Exercies 7 - Alon Goodman & Ran Hassid

Q1

a

```
# ex6data1 <- read.csv("C:/Users/Alon/Desktop/Studies/Statistics/Statistical_Computin
g/Exercises/HW6/ex6data1.csv")

ex7q1 <- read.csv("~/Desktop/Ran/D year/semester b/hishov statisti/exercies/HW7/ex7q
1.csv")
```

א. 1500 ערים נסקרו.

n_i מס' הערים שבהן היו i מקרים.

X_j - כמות המקרים שהיו j בעיר

$X_j \sim Pois(\lambda)$

$$L(x,\lambda) = \prod_{j=1}^{n} \frac{\lambda^{x_j} e^{-\lambda}}{x_j!} = \frac{\lambda^{\sum x_j} e^{-n\lambda}}{\prod x_j!}$$

$$l(x,\lambda) = \ln(L(n,\lambda)) = \sum x_j \cdot \ln\lambda - n\lambda - \sum \ln(x_j!)$$

$$\frac{\partial l(x,\lambda)}{\partial\lambda} = \frac{\sum x_j}{\lambda} - n \qquad \frac{\partial^2 l(x,\lambda)}{\partial\lambda} = \frac{-\sum x_j}{\lambda^2} < 0 \Rightarrow \hat{\lambda}_{MLE} = \frac{\sum x_j}{n}$$

לכ: בגלל הנתונים המקובצים וכמות המקרים שנבדקו הם $0,1,\dots 16$

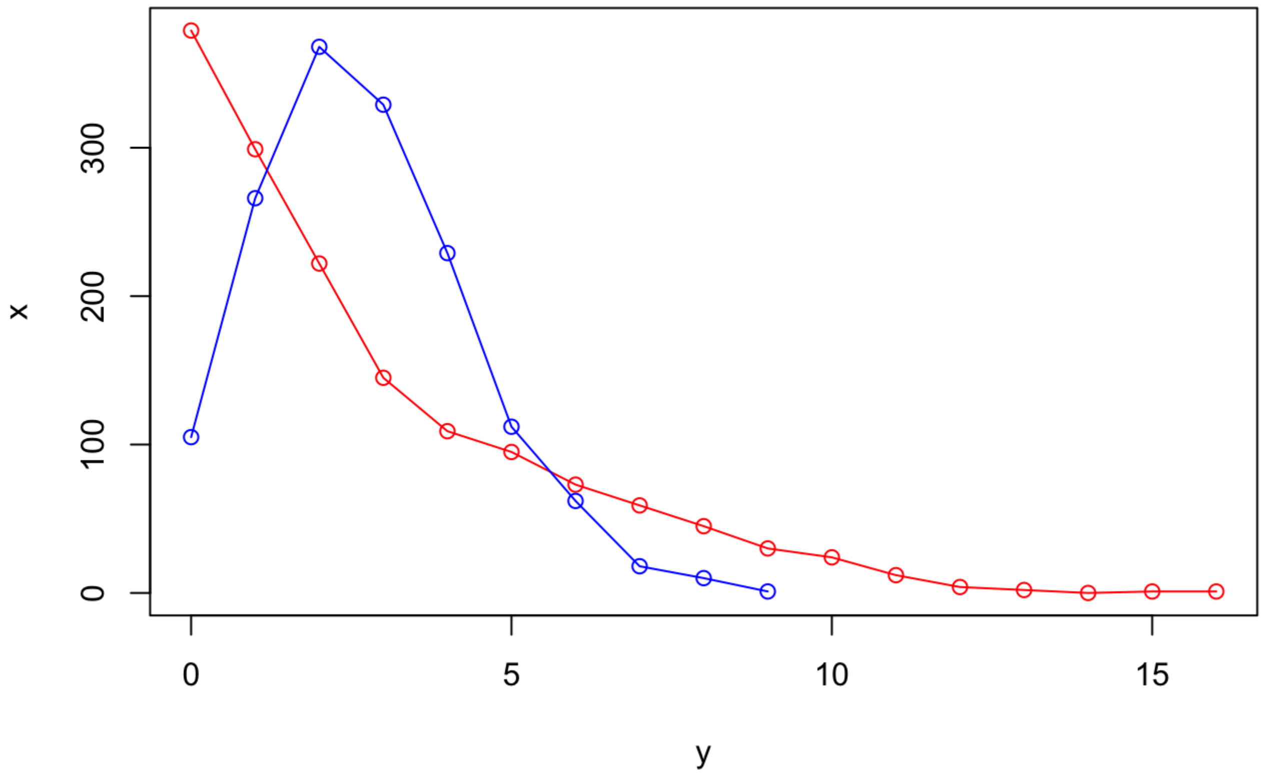
לכן $$\hat{\lambda}_{MLE} = \frac{\sum_{i=0}^{16} i \cdot n_i}{1500}$$

```
lambda_hat.a <-  sum((ex7q1$frequencies..ni.*ex7q1$number.of.cases..i.)/1500)
cat(sprintf("lambda (MLE) = %s", lambda_hat.a) )
```

```
## lambda (MLE) = 2.698
```

```
x <- ex7q1$frequencies..ni
y <- ex7q1$number.of.cases..i.

plot(y,x,type = "o",col = "red")
y_p <- rpois(n = 1500,lambda = lambda_hat.a)
table_y_p <- as.vector(table(y_p))
y_p_unique <- sort(unique((y_p)))
lines(y_p_unique,table_y_p, type = "o",col= "blue")
```



b

ב. נחלק לג קבוצות:

קבוצה I : מקומות 0 . סיכוי α להיות שייך לקבוצה זו

קבוצה II : התפלגות פואסון עם λ_1 סיכוי β להיות שייך לקבוצה זו

קבוצה III : התפלגות פואסון עם λ_2 סיכוי $1-\alpha-\beta$ להיות שייך לקבוצה זו

(i)
$$P(X_j=0)=\alpha+\beta\cdot P_r(pois(\lambda_1)=0)+(1-\alpha-\beta)\cdot P_r(pois(\lambda_2)=0)$$

