Homework 1

Santiago Ruiz, Nikita Karetnikov, Felix Ubl

2025-04-18

1 Exercise 4

It is assumed that the lifetime of light bulbs follows an exponential distribution with parameter λ. To estimate λ, n light bulbs were tested until they all failed. Their failure times were recorded as $u_1, ..., u_n$. In a separate experiment, m bulbs were tested, but the individual failure times were not recorded. Only the number of bulbs, r, that had failed at time t was recorded. The missing data are the failure times of the bulbs in the second experiment, $v_1, ..., v_m$.

1.1 Determine the complete-data log-likelihood.

The complete data log-likelihood is defined as follows:

$$x \sim \text{Exp}(\lambda), \quad f(x) = \lambda e^{-\lambda x}$$

$$\text{Likelihood}(x) = \prod_{i=1}^{n+m} \lambda e^{-\lambda x_i}$$

$$\text{Log-Likelihood}(x) = \sum_{i=1}^{n+m} \log(\lambda e^{-\lambda x_i}) = \sum_{i=1}^{n+m} \log(\lambda) - \lambda x_i$$

1.2 Determine the conditional means

$$E[x \mid x \leq t, \lambda_0] = \int_{x=0}^{t} f(x) \cdot x \, dx = \int_{x=0}^{x=t} (\lambda e^{-\lambda x}) \cdot x \, dx$$

$$E[x \mid x \leq t, \lambda_0] = \frac{e^{-\lambda x}(\lambda x + 1)}{\lambda}\Big|_0^t$$

$$E[x \mid x \leq t, \lambda_0] = \frac{e^{-\lambda t}(\lambda t + 1) - 1}{\lambda}$$

Given that $E[x \mid x = t] = t \cdot Pois(k = r)$, the latter is the probability of a $Pois(\lambda)$ with the same parameter as the exponential distribution. Such probability is given by:

$$E[x \mid x = t] = t \cdot P(x = r) = \frac{\lambda^r e^{-\lambda}}{r!}$$

$$E[x \mid x > t, \lambda_0] = E[x \mid x \geq t, \lambda_0] - E[x \mid x = t, \lambda_0] = \frac{e^{-\lambda x}(\lambda x + 1)}{\lambda}\Big|_t^\infty - E[x \mid x = t]$$

$$E[x \mid x > t, \lambda_0] = -\frac{e^{-\lambda t}(\lambda t + 1)}{\lambda} - \frac{\lambda^r e^{-\lambda}}{r!}$$

Let us formulate the latent variable as $Y = I(x \leq t)$, so that when we take the expectation over the complete log-likelihood with respect to the conditional distribution of y given the observed data, and paramenter we have:

$$E[x|Y = y] = y \left(\frac{e^{-\lambda t}(\lambda t + 1) - 1}{\lambda} \right) + (1 - y) \left(\frac{e^{-\lambda t}(\lambda t + 1)}{\lambda} - \frac{\lambda^r e^{-\lambda}}{r!} \right)$$

1.3 Determine the E- and M-step of the EM algorithm.

The EM algorithm consists of two steps: the E-step and the M-step.

The E-step computes the expected value of the complete-data log-likelihood given the observed data and the current estimate of the parameters. The M-step maximizes this expected value with respect to the parameters.

In the E-Step we compute the value of

$$\lambda$$

based on its previous value then we update in the M-step.

$$E[\text{Log-Likelihood}(x)|Y = y, \lambda_k] = N \log(\lambda_k) - \lambda_k \sum_{i=1}^{n+m} E[x_i|Y = y, \lambda_k]$$

M step:

The value λ_{k+1} that maximizes the expected log-likelihood is given by:

$$\lambda_{k+1} = \frac{N}{\sum_{i=1}^{n+m} E[x_i|Y = y, \lambda_k]}$$

This can be better writen as :

```r
n <- 100; m <- 20; t <- 3
set.seed(1234)

# We know from the simulation that lambda is equal to 2

u <- rexp(n, 2)
v <- rexp(m, 2)
r <- sum(v <= t) # All the values are observed

# We know only the values of u

# Let us write the log-likelihood function

loglike <- function(lambda, u) {
    logl <- n * log(lambda) - lambda * sum(u)
    return(-logl) # Negate to minimize
}

lambda_opt <- optimize(loglike, interval = c(0.1, 10), u = u)$minimum
lambda_opt_2 <- n/sum(u)
```

```r
# Given the interval and the observations

# E-Step: (We will use the value obtained from the optimization)
e_step <- function(lambda, u, t, r) {
  e_x <- numeric(length(u))
  for (i in 1:length(u)) {
    if (u[i] <= t) {
      e_x[i] <-  ((exp(-lambda * t) * (lambda * t + 1) - 1) / lambda)
    } else {
      e_x[i] <- -((exp(-lambda * t) * (lambda * t + 1)) / lambda) -
                  (lambda^r * (exp(-lambda)) / factorial(r))
    }
  }
  return(sum(e_x))
}

e_step(lambda_opt, u, t, r)
```

```
## [1] -48.07486
```

```r
# M-Step: (We will use the value obtained from the optimization)

m_step <- function(e_x) {
  lambda_new <- n / e_x
  return(lambda_new)
}

# E-M algorithm
lambda <- lambda_opt_2 # Initial guess for lambda

for (i in 1:100) {
  e_x <- e_step(lambda, u, t, r)
  lambda <- m_step(e_x)
  print(lambda)
}
```

```
## [1] -2.080092
## [1] 0.0007734798
## [1] -287.7467
## [1] 0
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
```

```
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
```

```
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
## [1] NaN
```

2 Exercise 5

A multivariate numeric data set with missing values is given. We assume that the data come from a multivariate normal distribution and we want to estimate the parameters using maximum-likelihood estimation with a general purpose optimizer.

2.1 Specify the log-likelihood for a single observation y_i. Assume for y_i that the first d_1 variables are observed and that the next d_2 variables are missing, i.e.,