HW7

Zeqiu.Yu

2022-11-14

#Homework 7
##4.2

```r
input1 <- read.table("./hivrisk.dat", header = TRUE)
input1
```

```
##    encounters frequency
## 1           0       379
## 2           1       299
## 3           2       222
## 4           3       145
## 5           4       109
## 6           5        95
## 7           6        73
## 8           7        59
## 9           8        45
## 10          9        30
## 11         10        24
## 12         11        12
## 13         12         4
## 14         13         2
## 15         14         0
## 16         15         1
## 17         16         1
```

(a.)
According to the question, we get following vectors:
$x = (n_0, n_1, ..., n_{16})$, $y = (n_{z,0},\ \ n_{t,0}, n_{t,1}, ...n_{t,16},\ \ n_{p,0}, n_{p,1}, ...n_{p,16})$ with $\theta_t = (\alpha^{(t)}, \beta^{(t)}, \mu^{(t)}, \lambda^{(t)})$.
x is the observation, and y is the complete data. The mapping from the complete data to the observed data is:
$x = M(y) = (n_{z,0} + n_{t,0} + n_{p,0},\ \ n_{t,1} + n_{p,1},\ \ ...\ n_{t,16} + n_{p,16})$.
Now, the likelihood function of X is given as follows:

$$L(\theta|x) \propto \Pi_{i=0}^{16}[\frac{\pi_i(\theta)}{i!}]^{n_i}$$

We can then get the complete likelihood function using multinomial as follows:

$$L(\theta|Y) = (^{N}_{n_{z,0},\ \ n_{t,0}, n_{t,1}, ...n_{t,16},\ \ n_{p,0}, n_{p,1}, ...n_{p,16}})\alpha^{n_{z,0}}\{\Pi_{i=0}^{16} p_{t,i}^{n_{t,i}}\}\{\Pi_{i=0}^{16} p_{p,i}^{n_{p,i}}\}$$

$p_{t,i} = \beta\frac{\mu^i}{i!}e^{-\mu}$, $p_{p,i} = (1-\alpha-\beta)\frac{\lambda^i}{i!}e^{-\lambda}$,
After taking the logarithm and getting log-likelihood function of the complete:

$$logL(\theta|Y) = n_{z,0}log(\alpha) + \sum_{i=0}^{16} n_{t,i}logp_{t,i} + \sum_{i=0}^{16} n_{p,i}logp_{p,i} + log(^{N}_{n_{z,0},\ n_{t,0},n_{t,1},...n_{t,16},\ n_{p,0},n_{p,1},...n_{p,16}})$$

.

W.L.O.G, we calculate the expectation of $n_{z,0}$, $n_{t,i}$ and $n_{p,i}$ with respect to $x, \theta^{(t)}$.

$\hat{n}_{z,0} = E\{n_{z,0}|x, \theta^{(t)}\} = n_0 \frac{\alpha^{(t)}}{\pi_0(\theta^{(t)})}$

$\hat{n}_{t,i} = E\{n_{t,i}|x, \theta^{(t)}\} = n_i \frac{\beta^{(t)} \frac{\mu^{(t)i}}{i!} e^{-\mu^{(t)}}}{\pi_i(\theta^{(t)})}$

$\hat{n}_{p,i} = E\{n_{p,i}|x, \theta^{(t)}\} = n_i \frac{(1-\alpha^{(t)}-\beta^{(t)}) \frac{\lambda^{(t)i}}{i!} e^{-\lambda^{(t)}}}{\pi_i(\theta^{(t)})}$

Input these values, we have:

$$Q(\theta|\theta^t) = \hat{n}_{z,0}log(\alpha) + \sum_{i=0}^{16} \hat{n}_{t,i}logp_{t,i} + \sum_{i=0}^{16} \hat{n}_{p,i}logp_{p,i} + \eta(\alpha + \beta + \gamma - 1)$$

Hence, take the partial derivatives and with repect to $\alpha, \beta, \eta, \gamma$, and set these derivatives to be zero. We have the following results:

$$\eta = N$$

$$\hat{\alpha}^{(t+1)} = \frac{n_{z,0}}{N} = \frac{n_0 z_0(\theta^{(t)})}{N}$$

$$\hat{\beta}^{(t+1)} = \frac{\sum_{i=0}^{16} n_{t,i}}{N} = \frac{\sum_{i=0}^{16} n_i t_i(\theta^{(t)})}{N}$$

$$\hat{\eta}^{(t+1)} = \frac{\sum_{i=0}^{16} i \times n_{t,i}}{\sum_{i=0}^{16} n_{t,i}} = \frac{\sum_{i=0}^{16} i n_i t_i(\theta^{(t)})}{\sum_{i=0}^{16} n_i t_i(\theta^{(t)})}$$

$$\hat{\eta}^{(t+1)} = \frac{\sum_{i=0}^{16} i \times n_{p,i}}{\sum_{i=0}^{16} n_{p,i}} = \frac{\sum_{i=0}^{16} i n_i p_i(\theta^{(t)})}{\sum_{i=0}^{16} n_i p_i(\theta^{(t)})}$$

