nu=3"))  +
  geom_line(aes(y = X3, color="nu=5")) +
  geom_line(aes(y = X4, color="nu=inf")) +
  labs(x = "x", y = "F(x)", color = "Legend")  +
  scale_color_manual(values = colors) +
  ggtitle("Cumulative Distribution Function")

ggarrange(plt.pdfs, plt.cdfs,
          ncol = 1, nrow = 2)
```



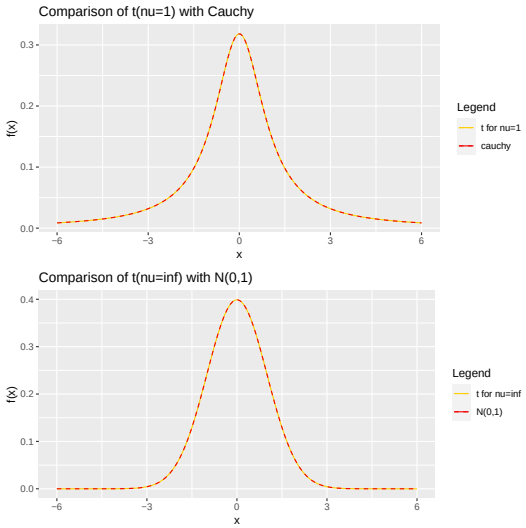
Special cases of the t-Distribution:

Furthermore, another property of the t-Distribution is that with higher values of ν (which approach infinity) the distribution approaches the standard normal distribution $\mathcal{N}(\mu, \sigma^2)$ [1], which has zero mean and a standard deviation of one. On the other hand, when $\nu = 1$, t-Distribution is equal to the Cauchy Distribution [1].

The p.d.f of Cauchy and Student's t (with $\nu = 1$), as well as the p.d.f of $\mathcal{N}(\mu, \sigma^2)$ and Student's t are plotted in R. The plots below show that the densities of the two pairs are identical:

```
# Special cases of t-distribution
library(ggplot2)
library(ggpubr) # Package for ggarrange

x <- seq(-6, 6, len = 200)

# Plotting t=1  with cauchy
pdfs <- cbind(dt(x, df=1), dcauchy(x))
colors <- c("t for nu=1" = "#FFCE03",
    "cauchy" = "#F00505")
plt1 <- ggplot(data.frame(pdfs), aes(x=x)) +
  geom_line(aes(y = X1, color="t for nu=1")) +
  geom_line(aes(y = X2, color="cauchy"),
    linetype = "dashed")  +
  labs(x = "x", y = "f(x)", color = "Legend")  +
  scale_color_manual(values = colors) +
  ggtitle("Comparison of t(nu=1) with Cauchy")
```

```
# Plotting t=inf with Normal(0,1)
pdfs <- cbind(dt(x, df=2 ^ .Machine$double.digits),
    dnorm(x))
colors <- c("t for nu=inf" = "#FFCE03",
    "N(0,1)" = "#F00505")
plt2 <- ggplot(data.frame(pdfs), aes(x=x)) +
  geom_line(aes(y = X1, color="t for nu=inf")) +
  geom_line(aes(y = X2, color="N(0,1)"),
    linetype = "dashed")  +
  labs(x = "x", y = "f(x)", color = "Legend")  +
  scale_color_manual(values = colors) +
  ggtitle("Comparison of t(nu=inf) with N(0,1)")

ggarrange(plt1, plt2,
          ncol = 1, nrow = 2)
```



Mean and Variance of the t-Distribution:

Student's t-Distribution has mean value

$$E(X) = \begin{cases} 0, & \nu > 1 \\ undefined, & \nu = 1 \end{cases}$$

and Variance

$$Var(X) = \begin{cases} \frac{\nu}{(\nu-2)}, & \nu > 2 \\ \infty, & \nu = 2 \\ undefined, & \nu = 1 \end{cases}$$

The mean and variance are undefined for $\nu = 1$, since the distribution is identical to the Cauchy, which does not have any finite moments [2].

In order to estimate the mean and the variance in R, simulation experiments were performed.

Firstly, a simulation was performed for different values of $\nu > 1$, where the sample mean was compared with the theoretical value of the actual mean. For each simulation a very large sample was drawn:

```
# Simulating to compare sample mean with
# actual theorytical mean
set.seed(420)

# Choosing very large sample size
# for more accurate results
sample.size <- 5000000

# The theorytical value of E(X) is 0

# For degrees of freedom = 2
nu <- 2
x.bar <- mean(rt(sample.size, df=nu))
x.bar
```

```
## [1] -0.00177187

# For degrees of freedom = 10
nu <- 10
x.bar <- mean(rt(sample.size, df=nu))
x.bar

## [1] -0.000435226

# For degrees of freedom = Inf
nu <- Inf
x.bar <- mean(rt(sample.size, df=nu))
x.bar

## [1] -0.0001389596
```

As it was observed, by running different simulations with differnt specified ν values, each time the sample mean was very close to the theorytical value of the mean which is 0.

Following, the same procedure was followed in order to compare the sample variance that occurred from the simulation with the actual variance of the distribution for different values of $\nu > 2$:

```
# Simulating to compare sample variance with
# actual theoretical variance
set.seed(420)

# Choosing very large sample size
# for more accurate results
sample.size <- 5000000

# For degrees of freedom = 5
nu <- 5
# The theorytical value of the variance
actual.variance <- nu/(nu-2)
actual.variance

## [1] 1.666667

sample.var <- var(rt(sample.size, df=nu))
sample.var

## [1] 1.667823

# For degrees of freedom = 25
nu <- 25
# The theorytical value of the variance
actual.variance <- nu/(nu-2)
actual.variance

## [1] 1.086957

sample.var <- var(rt(sample.size, df=nu))
sample.var

## [1] 1.086939
```

As it was seen, by running different simulations with different specified ν values, each time the sample variance was almost identical to the theoretical value of the variance, which is $\frac{\nu}{(\nu-2)}$.

Problem 2.2: Chi-Square Distribution

The Chi-Square distribution is a continuous distribution and a special case of the Gamma Distribution. In addition, it is one of the most extensively used probability distributions in inferential statistics, particularly in hypothesis testing and confidence intervals.

Probability Density and Cumulative Distribution Functions:

A Chi-Square χ^2 random variable X with $k \in \mathbb{N}^+$ degrees of freedom has Probability Density Function (p.d.f.):

$$f(X) = \begin{cases} \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-x^2}, & x > 0 \\ 0, & otherwise \end{cases}$$

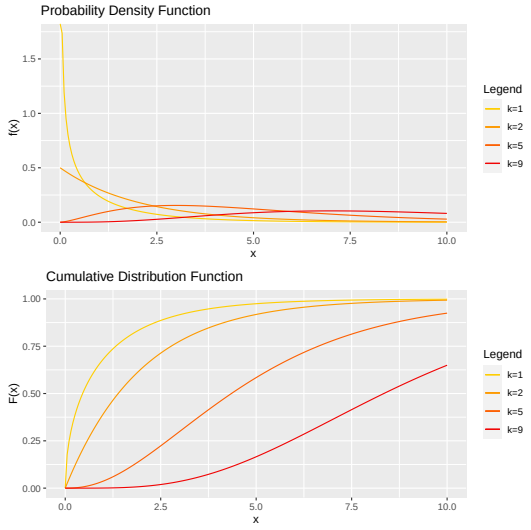
and Cumulative Distribution Function

$$F(X) = \begin{cases} F(x) = \frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

where Γ and γ are the upper and lower incomplete Gamma functions.

The following plots in R, shows the χ^2 p.d.f and c.d.f. for different values of k :

