STA 6351, Report.1.12

Carson Slater *Baylor University*

1.12

- Recall from Report 1.11 we had starting values

$$\lambda^{(0)} = 1.2000, \quad \mu^{(0)} = 1.8000, \quad \sigma^{2(0)} = 0.9000,$$

  and after one Fisher-scoring step

$$\lambda^{(1)} = 0.8915, \quad \mu^{(1)} = 1.6041, \quad \sigma^{2(1)} = 0.7999.$$

- For a second iteration, again take $n_{\max} = \lfloor \lambda^{(1)} + 8\sqrt{\lambda^{(1)}} \rfloor = 10$. Evaluate

$$w_{in} = \Pr(N_i = n \mid X_i; \lambda^{(1)}, \mu^{(1)}, \sigma^{2(1)})$$

  for all nonzero x_i.

- Again taking $\varphi \equiv \sigma^2$. The scores are computed and yield:

$$U_\lambda = 0.6975, \quad U_\mu = -0.0476, \quad U_\varphi = -0.7607.$$

- Complete-data expected (Fisher) information at the same point:

$$\mathcal{I}_{\lambda\lambda} = 9.7561, \quad \mathcal{I}_{\mu\mu} = 9.6938, \quad \mathcal{I}_{\varphi\varphi} = 9.8206, \quad \mathcal{I}_{\mu\varphi} = -0.0595,$$

  and cross-terms with λ are 0.

- For a second Fisher-scoring update, solve the 3×3 system (block-diagonal in λ and (μ, φ)) to obtain

$$\Delta\lambda = +0.0715, \quad \Delta\mu = -0.0054, \quad \Delta\varphi = -0.0775.$$

  Hence

$$\lambda^{(2)} = 0.9630, \quad \mu^{(2)} = 1.5987, \quad \sigma^{2(2)} = 0.7224.$$

- Check monotonicity of the log-likelihood: Using the observed likelihood

$$\ell(\lambda, \mu, \sigma^2) = \sum_{i=1}^{n} \log\left(e^{-\lambda}\mathbf{1}\{x_i = 0\} + \sum_{n \geq 1} e^{-\lambda}\frac{\lambda^n}{n!}\varphi(x_i; n\mu, n\sigma^2)\right),$$

  we obtain

$$\ell(\lambda^{(0)}, \mu^{(0)}, \sigma^{2(0)}) = -14.4727, \quad \ell(\lambda^{(1)}, \mu^{(1)}, \sigma^{2(1)}) = -14.0695, \quad \ell(\lambda^{(2)}, \mu^{(2)}, \sigma^{2(2)}) = -13.9759.$$

  The likelihood increases each step, as expected.

Solution

After computing the weights at $\boldsymbol{\theta}^{(1)} = (0.8915, 1.6041, 0.7999)$, we obtain the conditional expectations:

$$\mathbb{E}[N_i \mid x_i] = (0.000, 2.103, 1.127, 0.000, 2.218, 1.230, 1.076, 0.000).$$

The score functions at $\boldsymbol{\theta}^{(1)}$ are:

$$U_\lambda = 0.6975, \quad U_\mu = -0.0474, \quad U_\varphi = -0.7607.$$

The Fisher information matrix components are:

$$\mathcal{I}_{\lambda\lambda} = 9.7560, \quad \mathcal{I}_{\mu\mu} = 9.6935, \quad \mathcal{I}_{\varphi\varphi} = 9.8198, \quad \mathcal{I}_{\mu\varphi} = -0.0592.$$

Solving the Fisher-scoring system yields the increments:

$$\Delta\lambda = 0.0715, \quad \Delta\mu = -0.0054, \quad \Delta\varphi = -0.0775.$$

Therefore, the updated parameters are:

$$\lambda^{(2)} = 0.9630, \quad \mu^{(2)} = 1.5987, \quad \sigma^{2(2)} = 0.7224.$$

The observed log-likelihoods confirm monotone increase:

$$\ell(\boldsymbol{\theta}^{(0)}) = -14.4727,$$
$$\ell(\boldsymbol{\theta}^{(1)}) = -14.0695,$$
$$\ell(\boldsymbol{\theta}^{(2)}) = -13.9759.$$

Table 1: Updated responsibilities and conditional expectations at $\boldsymbol{\theta}^{(2)}$

i	x_i	$\mathbb{E}[N_i \mid x_i]$	Top 4 Responsibilities
2	3.9	2.1581	n=2: 0.7296, n=3: 0.1885, n=1: 0.0651, n=4: 0.0161
3	1.5	1.1252	n=1: 0.8818, n=2: 0.1115, n=3: 0.0065, n=4: 0.0003
5	4.2	2.2659	n=2: 0.7063, n=3: 0.2415, n=1: 0.0272, n=4: 0.0237
6	2.1	1.2404	n=1: 0.7765, n=2: 0.2074, n=3: 0.0154, n=4: 0.0007
7	0.7	1.0706	n=1: 0.9327, n=2: 0.0641, n=3: 0.0030, n=4: 0.0001

Checks & comments.

