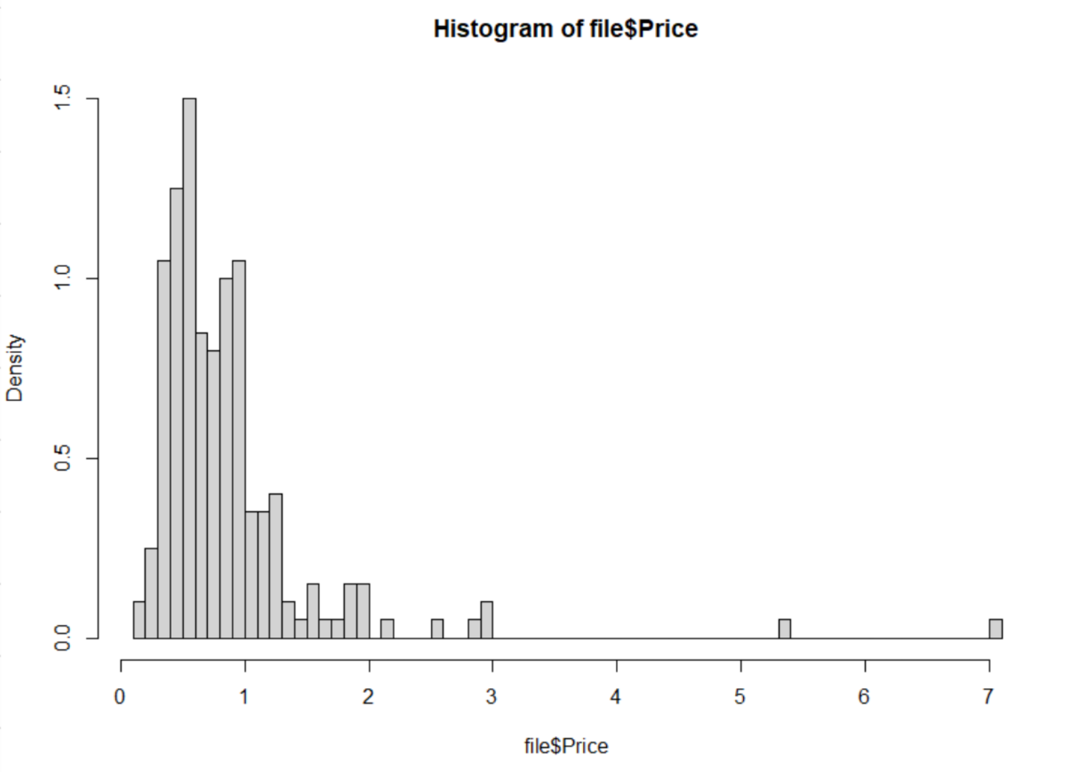
עבודת בית 1

ת.ז. 1: 212699581

ת.ז. 2: 211709597

שאלה 1:

א)