$i\neq 0$
$$P(X_j=i)=\beta\cdot P_r(pois(\lambda_1)=i)+(1-\alpha-\beta)\cdot P_r(pois(\lambda_2)=i)$$

(ii) נסמן ב n_0 את מס' הנשים שקיבלו 0

n_i את מס' הנשים שקיבלו i

$$L(\lambda_1,\lambda_2)_{obs}=\prod_{j=1}^{1000}P(X_j=x_j)=P(X_j=0)^{n_0}\cdot\prod_{i=1}^{16}P(X_j=i)^{n_i}=$$

$$\frac{\lambda^{x_j}e^{-\lambda}}{x_j!}$$

$$=\left(\alpha+\beta\cdot P_r(pois(\lambda_1)=0)+(1-\alpha-\beta)\cdot P_r(pois(\lambda_2)=0)\right)^{n_0}\cdot$$
$$\prod_{i=1}^{16}\left(\beta\cdot P_r(pois(\lambda_1)=i)+(1-\alpha-\beta)\cdot P_r(pois(\lambda_2)=i)\right)^{n_i}=$$
$$=\left[\left(\alpha+\beta\cdot e^{-\lambda_1}+(1-\alpha-\beta)\cdot e^{-\lambda_2}\right)\right]^{n_0}\cdot\prod_{i=1}^{16}\left[\beta\left(\frac{\lambda_1^i\cdot e^{-\lambda_1}}{i!}\right)+(1-\alpha-\beta)\cdot\left(\frac{\lambda_2^i\cdot e^{-\lambda_2}}{i!}\right)\right]=$$
$$=\prod_{i=0}^{16}\left(\alpha\cdot\mathbb{1}\{i=0\}+\frac{\beta\lambda_1^i e^{-\lambda_1}}{i!}+\frac{(1-\alpha-\beta)\lambda_2^i e^{-\lambda_2}}{i!}\right)^{n_i}$$

קשה למצוא אותה בצורה מקסימלית כי לוג הנראות שלה נראה לנו קשה

ולכן צריך מערכת משוואות לפי α,β, λ_1,λ_2. לכן עדיף לא להשתמש בנגזרות אלא

לעשות EM.

C

4. כשנתון $X_1,\dots,X_{1500}$ מס' הכוסות של אחד עובדי סוני

יש 3 קבוצות

קבוצה I : מקובעת 0 . סיכוי α

קבוצה II : מתפלגת פואסון עם λ_1 סיכוי β

קבוצה III : מתפלגת פואסון עם λ_2 סיכוי $1-\alpha-\beta$

הנראות החלקית היא

$$L(X,\alpha,\beta,\lambda_1,\lambda_2)=\prod_{j=1}^{1500}\alpha\cdot\mathbb{1}\{x_j=0\}+\beta\cdot\frac{\lambda_1^{x_j}e^{-\lambda_1}}{x_j!}+(1-\alpha-\beta)\cdot\frac{\lambda_2^{x_j}e^{-\lambda_2}}{x_j!}$$

$$l(X,\alpha,\beta,\lambda_1,\lambda_2)=\sum_{j=1}^{1500}\ln\left\{\alpha\cdot\mathbb{1}\{x_j=0\}+\beta\cdot\frac{\lambda_1^{x_j}e^{-\lambda_1}}{x_j!}+(1-\alpha-\beta)\cdot\frac{\lambda_2^{x_j}e^{-\lambda_2}}{x_j!}\right\}$$

ניתן להגדיר Z בעזרת כמ"מ המציינים שמתאר לכל הקבוצות הנ"ל

Z_1 - מתפלג מנוון שקול אפס בנקודה 0

Z_2 - מתפלג $Pois(\lambda_1)$

Z_3 מתפלג $Pois(\lambda_2)$

$\{Z_i\}_{i=1}^{1500}$ לאיזו קבוצה X_j שייך. $\Leftarrow$ $Z_1 \sim multinomial(\alpha,\beta,1-\alpha-\beta)$

לכן בנראות המלאה ניתן לכתוב באופן הבא

$$L(X,\alpha,\beta,\lambda_1,\lambda_2)=\prod_{j=1}^{1500}\alpha^{\mathbb{1}\{Z_j=1\}}\cdot\left(\beta\cdot\frac{\lambda_1^{x_j}e^{-\lambda_1}}{x_j!}\right)^{\mathbb{1}\{Z_j=2\}}\cdot\left((1-\alpha-\beta)\cdot\frac{\lambda_2^{x_j}e^{-\lambda_2}}{x_j!}\right)^{\mathbb{1}\{Z_j=3\}}$$

$$l(X,\alpha,\beta,\lambda_1,\lambda_2)=\sum_{j=1}^{1500}\mathbb{1}\{Z_j=1\}\cdot\ln\alpha+\mathbb{1}\{Z_j=2\}\cdot\ln\left(\frac{\beta\lambda_1^{x_j}e^{-\lambda_1}}{x_j!}\right)+\mathbb{1}\{Z_j=3\}\cdot\ln\left(\frac{(1-\alpha-\beta)\lambda_2^{x_j}e^{-\lambda_2}}{x_j!}\right)$$

מה שחסר לנו הם הנתונים לאיזו קבוצה X_j שייך.

Screenshot

שלב E :

נחשב את הפונקציה $Q=\underbrace{E_{\alpha^{(k)},\beta^{(k)},\lambda_1^{(k)},\lambda_2^{(k)}}}_{}\left[l(X,Z,\alpha,\beta,\lambda_1,\lambda_2)\mid h_i\right]$

בתנאי בשלב $\theta^{(k)}$

$$Q=E_{\theta^{(k)}}\left[\sum_{j=1}^{1500}\mathbb{1}\{Z_j=1\}\cdot\ln\alpha+\mathbb{1}\{Z_j=2\}\cdot\ln\left(\frac{\beta\lambda_1^{x_j}e^{-\lambda_1}}{x_j!}\right)+\mathbb{1}\{Z_j=3\}\cdot\ln\left(\frac{(1-\alpha-\beta)\lambda_2^{x_j}e^{-\lambda_2}}{x_j!}\right)\Big| h_i\right]$$

המשתנה בעצם היחיד שיש הוא Z_j וכל שאר הפרמטרים הם קבועים.