```
# Different pdf and cdf plots
library(ggplot2)
library(ggpubr) # Package for ggarrange

x <- seq(0, 10, len = 200)
pdfs <- cbind(dchisq(x, df=1), dchisq(x, df=2),
              dchisq(x, df=5), dchisq(x, df=9))

cdfs <- cbind(pchisq(x, df=1), pchisq(x, df=2),
              pchisq(x, df=5), pchisq(x, df=9))

colors <- c("k=1" = "#FFCE03","k=2" = "#FD9A01",
            "k=5" = "#FD6104", "k=9" = "#F00505")

plt.pdfs <- ggplot(data.frame(pdfs), aes(x=x)) +
  geom_line(aes(y = X1, color="k=1")) +
  geom_line(aes(y = X2, color="k=2"))  +
  geom_line(aes(y = X3, color="k=5")) +
  geom_line(aes(y = X4, color="k=9")) +
  labs(x = "x", y = "f(x)", color = "Legend")  +
  scale_color_manual(values = colors) +
  ggtitle("Probability Density Function")

plt.cdfs <- ggplot(data.frame(cdfs), aes(x=x)) +
  geom_line(aes(y = X1, color="k=1")) +
  geom_line(aes(y = X2, color="k=2"))  +
  geom_line(aes(y = X3, color="k=5")) +
  geom_line(aes(y = X4, color="k=9")) +
  labs(x = "x", y = "F(x)", color = "Legend")  +
  scale_color_manual(values = colors) +
  ggtitle("Cumulative Distribution Function")

ggarrange(plt.pdfs, plt.cdfs,
          ncol = 1, nrow = 2)
```



Mean and Variance of the Chi-Square Distribution:

It has mean value equal to the number of degrees of freedom ($\mu = k$) and its variance is equal to two times the number of degrees of freedom ($\sigma^2 = 2k$).

In order to estimate the mean and the variance in R, simulation experiments were performed.

Firstly, a simulation was performed for different values of k, where the sample mean was compared with the theoretical value of the actual mean. For each simulation a very large sample was drawn:

```
# Simulating to compare sample mean with
# actual theoretical mean
set.seed(420)

# Choosing very large sample size
# for more accurate results
sample.size <- 5000000

# The theoretical value of E(X) is 0

# For k = 2
k <- 2
x.bar <- mean(rchisq(sample.size, df=k))
x.bar

## [1] 1.999856

# For k = 5
k <- 5
x.bar <- mean(rchisq(sample.size, df=k))
x.bar

## [1] 5.000394
```

As it was observed, by running different simulations with different specified k values, each time the sample mean was very close to the theoretical value of the mean which is equal to k.

Following, the same procedure was followed in order to compare the sample variance that occurred from the simulation with the actual variance of the distribution for different values of k:

```
# Simulating to compare sample variance with
# actual theoretical variance
set.seed(420)

# Choosing very large sample size
# for more accurate results
sample.size <- 5000000

# For k = 5
k <- 5
# The theoretical value of the variance
actual.variance <- 2*k
actual.variance

## [1] 10

sample.var <- var(rchisq(sample.size, df=k))
sample.var

## [1] 9.990072

# For k = 25
k <- 25
# The theoretical value of the variance
actual.variance <- 2*k
actual.variance

## [1] 50

sample.var <- var(rchisq(sample.size, df=k))
sample.var

## [1] 49.99084
```

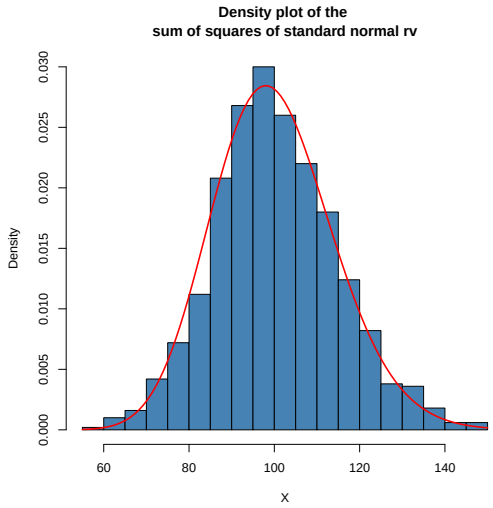
As it was seen, by running different simulations with different specified k values, each time the sample variance was almost identical to the theoretical value of the variance, which is $2k$.

The distribution of a sum of the squares i.i.d. standard normal random variables

Last but not least, one of the most crucial properties of the distribution is that it is the distribution of a sum of the squares of n identically distributed(i.i.d.) standard normal random variables. If $Z_1, Z_2, ..., Z_n$ are independent i.i.d. standard normal random variables, then it is stated that $\sum_{i=1}^{n}(Z_i)^2 \sim \chi_n^2$

Simulating to compare the density plots of the sum of squares of $Z_i$for$i \in \{1, ..., n\}$ with the density plot of $^2(n)$

