



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**UNA APLICACIÓN DEL ALGORITMO EM PARA UN
PROBLEMA DE DATOS CENSURADOS CON
IDENTIFICACIÓN PARCIAL**

T E S I S

QUE PARA OBTENER EL TÍTULO DE :

A C T U A R I O

P R E S E N T A :

MIGUEL ÁNGEL CHONG RODRÍGUEZ



FACULTAD DE CIENCIAS
UNAM

TUTOR : DR. ALBERTO CONTRERAS CRISTÁN

2006



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno

Chong
Rodríguez
Miguel Angel
52 07 52 98
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuaría
097437189

2. Datos del tutor

Dr
Alberto
Contreras
Cristán

3. Datos del sinodal 1

M en C
Leticia Eugenia
Gracia-Medrano
Valdelamar

4. Datos del sinodal 2

Mat
Eunice
Campirán
García

5. Datos del sinodal 3

Mat
Margarita Elvira
Chávez
Cano

6. Datos del sinodal 4

Act
Jaime
Vázquez
Alamilla

7. Datos del trabajo escrito

Una aplicación del algoritmo EM para un problema de datos
censurados con identificación parcial
129 p
2006

Dedicatoria

Dedico este trabajo a mi pequeña gran familia:

*A mi madre, Margarita,
a la memoria de Alfredo Vega Fernández,
a mis hermanos, Rosalia, Ramón y Margarita
y a mi tía Rosalia.
Para ellos, con todo mi amor.*

Agradecimientos

Al Dr. Alberto Contreras Cristán por todo su apoyo, tiempo, paciencia y generosidad. Sin sus atinados consejos este trabajo no hubiera llegado a buen puerto. Y también por ser un ejemplo para mi como docente y persona.

A la M. en C. Leticia Eugenia Gracia-Medrano, a la Mat. Margarita Elvira Chávez Cano, a la Mat. Eunice Campirán García y al Act. Jaime Vázquez Alamilla por ayudarme a mejorar este trabajo con sus valiosas observaciones.

A mi familia, por sus cuidados, consejos y todo lo que hemos vivido juntos que nos han hecho crecer.

Al buen Benjamín por ser mi "sensei" entre otras muchas cosas en el manejo de \LaTeX .

A todos mis amigos que han compartido parte de su tiempo conmigo, que me oyeron, que mutuamente nos divertimos y nos dimos consuelo y ánimos; todos ustedes son una buena razón de vida.

A mi muy querida Universidad Nacional Autónoma de México y a sus maestros, que me formaron y me dieron una visión distinta de la vida.

A los maestros que confiaron en mi para que fuera su ayudante y así aprender en uno de los oficios más nobles, ser maestro.

Cuadro 1: **Notación**

\mathbb{R}	Números reales
\mathbb{R}^+	Números reales positivos
\mathbb{Z}	Números enteros
$f_X(\cdot)$	Función de densidad
$\phi(u)$	Función de densidad normal estándar
$f_{X Y}(x y)$	Función de densidad condicional
$F_X(\cdot)$	Función de distribución
$\mathbb{1}_A(x)$	Función indicadora sobre el conjunto A
$N(\mu, \sigma)$	Distribución Normal
$\Phi(u)$	Función de distribución de una $N(0, 1)$
\mathcal{F}	σ -álgebra
$\mathbb{E}[\cdot]$	Esperanza
$\mathbb{E}[\cdot \mathcal{F}]$	Esperanza condicional
\mathbb{P}, π	Medidas de probabilidad
Ω	Espacio muestral
\sim	Se distribuye

Índice general

Introducción	1
1. Estadística Bayesiana	3
1.1. Introducción	3
1.2. El paradigma Bayesiano	4
1.3. Distribución <i>a priori</i>	5
1.4. Distribución predictiva <i>a priori</i>	6
1.5. Distribución <i>posterior</i> o final	6
1.6. Distribución predictiva posterior	8
1.7. Estimación por intervalos	13
2. El Algoritmo EM	17
2.1. Introducción.	17
2.2. Descripción del algoritmo	19
2.3. Algunas aplicaciones del algoritmo EM	20
3. Un problema de datos censurados con identificación parcial	34
3.1. Introducción.	34
3.2. Un algoritmo EM	38
3.3. Desviación estándar en el contexto del algoritmo EM.	44
3.4. Instrumentación.	48
3.5. Cobertura con datos simulados y $k = 2$	50
3.6. Datos simulados con $k = 3$	51

Conclusiones	54
Apéndice A.	56
Apéndice B.	62
B.1. Cálculos del paso E del Capítulo 3	64
B.2. Cálculos del paso M del Capítulo 3	74
B.3. Cálculos de la matriz Hessiana	85
Apéndice C.	103
Bibliografía	128

Introducción

El objetivo de este trabajo es mostrar el uso del algoritmo EM propuesto por Dempster, Laird y Rudin (1977), el cual tiene el propósito de encontrar estimaciones máximo verosímiles cuando hay datos faltantes y es útil para dar una solución alternativa al problema de hacer inferencia estadística sobre los parámetros de una mezcla de distribuciones de datos censurados con identificación parcial usando variables latentes.

Una aplicación concreta a este tipo de problemas lo podemos encontrar en un estudio de cáncer a N pacientes con este mal. Supongamos que dichos pacientes están bajo un tratamiento médico que durará un periodo de tiempo C , conocido éste como umbral de censura. Supongamos también que los decesos pueden ocurrir por dos causas ($k = 2$); muerte por cáncer o por una causa distinta a esta enfermedad. Por la forma en que está planteado el problema puede haber dos posibles escenarios:

a) Los N pacientes murieron antes del tiempo C ya sea por causa del cáncer u otra causa. En este caso no hay datos censurados ya que los tiempos de los decesos son todos conocidos y por lo tanto hacer inferencia sobre los parámetros de la mezcla de distribuciones que describen el fenómeno no es complicado.

b) El segundo escenario es que una cantidad $m < N$ de muertes se dieron antes de llegar al tiempo C por cualquiera de las dos causas, y por lo tanto los restantes $N - m$ pacientes continuaron vivos después del tiempo C . En este contexto tenemos $N - m$ datos censurados

Problemas de este estilo son los que podemos resolver usando el algoritmo EM, suponiendo que los $N - m$ datos censurados son variables latentes.

Este trabajo se divide en tres capítulos y tres apéndices. En el primer capítulo se da una breve introducción a la forma en que se hace inferencia puntual y por intervalos desde el punto de vista Bayesiano, a la vez que se hace una comparación con la forma en que se hace inferencia también puntual

y por intervalos pero desde el punto de vista clásico.

En el segundo capítulo se describe el algoritmo EM y se ilustra su aplicación en dos ejemplos, el primero trata de la unión de dos clases genéticas que al separarlas da como resultado un problema con datos faltantes. El otro ejemplo, es el uso del algoritmo EM en un problema de análisis de regresión simple con datos censurados para tiempos de falla de motores sometidos a distintas temperaturas.

En el tercer capítulo se estudia un problema como el expuesto en el escenario b) descrito anteriormente, y se describe una forma de aplicar el algoritmo EM para hacer estimaciones puntuales sobre los parámetros de una mezcla finita de distribuciones ($k = 2$ ó 3) de datos censurados con identificación parcial.

También en el tercer capítulo se describe una forma de obtener de manera numérica la matriz de varianza-covarianza del vector de parámetros y así obtener la desviación estándar de cada parámetro y como consecuencia poder construir el intervalo de confianza asociado a cada estimación puntual. Acto seguido, se simulan cien muestras con $k = 2$ para estudiar la cobertura de los intervalos de confianza. Finalmente se simula una muestra con $k = 3$ y se le construyen sus intervalos de confianza, en este ejercicio se puede ver como todos los valores verdaderos de los parámetros caen dentro del intervalo construido.

El contenido de los apéndices A y B son cálculos relacionados a lo expuesto en el segundo y tercer capítulo respectivamente. La razón por la cual se procedió de esta manera, es porque los cálculos son muy largos y desviaban el sentido de la exposición. El lector interesado en ver como se llegó a algún resultado puede consultarlo por separado.

Finalmente el apéndice C, muestra dos códigos de programas hechos en R para simular una muestra con datos censurados provenientes de una mezcla de distribuciones como se describe en el capítulo 3, y tres programas en *Mathematica* que calculan los estimadores puntuales y los intervalos de confianza bajo el contexto del algoritmo EM para cada uno de los parámetros. Todo esto en el caso de que la mezcla de distribución sea de tamaño $k = 2$ ó $k = 3$.

Capítulo 1

Estadística Bayesiana

1.1. Introducción

La estadística Bayesiana o metodología Bayesiana debe su nombre al ministro presbiteriano Thomas Bayes (1702-1761) quien fue elegido como Miembro de la Royal Society en 1742 por la publicación de sus dos únicas obras, una de ellas la tituló *An Essay towards solving a Problem in the Doctrine of chance*; en la cual demuestra el resultado que hoy en día conocemos como el Teorema de Bayes. Este teorema se enuncia a continuación en un lenguaje matemático moderno:

Teorema 1.1 *Teorema de Bayes* Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Supongamos que tenemos una sucesión de eventos $\{B_n\} \in \mathcal{F}$ tal que $\{B_n\}$ es una partición de Ω , donde se cumple:

- (i) $\Omega = \cup_n B_n$
- (ii) $B_i \cap B_j = \phi \quad \forall i \neq j$.

Sea A otro evento tal que $\mathbb{P}(A) > 0$, entonces

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_j \mathbb{P}(A|B_j)\mathbb{P}(B_j)} \quad , \text{ para alguna } i = 1, 2, \dots, n.$$

Resulta muy interesante que a partir de este resultado sencillo pero crucial y la esquematización fundamentada en la *teoría de decisiones* dada por

algunos autores como Dennis V. Lindley, se haya desarrollado la metodología Bayesiana la cual hoy día es una alternativa muy usada para resolver problemas de inferencia estadística. A continuación se expondrá *grosso modo* el paradigma Bayesiano y sus diferencias con respecto al enfoque clásico.

1.2. El paradigma Bayesiano

Tanto la estadística Bayesiana como la clásica o frecuentista son herramientas para un tomador de decisiones bajo condiciones de incertidumbre. Cuando se estudia un fenómeno aleatorio, la incertidumbre en datos provenientes de este, debe ser medida a través de la probabilidad. Es decir, al fenómeno aleatorio se le asocia un modelo probabilístico apropiado. Es común suponer que este modelo depende de un vector o un escalar que se conoce como parámetro θ y que pertenece al espacio parametral Θ .

Desde el punto de vista de la estadística frecuentista o clásica, nuestro problema de decisión a resolver es encontrar una distribución de probabilidad completamente especificada de entre un conjunto infinito de distribuciones de probabilidad. Como dijimos antes, podemos pensar en que cada distribución depende de un valor del parámetro θ , entonces podemos denotar nuestro conjunto de distribuciones posibles como

$$\mathcal{P} = \{\mathbb{P}(X|\theta) : \theta \in \Theta\}.$$

Entonces desde un punto de vista clásico, nuestro problema es equivalente a encontrar un $\theta \in \Theta$ “adecuado” para un conjunto de datos \mathbf{x} .

Desde el punto de vista de la estadística Bayesiana, podemos pensar en cuantificar la incertidumbre en la toma de decisiones proponiendo distribuciones de probabilidad no solamente para datos provenientes de nuestro fenómeno, sino también para el valor de θ . En algunos textos (por ejemplo Bernardo y Smith (1994)) se enuncia el *teorema de representación de de Finetti*, el cual justifica (bajo algunos supuestos) la existencia de una distribución de probabilidad para θ . Esta distribución se puede interpretar como la cuantificación de nuestro conocimiento inicial respecto a θ . Las conclusiones acerca del parámetro θ , así como acerca de datos no observados $\tilde{\mathbf{x}}$, se hacen en términos de probabilidades condicionales en el valor de datos observados \mathbf{x} . Así por ejemplo, son de interés $\mathbb{P}(\theta|\mathbf{x})$ o $\mathbb{P}(\tilde{\mathbf{x}}|\mathbf{x})$.

Así, el aspecto de condicionar en el valor de datos observados diferencia a la estadística Bayesiana de la estadística clásica. Notemos que en esta última,

el objetivo es la evaluación de un procedimiento que se utilizó para estimar θ (o $\tilde{\mathbf{x}}$) a partir de $\mathbb{P}(\mathbf{x}|\theta)$.

A lo largo de este trabajo usaremos $\mathbb{P}(\cdot|\cdot)$ para denotar densidades de probabilidad con los argumentos determinados por el contexto, al igual que en el caso de $\mathbb{P}(\cdot)$ que denota una densidad marginal. Los términos “distribución” y “densidad” se emplearán indistintamente, salvo en los casos donde se indique explícitamente.

1.3. Distribución *a priori*

Como se afirmó en la sección anterior consideraremos a θ como una variable aleatoria.

La distribución *a priori* o inicial, es una distribución de probabilidad sobre el parámetro θ que denotaremos por $\pi(\theta)$ y que refleja la cantidad de incertidumbre o certidumbre que subjetivamente tenemos de θ .

El término distribución *a priori* implica la traducción de la información inicial que tiene una persona sobre θ en una distribución de probabilidad sobre este parámetro. Aún antes de efectuar cualquier muestreo, naturalmente la inferencia sobre las observaciones del fenómeno aleatorio serán mejores en la medida en que el conocimiento de la persona sobre el verdadero valor del parámetro sea más preciso. Aunque más adelante se discutirán otras formas de obtener una distribución *a priori* con información vaga sobre el fenómeno aleatorio.

A continuación se presenta lo que entenderemos como muestra:

Definición 1.1 Sea $\mathbf{x} = (X_1, X_2, \dots, X_n)$ una muestra de tamaño n , proveniente del modelo probabilístico $\mathbb{P}(X|\theta)$. Supondremos la siguiente estructura de independencia condicional al valor de θ

$$\begin{aligned}\mathbb{P}(\mathbf{x}|\theta) &= \mathbb{P}(X_1, \dots, X_n|\theta) \\ &= \mathbb{P}(X_1|\theta) \cdot \dots \cdot \mathbb{P}(X_n|\theta) \\ &= l(\mathbf{x}|\theta).\end{aligned}$$

Llamamos a $l(\mathbf{x}|\theta)$ la función de verosimilitud de los datos.

1.4. Distribución predictiva *a priori*

Es claro que si nosotros conociéramos el valor real de θ entonces podríamos hacer predicciones sobre observaciones futuras del fenómeno aleatorio usando el modelo probabilístico $\mathbb{P}(\cdot|\theta)$, pero en general θ es desconocido. Sin embargo si usamos la distribución *a priori* que es la información que tenemos acerca del parámetro θ , sí podemos saber la probabilidad de que se genere una muestra $\tilde{\mathbf{x}}$.

Primero notemos que al tener una función de distribución *a priori* $\pi(\theta)$ y al utilizar el resultado de la probabilidad total podemos calcular

$$\mathbb{P}(\tilde{\mathbf{x}}, \theta) = \mathbb{P}(\tilde{\mathbf{x}}|\theta)\pi(\theta).$$

Al obtener la marginal con respecto a θ tenemos la probabilidad de generar una muestra $\tilde{\mathbf{x}}$

$$\begin{aligned}\mathbb{P}(\tilde{\mathbf{x}}) &= \int_{\theta \in \Theta} \mathbb{P}(\tilde{\mathbf{x}}, \theta) d\theta \\ &= \int_{\theta \in \Theta} \mathbb{P}(\tilde{\mathbf{x}}|\theta)\pi(\theta) d\theta.\end{aligned}$$

A $\mathbb{P}(\tilde{\mathbf{x}})$ se le conoce como la función de distribución predictiva *a priori*.

Antes de continuar destacamos que la distribución predictiva *a priori* se basa exclusivamente en la información *a priori* reflejada en $\pi(\theta)$, y que hasta el momento no hemos usado una muestra de valores observados para hacer predicciones. Esta es una ventaja de la estadística Bayesiana la cual se vale de cualquier información disponible, esté esta basada en un juicio personal o en experiencias, mientras que de la teoría clásica es principalmente empírica y sólo emplea información de la muestra como la base para estimar e inferir.

1.5. Distribución *posterior* o final

Ahora bien si podemos tener acceso a una muestra \mathbf{x} de datos observados y utilizando el Teorema de Bayes obtenemos

$$\begin{aligned}\pi(\theta|\mathbf{x}) &= \frac{\mathbb{P}(\theta, \mathbf{x})}{\mathbb{P}(\mathbf{x})} \\ &= \frac{\mathbb{P}(\mathbf{x}|\theta)\pi(\theta)}{\int_{\Theta} \mathbb{P}(\mathbf{x}|\theta)\pi(\theta) d\theta} \\ &= \frac{l(\mathbf{x}|\theta)\pi(\theta)}{\int_{\Theta} l(\mathbf{x}|\theta)\pi(\theta) d\theta}.\end{aligned}$$

A $\pi(\theta|\mathbf{x})$ se le conoce como la distribución *posterior* o final de θ .

De la función de densidad *posterior* podemos notar lo siguiente:

- (a) La cantidad $K = \frac{1}{\mathbb{P}(\mathbf{x})} = \frac{1}{\int_{\Theta} \mathbb{P}(\mathbf{x}|\theta)\pi(\theta) d\theta}$ es una constante con respecto a θ , la cual garantiza que $\int_{\Theta} \pi(\theta|\mathbf{x}) d\theta = 1$. Por lo tanto, la distribución de probabilidad *posterior* es proporcional a la multiplicación de la verosimilitud por la distribución inicial, y lo denotaremos con el símbolo \propto de la siguiente manera:

$$\pi(\theta|\mathbf{x}) \propto l(\mathbf{x}|\theta)\pi(\theta).$$

- (b) Tanto la distribución *posterior* como la *a priori* son distribuciones de probabilidad sobre θ , pero la primera toma en cuenta tanto la información que dan los datos observados respecto a θ , así como la información *a priori* ó inicial con respecto a θ .
- (c) Para obtener la distribución *posterior* se necesita de una muestra que puede ser tan pequeña como un solo dato.
- (d) Este razonamiento induce un ciclo, puesto que la distribución *posterior* pasa a ser la nueva distribución *a priori*, si es posible obtener más datos observados bajo el mismo procedimiento, podemos generar una nueva distribución *posterior*.
- (e) Si $\pi(\theta)$ es una constante entonces

$$\pi(\theta|\mathbf{x}) \propto l(\mathbf{x}|\theta). \tag{1.1}$$

y como consecuencia al encontrar el estimador máximo verosimil estaríamos encontrando la moda de la distribución posterior.

1.6. Distribución predictiva posterior

Naturalmente podemos hacer predicción sobre observaciones futuras del fenómeno aleatorio ahora que contamos con más elementos como: una muestra \mathbf{x} , un modelo probabilístico $\mathbb{P}(\cdot|\theta)$ y una distribución *posterior*. Veamos como estos elementos inducen una distribución conjunta para \tilde{x} y θ , condicionada a la muestra \mathbf{x} .

$$\begin{aligned}\pi(\tilde{x}, \theta|\mathbf{x}) &= \frac{\mathbb{P}(\tilde{x}, \theta, \mathbf{x})}{\mathbb{P}(\mathbf{x})} \\ &= \frac{\mathbb{P}(\tilde{x}|\theta, \mathbf{x})\mathbb{P}(\theta, \mathbf{x})}{\mathbb{P}(\mathbf{x})} \\ &= \frac{\mathbb{P}(\tilde{x}|\theta, \mathbf{x})\mathbb{P}(\theta|\mathbf{x})\mathbb{P}(\mathbf{x})}{\mathbb{P}(\mathbf{x})} \\ &= \mathbb{P}(\tilde{x}|\theta, \mathbf{x})\mathbb{P}(\theta|\mathbf{x}).\end{aligned}\tag{1.2}$$

Notemos que $\mathbb{P}(\tilde{x}|\theta, \mathbf{x}) = \mathbb{P}(\tilde{x}|\theta)$ porque \tilde{x} y \mathbf{x} son independientes dado el valor de θ .

Obtengamos la distribución marginal de \tilde{x} a partir de la función de distribución conjunta $\pi(\tilde{x}, \theta|\mathbf{x})$. De la ecuación (1.2) tenemos que

$$\begin{aligned}\mathbb{P}(\tilde{x}|\mathbf{x}) &= \int_{-\infty}^{\infty} \pi(\tilde{x}, \theta|\mathbf{x}) d\theta \\ &= \int_{-\infty}^{\infty} \mathbb{P}(\tilde{x}|\theta)\mathbb{P}(\theta|\mathbf{x}) d\theta.\end{aligned}$$

A la función de distribución $\mathbb{P}(\tilde{x}|\mathbf{x})$ se le conoce como distribución predictiva *posterior*.

Con el fin de mostrar en la práctica lo que hasta ahora se ha escrito se expone un ejemplo cuyo objetivo es dar pie al concepto de *familias conjugadas*.

Ejemplo 1.1

Supongamos que tenemos la siguiente familia paramétrica

$$\mathcal{P} = \{\mathbb{P}(X|\theta) = \text{Uniforme}(X|0, \theta) : \theta \in \mathbb{R}^+\}.$$

Por otro lado dos personas por separado proponen una distribución *a priori* que a su juicio refleja la información subjetiva que tienen de θ ; denotemos a:

$\pi_i(\theta)$ como la distribución *a priori* que propuso la persona i con $i = 1, 2$, donde

$$\pi_1(\theta) = \frac{2(4 - \theta)}{9} \mathbb{1}_{(1,4)}(\theta).$$

$\pi_2(\theta) = \text{Pareto}(\theta|a, b) = \frac{ba^b}{\theta^{b+1}} \mathbb{1}_{(\theta \geq a)}(\theta)$, con $b > 0$. Más aún, la segunda persona según su experiencia cree que los valores de los parámetros a y b que ajustan bien al fenómeno aleatorio son 1 y 2.5 respectivamente.

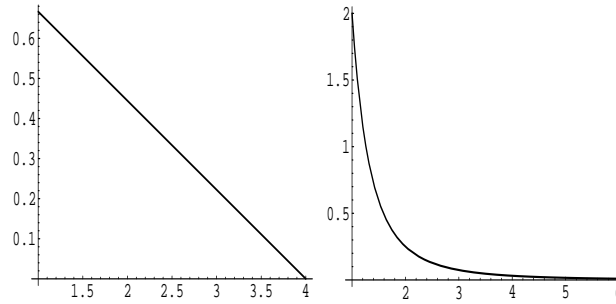


Figura 1.1: Distribuciones $\pi_1(\theta)$ y $\pi_2(\theta)$ respectivamente.

Obtengamos en ambos casos:

- (a) La distribución predictiva *a priori*
- (b) La distribución *posterior*

Solución

Primero hagamos unos cálculos que nos servirán en ambos casos

$$\mathbb{P}(\mathbf{x}|\theta) = \prod_{j=1}^n \mathbb{P}(X_j|\theta),$$

Como $\mathbb{1}_{(a_1, b_1)}(x) \cdot \dots \cdot \mathbb{1}_{(a_n, b_n)}(x) = \mathbb{1}_{(a_1, b_1) \cap \dots \cap (a_n, b_n)}(x)$ entonces

$$\mathbb{P}(\mathbf{x}|\theta) = \theta^{-n} \mathbb{1}_{(0,\theta)}(x_{(n)}). \quad (1.3)$$

Donde $x_{(n)}$ es la observación mayor o el estadístico de orden n-ésimo. Además notemos que

$$\mathbb{1}_{(0,\theta)}(x) = \left\{ \begin{array}{ll} 1 & \text{si } 0 < x < \theta \\ 0 & \text{en otro caso} \end{array} \right\} = \mathbb{1}_{(x,\infty)}(\theta).$$

Por lo tanto, la ecuación (1.3) la podemos reescribir de la siguiente manera

$$\theta^{-n} \mathbb{1}_{(x_{(n)},\infty)}(\theta).$$

Calculemos (a), por definición tenemos que:

$$\begin{aligned} \mathbb{P}_1(\mathbf{x}) &= \int_0^\infty \mathbb{P}(\mathbf{x}|\theta) \pi_1(\theta) d\theta \\ &= \int_0^\infty \theta^{-n} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \cdot \frac{2(4-\theta)}{9} \mathbb{1}_{(1,4)}(\theta) d\theta \\ &= \frac{2}{9} \int_0^\infty \frac{4-\theta}{\theta^n} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \mathbb{1}_{(1,4)}(\theta) d\theta \\ &= \frac{2}{9} \int_{\max\{x_{(n)},1\}}^4 4\theta^{-n} - \theta^{-n+1} d\theta \\ &= \begin{cases} \frac{2}{9} \{4 \ln(\theta) - \theta\} \Big|_{\max\{x_{(n)},1\}}^4 & \text{si } n = 1 \\ \frac{2}{9} \left\{ 4 \frac{\theta^{-n+1}}{-n+1} - \frac{\theta^{-n+2}}{-n+2} \right\} \Big|_{\max\{x_{(n)},1\}}^4 & \text{si } n > 1 \end{cases} \\ &= \begin{cases} \frac{2}{9} \{4 [\ln(4) - \ln(\max\{x_{(n)},1\})] - 4 + \max\{x_{(n)},1\}\} & \text{si } n = 1 \\ \frac{2}{9} \left\{ 4 \left[\frac{4^{-n+1}}{-n+1} - \frac{\max\{x_{(n)},1\}^{-n+1}}{-n+1} \right] - \frac{4^{-n+2}}{-n+2} + \frac{\max\{x_{(n)},1\}^{-n+2}}{-n+2} \right\} & \text{si } n > 1 \end{cases} . \end{aligned}$$

$$\begin{aligned}
 \mathbb{P}_2(\mathbf{x}) &= \int_0^\infty \mathbb{P}(\mathbf{x}|\theta)\pi_2(\theta) d\theta \\
 &= \int_0^\infty \theta^{-n} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \frac{ba^b}{\theta^{b+1}} \mathbb{1}_{(\theta \geq a)}(\theta) d\theta \\
 &= ba^b \int_0^\infty \frac{1}{\theta^{b+n+1}} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \mathbb{1}_{(\theta \geq a)}(\theta) d\theta \\
 &= ba^b \int_{\max\{x_{(n)},a\}}^\infty \theta^{-b-n-1} d\theta \\
 &= -ba^b \frac{\theta^{-b-n}}{b+n} \Big|_{\max\{x_{(n)},a\}}^\infty \\
 &= ba^b \frac{(\max\{x_{(n)},a\})^{-b-n}}{b+n}.
 \end{aligned}$$

(b) En cuanto a las distribuciones posteriores tenemos:

$$\begin{aligned}
 \pi_1(\theta|\mathbf{x}) &\propto \mathbb{P}(\mathbf{x}|\theta)\pi_1(\theta) \\
 &= \theta^{-n} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \frac{2(4-\theta)}{9} \mathbb{1}_{(1,4)}(\theta) \\
 &= \frac{2(4-\theta)}{9\theta^n} \mathbb{1}_{(\max\{x_{(n)},1\},4)}(\theta).
 \end{aligned}$$

$$\begin{aligned}
 \pi_2(\theta|\mathbf{x}) &\propto \mathbb{P}(\mathbf{x}|\theta)\pi_2(\theta) \\
 &= \theta^{-n} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \frac{ba^b}{\theta^{b+1}} \mathbb{1}_{(\theta \geq a)}(\theta) \\
 &= \frac{ba^b}{\theta^{b+n+1}} \mathbb{1}_{(x_{(n)},\infty)}(\theta) \mathbb{1}_{(\theta \geq a)}(\theta) \\
 &= \frac{ba^b}{\theta^{b+n+1}} \mathbb{1}_{(\max\{x_{(n)},a\},\infty)}(\theta).
 \end{aligned}$$

El lector debe notar que la estructura algebraica o *kernel* de $\pi_1(\theta|\mathbf{x})$ y $\pi_1(\theta)$ no es la misma (mientras la *a priori* es la ecuación de una recta

con respecto a θ la *posterior* no lo es), en cambio $\pi_2(\theta|\mathbf{x})$ y $\pi_2(\theta)$ sí tienen el mismo *kernel*, aún más, si multiplicamos la *posterior* por una constante apropiada ¹ obtendremos que dicha *posterior* es una distribución Pareto con nuevos parámetros $a' = \max\{x_{(n)}, a\}$ y $b' = b + n$.

Otra ventaja notoria es que hasta el momento el tratamiento algebraico ha sido más sencillo con $\pi_2(\theta)$ que con $\pi_1(\theta)$, estas dos características nos inducen a definir el concepto de familias conjugadas.

Definición 1.2 Sea $\mathcal{P} = \{\mathbb{P}(\cdot|\theta) : \theta \in \Theta\}$ una familia paramétrica. Se dice que una colección de distribuciones de probabilidad \mathcal{G} , es una familia conjugada para \mathcal{P} si ocurre que:

$$\forall \mathbb{P}(\cdot|\theta) \in \mathcal{P} \quad \text{y} \quad \pi(\theta) \in \mathcal{G} \quad \text{entonces} \quad \pi(\theta|\mathbf{x}) \in \mathcal{G}$$

El cuadro 1 muestra algunos ejemplos de familias conjugadas.