לכן בעבור התוחלת צריך לחשב רק את $E[\mathbb{1}\{Z_j=1/2/3\}\mid h_i]$ וזה שווה להסתברות בהינתן x_j

$$E_{\theta^{(k)}}[\mathbb{1}\{z_j=1\}|x_j] = P(z_j=1|x_j) =$$

$$\frac{\alpha^{k}\cdot\mathbb{1}(x_j=0)}{\alpha^{(k)}\mathbb{1}\{x_j=0\}+\frac{\beta^{(k)}\lambda_1^{(k)x_j}e^{-\lambda_1^{(k)}}}{x_j!}+\frac{1-\alpha^{(k)}-\beta^{(k)}\cdot\lambda_2^{(k)x_j}e^{-\lambda_2^{(k)}}}{x_j!}} \equiv C_{1j}^{(k)}$$

$$E_{\theta^{(k)}}[\mathbb{1}\{z_j=2\}|x_j] = P(z_j=2|x_j) =$$

$$\frac{\frac{\beta^{(k)}\lambda_1^{(k)x_j}e^{-\lambda_1^{(k)}}}{x_j!}}{\alpha^{(k)}\mathbb{1}\{x_j=0\}+\frac{\beta^{(k)}\lambda_1^{(k)x_j}e^{-\lambda_1^{(k)}}}{x_j!}+\frac{1-\alpha^{(k)}-\beta^{(k)}\cdot\lambda_2^{(k)x_j}e^{-\lambda_2^{(k)}}}{x_j!}} \equiv C_{2j}^{(k)}$$

$$E_{\theta^{(k)}}[\mathbb{1}\{z_j=3\}|x_j] = P(z_j=3|x_j) =$$

$$\frac{\frac{1-\alpha^{(k)}-\beta^{(k)}\cdot\lambda_2^{(k)x_j}e^{-\lambda_2^{(k)}}}{x_j!}}{\alpha^{(k)}\mathbb{1}\{x_j=0\}+\frac{\beta^{(k)}\lambda_1^{(k)x_j}e^{-\lambda_1^{(k)}}}{x_j!}+\frac{1-\alpha^{(k)}-\beta^{(k)}\cdot\lambda_2^{(k)x_j}e^{-\lambda_2^{(k)}}}{x_j!}} \equiv C_{3j}^{(k)}$$

$$Q=\sum_{j=1}^{1500}C_{1j}^{(k)}\ln(\alpha)+\sum_{j=1}^{1500}C_{2j}^{(k)}\cdot\left(\ln(\beta)+x_j\ln\lambda_1-\lambda_1-\ln(x_j!)\right)+\sum_{j=1}^{1500}C_{3j}^{(k)}\cdot\left(\ln(1-\alpha-\beta)+x_j\ln\lambda_2-\lambda_2-\ln(x_j!)\right)$$

שלב M - נרצה למקסם את Q לפי $\alpha, \beta, \lambda_1, \lambda_2$

$$\frac{\partial Q}{\partial\alpha}=\sum_{j=1}^{1500}\frac{C_{1j}^{(k)}}{\alpha}-\sum_{j=1}^{1500}\frac{C_{3j}^{(k)}}{1-\alpha-\beta}=0 \Rightarrow (1-\alpha-\beta)\cdot\sum_{j=1}^{1500}C_{1j}^{(k)}-\alpha\sum_{j=1}^{1500}C_{3j}^{(k)}=0$$

$$\Rightarrow \alpha^{k+1}=\frac{(1-\beta^{k+1})\sum_{j=1}^{1500}C_{1j}^{(k)}}{\sum_{j=1}^{1500}C_{1j}^{(k)}+C_{3j}^{(k)}}$$

$$\frac{\partial^2 Q}{\partial\alpha^2}=\sum\frac{-C_{1j}^{(k)}}{\alpha^2}-\sum\frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2}<0 \quad ,\frac{\partial^2 Q}{\partial\alpha\partial\lambda_1}=0 \quad ,\frac{\partial^2 Q}{\partial\alpha\partial\lambda_2}=0 \quad ,\frac{\partial^2 Q}{\partial\alpha\partial\beta}=-\sum\frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2}$$

$$\frac{\partial Q}{\partial \beta} = \sum_{j=1}^{1500} \frac{C_{2j}^{(k)}}{\beta} - \sum \frac{C_{3j}^{(k)}}{1-\alpha-\beta} = 0 \qquad (1-\alpha-\beta)\cdot\sum_{j=1}^{1500} C_{2j}^{(k)} - \beta \sum_{j=1}^{1500} C_{3j}^{(k)} = 0$$

$$\beta^{k+1} = \frac{\left(1-\alpha^{k+1}\right)\sum_{j=1}^{1500} C_{2j}^{(k)}}{\sum_{j=1}^{1500} C_{2j}^{(k)} + C_{3j}^{(k)}}$$

$$\frac{\partial^2 Q}{\partial^2 \beta} = \sum_{j=1}^{1500} \frac{-C_{2j}^{(k)}}{\beta^2} - \sum \frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2} < 0 \qquad \frac{\partial^2 Q}{\partial \beta \partial \lambda_1} = 0 \qquad \frac{\partial^2 Q}{\partial \beta \partial \lambda_2} = 0 \qquad \frac{\partial^2 Q}{\partial \beta \partial \alpha} = -\sum \frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2}$$

$$\frac{\partial Q}{\partial \lambda_1} = \sum_{j=1}^{1500} \frac{C_{2j}^{(k)} X_j}{\lambda_1} - \sum_{j=1}^{1500} C_{2j}^{(k)} = 0 \qquad \lambda_1^{k+1} = \frac{\sum_{j=1}^{1500} C_{2j}^{(k)} X_j}{\sum_{j=1}^{1500} C_{2j}^{(k)}}$$