```
# Number of repetitions
reps <- 1000
# Set degrees of freedom of a chi-Square
#Distribution
df <- 100
# Sample of 1000 column vectors which
#follow N(0,1)
Z <- replicate(reps, rnorm(df))
# column sums of squares
X <- colSums(Z^2)

# histogram of column sums of squares
hist(X, freq = F, col = "steelblue",breaks=30,
ylab = "Density", main = "Density plot of the
sum of squares of standard normal rv")

# add theoretical density
curve(dchisq(x, df = df), type = 'l', lwd = 2,
col = "red", add = T)
```



Problem 3

It is given that there are three machines $\{M1, M2, M3\}$ which produce manufactured items. It is also given that machine $M1$ produces 20% of the total items, while machines $M2$ and $M3$ produce 30% and 50% of the manufactured items, respectively. At the same time, 1%, 2% and 3% of the manufactured items of machines $M1$, $M2$ and $M3$ are defective.

To determine the probability that a randomly selected item was manufactured by the machine $M2$ given that it is defective, both the conditional probability and the Bayes' theorem are going to be utilized.

The conditional probability [3] of an event A given B has already occurred, is given by:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2}$$

Bayes' theorem [4] is stated as the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3}$$

Let,

$A = \{\text{Item produced by } M1\}$ $B = \{\text{Item produced by } M2$
$C = \{\text{Item produced by } M3\}$ $D = \{\text{Item is defective}\}$

Moreover, it is given that:

$P(A) = 0.2, \quad P(B) = 0.3, \quad P(C) = 0.5$
$P(D|A) = 0.01, \quad P(D|B) = 0.02, \quad P(D|C) = 0.03$

Firstly, $P(D)$ will be calculated by using the Conditional Probability of the event D given A,B and C, respectively.

$$P(D) = P(D \cap A) + P(D \cap B) + P(D \cap C)$$
$$\Rightarrow P(D) = P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C)$$
$$= 0.01 * 0.2 + 0.02 * 0.3 + 0.03 * 0.5 = 0.023$$

Next, since the goal is to calculate the probability that a randomly chosen item was produced by the machine $M2$, given that is defective, the objective is to calculate $P(B|D)$. Thus, by taking advantage of the Bayes theorem (3):

$$P(B|D) = \frac{P(B \cap D)}{P(D)} = \frac{P(D|B)P(B)}{P(D)} = \frac{0.02 * 0.3}{0.023}$$
$$= \frac{0.0006}{0.023} \approx 0.26$$

```
# The following script
# calculates the probability of B given D

# Probab. of items being produced by each machine
p_item_produced_1 <- 0.2
p_item_produced_2 <- 0.3
p_item_produced_3 <- 0.5

# The proportions of def. items of each machine
p_def_given1 <- 0.01
p_def_given2 <- 0.02
p_def_given3 <- 0.03

p_def <- p_item_produced_1 * p_def_given1 +
      p_item_produced_2 * p_def_given2 +
      p_item_produced_3 * p_def_given3