Cuadro 1.1: Familias conjugadas

Familia paramétrica	Parámetro a inferir	Familia conjugada
Bernoulli($X \theta$)	$\theta \in (0,1)$	Beta($\theta \alpha, \beta$)
Poisson($X \theta$)	$\theta \in \mathbb{R}^+$	Gama($\theta \alpha, \beta$)
Geométrica($X \theta$)	$\theta \in (0,1)$	Beta($\theta \alpha, \beta$)
Exponencial($X \theta$)	$\theta \in \mathbb{R}^+$	Gama($\theta \alpha, \beta$)
Uniforme ($X 0, \theta$)	$\theta \in \mathbb{R}^+$	Pareto($\theta \alpha, \beta$)
Normal($X \theta, \lambda$)	$\theta \in \mathbb{R}$	Normal($\theta \alpha, \beta$)
Normal($X \theta, \lambda$)	$\lambda \in \mathbb{R}^+$	Gama($\theta \alpha, \beta$)
Normal($X \theta, \lambda$)	$\theta \in \mathbb{R}$ y $\lambda \in \mathbb{R}^+$	Normal-Gama($\theta \alpha, \beta$)

Pese que en la estadística Bayesiana se permite incorporar el conocimiento respecto a θ a través de la asignación de una distribución inicial, este paso puede ser subjetivo. Existe en la literatura Bayesiana una discusión respecto a como debería hacerse esta asignación, en particular se ha desarrollado el concepto de *distribución no informativa* (véase, Bernardo y Smith (1994) y Gelman et al (1995)).

Como su nombre lo indica, la idea en este tipo de asignación es proponer una distribución que no se basa en la información que subjetivamente tiene el

¹ $K = \frac{(b+n) \max\{x_{(n)}, a\}^{b+n}}{ba^b}$.

investigador sobre el fenómeno sino de la evidencia que da una muestra. Hay varias formas de lograr este fin, algunos métodos muy simples son suponer una distribución con una varianza muy grande, la otra forma es proponer una distribución donde sea equiprobable tomar el valor del parámetro, pero la forma por excelencia de obtener una distribución no informativa en el caso unidimensional, es la regla de Jeffreys. Este método propone distribuciones iniciales que requieren de verificar las condiciones de regularidad que se mencionan a continuación:

i) El modelo paramétrico $f(x; \theta)$ para la distribución de la población es tal que el soporte de f no depende de θ .

ii) La función $\ln(l(\mathbf{x}; \theta))$ es dos veces diferenciable con continuidad, es decir, de clase C^2 .

iii) Las operaciones de derivación e integración o suma en caso discreto son intercambiables.

Definición 1.3 Sea un modelo paramétrico $\mathbb{P}(X|\theta)$, $\pi(\theta)$ es *no informativa* de Jeffreys si:

$$\pi(\theta) \propto \sqrt{I(\theta)} \quad \theta \in \Theta.$$

$$\text{donde } I(\theta) = \mathbb{E} \left[\frac{\partial \ln(l(\mathbf{x}|\theta))}{\partial \theta} \right]^2.$$

Un inconveniente de la regla de Jeffreys es que puede producir distribuciones que al integrarlas sobre su rango no integren uno sino infinito, a este tipo de distribuciones se les conoce como impropias.

1.7. Estimación por intervalos

A continuación se explicará brevemente las diferencias entre intervalo de confianza en la estadística clásica y un intervalo de probabilidad bajo el contexto de la estadística Bayesiana.

Comencemos con el enfoque clásico; una vez que se tiene un estimador puntual $\hat{\theta}$ dada una muestra \mathbf{x} en particular, no se sabe que tan cercano está el estimador del verdadero valor del parámetro θ ; esta es la razón por la cual es deseable acompañar la estimación puntual de alguna medida de error asociado a esta estimación, es decir, asociar a cada estimación puntual del parámetro un intervalo

$$[\underline{\theta}(\mathbf{x}), \bar{\theta}(\mathbf{x})],$$

donde $\underline{\theta}(\mathbf{x})$ y $\bar{\theta}(\mathbf{x})$ son el límite inferior y superior respectivamente del intervalo, y ambos son funciones de la información muestral.

Además, una medida que nos refleje la *confianza* que tenemos acerca de que el verdadero valor del parámetro pertenezca a dicho intervalo.

$$\mathbb{P}(\underline{\theta}(\mathbf{x}) \leq \theta \leq \bar{\theta}(\mathbf{x})) = 1 - \alpha, \quad (1.4)$$

donde a $1 - \alpha$ se le conoce como el nivel de confianza.

Observación 1

- (i) Una vez fijado el nivel de confianza, los límites del intervalo de confianza varían aleatoriamente según la muestra seleccionada, por lo tanto la longitud es aleatoria.
- (ii) La expresión (1.4) **no** debe entenderse como; la probabilidad de que el parámetro θ tome algún valor entre $\underline{\theta}(\mathbf{x})$ y $\bar{\theta}(\mathbf{x})$ es igual a $1 - \alpha$; ya que:
 - (a) En la práctica, el parámetro θ siempre será desconocido.
 - (b) Desde el punto de vista de la estadística clásica, en la expresión (1.4) las variables aleatorias son $\underline{\theta}(\mathbf{x})$ y $\bar{\theta}(\mathbf{x})$ y no el parámetro θ .
- (iii) Una forma de interpretar (1.4) es; $1 - \alpha$ es la probabilidad de que el intervalo $[\underline{\theta}(\mathbf{x}), \bar{\theta}(\mathbf{x})]$ incluya el verdadero valor del parámetro. O en otras palabras, si consideramos un número grande de muestras del mismo tamaño y calcularemos $\underline{\theta}(\mathbf{x})$ y $\bar{\theta}(\mathbf{x})$ para cada muestra entonces se tendrá que aproximadamente en el $(1 - \alpha) \times 100\%$ de los intervalos resultantes estará incluido el verdadero valor de θ . En consecuencia a $[\underline{\theta}(\mathbf{x}), \bar{\theta}(\mathbf{x})]$ se le conoce como intervalo de confianza con nivel de confianza del $(1 - \alpha) \times 100\%$.

Definamos primero lo que entenderemos por un intervalo de probabilidad en el contexto del enfoque Bayesiano.

Definición 1.4

Un intervalo de probabilidad al $100(1 - \alpha)\%$ para θ es un subconjunto $C = [L(\mathbf{x}), U(\mathbf{x})]$ de Θ tal que

$$1 - \alpha = \mathbb{P}(C|\mathbf{x}) = \begin{cases} \int_C \pi(\theta|\mathbf{x}) & \text{caso continuo} \\ \sum_{\theta \in C} \pi(\theta|\mathbf{x}) & \text{caso discreto.} \end{cases}$$

La razón por la cual en el enfoque Bayesiano sí se puede hablar de intervalos de probabilidad es que el parámetro θ es una variable aleatoria; a la cual se le asigna una distribución que ya puede ser la *a priori* $\pi(\theta)$, en el caso en que no se tenga muestra alguna; o la *a posteriori* $\pi(\theta|\mathbf{x})$ si se cuenta con información muestral. Por lo tanto, tiene sentido preguntarse por el evento $\theta \in [L(\mathbf{x}), U(\mathbf{x})]$.

Un método muy empleado para encontrar los límites del intervalo de probabilidad $L(\mathbf{x})$ y $U(\mathbf{x})$ de forma numérica es trazar una recta paralela al eje de las abscisas, $\theta = u$ con $u > 0$, la cual iremos bajando de tal forma que intersecte en dos puntos a la función de densidad $\pi(\theta|\mathbf{x})$ (si la distribución *posterior* es unimodal), a la proyección de estos puntos sobre el eje de las abscisas les llamaremos a y b ; este proceso de bajar la recta con pendiente cero terminará cuando $\int_a^b \pi(\theta|\mathbf{x})d\theta = 1 - \alpha$, entonces $a = L(\mathbf{x})$ y $b = U(\mathbf{x})$.

Regresando al Ejemplo 1.1, supongamos que tenemos una muestra aleatoria de tamaño $n = 5$ del fenómeno, $\mathbf{x} = \{1.21, 1.40, 1.10, 2.12, 3.67\}$.

A continuación se obtendrá un intervalo de probabilidad al 95% para θ sobre las funciones de densidad $\pi_1(\theta|\mathbf{x})$ y $\pi_2(\theta|\mathbf{x})$.

Primero notemos que; $\max\{x_{(n)}, 1\} = 3.67$ y $\max\{x_{(n)}, a\} = 3.67$, ya que $a = 1$, entonces

$$\pi_1(\theta|\mathbf{x}) = C_1 \frac{2(4 - \theta)}{9\theta^5} \mathbb{1}_{(3.67, 4)}(\theta),$$

$$\pi_2(\theta|\mathbf{x}) = \frac{(b + 5)3.67^{b+5}}{\theta^{b+5+1}} \mathbb{1}_{(3.67, \infty)}(\theta),$$

sustituyendo el valor de $b = 2.5$

$$\pi_2(\theta|\mathbf{x}) = \frac{C_2}{\theta^{8.5}} \mathbb{1}_{(3.67, \infty)}(\theta),$$

donde $C_1 = 63384.93$ y $C_2 = 128842$ son las constantes de proporcionalidad que garantizan que tanto $\pi_1(\theta|\mathbf{x})$ como $\pi_2(\theta|\mathbf{x})$ integren uno.

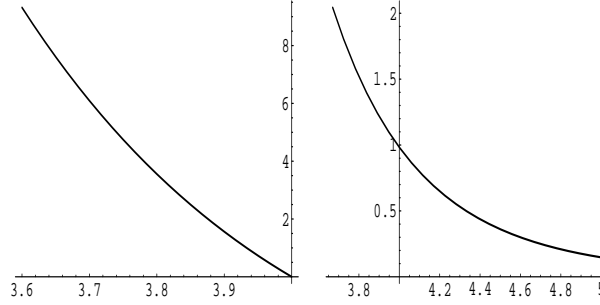


Figura 1.2: Distribuciones $\pi_1(\theta|\mathbf{x})$ y $\pi_2(\theta|\mathbf{x})$ respectivamente.

Como podemos ver en la figura (1.2) ambas funciones son estrictamente decrecientes, por lo tanto, al trazar una recta horizontal en los dos casos el límite inferior será $L(\mathbf{x}) = 3.67$. Solo nos resta encontrar $U(\mathbf{x})$ para cada distribución o de manera equivalente.

$$\int_{3.67}^{U_1} \pi_1(\theta|\mathbf{x}) d\theta = 0.95 \quad , \quad \int_{3.67}^{U_2} \pi_2(\theta|\mathbf{x}) d\theta = 0.95 \quad ,$$

o

$$\int_{3.67}^{U_1} C_1 \frac{2(4-\theta)}{9\theta^5} d\theta = 0.95 \quad , \quad \int_{3.67}^{U_2} \frac{C_2}{\theta^{8.5}} d\theta = 0.95 \quad ,$$

calculando las integrales y despejando U_1 y U_2

$$\frac{U_1^{-3}}{3} - U_1^{-4} = \frac{0.95 * 9}{2C_1} - 3.67^{-4} + \frac{3.67^{-3}}{3} \quad , \quad U_2 = \left(3.67^{-7.5} - \frac{0.95 * 7.5}{C_2} \right)^{-\frac{1}{7.5}} \quad ,$$

al resolver la expresión del lado izquierdo de manera numérica tenemos que $U_1 = 3.91577$; y al calcular la ecuación del lado derecho obtenemos que $U_2 = 5.47187$, por lo tanto, el intervalo con 0.95 de probabilidad para θ en el caso de la distribución $\pi_1(\theta|\mathbf{x})$ es $[3.67, 3.9177]$ mientras que para la distribución $\pi_2(\theta|\mathbf{x})$ es $[3.67, 5.47187]$.

Capítulo 2

El Algoritmo EM

2.1. Introducción.

No nos parecería descabellado pensar que la observación es parte del proceso de aprendizaje, de hecho esta es la forma que tenemos para darnos una idea de la “realidad objetiva”, muestra de ello es que la observación forma parte de uno de los pasos del método científico, más aún, gracias a ella nosotros podemos tomar decisiones. Sin embargo, puede ocurrir que durante la investigación de una población o fenómeno aleatorio existan datos que por una u otra razón no puedan ser observados y esto daría lugar a un problema con datos faltantes.

Los datos faltantes pueden surgir por varias causas, entre otras:

- (a) La imposibilidad de obtener la información, como por ejemplo, si uno quisiera hacer inferencia sobre una especie en peligro de extinción, en este caso es imposible obtener una muestra más grande que el tamaño de la población existente. Otro ejemplo es, cuando generar una observación tiene un costo muy alto y se cuenta con pocos recursos.
- (b) Al diseño del mismo experimento, como por ejemplo, supongamos que tenemos una cantidad N de focos, los cuales se pondrán en funcionamiento simultáneamente durante un período de tiempo C . Supongamos que una cantidad m de dichos focos se fundieron antes de llegar al tiempo C , estos son valores fijos u observados; mientras que los restantes $N-m$ focos seguían funcionando, y podrían seguirlo haciendo un tiempo desconocido mayor a C . Estos son valores faltantes que han pasado el umbral C .

- (c) Al extravío o no registro de la información de tal suerte que estos no pueden ser recuperados. Como pasa en las lecturas de secuencias de A.D.N. en las que en general hay datos faltantes.

Para resolver problemas de este estilo existe un algoritmo iterativo propuesto por Dempster, Laird y Rudin (1977), cuyo objetivo es encontrar la moda de la distribución posterior o de manera equivalente el estimador máximo verosímil en el caso en que se proponga una distribución *a priori* constante; el método consta de dos pasos: paso **E** (*Esperanza*) y el paso **M** (*Maximización*) y por tanto es conocido como el algoritmo EM.

Este método en lugar de hacer difíciles maximizaciones y simulaciones, propone “aumentar” los datos faltantes imputándoles variables aleatorias, las cuales llamamos variables latentes que nos faciliten el cálculo de la maximización de la función de verosimilitud aumentada o la distribución posterior aumentada con dichas variables latentes. Como se verá en los ejemplos, siempre resulta más fácil tratar con la verosimilitud aumentada.

De esta forma se trata de relacionar un problema de datos faltantes a uno con datos completos. O en palabras de los autores:

“ El término datos incompletos en su forma general implica la existencia de dos espacios muestrales \mathcal{Y} y \mathcal{Z} y una función de \mathcal{Z} a \mathcal{Y} . Los valores observados \mathbf{Y} son realizaciones de \mathcal{Y} . Los correspondientes \mathbf{Z} pertenecen a \mathcal{Z} los cuales no son observados directamente, sino sólo indirectamente a través de \mathbf{Y} . . . Nosotros suponemos una función $\mathbf{Z} \rightarrow \mathbf{Y}(\mathbf{Z})$ de \mathcal{Z} a \mathcal{Y} , y que \mathbf{Z} es conocido bajo la relación $\mathcal{Z}(\mathbf{Y})$, el subconjunto de \mathcal{Z} es determinado por la ecuación $\mathbf{Y}=\mathbf{Y}(\mathbf{Z})$, donde \mathbf{Y} son valores observados y \mathbf{Z} son los datos completados.

Postulamos una familia de densidades $f(\mathbf{Z}|\phi)$ que depende de los parámetros ϕ y derivada su correspondiente familia de densidades $g(\mathbf{Y}|\phi)$. La especificación completa de los datos $f(\cdot|\cdot)$ está relacionada a la especificación de los datos incompletos $g(\cdot|\cdot)$ por

$$g(\mathbf{Y}|\phi) = \int_{\mathcal{Z}(\mathbf{Y})} f(\mathbf{Z}|\phi) d\mathbf{Z}.$$

El algoritmo EM tiene como objetivo encontrar el valor de ϕ que maximice $g(\cdot|\cdot)$ dada una muestra \mathbf{Y} observada.” Dempster, Laird y Rudin (1977).

2.2. Descripción del algoritmo

El algoritmo es un método iterativo para encontrar la moda de la distribución posterior o el estimador máximo verosímil para datos incompletos. Como ya lo habíamos dicho, cada iteración consiste de dos pasos: el *paso E* (paso de la esperanza) y el *paso M* (paso de la maximización). Denotemos por θ^u la hipótesis actual para la moda de la distribución posterior de nuestro interés $\mathbb{P}(\theta|\mathbf{x})$ y sea $\mathbb{P}(\theta|\mathbf{x}, \mathbf{z})$ la distribución posterior aumentada, donde \mathbf{x} son los valores observados y \mathbf{z} los datos aumentados (variables latentes). Por último definamos a $\mathbb{P}(\mathbf{z}|\mathbf{x}; \theta^u)$ como la distribución predictiva condicional de los datos latentes \mathbf{z} dada la hipótesis actual de la moda posterior.

- **Entrada:** Una hipótesis actual para la moda de la distribución posterior de nuestro interés θ^u .
- **Salida:** La moda de la distribución posterior o el estimador obtenido a través del algoritmo EM θ^n .

Paso E Consiste en calcular

$$Q(\theta, \theta^u) = \int_{\mathbf{z}} \mathbb{P}(\mathbf{z}|\mathbf{x}; \theta^u) \log \mathbb{P}(\theta|\mathbf{x}, \mathbf{z}) d\mathbf{z}$$

es decir, la esperanza de $\log \mathbb{P}(\theta|\mathbf{x}, \mathbf{z})$ con respecto a $\mathbb{P}(\mathbf{z}|\mathbf{x}; \theta^u)$.

Paso M Consiste en maximizar la función $Q(\theta, \theta^u)$ con respecto a θ para obtener θ^{u+1} . Después de esto, regresamos al **Paso E**.

El algoritmo se itera hasta que $\|\theta^{u+1} - \theta^u\|$ o $|Q(\theta^{u+1}, \theta^u) - Q(\theta^u, \theta^u)|$ sea suficientemente pequeño.

2.3. Algunas aplicaciones del algoritmo EM

A continuación se presentan dos ejercicios donde se ilustran estas ideas al usar el algoritmo EM.

Ejemplo 2.1

Supongamos que 197 animales están separados en cuatro categorías genéticas de la siguiente forma

$$\mathbf{Y} = (y_1, y_2, y_3, y_4) = (125, 18, 20, 34).$$

Una variable aleatoria que describe el experimento de manera natural es una distribución multinomial, donde falta determinar p_1, p_2, p_3 y p_4 que son las probabilidades de caer en las categorías y_1, y_2, y_3 ó y_4 respectivamente, para tener el modelo multinomial probabilístico completamente especificado.

Supongamos que se propone al siguiente vector de probabilidades en función de θ

$$(p_1, p_2, p_3, p_4) = \left(\frac{1}{2} + \frac{\theta}{4}, \frac{1-\theta}{4}, \frac{1-\theta}{4}, \frac{\theta}{4} \right),$$

donde $\sum_{i=1}^4 p_i = 1$ y $0 \leq \theta \leq 1$.

El cual tiene sentido, ya que uno esperaría en función de lo observado, que las categorías y_2 y y_3 tengan la misma frecuencia de ocurrencia, mientras que, la categoría y_1 sea la que tiene mayor incidencia.

Notemos que hasta el momento no tenemos datos faltantes, y este problema se puede resolver de manera sencilla buscando el estimador máximo verosímil para θ de la siguiente función

$$f(\mathbf{Y} | \theta) = \frac{(y_1 + y_2 + y_3 + y_4)!}{y_1!y_2!y_3!y_4!} \left(\frac{1}{2} + \frac{\theta}{4} \right)^{y_1} \left(\frac{1-\theta}{4} \right)^{y_2} \left(\frac{1-\theta}{4} \right)^{y_3} \left(\frac{\theta}{4} \right)^{y_4}.$$

Por alguna razón las personas que hacen el experimento se dan cuenta que la primera categoría y_1 se puede dividir en dos, como resultado se tendrán cinco categorías distintas, es decir

$$\mathbf{Z} = (Z_1, Z_2, z_3, z_4, z_5) = (Z_1, Z_2, 18, 20, 34),$$

tal que $Z_1 + Z_2 = 125$.

En este caso, ya tenemos un problema con datos faltantes porque no sabemos con certeza cuantas observaciones pertenecen a Z_1 ó Z_2 .

Supongamos que se les asignan a las nuevas categorías Z_1 y Z_2 las siguientes probabilidades $\frac{1}{2}$ y $\frac{\theta}{4}$ respectivamente, nuevamente la distribución del experimento es una multinomial

$$g(\mathbf{Z}|\theta) = \frac{(Z_1 + Z_2 + z_3 + z_4 + z_5)!}{Z_1!Z_2!z_3!z_4!z_5!} \left(\frac{1}{2}\right)^{Z_1} \left(\frac{\theta}{4}\right)^{Z_2} \left(\frac{1-\theta}{4}\right)^{z_3} \left(\frac{1-\theta}{4}\right)^{z_4} \left(\frac{\theta}{4}\right)^{z_5},$$

al factorizar

$$= \frac{(Z_1 + Z_2 + z_3 + z_4 + z_5)!}{Z_1!Z_2!z_3!z_4!z_5!} \left(\frac{1}{4}\right)^{\frac{Z_1}{2} + Z_2 + z_3 + z_4 + z_5} \theta^{Z_2 + z_5} (1 - \theta)^{z_3 + z_4}.$$

Como los primeros dos factores son constantes con respecto a θ

$$g(\mathbf{Z}|\theta) \propto \theta^{Z_2 + z_5} (1 - \theta)^{z_3 + z_4}. \quad (2.1)$$

Antes de comenzar con el algoritmo notemos que:

- (1) $g(\mathbf{Z}|\theta)$ es precisamente la función de verosimilitud de una distribución multinomial. Por lo tanto, al maximizar (2.1) encontraremos el estimador máximo verosímil para θ

$$l_g(\mathbf{Z}|\theta) \propto g(\mathbf{Z}|\theta).$$

- (2) Si proponemos una distribución *a priori* constante en todo el rango de θ (el intervalo abierto $(0,1)$), es decir, una distribución no informativa. Y con la observación hecha en (1.1) tendremos que (2.1) es proporcional a la distribución *posterior* aumentada

$$\pi(\theta|\mathbf{Z}) \propto l_g(\mathbf{Z}|\theta).$$

- (3) Para facilitar los cálculos a la hora de maximizar trabajaremos con el logaritmo (2.1), quedando

$$\begin{aligned} L_g(\mathbf{Z}|\theta) &= \log(\pi(\theta|\mathbf{Z})) \\ &\propto (Z_2 + z_5) \log \theta + (z_3 + z_4) \log(1 - \theta). \end{aligned} \quad (2.2)$$

Paso E

Comencemos con el algoritmo tomando la esperanza de (2.2) condicionada al valor actual del parámetro θ^i y a la muestra observada \mathbf{Y} , a esta esperanza la denotaremos como $Q(\theta, \theta^i)$.

$$Q(\theta, \theta^i) = \mathbb{E} [L_g(\mathbf{Z}|\theta)|\theta^i, \mathbf{Y}]$$

$$= \mathbb{E} [(Z_2 + z_5) \log \theta + (z_3 + z_4) \log(1 - \theta)|\theta^i, \mathbf{Y}],$$

por la linealidad de la esperanza condicional

$$= (\mathbb{E} [Z_2|\theta^i, \mathbf{Y}] + z_5) \log \theta + (z_3 + z_4) \log(1 - \theta). \quad (2.3)$$

Antes de empezar con el paso el M, veamos cómo queda $\mathbb{E} [Z_2|\theta^i, \mathbf{Y}]$, ya que este cálculo nos será útil más adelante.

Observación 2

- (a) La variable aleatoria Z_2 depende únicamente del valor y_1 y θ^i y no de y_2, y_3, y_4 debido a que $Z_1 + Z_2 = y_1$, entonces $\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] = \mathbb{E}[Z_2|\theta^i, y_1]$.
- (b) La variable aleatoria $Z_2|y_1$ tiene una distribución binomial con $n = y_1$, y probabilidad de éxito $\frac{\frac{\theta^i}{4}}{\frac{1}{2} + \frac{\theta^i}{4}} = \frac{\theta^i}{2 + \theta^i}$.

Luego entonces

$$\mathbb{E}[Z_2|y_1] = y_1 \frac{\theta^i}{2 + \theta^i}.$$

Paso M

Debemos maximizar $Q(\theta, \theta^i)$ con respecto a θ , y así encontrar θ^{i+1} ; al derivar con respecto a θ e igualar a cero (2.3), da como resultado:

$$\frac{\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_5}{\theta} = \frac{z_3 + z_4}{1 - \theta},$$

de donde obtenemos

$$(1 - \theta)(\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_5) = \theta(z_3 + z_4).$$

O equivalentemente

$$\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_5 = \theta(\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_3 + z_4 + z_5).$$

Finalmente despejando a θ

$$\theta^{i+1} = \frac{\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_5}{\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_3 + z_4 + z_5}.$$

Al sustituir la esperanza tenemos que

$$\theta^{i+1} = \frac{y_1 \frac{\theta^i}{2+\theta^i} + z_5}{y_1 \frac{\theta^i}{2+\theta^i} + z_3 + z_4 + z_5}.$$

Notemos que θ^{i+1} es función sólo de θ^i y todos los demás términos son constantes.

Afirmamos que este punto crítico es máximo puesto que

$$\left. \frac{\partial^2 Q(\theta, \theta^i)}{\partial \theta^2} \right|_{\theta=\theta^{i+1}} = - \frac{(\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_3 + z_4 + z_5)(z_3 + z_4)}{\mathbb{E}[Z_2|\theta^i, \mathbf{Y}] + z_5} < 0.$$

En efecto, la última expresión es menor que cero ya que z_3, z_4 y z_5 son no negativos por ser observaciones de la muestra; mientras que $\mathbb{E}[Z_2|\theta^i, \mathbf{Y}]$ es mayor a cero por ser la esperanza de una distribución de conteo como lo es la binomial. Luego entonces, tanto el numerador y el denominador son positivos y al multiplicar por un signo negativo en general la última expresión es negativa.

Como consecuencia decimos que θ^{i+1} es el estimador máximo verosímil o el valor del parámetro actualizado. A continuación se presenta una tabla donde se dan tres valores iniciales distintos a θ_0 . En los tres casos notamos una rápida convergencia.

Cuadro 2.1:

i	θ^i	$ \theta^i - \theta^{i-1} $	θ^i	$ \theta^i - \theta^{i-1} $	θ^i	$ \theta^i - \theta^{i-1} $
0	0.5	-	0.1	-	0.9	-
1	0.6082474	0.1082470	0.5125229	0.4125230	0.6570184	-0.2429820
2	0.6243211	0.0160740	0.6102501	0.0977270	0.6307438	-0.0262750
3	0.6264889	0.0021680	0.6245940	0.0143440	0.6273408	-0.0034030
4	0.6267773	0.0002880	0.6265252	0.0019310	0.6268904	-0.0004500
5	0.6268156	0.0000380	0.6267822	0.0002570	0.6268306	-0.0000600
6	0.6268207	0.0000050	0.6268163	0.0000340	0.6268227	-0.0000080
7	0.6268214	0.0000010	0.6268208	0.0000050	0.6268217	-0.0000010
8	0.6268215	0.0000000	0.6268214	0.0000010	0.6268215	0.0000000

Si proponemos a $\hat{\theta} = 0.6268215$ como el valor del estimador de θ entonces las probabilidades de caer en cada categoría son

$$(p_1, p_2, p_3, p_4, p_5) = (0.5, 0.156705, 0.093294, 0.093294, 0.156705).$$

Ejemplo 2.2 (Análisis de regresión simple con valores censurados)

Consideremos los datos que se encuentran en Schmee y Hahn (1979). Estos datos corresponden a grupos de diez motores que se probaron a cuatro temperaturas distintas, 150°, 170°, 190° y 220° grados centígrados, los tiempos de falla fueron medidos en horas y se indica en el cuadro 2.2.

Con estos datos queremos estimar el siguiente modelo de regresión lineal:

$$t_i = \beta_0 + \beta_1 v_i + \sigma \epsilon_i, \quad (2.4)$$

donde:

(i) $\epsilon_i \sim N(0,1)$

Cuadro 2.2: Tiempos de falla

150°	170°	190°	220°
8064*	1764	408	408
8064*	2772	408	408
8064*	3444	1344	504
8064*	3542	1344	504
8064*	3780	1440	504
8064*	4860	1680*	528*
8064*	5196	1680*	528*
8064*	5448*	1680*	528*
8064*	5448*	1680*	528*
8064*	5448*	1680*	528*

(*) significa que el motor fue desconectado y llevaba funcionando la cantidad de tiempo que se indica.

$$(ii) v_i := \frac{1000}{temperatura + 273.2}$$

(iii) t_i es el logaritmo base diez del tiempo de falla del i -ésimo motor.

Como consecuencia inmediata del supuesto (i) y del supuesto (2.4) $t_i \sim N(\beta_0 + \beta_1 v_i, \sigma^2)$. Renombremos a $\beta_0 + \beta_1 v_i$ como μ_i .