$$\frac{\partial^2 Q}{\partial^2 \lambda_1} = -\sum \frac{C_{2j}^{(k)} X_j}{\lambda_1^2} < 0 \qquad \frac{\partial^2 Q}{\partial \lambda_1 \partial \alpha} = 0 \qquad \frac{\partial^2 Q}{\partial \lambda_1 \partial \beta} = 0 \qquad \frac{\partial^2 Q}{\partial \lambda_1 \partial \lambda_2} = 0$$

$$\frac{\partial Q}{\partial \lambda_2} = \sum_{j=1}^{1500} \frac{C_{3j}^{(k)} X_j}{\lambda_2} - \sum_{j=1}^{1500} C_{3j}^{(k)} = 0 \qquad \lambda_2^{k+1} = \frac{\sum_{j=1}^{1500} C_{3j}^{(k)} X_j}{\sum_{j=1}^{1500} C_{3j}^{(k)}}$$

$$\frac{\partial^2 Q}{\partial^2 \lambda_2} = -\sum \frac{C_{3j}^{(k)} X_j}{\lambda_2^2} < 0 \qquad \frac{\partial^2 Q}{\partial \lambda_1 \partial \alpha} = 0 \qquad \frac{\partial^2 Q}{\partial \lambda_1 \partial \beta} = 0 \qquad \frac{\partial^2 Q}{\partial \lambda_1 \partial \lambda_2} = 0$$

שכינו הכסאן | כין

$$\begin{array}{cc} & \begin{array}{cccc} \alpha & \beta & \lambda_1 & \lambda_2 \end{array} \\ \begin{array}{c} \alpha \\ \beta \\ \lambda_1 \\ \lambda_2 \end{array} & \begin{pmatrix} -\sum \frac{C_{1j}^{(k)}}{\alpha^2} - \sum \frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2} & -\sum \frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2} & 0 & 0 \\ -\sum \frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2} & \sum_{j=1}^{1500} \frac{-C_{2j}^{(k)}}{\beta^2} - \sum \frac{C_{3j}^{(k)}}{(1-\alpha-\beta)^2} & 0 & 0 \\ 0 & 0 & -\sum \frac{C_{2j}^{(k)} X_j}{\lambda_1^2} & 0 \\ 0 & 0 & 0 & -\sum \frac{C_{3j}^{(k)} X_j}{\lambda_2^2} \end{pmatrix} \end{array}$$

ניתן לשים לב שנוסח השרשרת הני הינו מורכב קיים כי כל הנעלמים
בפועל יש תלויות ולכן נציב הנוסחאות גורמים שלו ונקבל
ה $\lambda_1^{k+1}, \lambda_2^{k+1}, \alpha^{k+1}, \beta^{k+1}$ המביאים למקסימום את Q.

$$\alpha^{k+1} = \frac{\left(1-\beta^{k+1}\right)\sum_{j=1}^{1500} C_{1j}^{(k)}}{\sum_{j=1}^{1500} C_{1j}^{(k)} + C_{3j}^{(k)}} =$$

$$\frac{\left(1 - \frac{\left(1-\alpha^{k+1}\right)\sum_{j=1}^{1500} C_{2j}^{(k)}}{\sum_{j=1}^{1500} C_{2j}^{(k)} + C_{3j}^{(k)}}\right)\sum C_{1j}^{(k)}}{\sum_{j=1}^{1500} C_{1j}^{(k)} + C_{3j}^{(k)}} = \frac{\frac{\left(\sum_{j=1}^{1500} \overset{C_2}{\overbrace{C_{2j}^{(k)}}} + \overset{C_3}{\overbrace{C_{3j}^{(k)}}} - \sum C_{2j}^{(k)} - \alpha^{k+1} \sum C_{2j}^{(k)}\right)\overset{C_1}{\overbrace{\sum C_{1j}^{(k)}}}}{\sum_{j=1}^{1500} C_{2j}^{(k)} + C_{3j}^{(k)}}}{\sum C_{1j}^{(k)} + C_{3j}^{(k)}} =$$

$$= \frac{\left(C_2 + C_3 - (1-\alpha) C_2\right) C_1}{\left(C_2 + C_3\right)\left(C_1 + C_3\right)} = \frac{C_1 C_2 + C_1 C_3 - C_2 C_1 \left(1-\alpha\right)^{k+1}}{\left(C_2 + C_3\right)\left(C_1 + C_3\right)} = \frac{C_1 C_2 + C_1 C_3 - C_2 C_1 (1-\alpha)^{k+1}}{C_2 C_1 + C_2 C_3 + C_3 C_1 + C_3 C_3}$$

$$\alpha^{k+1}\left(\cancel{C_2} C_1 + C_2 \cancel{C_3} + \cancel{C_3} C_1 + C_3 \cancel{C_3}\right) = \cancel{C_1 C_2} + C_1 \cancel{C_3} - \cancel{C_2 C_1} + \cancel{C_2 C_1 \alpha^{k+1}}$$

$$\alpha^{k+1} = \frac{C_1}{C_1 + C_2 + C_3}$$

$$C_1 + C_2 + C_3 = \sum_{j=1}^{1500} C_{1j}^{(k)} + C_{2j}^{(k)} + C_{3j}^{(k)} = \sum \frac{\alpha^{(k)} \mathbb{1}\{x_j = 0\} + \frac{\beta^{(k)} {\lambda_1^{(k)}}^{x_j} e^{-\lambda_1^{(k)}}}{x_j!} + \frac{\left(1-\alpha^{(k)}-\beta^{(k)}\right) {\lambda_2^{(k)}}^{x_j} e^{-\lambda_2^{(k)}}}{x_j!}}{\alpha^{(k)} \mathbb{1}\{x_j = 0\} + \frac{\beta^{(k)} {\lambda_1^{(k)}}^{x_j} e^{-\lambda_1^{(k)}}}{x_j!} + \frac{\left(1-\alpha^{(k)}-\beta^{(k)}\right) {\lambda_2^{(k)}}^{x_j} e^{-\lambda_2^{(k)}}}{x_j!}}$$

$$= \sum_{j=1}^{1500} 1 = 1500$$

$$\Rightarrow \alpha^{k+1} = \frac{\sum_{j=1}^{1500} C_{1j}^{(k)}}{1500}, \quad \beta^{k+1} = \frac{\sum_{j=1}^{1500} C_{2j}^{(k)}}{1500}$$