p_def_given_2 <- (p_item_produced_2*p_def_given2)/p_def

p_def_given_2

## [1] 0.2608696
```

Problem 4

Problem 4.1

Initially, it is given that the probability that $X = x$ is proportional to $(x + 1)(8 - x)$ for $x \in \{0, 1, 2, 3, 4, 5, 6, 7\}$. Therefore, the p.m.f. of X can be written as follows:

$$P(X = x) = c(x + 1)(8 - x) \quad (4)$$

From the probabilities axioms, it is well known that the sum of the probabilities of all outcomes must be equal to 1. Therefore, since $x \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ and from (4) the sum of the probabilities can be written as follows:

$$\sum_{x=0}^{7} c(x + 1)(8 - x) = 1 \quad (5)$$

By solving the equation (5), we get that c is equal to $\frac{1}{120}$. The value of constant c can also be computed by the following simple *for-loop* in R:

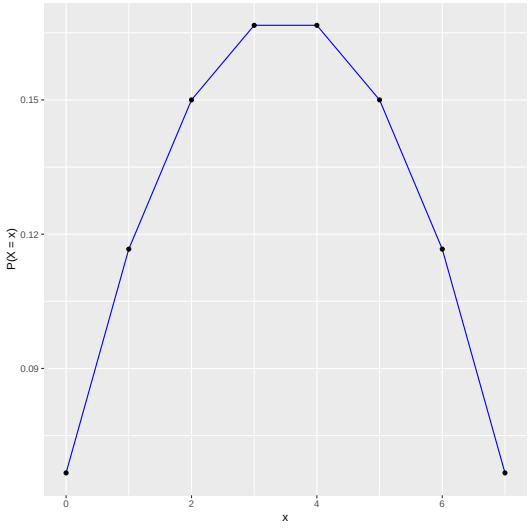
```
# The following script
# calculates the value of c for Problem 4
calculate.c <- function() {
  eq <- 1 # The initial equation equals to one
  sum <- 0 # Initialize sum to zero
  for (x in c(0:7)){
    sum <- sum + (x+1)*(8-x)
  }
  c <- eq/sum
}

# Will print the value of 0.008333333,
# which is equal to 1/120
print(calculate.c())

## [1] 0.008333333

# Confirmation
print(1/120)

## [1] 0.008333333
```

Finally, the p.m.f of X is the following:

$$P(X = x) = \begin{cases} \frac{1}{120}(x + 1)(8 - x), & x \in \{0, 1, 2, 3, 4, 5, 6, 7\} \\ 0, & Otherwise \end{cases}$$

The plotted p.m.f. is the following:

```
x <- seq(0,7)
data <- data.frame(x, (1/120)*(x+1)*(8-x))

library(ggplot2)
# Basic line plot with points
ggplot(data=data, aes(x=data[,1], y=data[,2]))+
  geom_line(col="blue")+
  geom_point()+
  labs(x="x", y="P(X = x)")
```



Problem 4.2

In order to find the probability that X will be at least 5 by simulating in R, the following procedure has been followed:

- Firstly, the p.m.f which was calculated-found in the first part of the problem was utilized in order to generate samples for the numbers of cars that could occur. That was settled by using the `sample()` and `replicate()` functions, in order to perform the experiment M times. Note that the parameter of probabilities for the `sample()` function is the calculated p.m.f.
- Following, the number of events A which the experiment had an outcome of more than 5 cars were counted

- Finally, the probability $P(X \geq 5)$ was estimated by the total number of events A over the total number of experiments M

```
# The function returns the calculated pmf
get.pmf <- function(){
  # Using the constant c (=1/120)
  # which was already calculated
  c <- calculate.c()
  probs <- c()
  for (x in c(0:7)){
    probs <- append(probs, c*(x+1)*(8-x))
  }
  return(probs)
}
# Replicate for M=100000 times
no.of.cars <- replicate(n=100000, sample(x=c(0:7),
                        size=1, prob=c(get.pmf())))
# Counting the events where X>=5
events.A <- 0
for (x in no.of.cars){
  if (x >= 5){
    events.A <- events.A + 1
  }
}
# Shoud be close to 1/3
prob <- events.A/length(no.of.cars)
prob