Si en nuestra muestra no existieran datos censurados, es decir, valores que exceden un cierto umbral c_i , entonces podríamos fácilmente calcular los estimadores de β_0, β_1 y σ^2 por máxima verosimilitud

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t_i - \mu_i)^2} = (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (t_i - \mu_i)^2}. \quad (2.5)$$

En general, si deseamos ajustar una regresión lineal a una muestra de tamaño n con m valores fijos denotados por t_i (los valores observados) y los restantes $n-m$ datos que son variables latentes $\{Z_i : i = m+1, m+2, \dots, n\}$ que cumplen con la característica $\{Z_i > c_i\}$. Podemos escribir la verosimilitud aumentada con las variables Z_i de la siguiente manera

$$L = (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^m (t_i - \mu_i)^2 - \frac{1}{2\sigma^2} \sum_{i=m+1}^n (Z_i - \mu_i)^2}. \quad (2.6)$$

Nuevamente usamos el hecho de que la distribución *posterior* es proporcional a la verosimilitud por la distribución *a priori*. De la misma manera que en el ejercicio anterior, si multiplicamos a (2.6) por una distribución *a priori* con las siguientes características; $\pi(\boldsymbol{\theta}) \propto 1$ donde $\boldsymbol{\theta} = (\beta_0, \beta_1, \sigma)$ y $\boldsymbol{\theta} \in \mathbb{R}^3$, es decir, que sea equiprobable tomar cualquier posible valor del parámetro, tendremos una distribución *posterior* aumentada, que es por construcción proporcional a (2.6). Para empezar con el algoritmo EM necesitamos calcular el logaritmo de la distribución *posterior* aumentada

$$C - n \ln \sigma - \frac{\sum_{i=1}^m (t_i - \mu_i)^2}{2\sigma^2} - \frac{\sum_{i=m+1}^n (Z_i - \mu_i)^2}{2\sigma^2}, \quad (2.7)$$

donde $C = -\frac{n}{2} \ln(2\pi)$.

Paso E

Sea $U = \{\beta_0, \beta_1, \sigma, Z_i > c_i\}$ de forma que a continuación denotaremos

$$\mathbb{E}[\cdot|U] = \mathbb{E}[\cdot|\beta_0, \beta_1, \sigma, Z_i > c_i].$$

Calculemos la esperanza de (2.7)

$$Q(\beta_0, \beta_1, \sigma; \beta_0^i, \beta_1^i, \sigma^i) = \mathbb{E} \left[C - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^m (t_i - \mu_i)^2 - \frac{1}{2\sigma^2} \sum_{i=m+1}^n (Z_i - \mu_i)^2 | U \right],$$

por la linealidad de la esperanza condicional y desarrollando el binomio al cuadrado tenemos que

$$\begin{aligned} Q(\beta_0, \beta_1, \sigma; \beta_0^i, \beta_1^i, \sigma^i) &= C - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^m (t_i - \mu_i)^2 \\ &\quad - \frac{1}{2\sigma^2} \mathbb{E} \left[\sum_{i=m+1}^n (Z_i^2 - 2Z_i\mu_i + \mu_i^2) | U \right] \\ &= C - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^m (t_i - \mu_i)^2 \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=m+1}^n \{ \mathbb{E} [Z_i^2 | U] - 2\mu_i \mathbb{E} [Z_i | U] + \mu_i^2 \}. \end{aligned} \quad (2.8)$$

La función (2.8) es la que debemos maximizar con respecto a β_0, β_1 y σ en el paso M, pero antes de hacer esto, notemos que

$$\mathbb{E}[Z_i|U] = \mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right) \text{ y además} \quad (2.9)$$

$$\mathbb{E}[Z_i^2|U] = \mu_i^2 + \sigma^2 + H\left(\frac{c_i - \mu_i}{\sigma}\right) \{\sigma(c_i + \mu_i)\}, \quad (2.10)$$

donde $H\left(\frac{c_i - \mu_i}{\sigma}\right) := \frac{\phi\left(\frac{c_i - \mu_i}{\sigma}\right)}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)}$ y

ϕ y Φ son las funciones de densidad y distribución de una normal estándar respectivamente.

La demostración de estos resultados se puede consultar en el Apéndice A.

Paso M

Maximicemos la función (2.8) con respecto a los parámetros desconocidos β_0, β_1 y σ .

$$\frac{\partial Q(\beta_0, \beta_1, \sigma; \beta_0^i, \beta_1^i, \sigma^i)}{\partial \beta_0} = 0.$$

Sustituyendo lo que vale la función, usando la linealidad de la derivada y despejando obtenemos

$$-\frac{\partial}{\partial \beta_0} \left(\frac{1}{2\sigma^2} \sum_{i=1}^m (t_i - \beta_0 - \beta_1 v_i)^2 \right) = \frac{\partial}{\partial \beta_0} \left(\frac{1}{2\sigma^2} \sum_{i=m+1}^n \{ \mathbb{E}[Z_i^2|U] - 2(\beta_0 + \beta_1 v_i) \mathbb{E}[Z_i|U] + (\beta_0 + \beta_1 v_i)^2 \} \right).$$

Al calcular la derivada, y multiplicar de ambos lados por σ^2

$$\sum_{i=1}^m (t_i - \beta_0 - \beta_1 v_i) = - \sum_{i=m+1}^n \{ \mathbb{E}[Z_i|U] - \beta_0 - \beta_1 v_i \}.$$

Al distribuir la suma y despejar a β_0 tenemos parámetro actualizado β_0^{i+1}

$$\beta_0^{i+1} = \sum_{i=1}^m \frac{t_i}{n} + \sum_{i=m+1}^n \frac{\mathbb{E}[Z_i|U]}{n} - \beta_1 \bar{v}, \quad (2.11)$$

donde $\bar{v} = \sum_{i=1}^m v_i + \sum_{i=m+1}^n v_i$.

Por último sustituimos la esperanza condicional

$$\beta_0^{i+1} = \sum_{i=1}^m \frac{t_i}{n} + \sum_{i=m+1}^n \frac{\mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right)}{n} - \beta_1 \bar{v}. \quad (2.12)$$

Notemos que β_0^{i+1} queda en términos de β_1 , para que β_0^{i+1} quede completamente determinado habrá que reemplazar β_1 por el valor del estimador β_1^{i+1} que encontraremos después.

Ahora maximicemos Q con respecto a β_1

$$\frac{\partial Q(\beta_0, \beta_1, \sigma; \beta_0^{i+1}, \beta_1^i, \sigma^i)}{\partial \beta_1} = 0,$$

sustituyendo por lo que vale la función Q

$$-\frac{1}{2\sigma^2} \frac{\partial}{\partial \beta_1} \left(\sum_{i=1}^m (t_i - \beta_0^{i+1} - \beta_1 v_i)^2 + \sum_{i=m+1}^n \{ \mathbb{E}[Z_i^2|U] - 2(\beta_0^{i+1} + \beta_1 v_i) \mathbb{E}[Z_i|U] + (\beta_0^{i+1} + \beta_1 v_i)^2 \} \right) = 0,$$

después de calcular las derivadas y al multiplicar toda la expresión por σ^2 tenemos

$$\sum_{i=1}^m v_i (t_i - \beta_0^{i+1} - \beta_1 v_i) + \sum_{i=m+1}^n v_i \{ \mathbb{E}[Z_i|U] - \beta_0^{i+1} - \beta_1 v_i \} = 0,$$

si distribuimos las sumas podemos escribir

$$\sum_{i=1}^m t_i v_i + \sum_{i=m+1}^n \mathbb{E}[Z_i|U] v_i - \beta_0^{i+1} \sum_{i=1}^n v_i - \beta_1 \sum_{i=1}^n v_i^2 = 0,$$

puesto que $\sum_{i=1}^n v_i^j = \sum_{i=1}^m v_i^j + \sum_{i=m+1}^n v_i^j$ para $j=1, 2$.

Sustituyendo (2.11) en la última expresión

$$\sum_{i=1}^m t_i v_i + \sum_{i=m+1}^n \mathbb{E}[Z_i|U] v_i - \left[\sum_{i=1}^m \frac{t_i}{n} + \sum_{i=m+1}^n \frac{\mathbb{E}[Z_i|U]}{n} - \beta_1 \bar{v} \right] \sum_{i=1}^n v_i - \beta_1 \sum_{i=1}^n v_i^2 = 0.$$

Notemos que esta última expresión es únicamente función de β_1 , ahora despejemos a β_1

$$\beta_1^{i+1} = \frac{\sum_{i=m+1}^n \frac{\mathbb{E}[Z_i|U]}{n} \sum_{i=1}^n v_i - \sum_{i=m+1}^n \mathbb{E}[Z_i|U] v_i + \sum_{i=1}^m \frac{t_i}{n} \sum_{i=1}^n v_i - \sum_{i=1}^m t_i v_i}{\bar{v} \sum_{i=1}^n v_i - \sum_{i=1}^n v_i^2},$$

si multiplicamos el lado derecho por $\frac{1}{\bar{v}}$ resulta

$$\beta_1^{i+1} = \frac{\bar{v} \sum_{i=m+1}^n \frac{\mathbb{E}[Z_i|U]}{n} - \sum_{i=m+1}^n \frac{\mathbb{E}[Z_i|U] v_i}{n} + \bar{v} \sum_{i=1}^m \frac{t_i}{n} - \sum_{i=1}^m \frac{t_i v_i}{n}}{\bar{v}^2 - \sum_{i=1}^n \frac{v_i^2}{n}}.$$

Por último se sustituirán las esperanzas condicionales quedando

$$\begin{aligned} \beta_1^{i+1} &= \frac{\bar{v} \sum_{i=m+1}^n \frac{\mu_i + \sigma H(\frac{c_i - \mu_i}{\sigma})}{n} - \sum_{i=m+1}^n \frac{(\mu_i + \sigma H(\frac{c_i - \mu_i}{\sigma})) v_i}{n}}{\bar{v}^2 - \sum_{i=1}^n \frac{v_i^2}{n}} \\ &+ \frac{\bar{v} \sum_{i=1}^m \frac{t_i}{n} - \sum_{i=1}^m \frac{t_i v_i}{n}}{\bar{v}^2 - \sum_{i=1}^n \frac{v_i^2}{n}} \end{aligned} \quad (2.13)$$

Una vez conocido β_1^{i+1} se sustituye en (2.12) y β_0^{i+1} queda determinado.

Para la maximización de σ^{i+1} se sustituirán los estimadores β_0^{i+1} y β_1^{i+1} calculados previamente. Definamos a $\mu_j^i = \beta_0^{i+1} + \beta_1^{i+1}v_j$.

$$\frac{\partial Q(\beta_0, \beta_1, \sigma; \beta_0^{i+1}, \beta_1^{i+1}, \sigma^i)}{\partial \sigma} = 0,$$

sustituyendo Q

$$\frac{\partial}{\partial \sigma} \left\{ C - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^m (t_i - \mu_i)^2 - \frac{1}{2\sigma^2} \sum_{i=m+1}^n \{ \mathbb{E} [Z_i^2|U] - 2\mu_i \mathbb{E} [Z_i|U] + \mu_i^2 \} \right\} = 0,$$

por la linealidad de la derivada

$$-\frac{n}{\sigma} - \frac{\partial}{\partial \sigma} \left(\frac{1}{2\sigma^2} \sum_{j=1}^m (t_j - \mu_j^i)^2 \right) - \frac{\partial}{\partial \sigma} \left(\frac{1}{2\sigma^2} \sum_{j=m+1}^n \{ \mathbb{E} [Z_j^2|U] - 2\mu_j^i \mathbb{E} [Z_j|U] + (\mu_j^i)^2 \} \right) = 0$$

$$-\frac{n}{\sigma} + 2\sigma^{-3} \left(\frac{1}{2} \sum_{j=1}^m (t_j - \mu_j^i)^2 \right) + 2\sigma^{-3} \left(\frac{1}{2} \sum_{j=m+1}^n \{ \mathbb{E} [Z_j^2|U] - 2(\mu_j^i) \mathbb{E} [Z_j|U] + (\mu_j^i)^2 \} \right) = 0.$$

De esta ecuación despejemos a σ

$$\frac{n}{\sigma} = \sigma^{-3} \sum_{j=1}^m (t_j - \mu_j^i)^2 + \sigma^{-3} \sum_{j=m+1}^n \{ \mathbb{E} [Z_j^2|U] - 2(\mu_j^i) \mathbb{E} [Z_j|U] + (\mu_j^i)^2 \}.$$

O equivalentemente

$$\sigma^2 = \frac{\sum_{j=1}^m (t_j - \mu_j^i)^2}{n} + \frac{\sum_{j=m+1}^n \{ \mathbb{E} [Z_j^2|U] - 2(\mu_j^i) \mathbb{E} [Z_j|U] + (\mu_j^i)^2 \}}{n}.$$

Sustituyendo las esperanzas

$$\sigma^2 = \frac{\sum_{j=1}^m (t_j - \mu_j^i)^2}{n} + \frac{\sum_{j=m+1}^n \left\{ \left(\mu_j^2 + \sigma^2 + \sigma(c_j + \mu_j)H\left(\frac{c_j - \mu_j}{\sigma}\right) \right) - 2\mu_j^i \left(\mu_j + \sigma H\left(\frac{c_j - \mu_j}{\sigma}\right) \right) + (\mu_j^i)^2 \right\}}{n}.$$

Observación 3

Si no tuvieramos datos faltantes, esto es, $m = n$ entonces los estimadores que hemos calculado en los tres casos se reducen a los estimadores máximo verosimiles o mínimo cuadráticos de β_0, β_1 y σ ¹. Es decir que

$$\beta_0^{i+1} = \sum_{i=1}^m \frac{t_i}{n} + \sum_{i=m+1}^n \frac{\mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right)}{n} - \beta_1 \bar{v},$$

se reduce a

$$\begin{aligned} \beta_0^{i+1} &= \sum_{i=1}^n \frac{t_i}{n} - \beta_1 \bar{v} \\ &= \bar{t} - \beta_1 \bar{v}. \end{aligned}$$

En el caso de β_1 tenemos que

$$\beta_1^{i+1} = \frac{\bar{v} \sum_{i=m+1}^n \frac{\mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right)}{n} - \sum_{i=m+1}^n \frac{(\mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right))v_i}{n}}{\bar{v}^2 - \sum_{i=1}^n \frac{v_i^2}{n}} + \frac{\bar{v} \sum_{i=1}^m \frac{t_i}{n} - \sum_{i=1}^m \frac{t_i v_i}{n}}{\bar{v}^2 - \sum_{i=1}^n \frac{v_i^2}{n}},$$

se reduce a

$$\beta_1^{i+1} = \frac{\bar{v} \bar{t} - \sum_{i=1}^n \frac{t_i v_i}{n}}{\bar{v}^2 - \sum_{i=1}^n \frac{v_i^2}{n}}.$$

¹Puesto que los estimadores de β_0, β_1 coinciden bajo estos dos métodos

Y para el caso de σ

$$\sigma^2 = \frac{\sum_{j=1}^m (t_j - \mu_j^i)^2}{n} + \frac{\sum_{j=m+1}^n \left\{ (\mu_j^2 + \sigma^2 + \sigma(c_j + \mu_j)H(\frac{c_j - \mu_j}{\sigma})) - 2\mu_j^i(\mu_j + \sigma H(\frac{c_j - \mu_j}{\sigma})) + (\mu_j^i)^2 \right\}}{n},$$

se reduce a

$$\sigma^2 = \frac{\sum_{j=1}^n (t_j - \mu_j)^2}{n}.$$

Esta observación nos muestra que es consistente lo que hemos calculado.

Una vez que ya tenemos los estimadores, lo podemos programar y para empezar las iteraciones debemos dar valores iniciales a β_0^0 , β_1^0 y σ^0 . Una manera que se optó para dar dichos valores iniciales fue correr una regresión lineal por mínimos cuadrados donde se tomó en cuenta únicamente los valores fijos u observados bajo este procedimiento obtuvimos que $\beta_0^0 = -5.15$, $\beta_1^0 = 3.82$ y $\sigma^0 = 0.219$. Acto seguido se hicieron una serie de iteraciones del algoritmo dando como resultado la siguiente tabla.

Cuadro 2.3:

i	β_0^i	β_1^i	σ^i
0	-5.15	3.82	0.219
1	-5.35152705	3.93098389	0.04805809
2	-5.59040665	4.05256721	0.04800388
3	-5.80404617	4.15810481	0.04833615
4	-5.97165015	4.23959046	0.04874269
5	-6.09461223	4.29875477	0.04909053
6	-6.18165575	4.34032201	0.04934930
7	-6.24203930	4.36899145	0.04952923
8	-6.28341516	4.38854661	0.04965004
9	-6.31153821	4.40178953	0.04972969
10	-6.33054567	4.41071347	0.04978172
11	-6.34333910	4.41670548	0.04981557
12	-6.35192301	4.42071797	0.04983756
13	-6.35766840	4.42339930	0.04985185
14	-6.36150647	4.42518814	0.04986115
15	-6.36406642	4.42637998	0.04986721
16	-6.36577174	4.42717322	0.04987117
17	-6.36690659	4.42770072	0.04987376
18	-6.36766117	4.42805125	0.04987546
19	-6.36816257	4.42828405	0.04987657
20	-6.36849554	4.42843859	0.04987730
\vdots	\vdots	\vdots	\vdots
50	-6.36915	4.42874	0.0498787

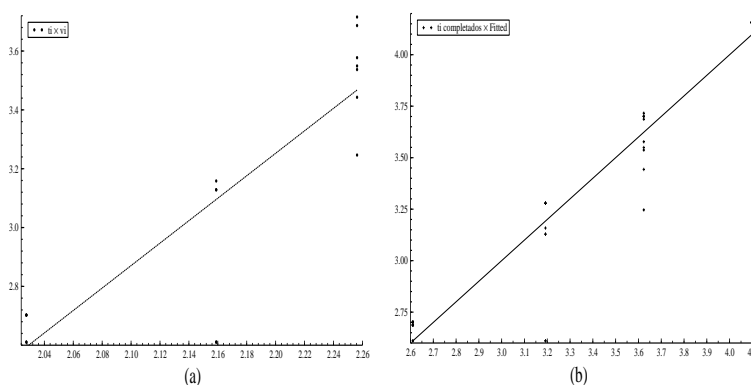


Figura 2.1: (a) Regresión con valores fijos y (b) con la muestra completada

Capítulo 3

Un problema de datos censurados con identificación parcial

3.1. Introducción.

El problema que se plantea en este capítulo tiene la siguiente descripción; supongamos que un artefacto está conformado de k componentes distintos. Para que el artefacto funcione es necesario que cada componente funcione. Nosotros podemos considerar que el tiempo de falla así descrito, es una muestra de una de las k distribuciones F_j con $j = 1, 2, \dots, k$, donde cada una de estas distribuciones está asociada a un tipo de falla.

Además supongamos que según el diseño del experimento, el artefacto se mantendrá funcionando a lo más un periodo de tiempo C , esto nos genera dos posibles escenarios:

- (i) Si el artefacto se descompone en un tiempo $t \leq C$, entonces sabremos el valor del tiempo de falla y de qué distribución provino.
- (ii) Mientras que si el artefacto funciona hasta el tiempo C , el experimento acabará y no sabremos cuanto tiempo adicional a C hubiera seguido funcionando ni tampoco cual de los componentes es el que hubiera originado la falla; lo único que sabremos es que su tiempo de falla es mayor a C .

Consideremos que podemos obtener N observaciones independientes de

tiempos de falla bajo el esquema anterior. Puesto que la identidad de la falla o la distribución F_j no es conocida para todas las observaciones, usaremos una mezcla de distribuciones para modelar la muestra completa.

Sea $G(\cdot; \boldsymbol{\theta}, \boldsymbol{\pi})$ la función de distribución que representa la mezcla de las k funciones de distribución $F_j \equiv F_j(\cdot; \theta_j)$, para $j \in \{1, 2, \dots, k\}$.

$$G(\cdot; \boldsymbol{\theta}, \boldsymbol{\pi}) = \sum_{j=1}^k \pi_j F_j(\cdot; \theta_j),$$

donde $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_k)$ es el vector de los parámetros y $\boldsymbol{\pi}' = (\pi_1, \dots, \pi_k)$ es el vector de los pesos de la mezcla. Notemos que tanto los componentes de $\boldsymbol{\theta}'$ y $\boldsymbol{\pi}'$ son todos desconocidos; lo único que sabemos es : $\pi_j \geq 0$ y $\sum_{j=1}^k \pi_j = 1$.

La información proporcionada por la muestra de tamaño N la podemos dividir en:

- (a) los datos observados, que a su vez los podemos agrupar de la siguiente manera

$$\{x_{1i}\}_{i=1}^{r_1}, \{x_{2i}\}_{i=1}^{r_2}, \dots, \{x_{ki}\}_{i=1}^{r_k},$$

donde, para cada $j \in \{1, 2, \dots, k\}$, $x_{j1}, x_{j2}, \dots, x_{jr_j}$ y r_j denota el número de observaciones las cuales fueron menores o iguales a C y fueron identificadas como provenientes de la distribución F_j , y además

- (b) la cantidad de datos censurados que pasaron el umbral C es

$$N - \sum_{j=1}^k r_j \equiv N - r$$

puesto que r denota la cantidad total de observaciones que fueron menores o iguales a C .

Denotemos a $L(\boldsymbol{\theta}, \boldsymbol{\pi}; \mathbf{x})$ como la función de verosimilitud para $\boldsymbol{\theta}$ y $\boldsymbol{\pi}$. Esta función es proporcional a

$$\prod_{j=1}^k \prod_{i=1}^{r_j} \{\pi_j f_j(x_{ji}; \theta_j)\} \times \{1 - G(C; \boldsymbol{\theta}, \boldsymbol{\pi})\}^{N-r}, \quad (3.1)$$

donde $f_j(\cdot; \theta_j)$ es la función de desidad de la correspondiente función de distribución $F_j(\cdot; \theta_j)$.

La expresión (3.1) la podemos reescribir de la siguiente forma

$$\begin{aligned}
 & \left\{ \prod_{j=1}^k \prod_{i=1}^{r_j} f_j(x_{ji}; \theta_j) \right\} \prod_{j=1}^k \pi_j^{r_j} \times \{1 - G(C; \boldsymbol{\theta}, \boldsymbol{\pi})\}^{N-r} \\
 &= \left\{ \prod_{j=1}^k \prod_{i=1}^{r_j} f_j(x_{ji}; \theta_j) \right\} \prod_{j=1}^k \left\{ \frac{\pi_j F_j(C; \theta_j)}{F_j(C; \theta_j)} \right\}^{r_j} \times \{1 - G(C; \boldsymbol{\theta}, \boldsymbol{\pi})\}^{N-r} \\
 &= \underbrace{\left\{ \prod_{j=1}^k \prod_{i=1}^{r_j} \frac{f_j(x_{ji}; \theta_j)}{F_j(C; \theta_j)} \right\}}_A \underbrace{\prod_{j=1}^k \{\pi_j F_j(C; \theta_j)\}^{r_j}}_B \times \{1 - G(C; \boldsymbol{\theta}, \boldsymbol{\pi})\}^{N-r}. \quad (3.2)
 \end{aligned}$$

El factor A de (3.2) es la función de densidad de los datos observados $\{x_{ji}\}_{i=1}^{r_j}$, para toda $j \in \{1, \dots, k\}$ condicionada al evento de que la observación fue menor o igual a C ; mientras que el factor B de la misma ecuación, corresponde a la probabilidad multinomial asociada a los conteos r_1, r_2, \dots, r_k y $N - r$.

En Mendenhall y Hader(1958) y Cox (1958) se han estudiado problemas de esta naturaleza. Más recientemente Díaz Francés (1998) trabajó este esquema modelando F_j como una distribución Weibull.

Siguiendo este último trabajo, supongamos que \tilde{F}_j se modela con una distribución Weibull con diferentes parámetros de forma y escala. Recordemos que si Y se distribuye Weibull con parámetros $\alpha, \beta > 0$, entonces su función de distribución está dada por

$$\tilde{F}_j(y; \alpha_j, \beta_j) = 1 - \exp \left\{ - \left(\frac{y}{\beta_j} \right)^{\alpha_j} \right\}. \quad (3.3)$$

Si consideramos la transformación $X = \log\left(\frac{Y}{C}\right)$

$$F_j(x; \alpha_j, \beta_j) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\log\left(\frac{Y}{C}\right) \leq x\right) = \mathbb{P}(Y \leq C e^x) = \tilde{F}_j(C e^x; \alpha_j, \beta_j) \text{ y}$$

al sustituir en (3.3) tenemos que

$$\begin{aligned}
 F_j(x; \alpha_j, \beta_j) &= 1 - \exp \left\{ - \left(\frac{C e^x}{\beta_j} \right)^{\alpha_j} \right\} \\
 &= 1 - \exp \left\{ - \left(\frac{\beta_j}{C} \right)^{-\alpha_j} e^{x\alpha_j} \right\} \\
 &= 1 - \exp \left\{ - e^{-\alpha_j \log \left(\frac{\beta_j}{C} \right)} e^{x\alpha_j} \right\},
 \end{aligned}$$

sea $\mu_j = \log \left(\frac{\beta_j}{C} \right)$, entonces

$$\begin{aligned}
 F_j(x; \alpha_j, \mu_j) &= 1 - \exp \left\{ - e^{-\alpha_j \mu_j} e^{x\alpha_j} \right\} \\
 &= 1 - \exp \left\{ - e^{\alpha_j(x - \mu_j)} \right\} \\
 &= 1 - \exp \left\{ - e^{\frac{x - \mu_j}{\sigma_j}} \right\}.
 \end{aligned}$$

Esta última expresión se tiene si se define a α_j como σ_j^{-1} , o equivalentemente

$$\begin{aligned}
 F_j(x; \mu_j, \sigma_j) &= 1 - \exp \left\{ - e^{(x - \mu_j)\sigma_j^{-1}} \right\} \\
 &= 1 - \exp \left\{ - e^{(x - \mu_j)e^{-\log(\sigma_j)}} \right\} \\
 &= 1 - \exp \left\{ - e^{(x - \mu_j)e^{-\varphi_j}} \right\}, \tag{3.4}
 \end{aligned}$$

donde $\varphi_j = \log(\sigma_j)$.

Esta última distribución es una distribución de valores extremos conocida en la literatura como Gumbel con parámetros de localización μ_j y escala φ_j . Una de las ventajas de esta transformación es que el nuevo umbral de censura es $C_0 = 0$, esto debido a la transformación propuesta, ya que $\log(C/C) = \log(1) = 0$. Por otra parte, la parametrización localización-escala es adecuada para trabajar con aproximación normal.

De la función de distribución que hemos obtenido en (3.4)

$$F_j(x; \mu_j, \varphi_j) = 1 - \exp \left\{ - e^{(x - \mu_j)e^{-\varphi_j}} \right\},$$

al derivar con respecto a x obtenemos la función de densidad

$$f_j(x; \mu_j, \varphi_j) = e^{-\varphi_j} e^{(x-\mu_j)e^{-\varphi_j}} \exp \left\{ -e^{(x-\mu_j)e^{-\varphi_j}} \right\}. \quad (3.5)$$

En Contreras-Cristán et al. (2002) se utiliza esta modelación para resolver el problema de datos con censura descrito al inicio de este capítulo. En particular en esta tesis trabajaremos un algoritmo EM para los datos de tiempos de falla estudiados originalmente en Mendelhall y Hader (1958).

3.2. Un algoritmo EM

Consideremos a Ξ como el vector de todos los parámetros, definido de la siguiente manera

$$\begin{aligned} \Xi &= (\boldsymbol{\theta}', \boldsymbol{\pi}') \\ &= ((\theta'_1, \theta'_2, \dots, \theta'_k), (\pi_1, \pi_2, \dots, \pi_k)), \end{aligned}$$

donde $\theta'_j = (\mu_j, \varphi_j)$.

Con el fin de plantear el algoritmo EM, comencemos por considerar a los datos aumentados como las parejas

$$z_1 = (x_{01}, i_1), z_2 = (x_{02}, i_2), \dots, z_{N-r} = (x_{0N-r}, i_{N-r}),$$

donde la primera entrada de cada pareja es un dato no observado, mientras que la segunda entrada es un número entero, el cual indica la población a la que corresponde la primera entrada. Entonces la verosimilitud aumentada con estos datos que denotaremos por $L(\mathbf{x}, \mathbf{z}|\Xi)$ es tal que

$$L(\mathbf{x}, \mathbf{z}|\Xi) \propto \prod_{j=1}^k \prod_{i=1}^{r_j} \pi_j f_j(x_{ji}; \theta_j) \times \prod_{l=1}^{N-r} \{\pi_{i_l} f_{i_l}(x_{0l}; \theta_{i_l})\}. \quad (3.6)$$

Por otra parte, de las ecuaciones (3.2) y (3.6) se sigue que la distribución predictiva para los datos aumentados es

$$\mathbb{P}(z_1, z_2, \dots, z_{N-r} | \mathbf{x}; \Xi) = \prod_{l=1}^{N-r} \left\{ \frac{\pi_{i_l} f_{i_l}(x_{0l}; \theta_{i_l})}{1 - G(0; \Xi)} \right\}, \quad (3.7)$$

para $z_l = (x_{0l}, i_l)$, con $i_l \in \{1, 2, \dots, k\}$, y $x_{0l} \geq 0$ para $l \in \{1, 2, \dots, N-r\}$.