בסך הכל העדכונים הם

$$\alpha^{k+1} = \sum_{j=1}^{1500} \frac{C_{1j}^{(k)}}{1500} \qquad \beta^{k+1} = \sum_{j=1}^{1500} \frac{C_{2j}^{(k)}}{1500}$$

$$\lambda_1^{k+1} = \frac{\sum_{j=1}^{1500} C_{2j}^{(k)} x_j}{\sum_{j=1}^{1500} C_{2j}^{(k)}} \qquad \lambda_2^{k+1} = \frac{\sum_{j=1}^{1500} C_{3j}^{(k)} x_j}{\sum_{j=1}^{1500} C_{3j}^{(k)}}$$

```
alpha_k1_numerator <- function(vec){
  lambda1.k = vec[1]
  lambda2.k = vec[2]
  alpha.k = vec[3]
  beta.k = vec[4]
  M1j.k = (alpha.k) / ((alpha.k)+(beta.k*exp(-lambda1.k))+((1-alpha.k-beta.k)*exp(-la
mbda2.k)))
  SUM.M1j.k = ex7q1[1,2] * M1j.k
  return(SUM.M1j.k)
}
```

```
beta_k1_numerator <- function(vec){
  lambda1.k = vec[1]
  lambda2.k = vec[2]
  alpha.k = vec[3]
  beta.k = vec[4]
  M2.k = rep(NA,17)
  for (i in 1:17){
    M2.k[i] = (beta.k*(lambda1.k^ex7q1[i,1])*exp(-lambda1.k))/
    ((alpha.k*as.numeric(ex7q1[i,1]==0))+(beta.k*(lambda1.k^ex7q1[i,1])*
    exp(-lambda1.k))+((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k)))
  }
  SUM.M2j.k = sum(ex7q1[,2]*M2.k)
  return(SUM.M2j.k)
}
```

```
lambda1_k1_denominator <- function(vec){
  lambda1.k = vec[1]
  lambda2.k = vec[2]
  alpha.k = vec[3]
  beta.k = vec[4]
  M2.k = rep(NA,17)
  for (i in 1:17){
    M2.k[i] = (beta.k*(lambda1.k^ex7q1[i,1])*exp(-lambda1.k))/
    ((alpha.k*as.numeric(ex7q1[i,1]==0))+(beta.k*(lambda1.k^ex7q1[i,1])*
    exp(-lambda1.k))+((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k)))
  }
  SUM.M2j.k = sum(ex7q1[,2]*M2.k)
  return(SUM.M2j.k)
}
```

```
lambda2_k1_denominator <- function(vec){
  lambda1.k = vec[1]
  lambda2.k = vec[2]
  alpha.k = vec[3]
  beta.k = vec[4]
  M3.k = rep(NA,17)
  for (i in 1:17){
    M3.k[i] = ((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k))/
    ((alpha.k*as.numeric(ex7q1[i,1]==0))+(beta.k*(lambda1.k^ex7q1[i,1])*
    exp(-lambda1.k))+((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k)))
  }
  SUM.M3j.k = sum(ex7q1[,2]*M3.k)
  return(SUM.M3j.k)
}
```

```
lambda1_k1_numerator <- function(vec){
  lambda1.k = vec[1]
  lambda2.k = vec[2]
  alpha.k = vec[3]
  beta.k = vec[4]
  M2.k = rep(NA,17)
  for (i in 1:17){
    M2.k[i] = (beta.k*(lambda1.k^ex7q1[i,1])*exp(-lambda1.k))/
    ((alpha.k*as.numeric(ex7q1[i,1]==0))+(beta.k*(lambda1.k^ex7q1[i,1])*
    exp(-lambda1.k))+((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k)))
  }
  M2.k_Xj = rep(NA,17)
  for(i in 1:17){
    M2.k_Xj[i] = (ex7q1[i,1]*M2.k[i])
  }
  SUM.M2j.k_Xj = sum(ex7q1[,2]*M2.k_Xj)
  return(SUM.M2j.k_Xj)
}
```

```
lambda2_k1_numerator <- function(vec){
  lambda1.k = vec[1]
  lambda2.k = vec[2]
  alpha.k = vec[3]
  beta.k = vec[4]
  M3.k = rep(NA,17)
  for (i in 1:17){
    M3.k[i] = ((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k))/
    ((alpha.k*as.numeric(ex7q1[i,1]==0))+(beta.k*(lambda1.k^ex7q1[i,1])*
    exp(-lambda1.k))+((1-alpha.k-beta.k)*(lambda2.k^ex7q1[i,1])*exp(-lambda2.k)))
  }
  M3.k_Xj = rep(NA,17)
  for(i in 1:17){
    M3.k_Xj[i] = (ex7q1[i,1]*M3.k[i])
  }
  SUM.M3j.k_Xj = sum(ex7q1[,2]*M3.k_Xj)
  return(SUM.M3j.k_Xj)
}
```

```
EM_algorithm <- function(epsilon, start){
  theta.k = start
  lambda1.kplus1 =  lambda1_k1_numerator(vec = theta.k) / lambda1_k1_denominator(vec
 = theta.k)
  lambda2.kplus1 =  lambda2_k1_numerator(vec = theta.k) / lambda2_k1_denominator(vec
 = theta.k)
  alpha.kplus1 = alpha_k1_numerator(vec = theta.k) / 1500
  beta.kplus1 = beta_k1_numerator(vec = theta.k) / 1500
  theta.kplus1 = c(lambda1.kplus1, lambda2.kplus1, alpha.kplus1, beta.kplus1)
  while(norm(theta.kplus1-theta.k, type = "2")>epsilon)
  {
    theta.k = theta.kplus1
    lambda1.kplus1 =  lambda1_k1_numerator(vec = theta.k) / lambda1_k1_denominator(ve
c = theta.k)
    lambda2.kplus1 =  lambda2_k1_numerator(vec = theta.k) / lambda2_k1_denominator(ve
c = theta.k)
    alpha.kplus1 = alpha_k1_numerator(vec = theta.k) / 1500
    beta.kplus1 = beta_k1_numerator(vec = theta.k) / 1500
    theta.kplus1 = c(lambda1.kplus1, lambda2.kplus1, alpha.kplus1, beta.kplus1)
  }

  return(cat(sprintf("alpha (EM) = %s", round(theta.kplus1[3],5)),
             sprintf("beta (EM) = %s", round(theta.kplus1[4],5)),
             sprintf("lambda1 (EM) = %s", round(theta.kplus1[1],5)),
             sprintf("lambda2 (EM) = %s", round(theta.kplus1[2],5)),
             sep = "\n"))
}
```

```
result <- EM_algorithm(epsilon = 0.0001,start = c(6.736,11.85,1/3,1/3))
```

```
## alpha (EM) = 0.12226
## beta (EM) = 0.56254
## lambda1 (EM) = 1.46809
## lambda2 (EM) = 5.93966
```