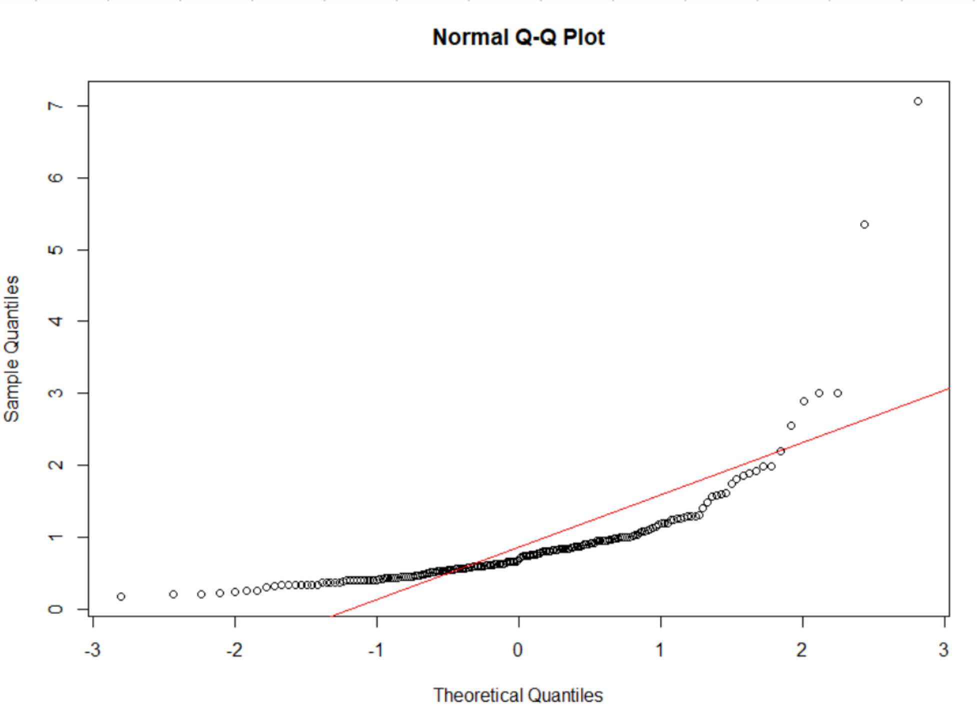


ההיסטוגרמה לא סימטרית סביב ערך מסויים אנו רואים
כי להיסטוגרמה יש זנב ימני

ב)

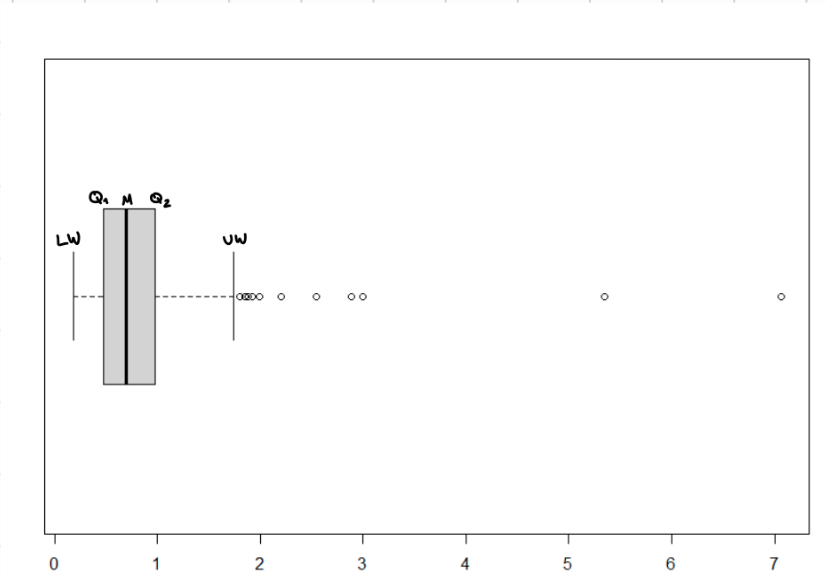
```
> mean(X)
[1] 0.860505
> median(X)
[1] 0.694
> Q1 = quantile(X, prob=0.25)
> Q3 = quantile(X, prob=0.75)
> Q1
  25%
0.4745
> Q3
 75%
0.98
> var(X)
[1] 0.5313651
> sd(X)
[1] 0.728948
> max(X) - min(X)
[1] 6.885
> r = Q3 - Q1
> r
  75%
0.5055
```

(2



ברור שזה לא ליניארי ולכן התצפיות לא מתפלגות נורמלית

(3




```
> step = 1.5* (Q3 - Q1)
> step
    75%
0.75825
>
> LF = Q1 - step
> UF = Q3 + step
>
> LF
     25%
-0.28375
> UF
    75%
1.73825
```

```
> UW
[1] 1.735
> LW
[1] 0.178
```

```
> outter = sum(X < LF | X > UF)
> outter
[1] 13
```

ערכים חריגים

אנו רואים כי ההתפלגות לא סימטרית
(מרווח Q_2 מחציון גדול ממרווח Q_1 מחציון)
וישנם כ 13 תצפיות חריגות מעל UW