En el Apéndice B se muestra que (3.7) es una densidad para z_1, z_2, \dots, z_{N-r} , es decir

$$\int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \prod_{l=1}^{N-r} \left\{ \frac{\pi_{i_l} f_{i_l}(x_{0l}; \theta_{i_l})}{1 - G(0; \Xi)} \right\} dx_{01} \dots dx_{0N-r} = 1. \quad (3.8)$$

Como en los dos ejemplos del algoritmo EM del segundo capítulo suponemos una distribución *a priori* vaga o no informativa $\mathbb{P}(\Xi) \propto 1$, entonces la correspondiente distribución *posterior* para Ξ es

$$\mathbb{P}(\Xi | \mathbf{x}) \propto L(\Xi; \mathbf{x}).$$

Notemos que de cambiar la especificación inicial $\mathbb{P}(\Xi)$ las expresiones que se derivan a continuación también cambiarán.

De sustituir (3.5) en (3.6) obtenemos la densidad de la posterior aumentada

$$\begin{aligned} \mathbb{P}(\Xi | \mathbf{z}; \mathbf{x}) &\propto \prod_{j=1}^k \left\{ \prod_{i=1}^{r_j} \pi_j e^{-\varphi_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \exp\{-e^{(x_{ji} - \mu_j) e^{-\varphi_j}}\} \right\} \\ &\times \prod_{l=1}^{N-r} \left\{ \pi_{i_l} e^{-\varphi_{i_l}} e^{(x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}}} \exp\{-e^{(x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}}}\} \right\}. \quad (3.9) \end{aligned}$$

Para el paso E, necesitamos calcular la esperanza de $\log\{\mathbb{P}(\Xi | \mathbf{z}; \mathbf{x})\}$ con respecto a la distribución predictiva aumentada (3.7) evaluada en el valor actual Ξ^u de Ξ .

Primero calculemos el logaritmo a (3.9)

$$\begin{aligned} \log \mathbb{P}(\Xi | \mathbf{z}; \mathbf{x}) &\propto \sum_{j=1}^k \sum_{i=1}^{r_j} \left\{ \log\{\pi_j\} - \varphi_j + (x_{ji} - \mu_j)e^{-\varphi_j} - e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \right\} \\ &+ \sum_{l=1}^{N-r} \left\{ \log\{\pi_{i_l}\} - \varphi_{i_l} + (x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}} - e^{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}} \right\}. \end{aligned}$$

Paso E

Necesitamos calcular la función

$$\begin{aligned} Q(\Xi, \Xi^u) &\equiv \mathbb{E} \{ \log\{\mathbb{P}(\Xi | \mathbf{x}, \mathbf{z})\} | \Xi^u; \mathbf{x} \} = \\ &\int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \log\{\mathbb{P}(\Xi | \mathbf{x}, \mathbf{z})\} \prod_{l=1}^{N-r} \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} dx_{01} \dots dx_{0N-r}. \end{aligned}$$

Para el siguiente cálculo, definamos al símbolo \triangleq como *igualdad salvo por sumandos que son constantes con respecto a Ξ*

$$\begin{aligned}
Q(\Xi, \Xi^u) &\triangleq \sum_{j=1}^k \sum_{i=1}^{r_j} \left\{ \log\{\pi_j\} - \varphi_j + (x_{ji} - \mu_j)e^{-\varphi_j} - e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \right\} \\
&+ \sum_{l=1}^{N-r} \sum_{i_l=1}^k \log\{\pi_{i_l}\} \left\{ \frac{\pi_{i_l}^u \{1 - F(0; \theta_{i_l}^u)\}}{1 - G(0; \Xi^u)} \right\} \\
&- \sum_{l=1}^{N-r} \sum_{i_l=1}^k \varphi_{i_l} \left\{ \frac{\pi_{i_l}^u \{1 - F(0; \theta_{i_l}^u)\}}{1 - G(0; \Xi^u)} \right\} \\
&+ \sum_{l=1}^{N-r} \sum_{i_l=1}^k \int_0^\infty (x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}} \left\{ \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} \right\} dx_{0l} \\
&- \sum_{l=1}^{N-r} \sum_{i_l=1}^k \int_0^\infty e^{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}} \left\{ \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} \right\} dx_{0l} \\
&= \sum_{j=1}^k \left\{ r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \log \pi_j \\
&- \sum_{j=1}^k \left\{ r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \varphi_j \\
&+ \sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} - \sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \\
&+ (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \right\} \\
&- (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\}, \tag{3.10}
\end{aligned}$$

donde

$$\begin{aligned}
m(\theta_j^u) &= e^{\varphi_j^u} \int_{l(\theta_j^u)}^\infty \log\{v\} e^{-v} dv + \mu_j^u \int_{l(\theta_j^u)}^\infty e^{-v} dv, \\
\nu(\theta_j^u; \varphi_j) &= e^{\mu_j^u e^{-\varphi_j}} \int_{l(\theta_j^u)}^\infty v e^{\varphi_j^u - \varphi_j} e^{-v} dv
\end{aligned}$$

y $l(\theta_j^u) = e^{-\mu_j^u e^{-\varphi_j^u}}$.

El lector interesado en ver con detalle los cálculos anteriores, los puede encontrar en el Apéndice B, esto con el fin de no desviar la exposición del problema.

Observemos que (3.10) la podemos expresar como la suma de dos funciones, una que es función exclusivamente del vector (π_1, \dots, π_k) y otra que es función de $(\varphi_1, \dots, \varphi_k; \mu_1, \dots, \mu_k)$, es decir

$$Q(\Xi, \Xi^u) \triangleq h(\pi_1, \dots, \pi_k) + g(\varphi_1, \dots, \varphi_k; \mu_1, \dots, \mu_k),$$

donde

$$h(\pi_1, \dots, \pi_k) = \sum_{j=1}^k \left\{ r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \log \pi_j$$

y

$$\begin{aligned} g(\varphi_1, \dots, \varphi_k; \mu_1, \dots, \mu_k) &= - \sum_{j=1}^k \left\{ r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \varphi_j \\ &+ \sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} - \sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\ &+ (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \right\} \\ &- (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\}. \end{aligned}$$

Paso M

Como antes se mencionó Q se puede expresar como la suma de dos funciones entonces para maximizar a Q con respecto a Ξ , es equivalente encontrar el siguiente máximo

$$\max_{\Xi}\{Q\} = \max_{\pi}\{h\} + \max_{\theta}\{g\}.$$

Puesto que la función h es función de $(\pi_1, \pi_2, \dots, \pi_k)$ y por definición del problema tenemos que $\sum_{i=1}^k \pi_i = 1$, se maximizó h usando multiplicadores de Lagrange. Alcanzando el máximo en el punto ¹

$$\pi_j = \frac{r_j + (N - r) \left\{ \frac{\pi_j^u (1 - F(0; \theta_j^u))}{1 - G(0; \Xi^u)} \right\}}{N}, \quad j = 1, 2, \dots, k. \quad (3.11)$$

Por otro lado, al diferenciar g con respecto a μ_j y φ_j , para $j \in \{1, 2, \dots, k\}$, el punto $(\mu_1, \dots, \mu_k; \varphi_1, \dots, \varphi_k)$ donde g se maximiza debe de ser tal que la derivada parcial de g con respecto a cada componente de θ se anula, es decir, si ∇_{θ} denota al vector gradiente de g con respecto a θ entonces el punto $(\mu_1, \dots, \mu_k; \varphi_1, \dots, \varphi_k)$ debe satisfacer

$$\nabla_{\theta} g = \mathbf{0}.$$

Luego entonces

$$\begin{aligned} \frac{\partial g}{\partial \mu_j} &= -r_j + \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j)e^{-\varphi_j}} - (N - r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \\ &+ (N - r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} = 0, \quad j = 1, \dots, k \end{aligned} \quad (3.12)$$

despejando μ_j de esta última expresión

$$\hat{\mu}_j = -e^{\varphi_j} \log \left\{ \frac{r_j + (N - r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)}}{\sum_{i=1}^{r_j} e^{x_{ji} e^{-\varphi_j}} + (N - r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)}} \right\} \quad (3.13)$$

¹La demostración se encuentra en el Apéndice B

y

$$\begin{aligned}
 \frac{\partial g}{\partial \varphi_j} &= - \left\{ r_j + (N - r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} e^{\varphi_j} - \sum_{i=1}^{r_j} (x_{ji} - \mu_j) \\
 &+ \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{(x_{ji} - \mu_j) e^{-\varphi_j}} - (N - r) \frac{\pi_j^u \{m(\theta_j^u) - \mu_j (1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} \\
 &- (N - r) \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} \\
 &= 0, \quad j = 1, 2, \dots, k,
 \end{aligned} \tag{3.14}$$

donde

$$\omega(\theta_j^u; \varphi_j) = \mu_j^u \nu(\theta_j^u; \varphi_j) + e^{\mu_j^u e^{-\varphi_j}} e^{\varphi_j^u} \int_{l(\theta_j^u)}^{\infty} v e^{\varphi_j^u - \varphi_j} \log\{v\} e^{-v} dv.$$

Los cálculos de la maximización de h y g el lector los puede encontrar en el Apéndice B.

La iteración del algoritmo, al calcular primero el paso E como en la ecuación (3.10) y después al hacer el paso M como en (3.11), (3.13) y (3.14) nos dará una sucesión de estimadores $\Xi^{(1)}, \Xi^{(2)}, \dots$ para Ξ .

Después de un número de iteraciones, detenemos el proceso cuando $\|\Xi^{u+1} - \Xi^u\|$ es suficientemente pequeño. El último valor Ξ^η obtenido al momento de detener las iteraciones será nuestro estimador puntual para Ξ .

Notemos que la solución numérica de la ecuación (3.14) no es trivial, es por eso que con ayuda de *Mathematica* se instrumentaron los pasos (3.10), (3.11), (3.13) y (3.14) para la obtención del estimador puntual de Ξ .

3.3. Desviación estándar en el contexto del algoritmo EM.

Como ya se mencionó, el resultado de aplicar el algoritmo EM nos da estimadores puntuales para la moda de la distribución posterior $\mathbb{P}(\Xi|\mathbf{x})$. Para que nosotros pudiéramos hacer intervalos de confianza con respecto a los

parámetros, necesitamos encontrar formas para calcular errores estándar con respecto a los estimadores puntuales que hemos encontrado.

Una vez obtenido un punto fijo Ξ^η via el algoritmo EM, la matriz de varianzas-covarianzas asintótica para Ξ^η se puede calcular de la siguiente manera

$$V(\Xi^\eta) = \left\{ -\frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} \right\}^{-1}. \quad (3.15)$$

En la práctica esto puede ser muy complicado de evaluar, es por eso que se han ideado métodos asintóticos. Discutiremos aquí uno propuesto por Meng y Rubin (1991).

Es fácil notar que

$$\mathbb{P}(\Xi|\mathbf{x}) = \frac{\mathbb{P}(\Xi|\mathbf{x}, \mathbf{z})\mathbb{P}(\mathbf{z}|\mathbf{x})}{\mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi)}.$$

Aplicando logaritmo a ambos lados

$$\log \mathbb{P}(\Xi|\mathbf{x}) = \log \mathbb{P}(\Xi|\mathbf{x}, \mathbf{z}) - \log \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) + K,$$

donde $K = \log \mathbb{P}(\mathbf{z}|\mathbf{x})$, la cual es una constante con respecto a Ξ .

Derivando dos veces con respecto a Ξ y multiplicando por menos uno

$$-\frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2} = -\frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x}, \mathbf{z})}{\partial \Xi^2} + \frac{\partial^2 \log \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi)}{\partial \Xi^2}.$$

Multiplicando por $\mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi)$ e integrando con respecto a \mathbf{z}

$$\begin{aligned} -\frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2} &= -\int_{\mathbf{z}} \frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x}, \mathbf{z})}{\partial \Xi^2} \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) d\mathbf{z} \\ &+ \int_{\mathbf{z}} \frac{\partial^2 \log \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi)}{\partial \Xi^2} \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) d\mathbf{z}, \end{aligned} \quad (3.16)$$

puesto que

$$\begin{aligned} - \int_{\mathbf{z}} \frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2} \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) d\mathbf{z} &= - \frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2} \int_{\mathbf{z}} \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) d\mathbf{z} \\ &= - \frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2}. \end{aligned}$$

En lo siguiente, supondremos que se dan las condiciones necesarias para intercambiar el orden en las operaciones de integral y derivada. Estas condiciones se conocen como condiciones de regularidad (véase Lehmann (1983)) y primordialmente se basan en el hecho de que el soporte de la distribución $F(\cdot, \Xi)$ no varía con el valor de Ξ .

De la definición de $Q(\cdot, \cdot)$ obtenemos que

$$-\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} = - \int_{\mathbf{z}} \frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x}, \mathbf{z})}{\partial \Xi^2} \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) d\mathbf{z}.$$

Por otra parte definamos

$$-\frac{\partial^2 H(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} = - \int_{\mathbf{z}} \frac{\partial^2 \log \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi)}{\partial \Xi^2} \mathbb{P}(\mathbf{z}|\mathbf{x}, \Xi) d\mathbf{z}.$$

Sustituyendo esto en (3.16) se sigue que

$$-\frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2} = \underbrace{\frac{-\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta}}_{(a)} - \underbrace{\frac{-\partial^2 H(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta}}_{(b)}. \quad (3.17)$$

A la ecuación (3.17) le podemos dar una interpretación, la parte (a) es la información completada, mientras que la parte (b) es la parte de la información perdida, es decir que

$$\text{inf. observada} = \text{inf. completada} - \text{inf. perdida}.$$

Por otro lado, como se dijo al principio del capítulo 2, el algoritmo EM lo podemos pensar como una función $M : \Omega \rightarrow \Omega$ que va del espacio de

parametral $\Omega \subset \mathbb{R}^d$ a el mismo, tal que $M(\Xi^u) = \Xi^{u+1}$. De tal manera que después de un número M de iteraciones, el algoritmo encuentra un punto fijo Ξ^η , el cual cumple lo siguiente; $M(\Xi^\eta) = \Xi^\eta$. Entonces podemos usar la sucesión $\Xi^0, \Xi^1, \dots, \Xi^m, m < \eta$, para diferenciar numéricamente $M(\Xi)$.

Queremos calcular $\frac{\partial M(\Xi)}{\partial \Xi}$. Tanner (1996) sugiere aproximar el elemento $(i, j), i, j \in \{1, 2, \dots, d\}$ en la matriz $\frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi=\Xi^\eta}$ usando un resultado de Meng y Rubin (1991) que dice que si M es diferenciable en Ξ^η entonces

$$\begin{aligned} \frac{\partial M_j(\Xi)}{\partial \Xi_i} \Big|_{\Xi=\Xi^\eta} &= \lim_{\Xi_i \rightarrow \Xi_i^\eta} \frac{M_j(\Xi_1^\eta, \dots, \Xi_i, \dots, \Xi_d^\eta) - M_j(\Xi^\eta)}{\Xi_i - \Xi_i^\eta} \\ &= \lim_{u \rightarrow \infty} \frac{M_j(\Xi_1^\eta, \dots, \Xi_i^u, \dots, \Xi_d^\eta) - \Xi_j^\eta}{\Xi_i^u - \Xi_i^\eta}, \end{aligned} \quad (3.18)$$

donde $\Xi^u = (\Xi_1^u, \dots, \Xi_i^u, \dots, \Xi_d^u)$ es el u -ésimo valor en la sucesión $\Xi^0, \Xi^1, \dots, \Xi^m$, con $m < \eta$.

Así entonces usando (3.18) y los valores $\Xi^0, \Xi^1, \dots, \Xi^m$, podemos aproximar los valores de las entradas de $\frac{\partial M_j(\Xi)}{\partial \Xi_i} \Big|_{\Xi=\Xi^\eta}$.

Este resultado se puede usar junto con la siguiente ecuación, cuya demostración puede hallarse en Dempster et. al. (1977)

$$\left\{ \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi=\Xi^\eta} \right\} \left\{ \frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} \right\} = \left\{ \frac{\partial^2 H(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} \right\} \quad (3.19)$$

Al sustituir (3.19) en (3.17)

$$\begin{aligned} \frac{-\partial^2 \log \mathbb{P}(\Xi | \mathbf{x})}{\partial \Xi^2} &= \frac{-\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\phi=\Xi^\eta} - \left\{ \left\{ \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi=\Xi^\eta} \right\} \left\{ \frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} \right\} \right\} \\ &= \left\{ \mathbf{I} - \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi=\Xi^\eta} \right\} \left\{ -\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi=\Xi^\eta} \right\}, \end{aligned} \quad (3.20)$$

donde \mathbf{I} es la matriz identidad de dimensión $3k$.

En el Apéndice B se muestra que la siguiente expresión es precisamente el inverso de (3.20)

$$\begin{aligned} \left\{ \frac{-\partial^2 \log \mathbb{P}(\Xi | \mathbf{x})}{\partial \Xi^2} \right\}^{-1} &= \left\{ -\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi = \Xi^\eta} \right\}^{-1} \left\{ I - \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi = \Xi^\eta} \right\}^{-1} \\ &\times \left\{ \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi = \Xi^\eta} \right\} + \left\{ -\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi = \Xi^\eta} \right\}^{-1} \end{aligned} \quad (3,21)$$

esto es precisamente la matriz de varianzas-covarianzas para Ξ^η a la que nos referimos en (3.15). Notemos que para calcular (3.21) basta encontrar $\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2}$ y evaluarla en $\Xi = \Xi^\eta$.

3.4. Instrumentación.

En Mendenhall y Hader (1958) se describe un ejemplo de datos censurados con identificación parcial con las siguientes características; $k = 2$, es decir, que uno puede identificar el género de la falla entre dos posibles causas o componentes. En este experimento se usó como umbral de censura $C = 630$ horas; con $N = 369$ observaciones de tiempos de falla de un receptor-transmisor de comunicación. De estos datos se registró lo siguiente: $r_1 = 107$, $r_2 = 218$ y los restantes $N - r = 44$ excedieron el umbral C ².

El punto fijo Ξ^η se calculó iterando el paso E (3.10) y después el paso M (3.11), (3.13) y (3.14) treinta veces, obteniendo la siguiente tabla

²Los datos el lector los puede encontrar al final del Apéndice B

Cuadro 3.1:

	μ_1	μ_2	φ_1	φ_2	π_1	π_2
Ξ^1	-0.66383	-0.81470	-0.11054	-0.12764	0.66445	0.33555
Ξ^2	-0.64095	-0.88125	-0.13766	-0.10275	0.68042	0.31958
Ξ^3	-0.63176	-0.91242	-0.13948	-0.11746	0.68498	0.31502
Ξ^4	-0.62336	-0.93753	-0.13703	-0.13460	0.68827	0.31173
Ξ^5	-0.61544	-0.95994	-0.13375	-0.15105	0.69115	0.30885
Ξ^6	-0.60835	-0.97967	-0.13051	-0.16636	0.69367	0.30633
Ξ^7	-0.60228	-0.99644	-0.12759	-0.18009	0.69578	0.30422
Ξ^8	-0.59729	-1.01015	-0.12509	-0.19185	0.69750	0.30250
Ξ^9	-0.59333	-1.02091	-0.12304	-0.20145	0.69884	0.30116
Ξ^{10}	-0.59031	-1.02905	-0.12141	-0.20894	0.69985	0.30015
Ξ^{11}	-0.58807	-1.03502	-0.12016	-0.21457	0.70059	0.29942
Ξ^{12}	-0.58644	-1.03929	-0.11922	-0.21865	0.70111	0.29889
Ξ^{13}	-0.58529	-1.04228	-0.11854	-0.22155	0.70148	0.29852
Ξ^{14}	-0.58448	-1.04435	-0.11805	-0.22358	0.70174	0.29827
Ξ^{15}	-0.58393	-1.04577	-0.11771	-0.22497	0.70191	0.29809
Ξ^{16}	-0.58355	-1.04673	-0.11747	-0.22592	0.70203	0.29797
Ξ^{17}	-0.58329	-1.04738	-0.11731	-0.22657	0.70211	0.29789
Ξ^{18}	-0.58311	-1.04782	-0.11720	-0.22700	0.70216	0.29784
Ξ^{19}	-0.58299	-1.04812	-0.11712	-0.22730	0.70220	0.29780
Ξ^{20}	-0.58291	-1.04832	-0.11707	-0.22749	0.70222	0.29778
Ξ^{21}	-0.58286	-1.04845	-0.11703	-0.22763	0.70224	0.29776
Ξ^{22}	-0.58282	-1.04854	-0.11701	-0.22772	0.70225	0.29775
Ξ^{23}	-0.58280	-1.04860	-0.11699	-0.22778	0.70226	0.29774
Ξ^{24}	-0.58278	-1.04864	-0.11698	-0.22782	0.70226	0.29774
Ξ^{25}	-0.58277	-1.04867	-0.11698	-0.22784	0.70227	0.29773
Ξ^{26}	-0.58276	-1.04869	-0.11697	-0.22786	0.70227	0.29773
Ξ^{27}	-0.58276	-1.04870	-0.11697	-0.22787	0.70227	0.29773
Ξ^{28}	-0.58276	-1.04871	-0.11697	-0.22788	0.70227	0.29773
Ξ^{29}	-0.58275	-1.04871	-0.11696	-0.22789	0.70227	0.29773
Ξ^7	-0.58275	-1.04872	-0.11696	-0.22789	0.70227	0.29773

En virtud de que tenemos forma de construir estimadores puntuales para cada componente de Ξ , así como para calcular aproximaciones de las varianzas asintóticas de estos estimadores, podemos usar estos ingredientes para construir intervalos de confianza usando teoría asintótica. En la literatura hay resultados referentes a la aproximación de la distribución posterior usando la distribución Normal (véase por ejemplo Bernardo y Smith, 1994 sec-

ción 5.3.2), recordemos además que al asumir una distribución inicial tal que $\mathbb{P}(\Xi) = 1$, la moda de la distribución posterior es equivalente al estimador máximo verosimil.

Así entonces el supuesto de que los estimadores obtenidos son asintóticamente normales está justificado.

Si Ξ_i^η denota la i -ésima componente de Ξ^η entonces un intervalo de confianza al $(1 - \alpha) \times 100\%$ para Ξ_i basado en la teoría asintótica Normal es

$$(\Xi_i^\eta - v_{ii} * \phi_{1-\alpha}, \Xi_i^\eta + v_{ii} * \phi_{1-\alpha}),$$

donde Ξ_i^η es la i -ésima componente de Ξ^η , $i \in \{1, 2, \dots, 3k\}$, $\phi_{1-\alpha}$ es el cuantil $(1 - \alpha) \times 100\%$ de una distribución Normal y v_{ii} es la raíz cuadrada de la i -ésima componente en la diagonal de la matriz $-\frac{\partial^2 \log \mathbb{P}(\Xi|\mathbf{x})}{\partial \Xi^2}$.

En la siguiente sección llevaremos a cabo un estudio de la cobertura de los intervalos producidos con esta metodología.

3.5. Cobertura con datos simulados y $k = 2$.

En esta sección se hará un estudio de la cobertura de los intervalos de confianza producidos, es decir, la proporción de intervalos que contienen al verdadero valor del parámetro.

Se simularon 100 muestras de mezclas censuradas donde el esquema de censura es como se describió en la introducción de este capítulo. Cada una es de tamaño $N = 369$ y con $C = 630$ como umbral de censura, donde el valor verdadero de los parámetros es

Cuadro 3.2: ($\pi_2 = 1 - \pi_1$)

Verdadero valor del parámetro	Ξ
μ_1	-1.138
μ_2	-0.635
φ_1	-0.327
φ_2	-0.171
π_1	0.33

A cada una de las 100 muestras simuladas, se le calcula su estimador puntual Ξ^η iterando el algoritmo EM treinta veces, después se le estima su

desviación estándar usando la matriz de varianzas covarianzas asintótica que se ha descrito en (3.21). Con la estimación puntual y su error estándar estimado construimos un intervalo de confianza al 95 % para cada componente de Ξ .

Para el caso $k = 2$, la matriz Hessiana $\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2}$ y el vector de parámetros $\Xi' = (\mu_1, \varphi_1, \mu_2, \varphi_2, \pi_1, \pi_2)$, nos queda

$$-\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} = \begin{bmatrix} Q_{1,1} & 0 & Q_{1,3} & 0 & 0 & 0 \\ 0 & Q_{2,2} & 0 & Q_{2,4} & 0 & 0 \\ Q_{1,3} & 0 & Q_{3,3} & 0 & 0 & 0 \\ 0 & Q_{2,4} & 0 & Q_{4,4} & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_{5,5} & 0 \\ 0 & 0 & 0 & 0 & 0 & Q_{6,6} \end{bmatrix} \quad (3.22)$$

Las dobles parciales $Q_{i,j}$ que son distintas de cero son expresiones muy extensas para mostrarlas aquí, el lector interesado en ver los cálculos los puede encontrar en el Apéndice B.

En la siguiente tabla se muestran los resultados de la cobertura al 95 % de la muestra que hemos generado

Cuadro 3.3:	
Parámetro	Ξ
μ_1	0.98
μ_2	0.96
φ_1	0.98
φ_2	0.95
π_1	0.98
π_2	1.00

Notemos que para todos los componentes de Ξ se cumple que al menos en 95 ocasiones de los 100 intervalos producidos para cada muestra y cada parámetro respectivamente el verdadero valor del parámetro queda dentro del intervalo.

3.6. Datos simulados con $k = 3$.

Por último, se simuló una muestra de tamaño $N = 500$ nuevamente con el esquema de censura descrito en la introducción del presente capítulo pero

esta vez de tres poblaciones y umbral de censura $C = 600$. Dicha muestra tiene las siguientes características: $r_1 = 115$, $r_2 = 153$, $r_3 = 194$ y $N - r = 38$.

Después de iterar treinta veces el algoritmo EM, encontramos los estimadores puntuales de cada parámetro, mismos que mostramos a continuación

Cuadro 3.4:

Parámetro	Valor verdadero del parámetro	Ξ^η
μ_1	-0.8680	-0.8009
μ_2	-0.6495	-0.7676
μ_3	-0.7496	-0.6905
φ_1	-0.1967	-0.1299
φ_2	-0.3737	-0.3674
φ_3	-0.3922	-0.4027
π_1	0.3229	0.3346
π_2	0.2434	0.2663
π_3	0.4337	0.3990

En la siguiente tabla se muestran los intervalos de confianza usando la aproximación asintótica de la matriz de varianzas covarianzas así como se definió en (3.21), donde la matriz $-\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2}$ tiene la siguiente forma

$$-\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} = \begin{bmatrix} Q_{1,1} & 0 & 0 & Q_{1,4} & 0 & 0 & 0 & 0 & 0 \\ 0 & Q_{2,2} & 0 & 0 & Q_{2,5} & 0 & 0 & 0 & 0 \\ 0 & 0 & Q_{3,3} & 0 & 0 & Q_{3,6} & 0 & 0 & 0 \\ Q_{1,4} & 0 & 0 & Q_{4,4} & 0 & 0 & 0 & 0 & 0 \\ 0 & Q_{2,5} & 0 & 0 & Q_{5,5} & 0 & 0 & 0 & 0 \\ 0 & 0 & Q_{3,6} & 0 & 0 & Q_{6,6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & Q_{7,7} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & Q_{8,8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Q_{9,9} \end{bmatrix} \quad (3.23)$$

Observemos que en todos los casos el valor verdadero del parámetro está contenido en el intervalo correspondiente.

Cuadro 3.5:

Parámetro	Valor verdadero del parámetro	Intervalo
μ_1	-0.8680	(-0.9746 , -0.6271)
μ_2	-0.6495	(-0.9273 , -0.6078)
μ_3	-0.7496	(-0.8160 , -0.5649)
φ_1	-0.1967	(-0.2693 , 0.0095)
φ_2	-0.3737	(-0.5358 , -0.1990)
φ_3	-0.3922	(-0.5283 , -0.2770)
π_1	0.3229	(0.2823 , 0.3869)
π_2	0.2434	(0.2195 , 0.3132)
π_3	0.4337	(0.3419 , 0.4562)

Conclusiones

Algunas ventajas que podemos notar al usar el algoritmo EM descrito en la sección 3.2. para datos de tiempos de falla de una mezcla de distribuciones censuradas son:

- Uno puede aprovechar la información tanto de los datos observados \mathbf{x}

$$\{x_{1i}\}_{i=1}^{r_1}, \{x_{2i}\}_{i=1}^{r_2}, \dots, \{x_{ki}\}_{i=1}^{r_k},$$

como de las variables latentes \mathbf{z}

$$z_1 = (x_{01}, i_1), z_2 = (x_{02}, i_2), \dots, z_{N-r} = (x_{0N-r}, i_{N-r}),$$

y así reducir el costo del experimento, ya que no se tendrá que invertir más que una cantidad determinada de recursos y no se tendrá que esperar hasta que el último artefacto se descomponga.

Comúnmente, solo las causa de falla i_1, i_2, \dots, i_{N-r} son consideradas para proceder o plantear el algoritmo EM.

No obstante, en Contreras-Cristán et al. (2003) se sugiere considerar tanto las causas de falla como los tiempos de falla $x_{01}, x_{02}, \dots, x_{0N-r}$, para poder establecer comparaciones con métodos Bayesianos.