Q2

a

שאלה 2

$X_i \sim N(\mu_1, \sigma^2) \quad Y_i \sim N(\mu_2, \sigma^2) \quad i=1\dots 10000$

$O_i = \max(X_i, Y_i) \quad i=1\dots 1000$

$i=1\dots m \quad O_i = X_i \quad \Leftarrow \quad X_i \geq Y_i$ ש מקרים $i=1\dots m$ נניח

$i=m+1,\dots,10000 \quad O_i = Y_i \quad \Leftarrow \quad Y_i > X_i$ ל מקרים $i=m+1\dots 10000$

א.

$$f_O(t) = P(O_i = t) = \begin{cases} P(X_i = t)\cdot P(Y_i \le t) & i=1\dots m \\ P(Y_i = t)\cdot P(X_i < t) & i=m+1\dots n \end{cases} =$$

$$= \begin{cases} f_{X_i}(x_i)\cdot F_Y(x_i) & i=1\dots m \\ f_{Y_i}(y_i)\cdot F_X(y_i) & i=m+1\dots n \end{cases}$$

$$L(X,Y,\mu_1,\mu_2) = \prod_{i=1}^{m} f_X(x_i)\cdot F_Y(x_i)\cdot \prod_{i=m+1}^{n} f_Y(y_i)\cdot F_X(y_i)$$

$$\ell = \sum_{i=1}^{m} \ln f_X(x_i) + \ln F_Y(x_i) + \sum_{i=m+1}^{n} \ln f_Y(y_i) + \ln F_X(y_i)$$

$$\frac{\partial \ell}{\partial \mu_1} = \sum_{i=1}^{m} \frac{\partial \ln f_X(x_i)}{\partial \mu_1} + \frac{\partial \ln F_Y(x_i)}{\partial \mu_1}$$

$$\frac{\partial \ln f_X(x_i)}{\partial \mu_1} = \frac{\partial \ln\left(\sqrt{2\pi\sigma^2}\cdot e^{-\frac{(x_i-\mu_1)^2}{2\sigma^2}}\right)}{\partial \mu_1} = \frac{\partial}{\partial \mu_1} - \frac{(x_i-\mu_1)^2}{2\sigma^2} = \frac{x_i-\mu_1}{\sigma^2}$$

$$\frac{\partial \ln F_Y(x_i)}{\partial \mu_1} = \frac{\frac{\partial}{\partial \mu_1} F_X(y_i)}{F_Y(x_i)} = \frac{\frac{\partial}{\partial \mu_1}\int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-\frac{(t-\mu_1)^2}{2\sigma^2}}\,dt}{F_Y(x_i)} = \begin{Bmatrix} s = \frac{t-\mu_1}{\sigma} \Rightarrow t = \mu_1 + s\sigma \\ ds = \frac{1}{\sigma}dt \end{Bmatrix}$$

$$= \frac{\frac{\partial}{\partial \frac{y_i-\mu_1}{\sigma}} \int_{-\infty}^{\frac{y_i-\mu_1}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}\,ds \cdot \frac{\partial \frac{y_i-\mu_1}{\sigma}}{\partial \mu_1}}{F_Y(x_i)} \underset{\downarrow}{=} \frac{-\frac{1}{\sigma}\cdot\frac{1}{\sqrt{2\pi}} e^{-\frac{(y_i-\mu_1)^2}{2\sigma^2}}}{F_Y(x_i)} = \frac{-f_X(y_i)}{F_Y(x_i)}$$

נשתמש בלייבניץ h תחתון

$$\frac{\partial \ell}{\partial \mu_1} = \sum_{i=1}^{m} \frac{x_i-\mu_1}{\sigma^2} - \sum_{i=1}^{n} \frac{f_X(O_i)}{F_Y(O_i)} = S_1 \quad \Rightarrow \quad \frac{\partial \ell}{\partial \mu_2} = \sum_{i=m+1}^{n} \frac{y_i-\mu_2}{\sigma^2} - \sum_{i=m+1}^{n} \frac{f_Y(O_i)}{F_X(O_i)} = S_2$$

לא ניתן לבודד את μ_1 ו μ_2 ולכן נפתור את המשוואות בעזרת שיטת NR.