## [1] 0.33381
```

In order to get the exact value, the probability $P(X \geq 5)$ has to be calculated. The occasions where the value of X is greater than 5 are only when the lane hold either 5, 6 or 7 cars.

Therefore,

$$\begin{aligned} P(X \geq 5) &= P(\{X = 5\} \cup \{X = 6\} \cup \{X = 7\}) \\ &= P(X = 5) + P(X = 6) + P(X = 7) \\ &= \frac{1}{120}(8 + 14 + 8) \\ &= \frac{1}{3} \end{aligned}$$

As it is shown, by comparing the actual value of $\frac{1}{3}$ and the outcome of the simulation 0.33425, the simulated results are pretty close to the actual value.

Problem 5

As $X \sim U(0, 5)$ the cumulative distribution function (cdf) of X is the following:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{5}, & x \in (0, 5) \\ 1, & x \geq 5 \end{cases}$$

So, $P(Y = 0) = P(X \leq 1) = F_X(1) = \frac{1}{5}$,
$P(Y = 5) = P(X \geq 3) = 1 - P(X \leq 3) = 1 - \frac{3}{5} = \frac{2}{5}$

Furthermore, the random variable Y is defined by:

$$Y = \begin{cases} 0, & X \leq 1 \\ 5, & X \geq 3 \\ X, & otherwise \end{cases}$$

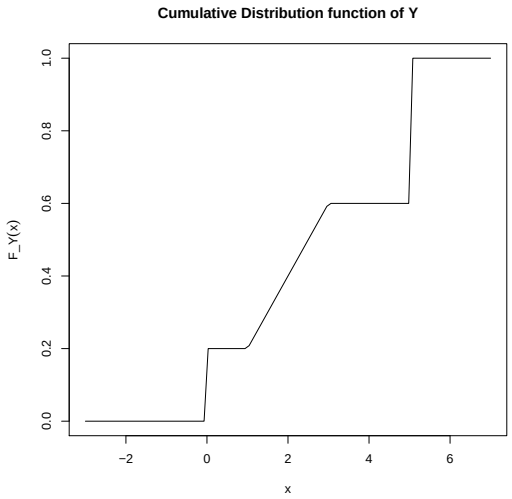
Moreover, we have that $\{0 < Y \leq 1\} = \emptyset$. Indeed, lets assume that $Y(\omega) \in (0, 1]$ for some $\omega \in \Omega$, then $Y(\omega) = X(\omega) \in (0, 1)$. Then, since $X(\omega) \in (0, 1)$, Y=0, which is a contradiction. Similarly, we can show that $\{3 \leq Y < 5\} = \emptyset$.

Therefore, The cumulative distribution function of Y is the following:

$$F_Y(x) = \begin{cases} P(Y \leq x) = 0, & x < 0 \\ P(Y < 0) + P(0 \leq Y \leq 1) & \\ = P(Y = 0) = \frac{1}{5} & 0 \leq x \leq 1 \\ P(Y \leq 1) + P(1 < X < x) & \\ = \frac{1}{5} + \frac{x}{5} - \frac{1}{5} = \frac{x}{5}, & 1 < x < 3 \\ P(Y \leq 3) + P(3 \leq Y \leq 5) & \\ = P(Y \leq 3) + P(3 < Y < 5) = \frac{3}{5}, & 3 \leq x \leq 5 \\ P(Y \leq 3) + P(Y = 5) = 1, & x \geq 5. \end{cases}$$

The plotted cdf of Y is the following:

```
# Creating the cdf of Y using the indicator function

cdf_Y <- function(x){
  r <- 0*I( x <0)+(1/5)*I(0 <= x)*I(x <= 1)+
  (x/5)*I(1 < x)*I(x <3)+
    (3/5)*I(3 <= x)*I(x <5)+1*I(x >= 5)
}

##We plot the cdf of Y in the interval [-3,7]

x<-seq(-3,7,length=100)
plot(x,cdf_Y(x),type='l',xlab='x',ylab=expression(F_Y(x)),
main='Cumulative Distribution function of Y')
```



References

[1] M. Ahsanullah, B. M. G. Kibria, and M. SHAKIL. *Normal and Student´s t Distributions and Their Applications*. 01 2014.

[2] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions, volume 2*, volume 289. John wiley & sons, 1995.

[3] A. N. Kolmogorov and A. T. Bharucha-Reid. *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications, 2018.

[4] A. Stuart and J. Ord. Kendall's advanced theory of statistics (edward arnold). *London, UK*, 1994.

[5] Wikipedia. Generalized hypergeometric function — Wikipedia, the free encyclopedia. `http://en.wikipedia.org/w/index.php?title=Generalized%20hypergeometric%20function&oldid=1031997283`, 2021. [Online; accessed 24-October-2021].