Esta especificación no vuelve más complicada la instrumentación del algoritmo.

- Obtener estimadores puntuales y por intervalo para los parámetros de una mezcla de distribuciones donde hay un esquema de censura como el descrito en la introducción del tercer capítulo, puede ser abordado de manera muy efectiva usando el algoritmo EM.

Esto lo pudimos constatar de tres maneras, la primera fue al usar una muestra de datos censurados con $k = 2$, tomada de Mendenhall y Hader

(1958). En esta muestra $N = 369$, $C = 630$ horas, $r_1 = 107$, $r_2 = 218$ y $N - r = N - r_1 - r_2 = 44$ ³. Estos datos también fueron usados en Díaz Francés (1998), y los resultados a los que llegaron en este artículo son los mismos a los que llegamos nosotros al iterar el algoritmo EM treinta veces⁴. El tiempo que tomó hacer esta estimación fue menor a siete minutos usando una computadora con un procesador Pentium IV.

La segunda forma fue al simular cien muestras censuradas con las mismas características que la de Mendenhall y Hader (1958) y al hacer un estudio de la cobertura de los intervalos de confianza para cada parámetro de la mezcla de distribuciones; en el Cuadro 3.3 podemos ver que la proporción de intervalos que contienen al verdadero valor del parámetro es cercana al nivel de confianza de los intervalos.

El tiempo que le tomó a la computadora calcular los estimadores puntuales y los intervalos de confianza para cada parámetro de la mezcla de distribuciones fue menor a quince minutos para cada muestra.

Por último, la tercera manera fue cuando se simuló una muestra con $k = 3$, es decir, cuando uno puede identificar el género de la falla entre tres posibles causas o componentes. A esta muestra se le calcularon sus estimadores puntuales y sus intervalos de confianza.⁵ Podemos notar que los estimadores puntuales son muy parecidos a los valores verdaderos de los parámetros y que estos últimos en todos los casos están contenidos en los intervalos de confianza.

Por lo tanto, concluimos que el algoritmo EM es una buena alternativa para hacer inferencia sobre los parámetros de una mezcla de distribuciones censurada con identificación parcial.

³Véase los cuadros B.1. y B.2.

⁴Véase el cuadro 3.1

⁵Véase el Cuadro 3.4 y el 3.5 respectivamente.

Apéndice A

El objetivo de este apéndice es calcular $\mathbb{E} [\epsilon_i \leq x | \epsilon_i > \frac{c_i - \mu_i}{\sigma}]$ y $\mathbb{E} [\epsilon_i^2 \leq x | \epsilon_i > \frac{c_i - \mu_i}{\sigma}]$ y así probar lo establecido en las ecuaciones (2.9) y (2.10) de la sección 2.3, pero antes hacemos tres observaciones que nos serán útiles para nuestro fin.

Observación 1

Si $\epsilon_i \sim N(0, 1)$, entonces la función de densidad condicional $f_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x)$ es igual a

$$\frac{\phi(x) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x)}{1 - \Phi(\frac{c_i - \mu_i}{\sigma})}.$$

Demostración

Por definición

$$F_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(u) = \mathbb{P}\left(\epsilon_i \leq u | \epsilon_i > \frac{c_i - \mu_i}{\sigma}\right) = \frac{\mathbb{P}\left(\frac{c_i - \mu_i}{\sigma} < \epsilon_i \leq u\right)}{\mathbb{P}\left(\epsilon_i > \frac{c_i - \mu_i}{\sigma}\right)} \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(u). \quad (\text{A.1})$$

Por otro lado

$$f_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x) = \lim_{h \rightarrow 0} \frac{F_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x + h) - F_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x)}{h},$$

usando (A.1) en el numerador podemos reescribir la expresión como

$$= \frac{1}{\mathbb{P}\left(\epsilon_i > \frac{c_i - \mu_i}{\sigma}\right)} \lim_{h \rightarrow 0} \frac{\mathbb{P}\left(\frac{c_i - \mu_i}{\sigma} < \epsilon_i \leq x + h\right) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x + h) - \mathbb{P}\left(\frac{c_i - \mu_i}{\sigma} < \epsilon_i \leq x\right) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x)}{h},$$

notemos que $\mathbb{P}\left(\epsilon_i > \frac{c_i - \mu_i}{\sigma}\right) = 1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)$, y además que

$$\mathbb{P}(x < \epsilon_i \leq x + h) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x + h) =$$

$$\mathbb{P}\left(\frac{c_i - \mu_i}{\sigma} < \epsilon_i \leq x + h\right) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x + h) - \mathbb{P}\left(\frac{c_i - \mu_i}{\sigma} < \epsilon_i \leq x\right) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x),$$

por lo tanto

$$f_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x) = \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \lim_{h \rightarrow 0} \frac{\mathbb{P}(x < \epsilon_i \leq x + h)}{h} \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x + h),$$

pero $\phi(x) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x) = \lim_{h \rightarrow 0} \frac{\mathbb{P}(x < \epsilon_i \leq x + h)}{h} \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x + h)$, entonces

$$f_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x) = \frac{\phi(x) \mathbb{1}_{(\frac{c_i - \mu_i}{\sigma}, \infty)}(x)}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)}.$$

■

Observación 2

Si $\phi(u) = \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{u^2}{2}}$ es la función de densidad de una variable aleatoria normal(0,1), entonces

- (a) $\phi'(u) = -u\phi(u)$ y además
- (b) $u^2\phi(u) = \phi(u) + \phi''(u)$.

Demostración

Al derivar $\phi(u)$ con respecto a u

$$\phi'(u) = -u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} = -u\phi(u).$$

Con respecto a (b) si calculamos $\phi''(u)$ tenemos

$$\phi''(u) = u^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} - \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} = u^2\phi(u) - \phi(u).$$

Luego entonces

$$u^2\phi(u) = \phi(u) + \phi''(u).$$

■

Observación 3

Calculemos los primeros dos momentos de la variable aleatoria ϵ_i condicionada a que $\{\epsilon_i > \frac{c_i - \mu_i}{\sigma}\}$

- (i) $\mathbb{E} \left[\epsilon_i \leq x \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] = H \left(\frac{c_i - \mu_i}{\sigma} \right).$
(ii) $\mathbb{E} \left[\epsilon_i^2 \leq x \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] = 1 + H \left(\frac{c_i - \mu_i}{\sigma} \right) \left(\frac{c_i - \mu_i}{\sigma} \right),$

donde $H \left(\frac{c_i - \mu_i}{\sigma} \right) := \frac{\phi \left(\frac{c_i - \mu_i}{\sigma} \right)}{1 - \Phi \left(\frac{c_i - \mu_i}{\sigma} \right)}.$

Demostración

Primero desarrollemos (i), por definición tenemos

$$\mathbb{E} \left[\epsilon_i \leq x \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] = \int_{-\infty}^{\infty} x f_{\epsilon_i \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x) dx,$$

de la observación 1 podemos escribir

$$= \int_{-\infty}^{\infty} x \frac{\phi(x)}{1 - \Phi \left(\frac{c_i - \mu_i}{\sigma} \right)} \mathbb{1}_{\left(\frac{c_i - \mu_i}{\sigma}, \infty \right)}(x) dx,$$

o equivalente

$$= \frac{1}{1 - \Phi \left(\frac{c_i - \mu_i}{\sigma} \right)} \int_{\frac{c_i - \mu_i}{\sigma}}^{\infty} x \phi(x) dx.$$

De la observación 2 (a) y al cambiar los límites de integración

$$= \frac{\int_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \phi(x) dx}{1 - \Phi \left(\frac{c_i - \mu_i}{\sigma} \right)},$$

y finalmente al aplicar el Teorema Fundamental del Cálculo

$$= \frac{\phi \left(\frac{c_i - \mu_i}{\sigma} \right) - 0}{1 - \Phi \left(\frac{c_i - \mu_i}{\sigma} \right)}$$

$$\mathbb{E} \left[\epsilon_i \leq x | \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] = H \left(\frac{c_i - \mu_i}{\sigma} \right).$$

Ahora se va a calcular el segundo momento

$$\begin{aligned} \mathbb{E} \left[\epsilon_i^2 \leq x | \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] &= \int_{-\infty}^{\infty} x^2 f_{\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma}}(x) dx \\ &= \int_{-\infty}^{\infty} x^2 \frac{\phi(x)}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \mathbb{1}_{\left(\frac{c_i - \mu_i}{\sigma}, \infty\right)}(x) dx \\ &= \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \int_{\frac{c_i - \mu_i}{\sigma}}^{\infty} x^2 \phi(x) dx, \end{aligned}$$

usando la observación 2 (b)

$$= \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \int_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \phi(x) + \phi''(x) dx,$$

por la linealidad de la integral

$$= \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \left\{ \int_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \phi(x) dx + \int_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \phi''(x) dx \right\},$$

resolviendo las integrales

$$\begin{aligned} &= \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \left\{ \left(1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right) \right) + \phi'(x) \Big|_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \right\}, \\ &= 1 + \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \left\{ \phi'(x) \Big|_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \right\}, \end{aligned}$$

nuevamente al usar la observación 2 (a)

$$= 1 + \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \left\{ -x\phi(x) \Big|_{\frac{c_i - \mu_i}{\sigma}}^{\infty} \right\}.$$

Al usar la regla de L'Hôpital podemos ver que $\lim_{x \rightarrow \infty} x\phi(x) = 0$, y por lo tanto solo basta con evaluar en el límite inferior

$$\begin{aligned}
&= 1 + \frac{1}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \left\{ \frac{c_i - \mu_i}{\sigma} * \phi\left(\frac{c_i - \mu_i}{\sigma}\right) \right\} \\
&= 1 + \frac{\phi\left(\frac{c_i - \mu_i}{\sigma}\right)}{1 - \Phi\left(\frac{c_i - \mu_i}{\sigma}\right)} \frac{c_i - \mu_i}{\sigma} \\
\mathbb{E} \left[\epsilon_i^2 \leq x | \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] &= 1 + H\left(\frac{c_i - \mu_i}{\sigma}\right) \left(\frac{c_i - \mu_i}{\sigma}\right).
\end{aligned}$$

■

Observación 4

Por último se verá que:

- (a) $\mathbb{E} [Z_i | U] = \mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right)$.
- (b) $\mathbb{E} [Z_i^2 | U] = \mu_i^2 + \sigma^2 + H\left(\frac{c_i - \mu_i}{\sigma}\right) \{\sigma(c_i + \mu_i)\}$.

Demostración

De (a) tenemos que

$$\mathbb{E} [Z_i | U] = \mathbb{E} [Z_i | \beta_0, \beta_1, \sigma, Z_i > c_i],$$

como $Z_i = \mu_i + \sigma \epsilon_i$, y $Z_i > c_i$ esto implica que $\epsilon_i > \frac{c_i - \mu_i}{\sigma}$

$$= \mathbb{E} \left[\mu_i + \sigma \epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right]$$

$$= \mu_i + \sigma \mathbb{E} \left[\epsilon_i | \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right]$$

$$\mathbb{E} [Z_i | U] = \mu_i + \sigma H\left(\frac{c_i - \mu_i}{\sigma}\right).$$

Ahora calculemos el segundo momento

$$\mathbb{E} [Z_i^2 | U] = \mathbb{E} [Z_i^2 | \beta_0, \beta_1, \sigma, Z_i > c_i]$$

$$= \mathbb{E} \left[(\mu_i + \sigma \epsilon_i)^2 \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right],$$

desarrollando el binomio y por linealidad de la esperanza tenemos

$$= \mu_i^2 + 2\sigma\mu_i \mathbb{E} \left[\epsilon_i \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] + \sigma^2 \mathbb{E} \left[\epsilon_i^2 \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right],$$

sustituyendo las esperanzas condicionales con lo que se obtuvo en la observación 3

$$= \mu_i^2 + 2\sigma\mu_i \left\{ H \left(\frac{c_i - \mu_i}{\sigma} \right) \right\} + \sigma^2 \left\{ 1 + H \left(\frac{c_i - \mu_i}{\sigma} \right) \left(\frac{c_i - \mu_i}{\sigma} \right) \right\},$$

factorizando

$$\mathbb{E} \left[\epsilon_i^2 \leq x \mid \epsilon_i > \frac{c_i - \mu_i}{\sigma} \right] = \mu_i^2 + \sigma^2 + H \left(\frac{c_i - \mu_i}{\sigma} \right) \{ \sigma(c_i + \mu_i) \}.$$

■

Apéndice B

Un resultado que utilizaremos en varias ocasiones es el siguiente

Proposición 1

$$\sum_{i_s=0}^k \int_0^{\infty} \pi_{i_s} f(x_{0s}; \theta_{i_s}) dx_{0s} = 1 - G(0; \Xi^u) \quad \forall s = \{1, 2, \dots, N-r\} \quad (\text{B.1})$$

Demostración

$$\begin{aligned} \sum_{i_s=0}^k \int_0^{\infty} \pi_{i_s} f(x_{0s}; \theta_{i_s}) dx_{0s} &= \sum_{i_s=0}^k \pi_{i_s} \int_0^{\infty} f(x_{0s}; \theta_{i_s}) dx_{0s} \\ &= \sum_{i_s=0}^k \pi_{i_s} \{1 - F(0; \theta_{i_s})\} \\ &= \sum_{i_s=0}^k \pi_{i_s} - \sum_{i_s=0}^k \pi_{i_s} F(0; \theta_{i_s}) \\ &= 1 - G(0; \Xi^u) \end{aligned}$$

■

Definamos $\lambda = \frac{1}{(1 - G(0; \Xi^u))^{N-r}}$.

A continuación se mostrará la afirmación hecha en (3.8), que dice

$$\lambda \int_0^\infty \cdots \int_0^\infty \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \prod_{l=1}^{N-r} \{\pi_{i_l} f(x_{0l}; \theta_{i_l})\} dx_{01} \cdots dx_{0N-r} = 1. \quad (\text{B.2})$$

Demostración

De la parte izquierda de (B.2), si intercambiamos las integrales por las sumas y desarrollamos el producto nos queda

$$\lambda \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \int_0^\infty \cdots \int_0^\infty \pi_{i_1} f(x_{01}; \theta_{i_1}) \cdots \pi_{i_{N-r}} f(x_{0N-r}; \theta_{i_{N-r}}) dx_{01} \cdots dx_{0N-r},$$

o equivalentemente

$$= \lambda \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \int_0^\infty \pi_{i_1} f(x_{01}; \theta_{i_1}) \cdots \int_0^\infty \pi_{i_{N-r}} f(x_{0N-r}; \theta_{i_{N-r}}) dx_{01} \cdots dx_{0N-r}$$

$$= \lambda \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \int_0^\infty \pi_{i_1} f(x_{01}; \theta_{i_1}) dx_{01} \cdots \int_0^\infty \pi_{i_{N-r}} f(x_{0N-r}; \theta_{i_{N-r}}) dx_{0N-r}$$

$$= \lambda \sum_{i_1=1}^k \int_0^\infty \pi_{i_1} f(x_{01}; \theta_{i_1}) dx_{01} \cdots \sum_{i_{N-r}=1}^k \int_0^\infty \pi_{i_{N-r}} f(x_{0N-r}; \theta_{i_{N-r}}) dx_{0N-r},$$

y por (B.1) tenemos

$$= \lambda \underbrace{\{1 - G(0; \Xi^u)\} \cdots \{1 - G(0; \Xi^u)\}}_{N-r \text{ veces}}$$

$$= \lambda \lambda^{-1} = 1.$$

■

B.1. Cálculos del paso E del Capítulo 3

Para el paso E necesitamos resolver

$$\lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \log \mathbb{P}(\Xi | \mathbf{z}; \mathbf{x}) \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r}, \quad (\text{B.3})$$

y recordemos que

$$\begin{aligned} \log \mathbb{P}(\Xi | \mathbf{z}; \mathbf{x}) &\propto \sum_{j=1}^k \sum_{i=1}^{r_j} \left\{ \log\{\pi_j\} - \varphi_j + (x_{ji} - \mu_j)e^{-\varphi_j} - e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \right\} \\ &+ \sum_{l=1}^{N-r} \left\{ \log\{\pi_{i_l}\} - \varphi_{i_l} + (x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}} - e^{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}} \right\}. \quad (\text{B.4}) \end{aligned}$$

Por la linealidad de la integral, empecemos por integrar el primer miembro de (B.4)

$$\begin{aligned} &\lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \left\{ \sum_{j=1}^k \sum_{i=1}^{r_j} \left\{ \log\{\pi_j\} - \varphi_j + (x_{ji} - \mu_j)e^{-\varphi_j} - e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \right\} \right\} \\ &\times \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r}, \end{aligned}$$

sacando de las integrales lo que no depende de x_{0l} para $l = 1, 2, 3, \dots, N - r$

$$\begin{aligned} &= \sum_{j=1}^k \sum_{i=1}^{r_j} \left\{ \log\{\pi_j\} - \varphi_j + (x_{ji} - \mu_j)e^{-\varphi_j} - e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \right\} \\ &\times \lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r}, \end{aligned}$$

por (B.2) sabemos que el segundo miembro de esta última ecuación es uno, entonces

$$= \sum_{j=1}^k \sum_{i=1}^{r_j} \left\{ \log\{\pi_j\} - \varphi_j + (x_{ji} - \mu_j)e^{-\varphi_j} - e^{(x_{ji} - \mu_j)e^{-\varphi_j}} \right\},$$

por último, distribuyendo la suma

$$\begin{aligned}
 &= \sum_{j=1}^k r_j \log\{\pi_j\} - \sum_{j=1}^k r_j \varphi_j \\
 &+ \sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} - \sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}}. \tag{B.5}
 \end{aligned}$$

Tomemos los dos primeros términos del segundo miembro de (B.4)

$$\lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \sum_{l=1}^{N-r} \{\log\{\pi_{i_l}\} - \varphi_{i_l}\} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r},$$

desarrollando la suma $\sum_{l=1}^{N-r} \{\log\{\pi_{i_l}\} - \varphi_{i_l}\}$

$$\begin{aligned}
 &\lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \{(\log\{\pi_{i_1}\} - \varphi_{i_1}) + \dots + (\log\{\pi_{i_{N-r}}\} - \varphi_{i_{N-r}})\} \\
 &\times \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r}.
 \end{aligned}$$

Por la linealidad de la integral podemos tomar únicamente uno de los sumandos, digamos $(\log\{\pi_{i_s}\} - \varphi_{i_s})$ con $s \in \{1, 2, \dots, N-r\}$ y resolvamos la integral

$$\lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k (\log\{\pi_{i_s}\} - \varphi_{i_s}) \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r},$$

desarrollando el producto

$$\begin{aligned}
 &= \lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k (\log\{\pi_{i_s}\} - \varphi_{i_s}) \\
 &\times \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) \dots \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) \dots \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{01} \dots dx_{0N-r},
 \end{aligned}$$

o equivalentemente

$$\begin{aligned}
 &= \lambda \int_0^\infty \cdots \int_0^\infty \sum_{i_1=1}^k \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) \cdots \sum_{i_s=1}^k (\log\{\pi_{i_s}\} - \varphi_{i_s}) \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) \cdots \\
 &\times \sum_{i_{N-r}=1}^k \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{01} \cdots dx_{0N-r},
 \end{aligned}$$

que es reescribible como

$$\begin{aligned}
 &= \lambda \int_0^\infty \sum_{i_{N-r}=1}^k \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{0N-r} \\
 &\times \int_0^\infty \sum_{i_s=1}^k (\log\{\pi_{i_s}\} - \varphi_{i_s}) \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) dx_{0s} \cdots \int_0^\infty \sum_{i_1=1}^k \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) dx_{01} \\
 &= \lambda \sum_{i_{N-r}=1}^k \int_0^\infty \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{0N-r} \\
 &\times \sum_{i_s=1}^k \int_0^\infty (\log\{\pi_{i_s}\} - \varphi_{i_s}) \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) dx_{0s} \cdots \sum_{i_1=1}^k \int_0^\infty \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) dx_{01},
 \end{aligned}$$

nuevamente usando (B.1)

$$= \lambda \{1 - G(0; \Xi^u)\} \cdots \sum_{i_s=1}^k \int_0^\infty (\log\{\pi_{i_s}\} - \varphi_{i_s}) \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) dx_{0s} \cdots \{1 - G(0; \Xi^u)\},$$

notemos que hay $N - r - 1$ términos $\{1 - G(0; \Xi^u)\}$ multiplicandose y de la definición de λ , es claro que $\lambda \{1 - G(0; \Xi^u)\}^{N-r-1} = \{1 - G(0; \Xi^u)\}^{-1}$, entonces

$$\begin{aligned}
 &= \frac{1}{\{1 - G(0; \Xi^u)\}} \sum_{i_s=1}^k (\log\{\pi_{i_s}\} - \varphi_{i_s}) \pi_{i_s}^u \int_0^\infty f(x_{0s}; \theta_{i_s}^u) dx_{0s} \\
 &= \sum_{i_s=1}^k (\log\{\pi_{i_s}\} - \varphi_{i_s}) \left\{ \frac{\pi_{i_s}^u \{1 - F(0; \theta_{i_s}^u)\}}{1 - G(0; \Xi^u)} \right\}.
 \end{aligned}$$

Recordemos que esto lo hemos hecho para un sumando, entonces al sumar sobre todos los valores que toma s

$$\begin{aligned} & \lambda \int_0^\infty \cdots \int_0^\infty \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \sum_{l=1}^{N-r} \{\log\{\pi_{i_l}\} - \varphi_{i_l}\} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \cdots dx_{0N-r} \\ &= \sum_{i_1=1}^k (\log\{\pi_{i_1}\} - \varphi_{i_1}) \left\{ \frac{\pi_{i_1}^u (1 - F(0; \theta_{i_1}^u))}{1 - G(0; \Xi^u)} \right\} + \cdots + \sum_{i_{N-r}=1}^k (\log\{\pi_{i_{N-r}}\} - \varphi_{i_{N-r}}) \left\{ \frac{\pi_{i_{N-r}}^u (1 - F(0; \theta_{i_{N-r}}^u))}{1 - G(0; \Xi^u)} \right\} \end{aligned}$$

si desarrollamos las sumas

$$= \left\{ \begin{array}{c} (\log\{\pi_1\} - \varphi_1) \left\{ \frac{\pi_1^u (1 - F(0; \theta_1^u))}{1 - G(0; \Xi^u)} \right\} \\ + \\ \vdots \\ + \\ (\log\{\pi_k\} - \varphi_k) \left\{ \frac{\pi_k^u (1 - F(0; \theta_k^u))}{1 - G(0; \Xi^u)} \right\} \end{array} \right\} + \cdots + \left\{ \begin{array}{c} (\log\{\pi_1\} - \varphi_1) \left\{ \frac{\pi_1^u (1 - F(0; \theta_1^u))}{1 - G(0; \Xi^u)} \right\} \\ + \\ \vdots \\ + \\ (\log\{\pi_k\} - \varphi_k) \left\{ \frac{\pi_k^u (1 - F(0; \theta_k^u))}{1 - G(0; \Xi^u)} \right\} \end{array} \right\}$$

sumando renglón por renglón lo podemos reescribir como

$$= (N - r)(\log\{\pi_1\} - \varphi_1) \left\{ \frac{\pi_1^u (1 - F(0; \theta_1^u))}{1 - G(0; \Xi^u)} \right\} + \cdots + (N - r)(\log\{\pi_k\} - \varphi_k) \left\{ \frac{\pi_k^u (1 - F(0; \theta_k^u))}{1 - G(0; \Xi^u)} \right\}$$

o de forma equivalente

$$= (N - r) \sum_{j=1}^{N-r} (\log\{\pi_j\} - \varphi_j) \left\{ \frac{\pi_j^u (1 - F(0; \theta_j^u))}{1 - G(0; \Xi^u)} \right\}. \quad (\text{B.6})$$

Tomemos el tercer sumando del segundo miembro de (B.4) para resolver

$$\lambda \int_0^\infty \cdots \int_0^\infty \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \sum_{l=1}^{N-r} \{(x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}}\} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \cdots dx_{0N-r},$$

$$\begin{aligned}
& \text{desarrollando la suma } \sum_{l=1}^{N-r} \{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}\} \\
&= \lambda \int_0^\infty \dots \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_s=1}^k \dots \sum_{i_{N-r}=1}^k \{(x_{01} - \mu_{i_1})e^{-\varphi_{i_1}} + \dots + (x_{0s} - \mu_{i_s})e^{-\varphi_{i_s}} + \dots \\
&+ (x_{0N-r} - \mu_{i_{N-r}})e^{-\varphi_{i_{N-r}}}\} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0s} \dots dx_{0N-r}. \tag{B.7}
\end{aligned}$$

Como lo hicimos antes, usando la linealidad de la integral, resolvamos la integral únicamente para el sumando $s \in \{1, 2, \dots, N-r\}$; y desarrollamos el producto

$$\begin{aligned}
&= \lambda \int_0^\infty \dots \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_s=1}^k \dots \sum_{i_{N-r}=1}^k (x_{0s} - \mu_{i_s})e^{-\varphi_{i_s}} \\
&\times \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) \dots \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) \dots \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{01} \dots dx_{0s} \dots dx_{0N-r}, \tag{B.8}
\end{aligned}$$

de manera equivalente

$$\begin{aligned}
&= \lambda \int_0^\infty \dots \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) \dots \sum_{i_s=1}^k (x_{0s} - \mu_{i_s})e^{-\varphi_{i_s}} \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) \dots \\
&\times \sum_{i_{N-r}=1}^k \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{01} \dots dx_{0s} \dots dx_{0N-r} \tag{B.9}
\end{aligned}$$

$$\begin{aligned}
&= \lambda \int_0^\infty \sum_{i_{N-r}=1}^k \pi_{i_{N-r}}^u f(x_{0N-r}; \theta_{i_{N-r}}^u) dx_{0N-r} \\
&\times \int_0^\infty \sum_{i_s=1}^k (x_{0s} - \mu_{i_s})e^{-\varphi_{i_s}} \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) dx_{0s} \cdot \dots \cdot \int_0^\infty \sum_{i_1=1}^k \pi_{i_1}^u f(x_{01}; \theta_{i_1}^u) dx_{01} \tag{B.10}
\end{aligned}$$

por (B.1) notamos que tenemos de nuevo una multiplicación de $N-r-1$ términos $\{1 - G(0; \Xi^u)\}$

$$= \lambda \{1 - G(0; \Xi^u)\}^{N-r-1} \int_0^\infty \sum_{i_s=1}^k (x_{0s} - \mu_{i_s})e^{-\varphi_{i_s}} \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) dx_{0s} \tag{B.11}$$

$$= \frac{1}{1 - G(0; \Xi^u)} \sum_{i_s=1}^k \int_0^\infty (x_{0s} - \mu_{i_s})e^{-\varphi_{i_s}} \pi_{i_s}^u f(x_{0s}; \theta_{i_s}^u) dx_{0s}. \tag{B.12}$$

Recordemos que esto se hizo para únicamente para el sumando s -ésimo, si lo hacemos para todos los sumandos tenemos que

$$\begin{aligned} & \lambda \int_0^\infty \cdots \int_0^\infty \sum_{i_1=1}^k \cdots \sum_{i_{N-r}=1}^k \sum_{l=1}^{N-r} \{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}\} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \cdots dx_{0N-r} \\ &= \sum_{l=1}^{N-r} \sum_{i_l=1}^k \underbrace{\int_0^\infty (x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}} \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} dx_{0l}}_{\alpha} . \end{aligned} \quad (\text{B.13})$$

Para resolver esta última expresión, primero calculemos α de (B.13)

$$\int_0^\infty (x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}} \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} dx_{0l} ,$$

por linealidad

$$\begin{aligned} &= \frac{e^{-\varphi_{i_l}} \pi_{i_l}^u}{1 - G(0; \Xi^u)} \left\{ \int_0^\infty x_{0l} f(x_{0l}; \theta_{i_l}^u) dx_{0l} - \mu_{i_l} \int_0^\infty f(x_{0l}; \theta_{i_l}^u) dx_{0l} \right\} \\ &= \frac{e^{-\varphi_{i_l}} \pi_{i_l}^u}{1 - G(0; \Xi^u)} \left\{ \underbrace{\int_0^\infty x_{0l} f(x_{0l}; \theta_{i_l}^u) dx_{0l}}_A - \mu_{i_l} \{1 - F(0; \theta_{i_l}^u)\} \right\} . \end{aligned} \quad (\text{B.14})$$

Calculemos aparte A de ecuación anterior, renombrando a x_{0l} como y , y recordemos que $f(y; \theta_{i_l}^u) = e^{-\varphi_{i_l}^u} e^{(y - \mu_{i_l}^u)e^{-\varphi_{i_l}^u}} \exp \left\{ -e^{(y - \mu_{i_l}^u)e^{-\varphi_{i_l}^u}} \right\}$

$$\int_0^\infty y f(y; \theta_{i_l}^u) dy = \int_0^\infty y e^{-\varphi_{i_l}^u} e^{(y - \mu_{i_l}^u)e^{-\varphi_{i_l}^u}} \exp \left\{ -e^{(y - \mu_{i_l}^u)e^{-\varphi_{i_l}^u}} \right\} dy . \quad (\text{B.15})$$

Propongamos el siguiente cambio de variable

$$x = e^{(y - \mu_{i_l}^u)e^{-\varphi_{i_l}^u}} \quad \text{entonces} \quad y = \mu_{i_l}^u + e^{\varphi_{i_l}^u} \log(x) \quad \text{y} \quad dy = \frac{e^{\varphi_{i_l}^u}}{x} dx ,$$

luego entonces los límites de integración cambian de la siguiente manera

si $0 < y < \infty$ entonces $l(\theta_{i_l}^u) < x < \infty$

donde $l(\theta_{i_l}^u) = e^{(-\mu_{i_l}^u)e^{-\varphi_{i_l}^u}}$.