שאלה 2

א)

```r
# Qes 2
# A

set.seed(9581)
sample1 = rexp(5, 0.2)
sample2 = rexp(20, 0.2)
sample3 = rexp(500, 0.2)

mean1 = mean(sample1)
mean2 = mean(sample2)
mean3 = mean(sample3)

var1 = var(sample1)
var2 = var(sample2)
var3 = var(sample3)

med1 = median(sample1)
med2 = median(sample2)
med3 = median(sample3)

p = 0.25 # Q1
Q_p1 = quantile(sample1, prob=p)
Q_p2 = quantile(sample2, prob=p)
Q_p3 = quantile(sample3, prob=p)

print(mean1)
print(var1)
print(med1)
print(Q_p1)

print(mean2)
print(var2)
print(med2)
print(Q_p2)

print(mean3)
print(var3)
print(med3)
print(Q_p3)
```

```
> print(mean1)
[1] 0.6191735
> print(var1)
[1] 0.3042535
> print(med1)
[1] 0.5495735
> print(Q_p1)
      25%
0.1111503
>
> print(mean2)
[1] 5.813065
> print(var2)
[1] 18.65971
> print(med2)
[1] 5.378358
> print(Q_p2)
     25%
2.429967
>
> print(mean3)
[1] 4.776981
> print(var3)
[1] 24.11499
> print(med3)
[1] 3.286302
> print(Q_p3)
   25%
1.3195
```

ב)

Sample 1 = (0.071, 0.1111, 0.549, 1.081, 1.282)

$n=5$

$$F_n = \frac{1}{n}\sum_{i=1}^{n} I(x_i \le x)$$

$$F_n(x) = \begin{cases} 0 & x < 0.071 \\ \frac{1}{5} & 0.071 \le x < 0.1111 \\ 2/5 & 0.1111 \le x < 0.549 \\ 3/5 & 0.549 \le x < 1.081 \\ 4/5 & 1.081 \le x < 1.282 \\ 1 & 1.282 \le x \end{cases}$$

$F_X(x) = 1 - e^{-0.2x}$ $\quad$ $X \sim Exp(0.2)$ $\quad$ (2

$E(X) = \frac{1}{0.2} = 5$

$Var(X) = \frac{1}{(0.2)^2} = 25$

$F_X(\xi_{0.5}) = 0.5 \Rightarrow 1 - e^{-0.2\xi_{0.5}} = 0.5$

$0.5 = e^{-0.2\xi}$

$-\ln(2) = -0.2\,\xi_{0.5}$

$p = 0.25$ $\quad$ $5\ln(2) = \xi_{0.5}$

$P(X \le \xi_p) = p = F(\xi_p) = 0.25$

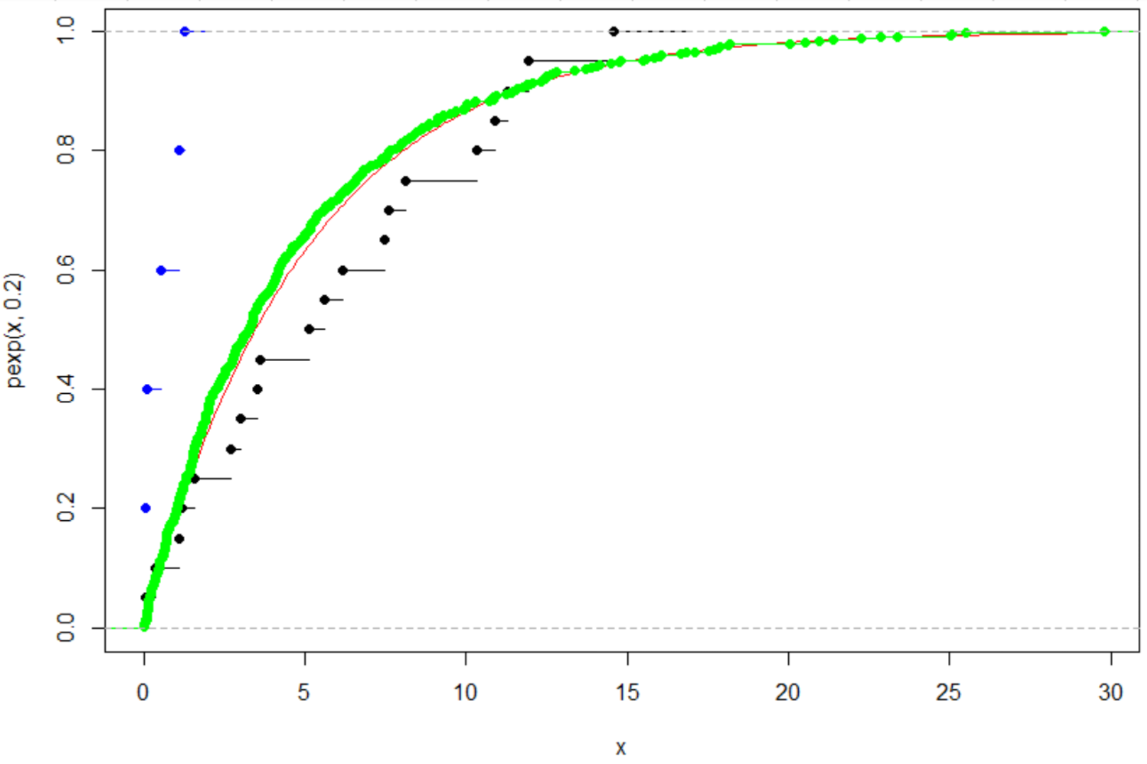
$1 - e^{-0.2\xi_p} = 0.25 \Rightarrow 0.75 = e^{-0.2\xi_p} = \ln(3/4) = -0.2\,\xi_p$