- The recomputed weights at $\boldsymbol{\theta}^{(2)}$ continue to show sensible behavior: larger observations like $x_5 = 4.2$ concentrate probability on $n = 2, 3$, while smaller observations like $x_7 = 0.7$ favor $n = 1$.

- The observed log-likelihood increases at each step:

$$\ell(\boldsymbol{\theta}^{(0)}) < \ell(\boldsymbol{\theta}^{(1)}) < \ell(\boldsymbol{\theta}^{(2)}),$$

  confirming that the Fisher-scoring algorithm is working correctly.

- The score U_λ has changed sign from negative at $\boldsymbol{\theta}^{(0)}$ to positive at $\boldsymbol{\theta}^{(1)}$, indicating that λ is now being pushed upward toward its optimal value.

- All three score components are approaching zero, suggesting convergence is near. Further iterations would continue until $\|\boldsymbol{U}(\boldsymbol{\theta}^{(t)})\| < \epsilon$ for some tolerance ϵ.

- The coupling between μ and φ through $\mathcal{I}_{\mu\varphi}$ remains small but nonzero, requiring the full 2×2 system solution for these parameters.

Appendix

```r
knitr::opts_chunk$set(
  dev = "cairo_pdf",
  fig.width = 5,
  fig.height = 5,
  fig.align = 'center',
  echo = FALSE,
  message = FALSE,
  warning = FALSE,
  error = FALSE,
  results = 'markup'
)

# Load required libraries
library("tidyverse")
library("patchwork")
library("glue")
library("scales", warn.conflicts = FALSE)
library("extrafont")
library("tinytex")
library("knitr")
library("tidyr")
library("latex2exp")
library("MASS")
library("kableExtra")

theme_set(theme_minimal(base_family = "Roboto Condensed"))

conflicted::conflicts_prefer(
  readr::col_factor(),
  purrr::discard(),
  dplyr::lag(),
  readr::parse_date(),
  kableExtra::group_rows(),
  dplyr::select
)
# Given data from Report 1.11
x <- c(0.0, 3.9, 1.5, 0.0, 4.2, 2.1, 0.7, 0.0)
n_obs <- length(x)
n_max <- 10

# Parameters after first Fisher-scoring iteration
lambda1 <- 0.8915
mu1 <- 1.6041
sigma2_1 <- 0.7999
```

```r
# Function to compute weights w_in
compute_weights <- function(x, lambda, mu, sigma2, n_max) {
  n_obs <- length(x)
  w <- matrix(0, nrow = n_obs, ncol = n_max)

  for (i in 1:n_obs) {
    if (x[i] == 0) {
      # Structural zero: all mass at n=0, represented as negligible for n>=1
      w[i, ] <- 0
    } else {
      for (k in 1:n_max) {
        # Poisson prior on N_i
        pois_part <- dpois(k, lambda)
        # Conditional normal likelihood
        norm_part <- dnorm(x[i], mean = k * mu, sd = sqrt(k * sigma2))
        w[i, k] <- pois_part * norm_part
      }
      # Normalize
      w[i, ] <- w[i, ] / sum(w[i, ])
    }
  }
  return(w)
}

# Function to compute E[N_i | x_i]
compute_E_N <- function(w, x, n_max) {
  n_obs <- length(x)
  E_N <- numeric(n_obs)
  for (i in 1:n_obs) {
    if (x[i] == 0) {
      E_N[i] <- 0
    } else {
      E_N[i] <- sum((1:n_max) * w[i, ])
    }
  }
  return(E_N)
}

# Compute weights and E[N_i | x_i] at theta(1)
w1 <- compute_weights(x, lambda1, mu1, sigma2_1, n_max)
E_N1 <- compute_E_N(w1, x, n_max)

# Compute scores at theta(1)
# U_lambda
U_lambda <- sum(-1 + E_N1 / lambda1)

# U_mu
U_mu <- 0
```

```r
for (i in which(x != 0)) {
  for (k in 1:n_max) {