את הנגזרות החלקיות נחשב באופן נומרי:

$$\frac{\partial S_i}{\partial \mu_j} = \frac{S_i(\mu_j+d) - S_i(\mu_j+d)}{2d}$$

ועדכון NR

$$\begin{pmatrix} \mu_1^{j+1} \\ \mu_2^{j+1} \end{pmatrix} = \begin{pmatrix} \mu_1^{(j)} \\ \mu_2^{(j)} \end{pmatrix} - H\left(\mu_1^{(j)}, \mu_2^{(j)}\right)\cdot\begin{pmatrix} S_1(\mu_1^{(j)}) \\ S_2(\mu_2^{(j)}) \end{pmatrix}$$

```
# ex6data1 <- read.csv("C:/Users/Alon/Desktop/Studies/Statistics/Statistical_Computin
g/Exercises/HW6/ex6data1.csv")

ex7q2 <- read.csv("~/Desktop/Ran/D year/semester b/hishov statisti/exercies/HW7/ex7q
2.csv")
```

```
grad_s1 <- function(mu1,sigma,o_vec,m){

  n <- length(o_vec)
  sum((o_vec[1:m])-mu1)/(sigma^2) - sum(dnorm(o_vec[(m+1):n],mu1,sigma)/pnorm(o_vec
[(m+1):n],mu1,sigma))
}


grad_s2 <- function(mu2,sigma,o_vec,m){

  n <- length(o_vec)
  sum((o_vec[(m+1):n])-mu2)/(sigma^2) - sum(dnorm(o_vec[1:m],mu2,sigma)/pnorm(o_vec[1
:m],mu2,sigma))
}
```

```
o_vec <- ex7q2$o

m <- 7659
sigma <- 2

mu1.j <- 0
mu2.j <- 0
mu1.j_1 <- 1
mu2.j_1 <- 1
eps_for_break <- 10^-5
while(norm(matrix(c(mu1.j,mu2.j),nrow = 2)-matrix(c(mu1.j_1,mu2.j_1),nrow = 2),type =
"2")>eps_for_break){
  mu1.j <- mu1.j_1
  mu2.j <- mu2.j_1
  s1 <- grad_s1(mu1.j,sigma,o_vec,m)
  s2 <- grad_s2(mu2.j,sigma,o_vec,m)
  h11 <- (grad_s1(mu1.j+0.000001,sigma,o_vec,m)-grad_s1(mu1.j-0.000001,sigma,o_vec,
m))/(2*0.000001)
  h22 <- (grad_s2(mu2.j+0.000001,sigma,o_vec,m)-grad_s2(mu2.j-0.000001,sigma,o_vec,
m))/(2*0.000001)

  H <- matrix(c(h11,0,0,h22),nrow = 2)
  mu1.j_1 <- (c(mu1.j,mu2.j) - solve(H)%*%c(s1,s2))[1]
  mu2.j_1 <- (c(mu1.j,mu2.j) - solve(H)%*%c(s1,s2))[2]
}
```

```
## mu1 (NR) = 4.06582
## mu2 (NR) = 1.99707
```

b

ב. בניסוי המלא

$$L(x,y,\mu_1,\mu_2) = \prod_{i=1}^{n} f_{X_i}(x) \cdot \prod_{i=1}^{n} f_{Y_i}(y) =$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{2\sigma^2}} \cdot \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{\sum_{i=1}^{n}(y_i-\mu_2)^2}{2\sigma^2}} =$$

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{2n} \cdot e^{-\left(\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2 + \sum_{i=1}^{n}(y_i-\mu_2)^2}{2\sigma^2}\right)}$$

לוג בניסוי המלא

$$\ell(x,y,\mu_1,\mu_2) = -2n\ln\left(\sqrt{2\pi\sigma^2}\right) - \frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{2\sigma^2} - \frac{\sum_{i=1}^{n}(y_i-\mu_2)^2}{2\sigma^2}$$

$$= -2n\ln\left(\sqrt{2\pi\sigma^2}\right) - \frac{\sum x_i^2}{2\sigma^2} + \frac{\mu_1\sum x_i}{\sigma^2} - \frac{n\mu_1^2}{2\sigma^2} - \frac{\sum y_i^2}{2\sigma^2} + \frac{\mu_2\sum y_i}{\sigma^2} - \frac{n\mu_2^2}{2\sigma^2}$$

$$Q(\theta,\theta^{(j)}) = E\left[\ell(x,y,\theta) \mid \underline{O};\theta^{(j)}\right] \qquad \theta = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$$

נחשב את התוחלות המותנות שתלויות ב μ_1 ו μ_2 כי רק הן משפיעות בגזירה

$$\underline{\underline{x_i^{(j)}}} = E[X_i \mid O_i] = \begin{cases} x_i & i=1\ldots m \\ E[X_i \mid Y=y_i] & i=m+1\ldots n \end{cases} = \begin{cases} x_i & i=1\ldots m \\ \mu_1^{(j)} - \sigma\dfrac{\phi\left(\frac{y_i-\mu_1^{j}}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu_1^{j}}{\sigma}\right)} & i=m+1\ldots n \end{cases}$$

$$\underline{\underline{y_i^{(j)}}} = E[Y_i \mid O_i] = \begin{cases} E[Y_i \mid X=x_i] & i=1\ldots m \\ y_i & i=m+1\ldots n \end{cases} = \begin{cases} \mu_2^{(j)} + \sigma\dfrac{\phi\left(\frac{x_i-\mu_2^{j}}{\sigma}\right)}{\Phi\left(\frac{x_i-\mu_2^{j}}{\sigma}\right)} & i=1\ldots m \\ y_i & i=m+1\ldots n \end{cases}$$

$$Q(\theta,\theta^{(j)}) = -2n\ln\left(\sqrt{2\pi\sigma^2}\right) - \frac{\sum E[X_i^2 \mid O_i]}{2\sigma^2} + \frac{\mu_1\sum E[X_i \mid O_i]}{\sigma^2} - \frac{n\mu_1^2}{2\sigma^2}$$
$$- \frac{\sum E[Y_i^2 \mid O_i]}{2\sigma^2} + \frac{\mu_2\sum E[Y_i \mid O_i]}{\sigma^2} - \frac{n\mu_2^2}{2\sigma^2} \propto$$