$\Rightarrow -5\ln(3/4) = \xi_p$

```
# D

x = seq(0, max(sample3), 0.1)
plot(x, pexp(x, 0.2), type='l', col='red')
lines(ecdf(sample1), col='blue')
lines(ecdf(sample2), col='black')
lines(ecdf(sample3), col='green')
```

C3



ה) בכלל שהמדגם לא מתאר האוכלוסיה עצמה
הוא מייצג את המדדים עם סטייה (אי דיוק בגלל
שמספר דגימות סופי)

עבור מדגמים גדולים מדדים אלה שואפים
למדדים של האוכלוסיה
עבור ממוצע מדגמי ותוחלת:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

לפי חוק המספרים הגדולים, לכל $\varepsilon > 0$

$$\lim_{n \to \infty} P(|\bar{X} - \mu| < \varepsilon) = 1$$

כלומר $\bar{X} \xrightarrow{P} \mu$

עבור התפלגות אמפירית:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \le x) \xrightarrow{P} E(I(X_i \le x)) = P(X_i \le x) = F_X(x)$$

שאלה 3

$$\forall d: \ F_X(\theta+d) = 1 - F_X(\theta-d)$$

$y = x - d$

$$P(X \le \theta+d) = P(X \ge \theta-d)$$

$$P(X \le \xi_p) = p \leftarrow \text{נתון}$$

I. נכון

ניקח $d = \theta - \xi_p$

$$P(X \le 2\theta - \xi_p) = P(X \le \theta + \theta - \xi_p) = P(X \le \theta + d)$$

$$= F_X(\theta+d) = 1 - F_X(\theta-d) = 1 - F_X(\theta - \theta + \xi_p)$$

$$= 1 - F_X(\xi_p) = 1 - p$$

II. נכון

$$1 - p \overset{\text{I מ}}{=} P(X \le \theta + \theta - \xi_p) = P(X - \theta \le \theta - \xi_p) =$$

$$= P(Y \le \theta - \xi_p) = P(Y \ge \xi_p - \theta)$$

$$= 1 - P(Y \le \xi_p - \theta)$$

$$\Rightarrow 1 - p = 1 - P(Y \le \xi_p - \theta)$$

$$\Rightarrow p = P(Y \le \xi_p - \theta)$$

III נכון

מהנתון

$$F_X(\theta+d) = 1 - F_X(\theta-d)$$

עבור $d=0$

$$F_X(\theta) = 1 - F_X(\theta)$$

$$\Rightarrow \quad 2F_X(\theta) = 1$$

$$\Rightarrow \quad F_X(\theta) = \frac{1}{2}$$

$$P(Y \le 0) = P(X - \theta \le 0) = P(X \le \theta) = F_X(\theta) = \frac{1}{2}$$