    U_mu <- U_mu + w1[i, k] * (x[i] - k * mu1) / sigma2_1
  }
}

# U_phi (where phi = sigma2)
U_phi <- 0
for (i in which(x != 0)) {
  for (k in 1:n_max) {
    U_phi <- U_phi +
      w1[i, k] *
        (-1 / (2 * sigma2_1) + (x[i] - k * mu1)2 / (2 * k * sigma2_12))
  }
}

# Compute Fisher information at theta(1)
I_lambda_lambda <- sum(E_N1) / lambda12

I_mu_mu <- 0
for (i in which(x != 0)) {
  for (k in 1:n_max) {
    I_mu_mu <- I_mu_mu + w1[i, k] * k / sigma2_1
  }
}

I_phi_phi <- 0
for (i in which(x != 0)) {
  for (k in 1:n_max) {
    I_phi_phi <- I_phi_phi +
      w1[i, k] * (1 / (2 * sigma2_12) + (x[i] - k * mu1)2 / (k * sigma2_13))
  }
}

I_mu_phi <- 0
for (i in which(x != 0)) {
  for (k in 1:n_max) {
    I_mu_phi <- I_mu_phi + w1[i, k] * (x[i] - k * mu1) / sigma2_12
  }
}

# Solve for increments (block-diagonal structure)
# Lambda update
Delta_lambda <- U_lambda / I_lambda_lambda

# (mu, phi) update - solve 2x2 system
I_mu_block <- matrix(c(I_mu_mu, I_mu_phi, I_mu_phi, I_phi_phi), nrow = 2)
U_mu_block <- c(U_mu, U_phi)
```

```r
Delta_mu_phi <- solve(I_mu_block, U_mu_block)

Delta_mu <- Delta_mu_phi[1]
Delta_phi <- Delta_mu_phi[2]

# Update parameters
lambda2 <- lambda1 + Delta_lambda
mu2 <- mu1 + Delta_mu
sigma2_2 <- sigma2_1 + Delta_phi

# Function to compute observed log-likelihood
compute_loglik <- function(x, lambda, mu, sigma2, n_max) {
  n_obs <- length(x)
  loglik <- 0

  for (i in 1:n_obs) {
    if (x[i] == 0) {
      # P(X_i = 0) = exp(-lambda)
      loglik <- loglik + log(exp(-lambda))
    } else {
      # P(X_i = x_i) = sum over n of P(N=n) * f(x_i | N=n)
      prob <- 0
      for (k in 1:n_max) {
        pois_part <- dpois(k, lambda)
        norm_part <- dnorm(x[i], mean = k * mu, sd = sqrt(k * sigma2))
        prob <- prob + pois_part * norm_part
      }
      loglik <- loglik + log(prob)
    }
  }
  return(loglik)
}

# Compute log-likelihoods
lambda0 <- 1.2000
mu0 <- 1.8000
sigma2_0 <- 0.9000

loglik0 <- compute_loglik(x, lambda0, mu0, sigma2_0, n_max)
loglik1 <- compute_loglik(x, lambda1, mu1, sigma2_1, n_max)
loglik2 <- compute_loglik(x, lambda2, mu2, sigma2_2, n_max)

# Recompute weights at theta(2) for display
w2 <- compute_weights(x, lambda2, mu2, sigma2_2, n_max)
E_N2 <- compute_E_N(w2, x, n_max)
# Create a table showing top weights for each nonzero observation at theta(2)
nonzero_idx <- which(x != 0)
weight_summary <- data.frame(
```

```r
  i = nonzero_idx,
  x_i = x[nonzero_idx],
  E_N = sprintf("%.4f", E_N2[nonzero_idx])
)

# For each observation, get top 4 weights
for (idx in 1:length(nonzero_idx)) {
  i <- nonzero_idx[idx]
  weights <- w2[i, ]
  top_indices <- order(weights, decreasing = TRUE)[1:4]
  top_weights <- sprintf("n=%d: %.4f", top_indices, weights[top_indices])
  weight_summary$top_weights[idx] <- paste(top_weights, collapse = ", ")
}

kable(
  weight_summary,
  col.names = c(
    "$i$",
    "$x_i$",
    "$\\mathbb{E}[N_i \\mid x_i]$",
    "Top 4 Responsibilities"
  ),
  caption = "Updated responsibilities and conditional expectations at $\\bolds
  booktabs = TRUE,
  escape = FALSE,
  format = "latex"
) %>%
  kable_styling(latex_options = c("hold_position"))
```