Sustituyendo esto en (B.15) tenemos que

$$\int_0^\infty y f(y; \theta_{i_l}^u) dy = \int_{l(\theta_{i_l}^u)}^\infty (\mu_{i_l}^u + e^{\varphi_{i_l}^u} \log(x)) e^{-\varphi_{i_l}^u} x e^{-x} \frac{e^{\varphi_{i_l}^u}}{x} dx,$$

simplicando tenemos que

$$= \int_{l(\theta_{i_l}^u)}^\infty (\mu_{i_l}^u + e^{\varphi_{i_l}^u} \log(x)) e^{-x} dx,$$

distribuyendo la multiplicación, renombremos al resultado de la siguiente manera

$$m(\theta_{i_l}^u) := \mu_{i_l}^u \int_{l(\theta_{i_l}^u)}^\infty e^{-x} dx + e^{\varphi_{i_l}^u} \int_{l(\theta_{i_l}^u)}^\infty \log(x) e^{-x} dx.$$

Sustituyendo en A de (B.14)

$$\int_0^\infty (x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}} \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} dx_{0l} = \frac{e^{-\varphi_{i_l}} \pi_{i_l}^u}{1 - G(0; \Xi^u)} \{m(\theta_{i_l}^u) - \mu_{i_l} \{1 - F(0; \theta_{i_l}^u)\}\}.$$

Recordemos que todo esto lo hemos hecho para encontrar una expresión que debemos sustituir en α de (B.13), entonces

$$\sum_{l=1}^{N-r} \sum_{i_l=1}^k \int_0^\infty (x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}} \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} dx_{0l} = \sum_{l=1}^{N-r} \sum_{i_l=1}^k \frac{e^{-\varphi_{i_l}} \pi_{i_l}^u \{m(\theta_{i_l}^u) - \mu_{i_l} \{1 - F(0; \theta_{i_l}^u)\}\}}{1 - G(0; \Xi^u)},$$

desarrollando una suma y después la otra tenemos que

$$= \sum_{i_1=1}^k \frac{e^{-\varphi_{i_1}} \pi_{i_1}^u \{m(\theta_{i_1}^u) - \mu_{i_1} \{1 - F(0; \theta_{i_1}^u)\}\}}{1 - G(0; \Xi^u)} + \dots + \sum_{i_{N-r}=1}^k \frac{e^{-\varphi_{i_{N-r}}} \pi_{i_{N-r}}^u \{m(\theta_{i_{N-r}}^u) - \mu_{i_{N-r}} \{1 - F(0; \theta_{i_{N-r}}^u)\}\}}{1 - G(0; \Xi^u)}$$

$$= \left\{ \begin{array}{c} \frac{e^{-\varphi_1^u} \pi_1^u \{m(\theta_1^u) - \mu_1 \{1 - F(0; \theta_1^u)\}\}}{1 - G(0; \Xi^u)} \\ + \\ \vdots \\ + \\ \frac{e^{-\varphi_k^u} \pi_k^u \{m(\theta_k^u) - \mu_k \{1 - F(0; \theta_k^u)\}\}}{1 - G(0; \Xi^u)} \end{array} \right\} + \dots + \left\{ \begin{array}{c} \frac{e^{-\varphi_1^u} \pi_1^u \{m(\theta_1^u) - \mu_1 \{1 - F(0; \theta_1^u)\}\}}{1 - G(0; \Xi^u)} \\ + \\ \vdots \\ + \\ \frac{e^{-\varphi_k^u} \pi_k^u \{m(\theta_k^u) - \mu_k \{1 - F(0; \theta_k^u)\}\}}{1 - G(0; \Xi^u)} \end{array} \right\},$$

sumando renglón por renglón

$$= (N-r) \left(\frac{e^{-\varphi_1^u} \pi_1^u \{m(\theta_1^u) - \mu_1 \{1 - F(0; \theta_1^u)\}\}}{1 - G(0; \Xi^u)} + \dots + \frac{e^{-\varphi_{N-r}^u} \pi_{N-r}^u \{m(\theta_{N-r}^u) - \mu_{N-r} \{1 - F(0; \theta_{N-r}^u)\}\}}{1 - G(0; \Xi^u)} \right),$$

por lo tanto

$$\begin{aligned} & \lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \sum_{l=1}^{N-r} \{ (x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}} \} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r} \\ &= (N-r) \sum_{j=1}^k \frac{e^{-\varphi_j^u} \pi_j^u \{m(\theta_j^u) - \mu_j \{1 - F(0; \theta_j^u)\}\}}{1 - G(0; \Xi^u)}. \end{aligned} \quad (\text{B.16})$$

Por último, tomemos el cuarto sumando de (B.4) para resolver

$$\lambda \int_0^\infty \dots \int_0^\infty \sum_{i_1=1}^k \dots \sum_{i_{N-r}=1}^k \sum_{l=1}^{N-r} \{ e^{(x_{0l} - \mu_{i_l}) e^{-\varphi_{i_l}}} \} \prod_{l=1}^{N-r} \pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u) dx_{01} \dots dx_{0N-r}, \quad (\text{B.17})$$

se puede demostrar de manera análoga como se hizo en los pasos (B.7), (B.8), (B.9), (B.10), (B.11) y (B.12), que (B.17) es igual a

$$\sum_{l=1}^{N-r} \sum_{i_l=1}^k \int_0^\infty e^{(x_{0l}-\mu_{i_l})e^{-\varphi_{i_l}}} \left\{ \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1-G(0; \Xi^u)} \right\} dx_{0l}. \quad (\text{B.18})$$

Como en el caso anterior primero resolvamos primero la integral

$$\int_0^\infty e^{(y-\mu_{i_l})e^{-\varphi_{i_l}}} \left\{ \frac{\pi_{i_l}^u f(y; \theta_{i_l}^u)}{1-G(0; \Xi^u)} \right\} dy = \frac{\pi_{i_l}^u}{1-G(0; \Xi^u)} \int_0^\infty e^{(y-\mu_{i_l})e^{-\varphi_{i_l}}} f(y; \theta_{i_l}^u) dy,$$

que al sustituir lo que vale $f(y; \theta_{i_l}^u)$

$$= \frac{\pi_{i_l}^u}{1-G(0; \Xi^u)} \int_0^\infty e^{(y-\mu_{i_l})e^{-\varphi_{i_l}}} e^{-\varphi_{i_l}^u} e^{(y-\mu_{i_l}^u)e^{-\varphi_{i_l}^u}} \exp\left\{-e^{(y-\mu_{i_l}^u)e^{-\varphi_{i_l}^u}}\right\} dy,$$

bajo el mismo cambio de variable propuesto anteriormente tenemos que

$$= \frac{\pi_{i_l}^u}{1-G(0; \Xi^u)} \int_{l(\theta_{i_l}^u)}^\infty \exp\left\{(\mu_{i_l}^u + e^{\varphi_{i_l}^u} \log(x) - \mu_{i_l})e^{-\varphi_{i_l}}\right\} e^{-\varphi_{i_l}^u} x e^{-x} \frac{e^{\varphi_{i_l}^u}}{x} dx,$$

simplificando y usando el hecho $e^{\varphi_{i_l}^u} \log(x) = \log(x e^{\varphi_{i_l}^u})$

$$= \frac{\pi_{i_l}^u}{1-G(0; \Xi^u)} \int_{l(\theta_{i_l}^u)}^\infty \exp\left\{(\mu_{i_l}^u + \log\{x e^{\varphi_{i_l}^u}\} - \mu_{i_l})e^{-\varphi_{i_l}}\right\} e^{-x} dx,$$

distribuyendo el producto y por propiedades de los exponentes

$$= \frac{\pi_{i_l}^u}{1-G(0; \Xi^u)} \int_{l(\theta_{i_l}^u)}^\infty \exp\{\mu_{i_l}^u e^{-\varphi_{i_l}}\} \exp\left\{e^{-\varphi_{i_l}} \log\left(x e^{\varphi_{i_l}^u}\right)\right\} \exp\{-\mu_{i_l} e^{-\varphi_{i_l}}\} e^{-x} dx,$$

como $\exp\{e^{-\varphi_{i_l}} \log\left(x e^{\varphi_{i_l}^u}\right)\} = \exp\{\log\left(x e^{\varphi_{i_l}^u} e^{-\varphi_{i_l}}\right)\}$, entonces

$$\begin{aligned} &= \frac{\pi_{i_l}^u e^{-\mu_{i_l} e^{-\varphi_{i_l}}}}{1-G(0; \Xi^u)} e^{\mu_{i_l}^u e^{-\varphi_{i_l}}} \int_{l(\theta_{i_l}^u)}^\infty \exp\{\log\left(x e^{\varphi_{i_l}^u} e^{-\varphi_{i_l}}\right)\} e^{-x} dx \\ &= \frac{\pi_{i_l}^u e^{-\mu_{i_l} e^{-\varphi_{i_l}}}}{1-G(0; \Xi^u)} e^{\mu_{i_l}^u e^{-\varphi_{i_l}}} \int_{l(\theta_{i_l}^u)}^\infty x e^{\varphi_{i_l}^u - \varphi_{i_l}} e^{-x} dx, \end{aligned}$$

definamos $\nu(\theta_{i_l}^u; \varphi_{i_l}) = e^{\mu_{i_l}^u e^{-\varphi_{i_l}}} \int_{l(\theta_{i_l}^u)}^{\infty} x e^{\varphi_{i_l}^u - \varphi_{i_l}} e^{-x} dx$, entonces

$$\int_0^{\infty} e^{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}} \left\{ \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} \right\} dx_{0l} = \frac{\pi_{i_l}^u e^{-\mu_{i_l} e^{-\varphi_{i_l}}}}{1 - G(0; \Xi^u)} \nu(\theta_{i_l}^u; \varphi_{i_l}).$$

Sustituyendo esta última en (B.18)

$$\sum_{l=1}^{N-r} \sum_{i_l=1}^k \int_0^{\infty} e^{(x_{0l} - \mu_{i_l})e^{-\varphi_{i_l}}} \left\{ \frac{\pi_{i_l}^u f(x_{0l}; \theta_{i_l}^u)}{1 - G(0; \Xi^u)} \right\} dx_{0l} = \sum_{l=1}^{N-r} \sum_{i_l=1}^k \frac{\pi_{i_l}^u e^{-\mu_{i_l} e^{-\varphi_{i_l}}}}{1 - G(0; \Xi^u)} \nu(\theta_{i_l}^u; \varphi_{i_l}),$$

desarrollando las sumas como se hizo antes

$$= (N-r) \sum_{j=1}^k \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j). \quad (\text{B.19})$$

Si sumamos (B.5),(B.6),(B.16) y (B.19)

$$\begin{aligned} & \sum_{j=1}^k r_j \log\{\pi_j\} - \sum_{j=1}^k r_j \varphi_j + \sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} - \sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\ & + (N-r) \sum_{j=1}^{N-r} (\log\{\pi_j\} - \varphi_j) \left\{ \frac{\pi_j^u (1 - F(0; \theta_j^u))}{1 - G(0; \Xi^u)} \right\} \\ & + (N-r) \sum_{j=1}^k \frac{e^{-\varphi_j^u} \pi_j^u \{m(\theta_j^u) - \mu_j \{1 - F(0; \theta_j^u)\}\}}{1 - G(0; \Xi^u)} \\ & + (N-r) \sum_{j=1}^k \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j), \end{aligned}$$

o

$$\begin{aligned}
&= \sum_{j=1}^k \left\{ r_j + (N-r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \log \pi_j \\
&- \sum_{j=1}^k \left\{ r_j + (N-r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \varphi_j \\
&+ \sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} - \sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\
&+ (N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \right\} \\
&- (N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\},
\end{aligned}$$

con esto queda demostrado (3.10).

B.2. Cálculos del paso M del Capítulo 3

Calculemos el Lagrangiano de la función

$$h(\pi_1, \dots, \pi_k) = \sum_{j=1}^k \left\{ r_j + (N-r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \log \pi_j,$$

restringido a

$$f(\pi_1, \dots, \pi_k) = \sum_{j=1}^k \pi_j = 1.$$

Debemos resolver

$$\nabla h = \lambda \nabla f.$$

Pero

$$\lambda \nabla f = (\lambda, \lambda, \dots, \lambda),$$

mientras que si $s \in \{1, 2, \dots, k\}$, entonces

$$\frac{\partial h}{\partial \pi_s} = \frac{r_s + (N - r) \left\{ \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_s}. \quad (\text{B.20})$$

Entonces para que dos vectores sean iguales, deben coincidir entrada a entrada, por lo tanto

$$\lambda = \frac{r_s + (N - r) \left\{ \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_s} \quad \forall s = 1, 2, \dots, k,$$

o

$$\lambda \pi_s = r_s + (N - r) \left\{ \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\} \quad \forall s = 1, 2, \dots, k. \quad (\text{B.21})$$

Para determinar el valor de λ sumemos sobre todo el conjunto $s = \{1, 2, \dots, k\}$ en la ecuación (B.21)

$$\lambda \sum_{s=1}^k \pi_s = \sum_{s=1}^k \left\{ r_s + (N - r) \left\{ \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\} \right\},$$

distribuyendo la suma y usando el hecho que $\sum_{s=1}^k \pi_s = 1$

$$\lambda = \sum_{s=1}^k r_s + (N - r) \left\{ \frac{\sum_{s=1}^k \pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\},$$

puesto que $\sum_{s=1}^k r_s = r$, y $\sum_{s=1}^k \pi_s^u \{1 - F(0; \theta_s^u)\} = 1 - G(0; \Xi^u)$ podemos escribir

$$\lambda = r + (N - r).$$

Entonces $\lambda = N$ y por tanto sustituyendo en (B.21) tenemos que

$$\pi_s = \frac{r_s + (N - r) \left\{ \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\}}{N}, \quad \forall s = 1, 2, \dots, k.$$

Por lo tanto queda demostrado (3.11).

Calculemos el gradiente de la función

$$\begin{aligned} g(\varphi_1, \dots, \varphi_k; \mu_1, \dots, \mu_k) &= - \sum_{j=1}^k \left\{ r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \varphi_j \\ &+ \sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} - \sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\ &+ (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \right\} \\ &- (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\}. \quad (\text{B.22}) \end{aligned}$$

Sea $s \in \{1, \dots, k\}$.

Del primer sumando de (B.22), es claro que

$$-\frac{\partial}{\partial \mu_s} \left(\sum_{j=1}^k \left\{ r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \varphi_j \right) = 0. \quad (\text{B.23})$$

Derivando parcialmente con respecto a μ_s el segundo sumando de (B.22)

$$\frac{\partial}{\partial \mu_s} \left(\sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} \right)$$

desarrollando una suma

$$= \frac{\partial}{\partial \mu_s} \left(\sum_{i=1}^{r_1} (x_{1i} - \mu_1) e^{-\varphi_1} + \dots + \sum_{i=1}^{r_s} (x_{si} - \mu_s) e^{-\varphi_s} + \dots + \sum_{i=1}^{r_k} (x_{ki} - \mu_k) e^{-\varphi_k} \right),$$

como solo un sumando depende de μ_s entonces

$$\begin{aligned} &= \frac{\partial}{\partial \mu_s} \sum_{i=1}^{r_s} (x_{si} - \mu_s) e^{-\varphi_s} \\ &= \sum_{i=1}^{r_s} \frac{\partial}{\partial \mu_s} (x_{si} - \mu_s) e^{-\varphi_s} \\ &= - \sum_{i=1}^{r_s} e^{-\varphi_s}, \end{aligned}$$

por lo que

$$\frac{\partial}{\partial \mu_s} \left(\sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} \right) = -r_s e^{-\varphi_s}. \quad (\text{B.24})$$

Derivando parcialmente con respecto a μ_s el tercer sumando de (B.22)

$$- \frac{\partial}{\partial \mu_s} \left(\sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \right)$$

y de manera análoga

$$= - \frac{\partial}{\partial \mu_s} \left(\sum_{i=1}^{r_1} e^{(x_{1i} - \mu_1) e^{-\varphi_1}} + \dots + \sum_{i=1}^{r_s} e^{(x_{si} - \mu_s) e^{-\varphi_s}} + \dots + \sum_{i=1}^{r_k} e^{(x_{ki} - \mu_k) e^{-\varphi_k}} \right)$$

$$\begin{aligned}
 &= -\frac{\partial}{\partial \mu_s} \sum_{i=1}^{r_s} e^{(x_{si}-\mu_s)e^{-\varphi_s}} \\
 &= -\sum_{i=1}^{r_s} \frac{\partial}{\partial \mu_s} e^{(x_{si}-\mu_s)e^{-\varphi_s}} \\
 &= \sum_{i=1}^{r_s} e^{(x_{si}-\mu_s)e^{-\varphi_s}} e^{-\varphi_s}
 \end{aligned}$$

finalmente

$$-\frac{\partial}{\partial \mu_s} \left(\sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} \right) = e^{-\varphi_s} \sum_{i=1}^{r_s} e^{(x_{si}-\mu_s)e^{-\varphi_s}}. \quad (\text{B.25})$$

Derivando parcialmente con respecto a μ_s el cuarto sumando de (B.22)

$$\frac{\partial}{\partial \mu_s} \left((N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \right\} \right)$$

es fácil notar que queda

$$= -(N-r) \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} e^{-\varphi_s}. \quad (\text{B.26})$$

Derivando parcialmente con respecto a μ_s el quinto sumando de (B.22)

$$\frac{\partial}{\partial \mu_s} \left(-(N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \right)$$

$$\begin{aligned}
 &= -(N-r) \sum_{j=1}^k \frac{\partial}{\partial \mu_s} \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \\
 &= (N-r) \left\{ \frac{\pi_s^u e^{-\mu_s e^{-\varphi_s}} e^{-\varphi_s}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\}
 \end{aligned}$$

es decir, que

$$\begin{aligned}
 &\frac{\partial}{\partial \mu_s} \left(-(N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \right) \\
 &= (N-r) \left\{ \frac{\pi_s^u e^{-\mu_s e^{-\varphi_s}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} e^{-\varphi_s}.
 \end{aligned} \tag{B.27}$$

Ahora saquemos las parciales con respecto a φ_s de la función (B.22), del primer sumando tenemos

$$\begin{aligned}
 &\frac{\partial}{\partial \varphi_s} - \left(\sum_{j=1}^k \left\{ r_j + (N-r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right\} \varphi_j \right) = \\
 &= - \left\{ r_s + (N-r) \left\{ \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\} \right\}.
 \end{aligned} \tag{B.28}$$

Derivando parcialmente con respecto a φ_s el segundo sumando de (B.22)

$$\begin{aligned}
 & \frac{\partial}{\partial \varphi_s} \left(\sum_{j=1}^k \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{-\varphi_j} \right) \\
 &= \frac{\partial}{\partial \varphi_s} \left(\sum_{i=1}^{r_1} (x_{1i} - \mu_1) e^{-\varphi_1} + \dots + \sum_{i=1}^{r_s} (x_{si} - \mu_s) e^{-\varphi_s} + \dots + \sum_{i=1}^{r_k} (x_{ki} - \mu_k) e^{-\varphi_k} \right) \\
 &= \sum_{i=1}^{r_s} \frac{\partial}{\partial \varphi_s} (x_{si} - \mu_s) e^{-\varphi_s} \\
 &= -e^{-\varphi_s} \sum_{i=1}^{r_s} (x_{si} - \mu_s). \tag{B.29}
 \end{aligned}$$

Derivando parcialmente con respecto a φ_s el tercer sumando de (B.22)

$$\begin{aligned}
 & -\frac{\partial}{\partial \varphi_s} \left(\sum_{j=1}^k \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \right) \\
 &= -\frac{\partial}{\partial \varphi_s} \left(\sum_{i=1}^{r_1} e^{(x_{1i} - \mu_1) e^{-\varphi_1}} + \dots + \sum_{i=1}^{r_s} e^{(x_{si} - \mu_s) e^{-\varphi_s}} + \dots + \sum_{i=1}^{r_k} e^{(x_{ki} - \mu_k) e^{-\varphi_k}} \right) \\
 &= -\sum_{i=1}^{r_s} \frac{\partial}{\partial \varphi_s} e^{(x_{si} - \mu_s) e^{-\varphi_s}} \\
 &= \sum_{i=1}^{r_s} e^{(x_{si} - \mu_s) e^{-\varphi_s}} (x_{si} - \mu_s) e^{-\varphi_s}
 \end{aligned}$$

$$= e^{-\varphi_s} \sum_{i=1}^{r_s} (x_{si} - \mu_s) e^{(x_{si} - \mu_s) e^{-\varphi_s}}. \quad (\text{B.30})$$

Derivando parcialmente con respecto a φ_s el cuarto sumando de (B.22)

$$\begin{aligned} & \frac{\partial}{\partial \varphi_s} \left((N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \right\} \right) \\ &= (N-r) \sum_{j=1}^k \left\{ \left\{ \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} \right\} \frac{\partial}{\partial \varphi_s} e^{-\varphi_j} \right\} \\ &= -(N-r) \left\{ \frac{\pi_s^u \{m(\theta_s^u) - \mu_s(1 - F(0; \theta_s^u))\}}{1 - G(0; \Xi^u)} \right\} e^{-\varphi_s} \end{aligned} \quad (\text{B.31})$$

Derivando parcialmente con respecto a φ_s el quinto sumando de (B.22)

$$\begin{aligned} & -\frac{\partial}{\partial \varphi_s} (N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \\ &= -(N-r) \sum_{j=1}^k \frac{\partial}{\partial \varphi_s} \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \\ &= -(N-r) \frac{\partial}{\partial \varphi_s} \left\{ \frac{\pi_s^u e^{-\mu_s e^{-\varphi_s}}}{1 - G(0; \Xi^u)} \nu(\theta_s^u; \varphi_s) \right\}, \end{aligned}$$

sacando lo que es constante con respecto a φ_s

$$= -\frac{(N-r)\pi_s^u}{1-G(0; \Xi^u)} \left\{ \frac{\partial}{\partial \varphi_s} \left\{ e^{-\mu_s e^{-\varphi_s}} \nu(\theta_s^u; \varphi_s) \right\} \right\},$$

por la definición de la derivada de un producto

$$= -\frac{(N-r)\pi_s^u}{1-G(0; \Xi^u)} \left\{ \frac{\partial}{\partial \varphi_s} \left\{ \nu(\theta_s^u; \varphi_s) \right\} e^{-\mu_s e^{-\varphi_s}} + \frac{\partial}{\partial \varphi_s} \left\{ e^{-\mu_s e^{-\varphi_s}} \right\} \nu(\theta_s^u; \varphi_s) \right\},$$

realizando la derivada del segundo sumando tenemos que

$$= -\frac{(N-r)\pi_s^u}{1-G(0; \Xi^u)} \left\{ e^{-\mu_s e^{-\varphi_s}} \frac{\partial}{\partial \varphi_s} \left\{ \nu(\theta_s^u; \varphi_s) \right\} + \mu_s e^{-\varphi_s} e^{-\mu_s e^{-\varphi_s}} \nu(\theta_s^u; \varphi_s) \right\},$$

factorizando $e^{-\mu_s e^{-\varphi_s}}$

$$\begin{aligned} & -\frac{\partial}{\partial \varphi_s} (N-r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1-G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \\ & = -\frac{(N-r)\pi_s^u e^{-\mu_s e^{-\varphi_s}}}{1-G(0; \Xi^u)} \left\{ \frac{\partial}{\partial \varphi_s} \left\{ \nu(\theta_s^u; \varphi_s) \right\} + \mu_s e^{-\varphi_s} \nu(\theta_s^u; \varphi_s) \right\} \quad (\text{B.32}) \end{aligned}$$

Por otro lado calculemos $\frac{\partial}{\partial \varphi_s} \left\{ \nu(\theta_s^u; \varphi_s) \right\}$, por definición

$$\begin{aligned} \frac{\partial}{\partial \varphi_s} \left\{ \nu(\theta_s^u; \varphi_s) \right\} & = \frac{\partial}{\partial \varphi_s} \left\{ e^{\mu_s^u e^{-\varphi_s}} \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} e^{-x} dx \right\} \\ & = e^{\mu_s^u e^{-\varphi_s}} \frac{\partial}{\partial \varphi_s} \left\{ \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} e^{-x} dx \right\} + \frac{\partial}{\partial \varphi_s} \left\{ e^{\mu_s^u e^{-\varphi_s}} \right\} \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} e^{-x} dx, \end{aligned}$$

al meter la parcial a la derivada y hacer la parcial del segundo sumando

$$= e^{\mu_s^u e^{-\varphi_s}} \int_{l(\theta_s^u)}^{\infty} \frac{\partial}{\partial \varphi_s} \left\{ x^{e^{\varphi_s^u - \varphi_s}} \right\} e^{-x} dx - \mu_s^u e^{-\varphi_s} \left\{ e^{\mu_s^u e^{-\varphi_s}} \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} e^{-x} dx \right\},$$

lo que está entre parentesis del lado derecho es precisamente $\nu(\theta_s^u; \varphi_s)$

$$= e^{\mu_s^u e^{-\varphi_s}} \int_{l(\theta_s^u)}^{\infty} \frac{\partial}{\partial \varphi_s} \left\{ x^{e^{\varphi_s^u - \varphi_s}} \right\} e^{-x} dx - \mu_s^u e^{-\varphi_s} \nu(\theta_s^u; \varphi_s),$$

puesto que $\frac{d a^z}{d v} = a^z \log(a) \frac{d z}{d v}$, con a constante y z función de v

$$\frac{\partial}{\partial \varphi_s} \{ \nu(\theta_s^u; \varphi_s) \} = -e^{\mu_s^u e^{-\varphi_s}} \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} \log(x) e^{\varphi_s^u - \varphi_s} e^{-x} dx - \mu_s^u e^{-\varphi_s} \nu(\theta_s^u; \varphi_s),$$

o equivalentemente

$$= -e^{-\varphi_s} \left\{ e^{\mu_s^u e^{-\varphi_s}} e^{\varphi_s^u} \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} \log(x) e^{-x} dx + \mu_s^u \nu(\theta_s^u; \varphi_s) \right\},$$

definamos $\omega(\theta_s^u; \varphi_s) = e^{\mu_s^u e^{-\varphi_s}} e^{\varphi_s^u} \int_{l(\theta_s^u)}^{\infty} x^{e^{\varphi_s^u - \varphi_s}} \log(x) e^{-x} dx + \mu_s^u \nu(\theta_s^u; \varphi_s)$

entonces

$$\frac{\partial}{\partial \varphi_s} \{ \nu(\theta_s^u; \varphi_s) \} = -e^{-\varphi_s} \omega(\theta_s^u; \varphi_s) \quad (\text{B.33})$$

sustituyendo y factorizando $e^{-\varphi_s}$ en (B.32) queda

$$\begin{aligned} & -\frac{\partial}{\partial \varphi_s} (N - r) \sum_{j=1}^k \left\{ \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \nu(\theta_j^u; \varphi_j) \right\} \\ & = -e^{-\varphi_s} \frac{(N - r) \pi_s^u e^{-\mu_s e^{-\varphi_s}}}{1 - G(0; \Xi^u)} \{ \mu_s \nu(\theta_s^u; \varphi_s) - \omega(\theta_s^u; \varphi_s) \}. \quad (\text{B.34}) \end{aligned}$$

Al sumar las ecuaciones (B.23), (B.24), (B.25), (B.26) y (B.27) e igualando a cero

$$\begin{aligned} & e^{-\varphi_s} \left\{ -r_s + \sum_{i=1}^{r_s} e^{(x_{si} - \mu_s) e^{-\varphi_s}} - (N - r) \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\} \\ & + e^{-\varphi_s} \left\{ (N - r) \frac{\pi_s^u \{ \nu(\theta_s^u; \varphi_s) \}}{1 - G(0; \Xi^u)} e^{-\mu_s e^{-\varphi_s}} \right\} = 0, \end{aligned}$$

donde $s \in \{1, \dots, k\}$, multiplicando todo por e^{φ_s} tenemos

$$\begin{aligned} & - r_s + \sum_{i=1}^{r_s} e^{(x_{si} - \mu_s)e^{-\varphi_s}} - (N - r) \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \\ & + (N - r) \frac{\pi_s^u \{\nu(\theta_s^u; \varphi_s)\}}{1 - G(0; \Xi^u)} e^{-\mu_s e^{-\varphi_s}} = 0, \quad \forall s = 1, 2, \dots, k. \end{aligned}$$

Por lo tanto queda probado (3.12).