(b.)

```
pi_function <- function(index = 0, parameter_vector ){
  # parameter_vector: c(alpha, beta, mu, lambda)
  alpha = parameter_vector[1]
  beta = parameter_vector[2]
  mu = parameter_vector[3]
  lambda = parameter_vector[4]
  if(index == 0){
    return(alpha + beta*exp(-mu) + (1-alpha-beta)*exp(-lambda))
  }else{
    return(beta*mu^index*exp(-mu) + (1-alpha-beta)*lambda^index*exp(-lambda))
  }
}

alpha.iteration <- function(input = input1, initial_param){
  i = 0
  alpha_t1 <- initial_param[1]
  N <-  1500
  n_0 <-  input$frequency[i+1]
  alpha_t2 <- n_0/N*alpha_t1/pi_function(index = i, initial_param)
  return(alpha_t2)
}
```

```
beta.iteration <- function(input = input1, initial_param){
  alpha_t1 <- initial_param[1]
  beta_t1 <- initial_param[2]
  mu_t1 <- initial_param[3]
  lambda_t1 <- initial_param[4]
  N <-  1500
  beta_t2 <- c()
  for(i in 0:dim(input)[1]-1){
    n_i =  input$frequency[i+1]
    beta_t2[i+1] <-n_i/N*beta_t1*mu_t1^i*exp(-mu_t1)/pi_function(index = i, initial_param)
  }
  return(sum(beta_t2))

}

mu.iteration <- function(input = input1, initial_param){
  alpha_t1 <- initial_param[1]
  beta_t1 <- initial_param[2]
  mu_t1 <- initial_param[3]
  lambda_t1 <- initial_param[4]
  N <-  1500
  numerator <- c()
  denoimnator <- c()
  for(i in 0:dim(input)[1]-1){
    n_i <-  input$frequency[i+1]
    numerator[i+1] <- (i*n_i*beta_t1*mu_t1^i*exp(-mu_t1)/pi_function(index = i, initial_param))
    denoimnator[i+1] <-  (n_i*beta_t1*mu_t1^i*exp(-mu_t1)/pi_function(index = i, initial_param))
  }
  return(sum(numerator)/sum(denoimnator))

}

lambda.iteration <- function(input = input1, initial_param){
  alpha_t1 <- initial_param[1]
  beta_t1 <- initial_param[2]
  mu_t1 <- initial_param[3]
  lambda_t1 <- initial_param[4]
  N <-  1500
  numerator <- c()
  denoimnator <- c()
  for(i in 0:dim(input)[1]-1){
    n_i <-  input$frequency[i+1]
    numerator[i+1] <-  (i*n_i*(1-alpha_t1 - beta_t1)*lambda_t1^i*exp(-lambda_t1)/pi_function(index = i,
    denoimnator[i+1] <- (n_i*(1-alpha_t1 - beta_t1)*lambda_t1^i*exp(-lambda_t1)/pi_function(index = i,
  }
  return(sum(numerator)/sum(denoimnator))

}
```

```
accuracy = 0.000001
initial_values <- c(0.33, 0.8, 5, 8)
while(TRUE){
```

```
  alpha_t2 <- alpha.iteration(input1, initial_values)
  beta_t2 <- beta.iteration(input1, initial_values)
  mu_t2 <- mu.iteration(input1, initial_values)
  lambda_t2 <- lambda.iteration(input1, initial_values)
  initial_values_before <- initial_values
  initial_values <- c(alpha_t2, beta_t2, mu_t2, lambda_t2)
  if(all(abs(initial_values- initial_values_before)<=accuracy) ){
    break;
  }else{
    next;
  }
}
print(sprintf("alpha:%.5f, beta:%.5f, mu:%.5f, lambda:%.5f, 1-alpha-beta: %.5f", initial_values[1], ini
```

```
## [1] "alpha:0.12216, beta:0.56254, mu:1.46747, lambda:5.93888, 1-alpha-beta: 0.31529"
```

(Bonus) (4.3)
(a.)
E-step:

$$x_{iM} = E(\hat{X_{iM}}|\theta^{(t)}, X_{i}\cancel{M}) = \mu_M + \Sigma^{-1}_{M\bar{M}}(X_{i\bar{M}} - \mu_{\bar{M}})$$

(b.)