(נשים לב שרק אלו שתלויים ב μ_1 ו μ_2 משפיעים בגזירה)

$$\propto \frac{\mu_1\sum x_i^{(j)}}{\sigma^2} - \frac{n\mu_1^2}{2\sigma^2} + \frac{\mu_2\sum y_i^{(j)}}{\sigma^2} - \frac{n\mu_2^2}{2\sigma^2}$$

$$\frac{\partial Q}{\partial\mu_1} = \frac{\sum x_i^{(j)}}{\sigma^2} - \frac{n\mu_1}{\sigma^2} \qquad \frac{\partial^2 Q}{\partial^2\mu_1} = -\frac{n}{\sigma^2} < 0 \qquad \frac{\partial^2 Q}{\partial\mu_1\partial\mu_2} = 0$$

$$\frac{\partial Q}{\partial\mu_2} = \frac{\sum y_i^{(j)}}{\sigma^2} - \frac{n\mu_2}{\sigma^2} \qquad \frac{\partial^2 Q}{\partial^2\mu_2} = -\frac{n}{\sigma^2} < 0 \qquad \frac{\partial^2 Q}{\partial\mu_2\partial\mu_1} = 0$$

$$\hat{\mu}_1^{(j+1)} = \frac{\sum x_i^{(j)}}{n} \qquad \mu_2^{(j+1)} = \frac{\sum y_i^{(j)}}{n}$$

נראה את מטריצת ההסיאן $\begin{pmatrix} -n/\sigma^2 & 0 \\ 0 & -n/\sigma^2 \end{pmatrix}$ מוגדרת שלילית ולכן קיבלנו מקסימום

```
# ex6data1 <- read.csv("C:/Users/Alon/Desktop/Studies/Statistics/Statistical_Computin
g/Exercises/HW6/ex6data1.csv")

ex7q2 <- read.csv("~/Desktop/Ran/D year/semester b/hishov statisti/exercies/HW7/ex7q
2.csv")
```

```
x_j_sum <- function(mu1_j,sigma,m,o_vec){
  n <- length(o_vec)
  sum.xj <-(sum(o_vec[1:m]) + sum((mu1_j - sigma * (dnorm( (o_vec-mu1_j) / sigma) /
                                  pnorm ( (o_vec-mu1_j) / sigma)))[(m+1):n]))
  return(sum.xj)
}

y_j_sum <- function(mu2_j,sigma,m,o_vec){
  n <- length(o_vec)
  sum.yj <-(sum((mu2_j - sigma * (dnorm( (o_vec-mu2_j) / sigma) /
                                  (pnorm ( (o_vec-mu2_j) / sigma))))[1:m]) + sum(o_v
ec[(m+1):n]))
  return(sum.yj)
}
```

```
o_vec.Q2 <- ex7q2$o

mu1.j <- mu2.j <- mean(o_vec.Q2) # first guess

m <- 7659
sigma <- 2
n <- length(o_vec.Q2 )

eps_for_break <- 10^-5

mus_j_1 <- c(mu1.j,mu2.j)

while (TRUE) {
  mus_j <- mus_j_1
  mu1.j_1 <- x_j_sum(mus_j[1],sigma,m,o_vec.Q2) / n
  mu1.j_2 <- y_j_sum(mus_j[2],sigma,m,o_vec.Q2) / n
  mus_j_1 <- c(mu1.j_1,mu1.j_2)

  if (max(abs(mus_j_1 - mus_j)) < eps_for_break) {break}
}

mu1.j_1
```

```
## [1] 4.065818
```

```
mu1.j_2
```

```
## [1] 1.99708
```

```
## mu1 (EM) = 4.06582
## mu2 (EM) = 1.99707
```

We can see that both algorithm gave us the same result.

C

$$L(O,\mu_1,\mu_2)=\prod_{i=1}^{n} f_{O_i}(o_i)=\prod_{i=1}^{m}\frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-\left(\frac{(o_i-\mu_1)^2}{2\sigma^2}\right)}\cdot\prod_{i=m+1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\left(\frac{(o_i-\mu_2)^2}{2\sigma^2}\right)}=$$

$$=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n\cdot e^{-\frac{\sum_{i=1}^{m}(o_i-\mu_1)^2}{2\sigma^2}}\cdot e^{-\frac{\sum_{i=m+1}^{n}(o_i-\mu_2)^2}{2\sigma^2}}$$

$$\ell(O,\mu_1,\mu_2)=-n\ln\sqrt{2\pi\sigma^2}-\frac{\sum_{i=1}^{m}(o_i-\mu_1)^2}{2\sigma^2}-\frac{\sum_{i=m+1}^{n}(o_i-\mu_2)^2}{2\sigma^2}$$

$$\frac{\partial\ell}{\partial\mu_1}=\frac{\sum_{i=1}^{m}o_i-m\mu_1}{\sigma^2}\qquad\frac{\partial^2\ell}{\partial^2\mu_1}=-\frac{m}{\sigma^2}<0\qquad\frac{\partial^2\ell}{\partial\mu_1\partial\mu_2}=0$$

$$\frac{\partial\ell}{\partial\mu_2}=\frac{\sum_{i=m+1}^{n}o_i-(n-m)\mu_2}{\sigma^2}\qquad\frac{\partial^2\ell}{\partial^2\mu_2}=-\frac{(n-m)}{\sigma^2}<0\qquad\frac{\partial^2\ell}{\partial\mu_2\partial\mu_1}=0$$

$$\hat{\mu}_1=\frac{\sum_{i=1}^{n}o_i}{m}\qquad\hat{\mu}_2=\frac{\sum_{i=m+1}^{n}o_i}{n-m}$$

נשים לב שההסיאן הוא $\begin{pmatrix}-\frac{m}{\sigma^2} & 0\\ 0 & -\frac{(n-m)}{\sigma^2}\end{pmatrix}$ מוגדרת שלילית כי אלכסון קיבלנו מקסימום

```
mu1_mle <- sum(o_vec[1:m])/m
mu2_mle <- sum(o_vec[(m+1):10000])/(10000-m)
```

```
## mu1 (MLE) = 4.63491
## mu2 (MLE) = 3.85768
```

We see that the estimates are skewed upwards.

We know that we saw only the maximum results from the two different distribution.

Since we only see the observations with the relatively high values, it is observed that the estimators will be tilted upwards.