Por otra parte si sumamos las ecuaciones (B.28), (B.29), (B.30), (B.31) y (B.34) e igualamos a cero

$$\begin{aligned} & - \left\{ r_s + (N - r) \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\} - e^{-\varphi_s} \sum_{i=1}^{r_s} (x_{si} - \mu_s) \\ & + e^{-\varphi_s} \sum_{i=1}^{r_s} (x_{si} - \mu_s) e^{(x_{si} - \mu_s)e^{-\varphi_s}} - (N - r) \frac{\pi_s^u \{m(\theta_s^u) - \mu_s(1 - F(0; \theta_s^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_s} \\ & - (N - r) \frac{\pi_s^u e^{-\varphi_s - \mu_s e^{-\varphi_s}}}{1 - G(0; \Xi^u)} \{ \nu(\theta_s^u; \varphi_s) \mu_s - \omega(\theta_s^u; \varphi_s) \} \\ & = 0, \end{aligned} \tag{B.35}$$

multiplicando por e^{φ_s} tenemos que

$$\begin{aligned} & - \left\{ r_s + (N - r) \frac{\pi_s^u \{1 - F(0; \theta_s^u)\}}{1 - G(0; \Xi^u)} \right\} e^{\varphi_s} - \sum_{i=1}^{r_s} (x_{si} - \mu_s) \\ & + \sum_{i=1}^{r_s} (x_{si} - \mu_s) e^{(x_{si} - \mu_s)e^{-\varphi_s}} - (N - r) \frac{\pi_s^u \{m(\theta_s^u) - \mu_s(1 - F(0; \theta_s^u))\}}{1 - G(0; \Xi^u)} \\ & - (N - r) \frac{\pi_s^u e^{-\mu_s e^{-\varphi_s}}}{1 - G(0; \Xi^u)} \{ \nu(\theta_s^u; \varphi_s) \mu_s - \omega(\theta_s^u; \varphi_s) \} \\ & = 0, \end{aligned}$$

por lo tanto (3.14) queda demostrado.

B.3. Cálculos de la matriz Hessiana

En esta sección nos daremos a la tarea de calcular la matriz Hessiana $-\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2}$ para (3.22) y (3.23).

Primero recordemos que habíamos notado que a la función Q la podemos reescribir como

$$Q(\Xi, \Xi^u) \triangleq h(\boldsymbol{\pi}) + g(\boldsymbol{\varphi}, \boldsymbol{\mu}),$$

donde $\Xi = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_k, \pi_1, \dots, \pi_k)$ y $\boldsymbol{\theta}_s = (\mu_s, \varphi_s)$ para $s = 1, \dots, k$.

Entonces debemos calcular

$$-\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} = -\frac{\partial^2 h(\boldsymbol{\pi})}{\partial \Xi^2} - \frac{\partial^2 g(\boldsymbol{\varphi}, \boldsymbol{\mu})}{\partial \Xi^2}$$

Para los siguientes cálculos supongamos que $s, j \in \{1, 2, \dots, k\}$.

Primero calculemos

$$-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \mu_s \partial \mu_j} = -\frac{\partial^2 h(\boldsymbol{\pi})}{\partial \mu_s \partial \mu_j} - \frac{\partial^2 g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \mu_s \partial \mu_j} = -\frac{\partial}{\partial \mu_s} \left(\frac{\partial g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \mu_j} \right)$$

la derivada $\frac{\partial g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \mu_j}$ ya la calculamos en (B.35), ahora derivemos con respecto a μ_s

$$\begin{aligned} & - \frac{\partial}{\partial \mu_s} \left(e^{-\varphi_j} \left\{ -r_j + \sum_{i=1}^{r_j} e^{(x_{ji} - \mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right) \\ & + \frac{\partial}{\partial \mu_s} \left(e^{-\varphi_j} \left\{ (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} \right\} \right), \end{aligned}$$

notemos que si $j \neq s$, entonces $-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \mu_s \partial \mu_j} = 0$.

Si $s = j$ entonces

$$\begin{aligned}
& -\frac{\partial^2 Q(\Xi, \Xi)}{\partial \mu_j^2} = \\
& -\frac{\partial}{\partial \mu_j} \left(e^{-\varphi_j} \left\{ -r_j + \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} \right) \\
& + \frac{\partial}{\partial \mu_j} \left(e^{-\varphi_j} \left\{ (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} \right\} \right) \\
& = -e^{-\varphi_j} \left\{ \sum_{i=1}^{r_j} \frac{\partial}{\partial \mu_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} + (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)} \frac{\partial}{\partial \mu_j} e^{-\mu_j e^{-\varphi_j}} \right\} \\
& = -e^{-\varphi_j} \left\{ -e^{-\varphi_j} \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - e^{-\varphi_j} (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} \right\} \\
& -\frac{\partial^2 Q(\Xi, \Xi)}{\partial \mu_j^2} = e^{-2\varphi_j} \left\{ \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} + (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} \right\}.
\end{aligned}$$

Si $k = 2$ podemos obtener las entradas de la matriz (3.22)

$$Q_{11} = e^{-2\varphi_1} \left\{ \sum_{i=1}^{r_1} e^{(x_{1i}-\mu_1)e^{-\varphi_1}} + (N-r) \frac{\pi_1^u \{\nu(\theta_1^u; \varphi_1)\}}{1 - G(0; \Xi^u)} e^{-\mu_1 e^{-\varphi_1}} \right\},$$

y

$$Q_{22} = e^{-2\varphi_2} \left\{ \sum_{i=1}^{r_2} e^{(x_{2i}-\mu_2)e^{-\varphi_2}} + (N-r) \frac{\pi_2^u \{\nu(\theta_2^u; \varphi_2)\}}{1 - G(0; \Xi^u)} e^{-\mu_2 e^{-\varphi_2}} \right\}.$$

Si $k = 3$, Q_{11} y Q_{22} son iguales que en el caso $k = 2$ y además la tercera entrada sobre la diagonal principal de (3.23) está dada por

$$Q_{33} = e^{-2\varphi_3} \left\{ \sum_{i=1}^{r_3} e^{(x_{3i}-\mu_3)e^{-\varphi_3}} + (N-r) \frac{\pi_3^u \{\nu(\theta_3^u; \varphi_3)\}}{1 - G(0; \Xi^u)} e^{-\mu_3 e^{-\varphi_3}} \right\}.$$

Antes de continuar, calculemos $\frac{\partial \omega(\theta_j^u; \varphi_j)}{\partial \varphi_j}$, ya que lo necesitaremos más adelante.

Por definición tenemos que

$$\begin{aligned} \frac{\partial \omega(\theta_j^u; \varphi_j)}{\partial \varphi_j} &= \frac{\partial}{\partial \varphi_j} \left(e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx + \mu_j^u \nu(\theta_j^u; \varphi_j) \right) \\ &= \frac{\partial}{\partial \varphi_j} \left(e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx \right) + \mu_j^u \frac{\partial \nu(\theta_j^u; \varphi_j)}{\partial \varphi_j}, \end{aligned}$$

de (B.33) se sigue

$$= \frac{\partial}{\partial \varphi_j} \left(e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx \right) - \mu_j^u e^{-\varphi_j} \omega(\theta_j^u; \varphi_j),$$

de la definición de la derivada del producto

$$\begin{aligned} &= e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \frac{\partial}{\partial \varphi_j} \left(\int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx \right) \\ &+ \frac{\partial}{\partial \varphi_j} \left(e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \right) \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx - \mu_j^u e^{-\varphi_j} \omega(\theta_j^u; \varphi_j), \end{aligned}$$

calculando las derivadas

$$\begin{aligned} &= -e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} (\log(x))^2 e^{\varphi_j^u - \varphi_j} e^{-x} dx \\ &- \mu_j^u e^{-\varphi_j} e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx - \mu_j^u e^{-\varphi_j} \omega(\theta_j^u; \varphi_j), \end{aligned}$$

definamos a $\Gamma(x, \varphi_j) = \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} (\log(x))^2 e^{\varphi_j^u - \varphi_j} e^{-x} dx$ y $\Lambda(x, \varphi_j) = \int_{l(\theta_j^u)}^{\infty} x^{e^{\varphi_j^u - \varphi_j}} \log(x) e^{-x} dx$, entonces podemos reescribir la última expresión como

$$\begin{aligned} \frac{\partial \omega(\theta_j^u; \varphi_j)}{\partial \varphi_j} &= -e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \Gamma(x, \varphi_j) \\ &- \mu_j^u e^{\varphi_j^u - \varphi_j + \mu_j^u e^{-\varphi_j}} \Lambda(x, \varphi_j) - \mu_j^u e^{-\varphi_j} \omega(\theta_j^u; \varphi_j). \end{aligned} \quad (\text{B.36})$$

Ahora calculemos

$$-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \varphi_s \partial \varphi_j} = -\frac{\partial^2 h(\pi)}{\partial \varphi_s \partial \varphi_j} - \frac{\partial^2 g(\mu, \varphi)}{\partial \varphi_s \partial \varphi_j} = -\frac{\partial}{\partial \varphi_s} \left(\frac{\partial g(\mu, \varphi)}{\partial \varphi_j} \right)$$

en (B.35) se mostró que $\frac{\partial g(\mu, \varphi)}{\partial \varphi_j}$ es igual a

$$\begin{aligned} &- \left\{ r_j + (N - r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\} - e^{-\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) \\ &+ e^{-\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{(x_{ji} - \mu_j) e^{-\varphi_j}} - (N - r) \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_j} \\ &- (N - r) \frac{\pi_j^u e^{-\mu_j e^{-\varphi_j}}}{1 - G(0; \Xi^u)} \{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} e^{-\varphi_j} \end{aligned}$$

si a esta última la derivamos con respecto a φ_s será cero si $s \neq j$, pero si $s = j$ tenemos que

$$\begin{aligned}
-\frac{\partial^2 g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \varphi_j^2} &= -e^{-\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) + e^{-\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\
&+ e^{-2\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j)^2 e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\
&- (N - r) \frac{\pi_j^u \{m(\theta_j^u) - \mu_j(1 - F(0; \theta_j^u))\}}{1 - G(0; \boldsymbol{\Xi}^u)} e^{-\varphi_j} \\
&+ (N - r) \left(\frac{\pi_j^u}{1 - G(0; \boldsymbol{\Xi}^u)} \right) \\
&\times \underbrace{\frac{\partial}{\partial \varphi_j} \left(\{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} e^{-\varphi_j - \mu_j e^{-\varphi_j}} \right)}_i. \quad (\text{B.37})
\end{aligned}$$

Calculemos aparte i de la última expresión

$$\begin{aligned}
&\frac{\partial}{\partial \varphi_j} \left(\{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} e^{-\varphi_j - \mu_j e^{-\varphi_j}} \right) \\
&= \{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} \frac{\partial}{\partial \varphi_j} (e^{-\varphi_j - \mu_j e^{-\varphi_j}}) \\
&+ \frac{\partial}{\partial \varphi_j} \left(\{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} \right) (e^{-\varphi_j - \mu_j e^{-\varphi_j}}),
\end{aligned}$$

calculando las derivadas

$$\begin{aligned}
&= - \{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} \left(e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \right) \\
&+ \left\{ \mu_j \frac{\partial \nu(\theta_j^u; \varphi_j)}{\partial \varphi_j} - \frac{\partial \omega(\theta_j^u; \varphi_j)}{\partial \varphi_j} \right\} (e^{-\varphi_j - \mu_j e^{-\varphi_j}}),
\end{aligned}$$

distribuyendo los productos

$$\begin{aligned}
&= \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j \nu(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&+ \mu_j \frac{\partial \nu(\theta_j^u; \varphi_j)}{\partial \varphi_j} e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&- \frac{\partial \omega(\theta_j^u; \varphi_j)}{\partial \varphi_j} e^{-\varphi_j - \mu_j e^{-\varphi_j}},
\end{aligned}$$

de (B.33) sabemos que $\frac{\partial}{\partial \varphi_s} \{ \nu(\theta_s^u; \varphi_s) \} = -e^{-\varphi_s} \omega(\theta_s^u; \varphi_s)$, entonces

$$\begin{aligned}
&= \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j \nu(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&- \frac{\partial \omega(\theta_j^u; \varphi_j)}{\partial \varphi_j} e^{-\varphi_j - \mu_j e^{-\varphi_j}}
\end{aligned}$$

al sustituir (B.36) en la última expresión tenemos que

$$\begin{aligned}
&= \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j \nu(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ \left\{ e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \Gamma(x, \varphi_j) + \mu_j^u e^{\varphi_j^u - \varphi_j + \mu_j^u e^{-\varphi_j}} \Lambda(x, \varphi_j) + \mu_j^u e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) \right\} e^{-\varphi_j - \mu_j e^{-\varphi_j}},
\end{aligned}$$

distribuyendo el producto

$$\begin{aligned}
&= \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j \nu(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \Gamma(x, \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ \mu_j^u e^{\varphi_j^u - \varphi_j + \mu_j^u e^{-\varphi_j}} \Lambda(x, \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ \mu_j^u e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}},
\end{aligned}$$

notemos que el tercer sumando y el sexto se pueden factorizar de la siguiente manera que

$$\begin{aligned}
& \frac{\partial}{\partial \varphi_j} \left(\{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} e^{-\varphi_j - \mu_j e^{-\varphi_j}} \right) \\
&= \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&- \mu_j \nu(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&+ (\mu_j^u - \mu_j) e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \Gamma(x, \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ \mu_j^u e^{\varphi_j^u - \varphi_j + \mu_j^u e^{-\varphi_j}} \Lambda(x, \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}},
\end{aligned}$$

sustituyendo esto en i de (B.37)

$$\begin{aligned}
-\frac{\partial^2 g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \varphi_j^2} &= -e^{-\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) + e^{-\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\
&+ e^{-2\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j)^2 e^{(x_{ji} - \mu_j) e^{-\varphi_j}} \\
&- (N - r) \frac{\pi_j^u \{ m(\theta_j^u) - \mu_j (1 - F(0; \theta_j^u)) \}}{1 - G(0; \boldsymbol{\Xi}^u)} e^{-\varphi_j} \\
&+ (N - r) \left(\frac{\pi_j^u}{1 - G(0; \boldsymbol{\Xi}^u)} \right) \\
&\times \left\{ \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \right. \\
&- \mu_j \nu(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \{ \mu_j e^{-\varphi_j} - 1 \} \\
&+ (\mu_j^u - \mu_j) e^{-\varphi_j} \omega(\theta_j^u; \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&+ e^{\varphi_j^u + \mu_j^u e^{-\varphi_j}} \Gamma(x, \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \\
&\left. + \mu_j^u e^{\varphi_j^u - \varphi_j + \mu_j^u e^{-\varphi_j}} \Lambda(x, \varphi_j) e^{-\varphi_j - \mu_j e^{-\varphi_j}} \right\}.
\end{aligned}$$

entonces en el caso en que $k = 2$ la tercera y cuarta entrada de las matriz (3.22) sobre la diagonal principal están dadas por

$$\begin{aligned}
Q_{3,3} &= -e^{-\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1) + e^{-\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1) e^{(x_{1i} - \mu_1)e^{-\varphi_1}} \\
&+ e^{-2\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1)^2 e^{(x_{1i} - \mu_1)e^{-\varphi_1}} \\
&- (N - r) \frac{\pi_1^u \{m(\theta_1^u) - \mu_1(1 - F(0; \theta_1^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_1} \\
&+ (N - r) \left(\frac{\pi_1^u}{1 - G(0; \Xi^u)} \right) \\
&\times \left\{ \omega(\theta_1^u; \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \{ \mu_1 e^{-\varphi_1} - 1 \} \right. \\
&- \mu_1 \nu(\theta_1^u; \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \{ \mu_1 e^{-\varphi_1} - 1 \} \\
&+ (\mu_1^u - \mu_1) e^{-\varphi_1} \omega(\theta_1^u; \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \\
&+ e^{\varphi_1^u + \mu_1^u e^{-\varphi_1}} \Gamma(x, \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \\
&\left. + \mu_1^u e^{\varphi_1^u - \varphi_1 + \mu_1^u e^{-\varphi_1}} \Lambda(x, \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \right\}.
\end{aligned}$$

$$\begin{aligned}
Q_{4,4} &= -e^{-\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2) + e^{-\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2) e^{(x_{2i} - \mu_2)e^{-\varphi_2}} \\
&+ e^{-2\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2)^2 e^{(x_{2i} - \mu_2)e^{-\varphi_2}} \\
&- (N - r) \frac{\pi_2^u \{m(\theta_2^u) - \mu_2(1 - F(0; \theta_2^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_2} \\
&+ (N - r) \left(\frac{\pi_2^u}{1 - G(0; \Xi^u)} \right) \\
&\times \left\{ \omega(\theta_2^u; \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \{ \mu_2 e^{-\varphi_2} - 1 \} \right. \\
&- \mu_2 \nu(\theta_2^u; \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \{ \mu_2 e^{-\varphi_2} - 1 \} \\
&+ (\mu_2^u - \mu_2) e^{-\varphi_2} \omega(\theta_2^u; \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \\
&+ e^{\varphi_2^u + \mu_2^u e^{-\varphi_2}} \Gamma(x, \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \\
&\left. + \mu_2^u e^{\varphi_2^u - \varphi_2 + \mu_2^u e^{-\varphi_2}} \Lambda(x, \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \right\}.
\end{aligned}$$

Si $k = 3$ la cuarta, quinta y sexta entrada de las matriz (3.23) sobre la diagonal principal están dadas por

$$\begin{aligned}
Q_{4,4} = & -e^{-\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1) + e^{-\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1) e^{(x_{1i} - \mu_1) e^{-\varphi_1}} \\
& + e^{-2\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1)^2 e^{(x_{1i} - \mu_1) e^{-\varphi_1}} \\
& - (N - r) \frac{\pi_1^u \{m(\theta_1^u) - \mu_1(1 - F(0; \theta_1^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_1} \\
& + (N - r) \left(\frac{\pi_1^u}{1 - G(0; \Xi^u)} \right) \\
& \times \left\{ \omega(\theta_1^u; \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \{ \mu_1 e^{-\varphi_1} - 1 \} \right. \\
& - \mu_1 \nu(\theta_1^u; \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \{ \mu_1 e^{-\varphi_1} - 1 \} \\
& + (\mu_1^u - \mu_1) e^{-\varphi_1} \omega(\theta_1^u; \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \\
& + e^{\varphi_1^u + \mu_1^u e^{-\varphi_1}} \Gamma(x, \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \\
& \left. + \mu_1^u e^{\varphi_1^u - \varphi_1 + \mu_1^u e^{-\varphi_1}} \Lambda(x, \varphi_1) e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} \right\}.
\end{aligned}$$

$$\begin{aligned}
Q_{5,5} = & -e^{-\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2) + e^{-\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2) e^{(x_{2i} - \mu_2) e^{-\varphi_2}} \\
& + e^{-2\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2)^2 e^{(x_{2i} - \mu_2) e^{-\varphi_2}} \\
& - (N - r) \frac{\pi_2^u \{m(\theta_2^u) - \mu_2(1 - F(0; \theta_2^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_2} \\
& + (N - r) \left(\frac{\pi_2^u}{1 - G(0; \Xi^u)} \right) \\
& \times \left\{ \omega(\theta_2^u; \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \{ \mu_2 e^{-\varphi_2} - 1 \} \right. \\
& - \mu_2 \nu(\theta_2^u; \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \{ \mu_2 e^{-\varphi_2} - 1 \} \\
& + (\mu_2^u - \mu_2) e^{-\varphi_2} \omega(\theta_2^u; \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \\
& + e^{\varphi_2^u + \mu_2^u e^{-\varphi_2}} \Gamma(x, \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \\
& \left. + \mu_2^u e^{\varphi_2^u - \varphi_2 + \mu_2^u e^{-\varphi_2}} \Lambda(x, \varphi_2) e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} \right\}.
\end{aligned}$$

$$\begin{aligned}
Q_{6,6} &= -e^{-\varphi_3} \sum_{i=1}^{r_3} (x_{3i} - \mu_3) + e^{-\varphi_3} \sum_{i=1}^{r_3} (x_{3i} - \mu_3) e^{(x_{3i} - \mu_3)e^{-\varphi_3}} \\
&+ e^{-2\varphi_3} \sum_{i=1}^{r_3} (x_{3i} - \mu_3)^2 e^{(x_{3i} - \mu_3)e^{-\varphi_3}} \\
&- (N - r) \frac{\pi_3^u \{m(\theta_3^u) - \mu_3(1 - F(0; \theta_3^u))\}}{1 - G(0; \Xi^u)} e^{-\varphi_3} \\
&+ (N - r) \left(\frac{\pi_3^u}{1 - G(0; \Xi^u)} \right) \\
&\times \left\{ \omega(\theta_3^u; \varphi_3) e^{-\varphi_3 - \mu_3 e^{-\varphi_3}} \{ \mu_3 e^{-\varphi_3} - 1 \} \right. \\
&- \mu_3 \nu(\theta_3^u; \varphi_3) e^{-\varphi_3 - \mu_3 e^{-\varphi_3}} \{ \mu_3 e^{-\varphi_3} - 1 \} \\
&+ (\mu_3^u - \mu_3) e^{-\varphi_3} \omega(\theta_3^u; \varphi_3) e^{-\varphi_3 - \mu_3 e^{-\varphi_3}} \\
&+ e^{\varphi_3^u + \mu_3^u e^{-\varphi_3}} \Gamma(x, \varphi_3) e^{-\varphi_3 - \mu_3 e^{-\varphi_3}} \\
&\left. + \mu_3^u e^{\varphi_3^u - \varphi_3 + \mu_3^u e^{-\varphi_3}} \Lambda(x, \varphi_3) e^{-\varphi_3 - \mu_3 e^{-\varphi_3}} \right\}.
\end{aligned}$$

Ahora calculemos

$$-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \pi_s \partial \pi_j} = -\frac{\partial^2 h(\boldsymbol{\pi})}{\partial \pi_s \partial \pi_j} - \frac{\partial^2 g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \pi_s \partial \pi_j} = -\frac{\partial}{\partial \pi_s} \left(\frac{\partial h(\boldsymbol{\pi})}{\partial \pi_j} \right),$$

la primera derivada $\frac{\partial h(\boldsymbol{\pi})}{\partial \pi_j}$ ya se calculó en (B.20), a esta, la debemos derivar parcialmente con respecto a π_s

$$= -\frac{\partial}{\partial \pi_s} \left(\frac{r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_j} \right),$$

notemos que si $j \neq s$ entonces $-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \pi_s \partial \pi_j} = 0$. Pero si $s = j$ tenemos que

$$-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \pi_j^2} = \frac{r_j + (N - r) \left\{ \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_j^2}.$$

Entonces si $k = 2$ la quinta y sexta entrada sobre la diagonal principal de las matriz (3.22) están dadas por

$$Q_{55} = \frac{r_1 + (N - r) \left\{ \frac{\pi_1^u \{1 - F(0; \theta_1^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_1^2},$$

y

$$Q_{66} = \frac{r_2 + (N - r) \left\{ \frac{\pi_2^u \{1 - F(0; \theta_2^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_2^2}.$$

Mientras que se k fuera igual a tres entonces la septima, octava y novena entradas sobre la diagonal principal de las matriz (3.23) están dadas por

$$Q_{77} = \frac{r_1 + (N - r) \left\{ \frac{\pi_1^u \{1 - F(0; \theta_1^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_1^2}$$

$$Q_{88} = \frac{r_2 + (N - r) \left\{ \frac{\pi_2^u \{1 - F(0; \theta_2^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_2^2},$$

y

$$Q_{99} = \frac{r_3 + (N - r) \left\{ \frac{\pi_3^u \{1 - F(0; \theta_3^u)\}}{1 - G(0; \Xi^u)} \right\}}{\pi_3^2}.$$

Ahora calculemos

$$-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \varphi_s \partial \mu_j} = -\frac{\partial^2 h(\pi)}{\partial \varphi_s \partial \mu_j} - \frac{\partial^2 g(\mu, \varphi)}{\partial \varphi_s \partial \mu_j} = -\frac{\partial}{\partial \varphi_s} \left(\frac{\partial g(\mu, \varphi)}{\partial \mu_j} \right)$$

de (B.35) sabemos que $\frac{\partial g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \mu_j}$ es igual a

$$-e^{-\varphi_j} \left\{ -r_j + \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \boldsymbol{\Xi}^u)} \right. \\ \left. + (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \boldsymbol{\Xi}^u)} e^{-\mu_j e^{-\varphi_j}} \right\},$$

a esta función, tenemos que derivarla con respecto a φ_s , pero si $s \neq j$ entonces $-\frac{\partial^2 Q(\boldsymbol{\Xi}, \boldsymbol{\Xi})}{\partial \varphi_s \partial \mu_j} = 0$.

Si $s = j$ entonces al derivar el producto

$$-\frac{\partial^2 Q(\boldsymbol{\Xi}, \boldsymbol{\Xi})}{\partial \varphi_j \partial \mu_j} = \\ -e^{-\varphi_j} \left\{ \frac{\partial}{\partial \varphi_j} \left(-r_j + \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \boldsymbol{\Xi}^u)} \right. \right. \\ \left. \left. + (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \boldsymbol{\Xi}^u)} e^{-\mu_j e^{-\varphi_j}} \right) \right\} \\ + e^{-\varphi_j} \left\{ -r_j + \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \boldsymbol{\Xi}^u)} \right. \\ \left. + (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \boldsymbol{\Xi}^u)} e^{-\mu_j e^{-\varphi_j}} \right\},$$

o equivalentemente

$$= -e^{-\varphi_j} \left\{ \frac{\partial}{\partial \varphi_j} \left(\sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} \right) + \frac{(N-r)\pi_j^u}{1 - G(0; \boldsymbol{\Xi}^u)} \frac{\partial}{\partial \varphi_j} \left(\nu(\theta_j^u; \varphi_j) e^{-\mu_j e^{-\varphi_j}} \right) \right\} \\ + e^{-\varphi_j} \left\{ -r_j + \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \boldsymbol{\Xi}^u)} \right. \\ \left. + (N-r) \frac{\pi_j^u \{\nu(\theta_j^u; \varphi_j)\}}{1 - G(0; \boldsymbol{\Xi}^u)} e^{-\mu_j e^{-\varphi_j}} \right\},$$

de (B.33) se sigue

$$\begin{aligned}
&= -e^{-\varphi_j} \left\{ -\sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} (x_{ji}-\mu_j)e^{-\varphi_j} \right. \\
&+ \left. \frac{(N-r)\pi_j^u}{1-G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} e^{-\varphi_j} \{ \nu(\theta_j^u; \varphi_j) \mu_j - \omega(\theta_j^u; \varphi_j) \} \right\} \\
&+ e^{-\varphi_j} \left\{ -r_j + \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \right. \\
&+ \left. (N-r) \frac{\pi_j^u \{ \nu(\theta_j^u; \varphi_j) \}}{1 - G(0; \Xi^u)} e^{-\mu_j e^{-\varphi_j}} \right\},
\end{aligned}$$

por último tenemos que al distribuir

$$\begin{aligned}
-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \varphi_j \partial \mu_j} &= -r_j e^{-\varphi_j} + e^{-\varphi_j} \sum_{i=1}^{r_j} e^{(x_{ji}-\mu_j)e^{-\varphi_j}} - (N-r) e^{-\varphi_j} \frac{\pi_j^u \{1 - F(0; \theta_j^u)\}}{1 - G(0; \Xi^u)} \\
&+ \frac{(N-r)\pi_j^u \{ \nu(\theta_j^u; \varphi_j) \}}{1 - G(0; \Xi^u)} \left\{ e^{-\varphi_j - \mu_j e^{-\varphi_j}} - \mu_j e^{-2\varphi_j - \mu_j e^{-\varphi_j}} \right\} \\
&+ e^{-2\varphi_j} \sum_{i=1}^{r_j} (x_{ji} - \mu_j) e^{(x_{ji}-\mu_j)e^{-\varphi_j}} + \frac{(N-r)\pi_j^u \omega(\theta_j^u; \varphi_j)}{1 - G(0; \Xi^u)} e^{-2\varphi_j - \mu_j e^{-\varphi_j}}.
\end{aligned}$$

Si estamos bajo el supuesto de que $k = 2$ entonces las entradas (1,3) y (2,4) de la matriz (3.22) quedan determinadas de la siguiente manera

$$\begin{aligned}
Q_{1,3} &= -r_1 e^{-\varphi_1} + e^{-\varphi_1} \sum_{i=1}^{r_1} e^{(x_{1i}-\mu_1)e^{-\varphi_1}} - (N-r) e^{-\varphi_1} \frac{\pi_1^u \{1 - F(0; \theta_1^u)\}}{1 - G(0; \Xi^u)} \\
&+ \frac{(N-r)\pi_1^u \{ \nu(\theta_1^u; \varphi_1) \}}{1 - G(0; \Xi^u)} \left\{ e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} - \mu_1 e^{-2\varphi_1 - \mu_1 e^{-\varphi_1}} \right\} \\
&+ e^{-2\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1) e^{(x_{1i}-\mu_1)e^{-\varphi_1}} + \frac{(N-r)\pi_1^u \omega(\theta_1^u; \varphi_1)}{1 - G(0; \Xi^u)} e^{-2\varphi_1 - \mu_1 e^{-\varphi_1}},
\end{aligned}$$

y

$$\begin{aligned}
Q_{2,4} &= -r_2 e^{-\varphi_2} + e^{-\varphi_2} \sum_{i=1}^{r_2} e^{(x_{2i}-\mu_2)e^{-\varphi_2}} - (N-r)e^{-\varphi_2} \frac{\pi_2^u \{1 - F(0; \theta_2^u)\}}{1 - G(0; \Xi^u)} \\
&+ \frac{(N-r)\pi_2^u \{\nu(\theta_2^u; \varphi_2)\}}{1 - G(0; \Xi^u)} \left\{ e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} - \mu_2 e^{-2\varphi_2 - \mu_2 e^{-\varphi_2}} \right\} \\
&+ e^{-2\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2) e^{(x_{2i}-\mu_2)e^{-\varphi_2}} + \frac{(N-r)\pi_2^u \omega(\theta_2^u; \varphi_2)}{1 - G(0; \Xi^u)} e^{-2\varphi_2 - \mu_2 e^{-\varphi_2}}.
\end{aligned}$$

Mientras que si $k = 3$ entonces las entradas (1,4), (2,5) y (3,6) de la matriz (3.23) quedan determinadas de la siguiente forma

$$\begin{aligned}
Q_{1,4} &= -r_1 e^{-\varphi_1} + e^{-\varphi_1} \sum_{i=1}^{r_1} e^{(x_{1i}-\mu_1)e^{-\varphi_1}} - (N-r)e^{-\varphi_1} \frac{\pi_1^u \{1 - F(0; \theta_1^u)\}}{1 - G(0; \Xi^u)} \\
&+ \frac{(N-r)\pi_1^u \{\nu(\theta_1^u; \varphi_1)\}}{1 - G(0; \Xi^u)} \left\{ e^{-\varphi_1 - \mu_1 e^{-\varphi_1}} - \mu_1 e^{-2\varphi_1 - \mu_1 e^{-\varphi_1}} \right\} \\
&+ e^{-2\varphi_1} \sum_{i=1}^{r_1} (x_{1i} - \mu_1) e^{(x_{1i}-\mu_1)e^{-\varphi_1}} + \frac{(N-r)\pi_1^u \omega(\theta_1^u; \varphi_1)}{1 - G(0; \Xi^u)} e^{-2\varphi_1 - \mu_1 e^{-\varphi_1}}
\end{aligned}$$

$$\begin{aligned}
Q_{2,5} &= -r_2 e^{-\varphi_2} + e^{-\varphi_2} \sum_{i=1}^{r_2} e^{(x_{2i}-\mu_2)e^{-\varphi_2}} - (N-r)e^{-\varphi_2} \frac{\pi_2^u \{1 - F(0; \theta_2^u)\}}{1 - G(0; \Xi^u)} \\
&+ \frac{(N-r)\pi_2^u \{\nu(\theta_2^u; \varphi_2)\}}{1 - G(0; \Xi^u)} \left\{ e^{-\varphi_2 - \mu_2 e^{-\varphi_2}} - \mu_2 e^{-2\varphi_2 - \mu_2 e^{-\varphi_2}} \right\} \\
&+ e^{-2\varphi_2} \sum_{i=1}^{r_2} (x_{2i} - \mu_2) e^{(x_{2i}-\mu_2)e^{-\varphi_2}} + \frac{(N-r)\pi_2^u \omega(\theta_2^u; \varphi_2)}{1 - G(0; \Xi^u)} e^{-2\varphi_2 - \mu_2 e^{-\varphi_2}},
\end{aligned}$$

y

$$\begin{aligned}
Q_{3,6} &= -r_3 e^{-\varphi_3} + e^{-\varphi_3} \sum_{i=1}^{r_3} e^{(x_{3i}-\mu_3)e^{-\varphi_3}} - (N-r)e^{-\varphi_3} \frac{\pi_3^u \{1 - F(0; \theta_3^u)\}}{1 - G(0; \Xi^u)} \\
&+ \frac{(N-r)\pi_3^u \{\nu(\theta_3^u; \varphi_3)\}}{1 - G(0; \Xi^u)} \left\{ e^{-\varphi_3 - \mu_3 e^{-\varphi_3}} - \mu_3 e^{-2\varphi_3 - \mu_3 e^{-\varphi_3}} \right\} \\
&+ e^{-2\varphi_3} \sum_{i=1}^{r_3} (x_{3i} - \mu_3) e^{(x_{3i}-\mu_3)e^{-\varphi_3}} + \frac{(N-r)\pi_3^u \omega(\theta_3^u; \varphi_3)}{1 - G(0; \Xi^u)} e^{-2\varphi_3 - \mu_3 e^{-\varphi_3}}.
\end{aligned}$$

Ahora calculemos

$$\begin{aligned} -\frac{\partial^2 Q(\Xi, \Xi)}{\partial \mu_s \partial \pi_j} &= -\frac{\partial^2 h(\boldsymbol{\pi})}{\partial \mu_s \partial \pi_j} - \frac{\partial^2 g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \mu_s \partial \pi_j} \\ &= -\frac{\partial}{\partial \mu_s} \left(\frac{\partial h(\boldsymbol{\pi})}{\partial \pi_j} \right) - \frac{\partial}{\partial \mu_s} \left(\frac{\partial g(\boldsymbol{\mu}, \boldsymbol{\varphi})}{\partial \pi_j} \right) = 0 + 0 = 0, \end{aligned}$$

para toda $s, j \in \{1, 2, \dots, k\}$.

Esto es claro, puesto que g es función de $\boldsymbol{\mu}$ y $\boldsymbol{\varphi}$ y al derivar con respecto a π_j es cero, mientras que $\frac{\partial h(\boldsymbol{\pi})}{\partial \pi_j}$ será función de π_j y al derivar después con respecto a μ_s nos da nuevamente cero.

De manera completamente análoga podemos concluir que

$$-\frac{\partial^2 Q(\Xi, \Xi)}{\partial \varphi_s \partial \pi_j} = 0.$$

Por último veamos que en efecto al multiplicar (3.20) por (3.21) obtenemos la matriz identidad de dimensión $3k$.

Recordemos que

$$\frac{-\partial^2 \log \mathbb{P}(\Xi | \boldsymbol{x})}{\partial \Xi^2} = \left\{ \boldsymbol{I} - \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi = \Xi^\eta} \right\} \left\{ -\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi = \Xi^\eta} \right\} \quad (\text{B.38})$$

y que

$$\begin{aligned} \left\{ \frac{-\partial^2 \log \mathbb{P}(\Xi | \boldsymbol{x})}{\partial \Xi^2} \right\}^{-1} &= \left\{ -\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi = \Xi^\eta} \right\}^{-1} \left\{ \boldsymbol{I} - \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi = \Xi^\eta} \right\}^{-1} \\ &\times \left\{ \frac{\partial M(\Xi)}{\partial \Xi} \Big|_{\Xi = \Xi^\eta} \right\} + \left\{ -\frac{\partial^2 Q(\Xi, \Xi^\eta)}{\partial \Xi^2} \Big|_{\Xi = \Xi^\eta} \right\}^{-1} \quad (\text{B.39}) \end{aligned}$$

Renombremos como

$$A = \left. \frac{\partial M(\Xi)}{\partial \Xi} \right|_{\Xi=\Xi^n} \quad \text{y} \quad B = - \left. \frac{\partial^2 Q(\Xi, \Xi^n)}{\partial \Xi^2} \right|_{\Xi=\Xi^n},$$

al sustituir A y B en (B.38) y (B.39), y multiplicar ambas expresiones tenemos que

$$\begin{aligned} & \{\mathbf{I} - A\} B \times \{B^{-1} \{\mathbf{I} - A\}^{-1} A + B^{-1}\} \\ &= \{\mathbf{I} - A\} B B^{-1} \{\mathbf{I} - A\}^{-1} A + \{\mathbf{I} - A\} B B^{-1} \\ &= \mathbf{I} A + \{\mathbf{I} - A\} \mathbf{I} \\ &= A + \mathbf{I} - A = \mathbf{I}. \end{aligned}$$

Por lo tanto, queda demostrado que (B.39) es el inverso de (B.38).

Cuadro B.1:

-3.6731	-0.9990	-1.2308	-1.2308	-0.4150	-2.7568	-0.6774	-2.4204
-0.4745	-2.1691	-0.6527	-1.9684	-1.7272	-0.6774	-2.7568	-0.0290
-0.4345	-1.3218	-2.7568	-2.9800	-1.6094	-0.9005	-2.7568	-0.1468
-0.7287	-0.1616	-1.3218	-1.4759	-1.7272	-1.4759	-2.0637	-1.2308
-1.1082	-1.0341	-1.8814	-1.7272	-0.4345	-1.2752	-1.8013	-0.8109
-0.9005	-0.3058	-1.3705	-0.1766	-0.1766	-0.0488	-0.6527	-0.1322
-0.2554	-2.1691	-2.0637	-1.6582	-2.7568	-1.8013	-1.4759	-0.9323
-2.3514	-3.6731	-2.4204	-1.2528	-0.7554	-1.3218	-1.0341	-1.2308
-0.5596	-2.1691	-0.9404	-1.5937	-0.2719	-1.6582	-0.9323	
-2.4204	-1.4218	-4.3663	-2.7568	-0.8399	-0.7287	-0.6527	
-1.3218	-1.3218	-1.0341	-1.1882	-1.2752	-0.9005	-1.4759	
-1.8814	-0.7828	-1.2308	-1.3218	-0.0758	-0.5596	-1.7272	
-0.0225	-0.8698	-2.9800	-1.4759	-1.4759	-0.8109	-3.2677	
-0.6527	-1.1082	-1.8013	-0.9005	-0.4951	-2.4204	-4.3663	
-2.0637	-3.6731	-1.7272	-1.5937	-2.2868	-1.3705	-0.4745	
-2.0637	-1.5937	-2.7568	-1.8814	-2.5745	-1.8013	-0.7554	
-0.1322	-0.9005	-0.8109	-0.2554	-2.7568	-0.7828	-0.5596	
-0.8399	-2.1691	-2.0637	-1.3705	-0.2554	-0.7554	-0.1036	
-2.1691	-1.5331	-2.9800	-0.0622	-0.6286	-0.1178	-0.4150	
-1.7272	-1.3218	-0.0225	-2.1691	-0.8698	-0.9005	-0.4345	
-1.2308	-1.1082	-0.0355	-1.3218	-2.4204	-1.8814	-0.2074	
-1.4218	-0.3589	-0.8698	-1.3705	-2.5745	-0.4543	-2.5745	
-1.0341	-2.7568	-0.9651	-0.3589	-0.4745	-0.0896	-1.3705	
-0.0896	-2.2868	-0.2719	-1.3705	-1.3705	-2.0637	-0.7554	
-0.4951	-1.7095	-1.4218	-0.2231	-0.4745	-0.6051	-1.8814	
-3.6731	-0.4543	-0.5821	-1.0341	-0.3409	-1.2752	-1.3218	
-1.1779	-2.7568	-1.0341	-0.7828	-3.6731	-1.2752	-1.1882	
-1.0704	-0.3409	-1.2752	-0.4951	-2.4204	-2.7568	-1.1082	
-1.6582	-2.4204	-1.5937	-2.9800	-0.0488	-1.3218	-0.6286	
-1.1082	-0.3960	-1.7272	-1.4218	-0.0622	-2.0637	-1.6582	

Tiempos de fallo del componente r_1 bajo la transformación $\log(\frac{x}{\sigma})$

Cuadro B.2:

-0.5376	-1.7272	-1.8814	-0.4745	-3.2677	-0.0758	-4.3663	-0.6774
-0.6051	-2.1691	-0.6774	-0.3960	-2.1150	-1.6582	-1.9684	-0.9990
-0.1178	-1.1082	-1.8013	-0.2887	-0.9005	-1.3218	-1.9684	-3.2677
-2.9800	-1.7272	-1.1474	-0.7027	-2.0637	-0.8109	-1.8013	-1.1082
-1.1474	-0.8399	-0.4543	-3.2677	-0.4745	-0.9651	-3.6731	-0.8399
-1.4759	-0.9005	-1.5331	-2.7568	-1.5331	-2.7568	-2.9800	-1.7272
-2.2868	-1.9684	-2.0637	-1.7272	-1.0341	-1.2308	-2.5745	-1.0704
-1.8013	-1.2308	-0.0225	-1.7095	-0.7287	-1.7272	-0.8399	-0.2887
-1.6582	-0.2074	-1.2308	-0.5596	-2.1691	-0.7287	-1.8013	-1.0704
-0.7287	-2.1691	-0.7828	-0.3589	-1.4218	-3.2677	-0.9005	
-1.4218	-1.6582	-0.2719	-4.3663	-1.4759	-0.3232	-2.4204	
-1.5331	-2.4204	-2.2868	-1.3218	-1.5937	-0.5376	-0.9323	
-0.9404	-1.3218	-4.3663	-1.9684	-3.6731	-2.5745	-1.2752	
-1.3218	-1.7095	-1.8013	-1.8814	-1.3705	-1.2752	-1.8013	

Tiempos de fallo del componente r_2 bajo la transformación $\log(\frac{x}{\bar{c}})$

Apéndice C

Primero hagamos una observación que nos será muy útil para la simulación de una muestra

Observación 4

Si $Y \sim Weibull(a, b)$ entonces

$$F_Y(y) = 1 - e^{-\frac{y^b}{a}} \quad (\text{C.1})$$

o también se puede ver como

$$F_Y(y) = 1 - e^{-\frac{y^\alpha}{\beta^\alpha}} \quad \text{donde } b = \alpha \text{ y } a = \beta^\alpha \quad (\text{C.2})$$

De (C.1) tenemos que si $u \in (0, 1)$

$$u = 1 - e^{-\frac{y^b}{a}}$$

entonces al despejar u tenemos que

$$Y = (-a \log(1 - u))^{\frac{1}{b}}.$$

Entonces de la transformación

$$X = \log\left(\frac{Y}{C}\right) \quad (\text{C.3})$$

al despejar Y

$$Y = C e^X \quad \text{y} \quad \frac{dy}{dx} = C e^X$$

donde la media y la varianza son

$$\mu = \log\left(\frac{\beta}{C}\right) \quad \text{y} \quad \sigma = \frac{1}{\alpha}.$$

De (C.2) sabemos que $b = \alpha$ y que $a^{\frac{1}{\alpha}} = \beta$, entonces al sustituir esto en las últimas ecuaciones y al despejar a y b tenemos que

$$b = \frac{1}{\sigma} \quad \text{y que} \quad a = C^{\frac{1}{\sigma}} e^{\frac{\mu}{\sigma}}.$$

Entonces dado C , σ y μ podemos producir una muestra proveniente del modelo (C.2), y después bajo la transformación dada en (C.3) podemos producir una muestra de la función de distribución $F_X(x) = 1 - e^{-e^{\frac{x-\mu}{\sigma}}}$.

A continuación se presenta dos códigos hechos en el lenguaje R, el primero es una función llamada *simula* que se empleó para simular una mezcla de dos poblaciones con distribución Weibull, donde Te es el tiempo de censura y N es el tamaño de la muestra; mientras que en el segundo programa se hizo una función llamada *sampler* que simula una muestra para ser usada por el algoritmo EM.

```
simula <- function(N,mu1,mu2,sig1,sig2,p1){
p2 <- 1-p1

Te <- 630

a1 <- exp(mu1/sig1)*Te ^ {1/sig1}
b1 <- 1/sig1

a2 <- exp(mu2/sig2)*Te ^ {1/sig2}
b2 <- 1/sig2

u <- runif(N)
uaux <- runif(N)
```

```

indexu <- {1:length(u)}

ind1 <- indexu[ u < p1 ]
u1 <- uaux[ind1]

ind2 <- indexu[ u >= p1 ]
u2 <- uaux[ind2]

t1 <- -a1*log(1-u1)
t1 <- t1^(1/b1)
len1 <- length(t1)

t2 <- -a2*log(1-u2)
t2 <- t2^(1/b2)
len2 <- length(t2)

t1 <- t1[ t1 <= Te]
len1 <- length(t1)

t2 <- t2[ t2 <= Te]
len2 <- length(t2)

t1 <- log(t1/Te)
t2 <- log(t2/Te)

tsamples <- list(pop1=t1, pop2=t2, leng1=len1, leng2=len2)
tsamples
}

sampler <- function(k){
n <- 369

Tsigma1 <- 0.7208386
Tsigma2 <- 0.8424424

Ttheta1 <- -1.1374

```

```

Ttheta2 <- -0.6353

Tp <- 0.33
set.seed(k)
tsample <- simula(n,Ttheta1,Ttheta2,Tsigma1,Tsigma2,Tp)

y1 <- tsample$pop1
y2 <- tsample$pop2

r1 <- tsample$leng1
r2 <- tsample$leng2

write(y1,file="c:/tesis/muestras para el cap 3/y1.80.txt",ncol=1)
write(y2,file="c:/tesis/muestras para el cap 3/y2.80.txt",ncol=1)

vect <- c(k,r1,r2,r1+r2,n,n-(r1+r2))
paste(c("k=", "r1=", "r2=", "r=", "N=", "N-r="),vect)
}

```

Los siguientes tres programas fueron hechos en *Mathematica*; el primero es una función llamada `Mf[]`, este programa necesita que se le de como entrada el vector de parámetros de la distribución posterior $\text{THETA}^{\{\Xi\}}$, y regresa $\text{THETA}^{\{\Xi+1\}}$ bajo el algoritmo EM.

```

Mf[mu1_, mu2_, lgsig1_, lgsig2_, pi1_, pi2_, xm1_, xm2_, nm_, lgsig01_, lgsig02_] :=
Module[{r1, r2, r, mu, phi, mu1, mu2, phi1, phi2, pi1, pi2, fi1, fi2},
r1 = Length[xm1]; r2 = Length[xm2]; r = r1 + r2;
summn1[phi_] := Sum[Exp[xm1 * Exp[-phi]][[i]], {i, 1, r1};
summn2[phi_] := Sum[Exp[xm2 * Exp[-phi]][[i]], {i, 1, r2};
summl1[mu_] := Sum[(xm1[[i]] - mu), {i, 1, r1};
summl2[mu_] := Sum[(xm2[[i]] - mu), {i, 1, r2};
summk1[phi_, mu_] := Sum[(xm1[[i]] - mu) * Exp[xm1 * Exp[-phi]][[i]], {i, 1, r1};
summk2[phi_, mu_] := Sum[(xm2[[i]] - mu) * Exp[xm2 * Exp[-phi]][[i]], {i, 1, r2};
alpha2 = Exp[-mu2 * Exp[-lgsig2]];
alpha1 = Exp[-mu1 * Exp[-lgsig1]];
FFF1 = Exp[-alpha1];
FFF2 = Exp[-alpha2];

```

$\Xi_{m12} = \pi_1 * FFF1 + \pi_2 * FFF2;$
 $um1 = -r1 - (nm - r) * (\pi_1 * FFF1 \Xi_{m12});$
 $um2 = -r2 - (nm - r) * (\pi_2 * FFF2 \Xi_{m12});$
 $uum1 = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{w, \alpha1, \text{Infinity}\}];$
 $uum2 = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{w, \alpha2, \text{Infinity}\}];$
 $vvm1 = \text{Integrate}[\text{Exp}[-w], \{w, \alpha1, \text{Infinity}\}];$
 $vvm2 = \text{Integrate}[\text{Exp}[-w], \{w, \alpha2, \text{Infinity}\}];$
 $mm2 = \text{Exp}[\text{lgsig2}] * uum2 + mu2 * vvm2;$
 $mm1 = \text{Exp}[\text{lgsig1}] * uum1 + mu1 * vvm1;$
 $fm1[t, \varphi] := t^{\frac{e^{\text{lgsig1}}}{e^\varphi}} * \text{Exp}[-t];$
 $fm2[t, \varphi] := t^{\frac{e^{\text{lgsig2}}}{e^\varphi}} * \text{Exp}[-t];$
 $gm1[\varphi] := \text{Integrate}[fm1[t, \varphi], \{t, \alpha1, \text{Infinity}\}];$
 $gm2[\varphi] := \text{Integrate}[fm2[t, \varphi], \{t, \alpha2, \text{Infinity}\}];$
 $\nu m1[\varphi] := \text{Exp}[\mu1 * \text{Exp}[-\varphi]] * gm1[\varphi];$
 $\nu m2[\varphi] := \text{Exp}[\mu2 * \text{Exp}[-\varphi]] * gm2[\varphi];$
 $hm1[t, \varphi] := t^{\frac{e^{\text{lgsig1}}}{e^\varphi}} * \text{Log}[t] * \text{Exp}[-t];$
 $hm2[t, \varphi] := t^{\frac{e^{\text{lgsig2}}}{e^\varphi}} * \text{Log}[t] * \text{Exp}[-t];$
 $\omega m1[\varphi] := \text{NIntegrate}[hm1[t, \varphi], \{t, \alpha1, \text{Infinity}\}];$
 $\omega m2[\varphi] := \text{NIntegrate}[hm2[t, \varphi], \{t, \alpha2, \text{Infinity}\}];$
 $\Phi m1[\varphi] := \mu1 * \nu m1[\varphi] + (\text{Exp}[\text{lgsig1}] * \text{Exp}[\mu1 * \text{Exp}[-\varphi]] * \omega m1[\varphi]);$
 $\Phi m2[\varphi] := \mu2 * \nu m2[\varphi] + (\text{Exp}[\text{lgsig2}] * \text{Exp}[\mu2 * \text{Exp}[-\varphi]] * \omega m2[\varphi]);$
 $fi1 = \frac{\text{lgsig01} + \text{lgsig1}}{2};$
 $\mu1 = -\text{Exp}[fi1] * \text{Log}\left[\frac{-um1}{\text{summn1}[fi1] + (nm - r) * \left(\frac{\pi_1 * \nu m1[fi1]}{\Xi_{m12}}\right)}\right];$
 $Bma1[\varphi] := um1 * \text{Exp}[\varphi] - \text{summl1}[\mu1] + \text{Exp}[-\mu1 * \text{Exp}[-\varphi]] * \text{summk1}[\varphi, \mu1] - \frac{(nm - r) * \pi_1}{\Xi_{m12}}$
 $* \text{Exp}[-\mu1 * \text{Exp}[-\varphi]] * \left(\left(\nu m1[\varphi] * \mu1\right) - \Phi m1[\varphi]\right) - (nm - r) * \left(\frac{\pi_1 * mm1 - \mu1 * \pi_1 * FFF1}{\Xi_{m12}}\right);$
 $\varphi1 = \varphi1 /. \text{FindRoot}[Bma1[\varphi1] == 0, \{\varphi1, \text{lgsig01}, \text{lgsig1}\}];$
 $fi2 = \frac{\text{lgsig02} + \text{lgsig2}}{2};$
 $\mu2 = -\text{Exp}[fi2] * \text{Log}\left[\frac{-um2}{\text{summn1}[fi2] + (nm - r) * \left(\frac{\pi_2 * \nu m2[fi2]}{\Xi_{m12}}\right)}\right];$
 $Bma2[\varphi] := um2 * \text{Exp}[\varphi] - \text{summl2}[\mu2] + \text{Exp}[-\mu2 * \text{Exp}[-\varphi]] * \text{summk2}[\varphi, \mu2] - \frac{(nm - r) * \pi_2}{\Xi_{m12}}$
 $* \text{Exp}[-\mu2 * \text{Exp}[-\varphi]] * \left(\left(\nu m2[\varphi] * \mu2\right) - \Phi m2[\varphi]\right) - (nm - r) * \left(\frac{\pi_2 * mm2 - \mu2 * \pi_2 * FFF2}{\Xi_{m12}}\right);$
 $\varphi2 = \varphi2 /. \text{FindRoot}[Bma1[\varphi2] == 0, \{\varphi2, \text{lgsig02}, \text{lgsig2}\}];$
 $\pi1 = \text{frac} - um1 nm;$

$\pi_2 = \text{frac-um2nm}; \{\mu_1, \mu_2, \varphi_1, \varphi_2, \pi_1, \pi_2\}$

El segundo programa es la función llamada DG[], este programa evalúa la matriz Hessiana $\frac{\partial^2 Q(\Xi, \Xi^n)}{\partial \Xi^2}$ evaluada en $\Xi = \Xi^n$ y la guarda en una variable llamada Q.

```

DG[mu1_, mu2_, lgsig1_, lgsig2_, pi1_, pi2_, xg1_, xg2_, ng_] := Module[{r1, r2, r, mu1, mu2, phi1,
phi2, pi1, pi2, a1, a2, FF1, FF2, ug1, ug2, uug1, uug2, vvg1, vvg2, mg1, mg2, Qg = Array[q,
{6, 6}]}, r1 = Length[xg1]; r2 = Length[xg2];
r = r1 + r2;
sumgl1[mu1_] := Sum[xg1[[i]] - mu1, {i, 1, r1};
sumgl2[mu2_] := Sum[xg2[[i]] - mu2, {i, 1, r2};
sumgk1[mu1_, phi1_] := Sum[xg1[[i]] - mu1 * Exp[(xg1[[i]] - mu1) * Exp[-phi1]], {i, 1, r1};
sumgk2[mu2_, phi2_] := Sum[xg2[[i]] - mu2 * Exp[(xg2[[i]] - mu2) * Exp[-phi2]], {i, 1, r2};
sumgo1[mu1_, phi1_] := Sum[xg1[[i]] - mu1)^2 * Exp[(xg1[[i]] - mu1) * Exp[-phi1]], {i, 1, r1};
sumgo2[mu2_, phi2_] := Sum[xg2[[i]] - mu2)^2 * Exp[(xg2[[i]] - mu2) * Exp[-phi2]], {i, 1, r2};
sumgn1[mu1_, phi1_] := Sum[Exp[(xg1[[i]] - mu1) * Exp[-phi1]], {i, 1, r1};
sumgn2[mu2_, phi2_] := Sum[Exp[(xg2[[i]] - mu2) * Exp[-phi2]], {i, 1, r2};
a1 = Exp[-mu1 * Exp[-lgsig1]];
a2 = Exp[-mu2 * Exp[-lgsig2]];
FF1 = Exp[-a1];
FF2 = Exp[-a2];
Xi_g12 = pi1*FF1 + pi2*FF2;
ug1 = -r1 - (ng - r) * (pi1*FF1 / Xi_g12);
ug2 = -r2 - (ng - r) * (pi2*FF2 / Xi_g12);
uug1 = Integrate[Log[w]*Exp[-w], {w, a1, Infinity}];
uug2 = Integrate[Log[w]*Exp[-w], {w, a2, Infinity}];
vvg1 = Integrate[Exp[-w], {w, a1, Infinity}];
vvg2 = Integrate[Exp[-w], {w, a2, Infinity}];
mg2 = Exp[lgsig2]*uug2 + mu2*vvg2;
mg1 = Exp[lgsig1]*uug1 + mu1*vvg1;
fg1[t_, phi1_] := t * (Exp[lgsig1] / Exp[phi1]) * Exp[-t];
fg2[t_, phi2_] := t * (Exp[lgsig2] / Exp[phi2]) * Exp[-t];
gg1[phi1_] := NIntegrate[fg1[t, phi1], {t, a1, Infinity}];

```

```

gg2[φ2_] := NIntegrate[fg2[t, φ2], {t, a2, Infinity}];
νg1[φ1_] := Exp[mu1 * Exp[-φ1]] * gg1[φ1];
νg2[φ2_] := Exp[mu2 * Exp[-φ2]] * gg2[φ2];
hg1[t_, φ1_] := t elgsig1/eφ1 * Exp[-t] * Log[t];
hg2[t_, φ2_] := t elgsig2/eφ2 * Exp[-t] * Log[t];
Λg1[φ1_] := NIntegrate[hg1[t, φ1], {t, a1, Infinity}];
Λg2[φ2_] := NIntegrate[hg2[t, φ2], {t, a2, Infinity}];
γg1[t_, φ1_] := t elgsig1/eφ1 * Exp[-t] * (Log[t])2;
γg2[t_, φ2_] := t elgsig2/eφ2 * Exp[-t] * (Log[t])2;
Γg1[φ1_] := NIntegrate[γg1[t, φ1], {t, a1, Infinity}];
Γg2[φ2_] := NIntegrate[γg2[t, φ2], {t, a2, Infinity}];
Dνg1[φ1_] := Exp[mu1 * Exp[-φ1]] * Exp[lgsig1 - φ1] * Λg1[φ1] + Exp[-φ1] * mu1 * νg1[φ1];
Dνg2[φ2_] := Exp[mu2 * Exp[-φ2]] * Exp[lgsig2 - φ2] * Λg2[φ2] + Exp[-φ2] * mu2 * νg2[φ2];
qg11[μ1_, φ1_] := Exp[-2 * φ1] * sumgn1[μ1, φ1] + (ng - r) * (π1*νg1[φ1]/Ξg12) * (Exp[-2 * φ1 - μ1 * Exp[-φ1]]);
qg12[μ1_, φ1_] := 0;
qg13[μ1_, φ1_] := -r1 * Exp[-φ1] + Exp[-2 * φ1] * sumgk1[μ1, φ1] + Exp[-φ1] * sumgn1[μ1, φ1] - (ng - r) * Exp[-φ1] * (π1*FF1/Ξg12) + (ng - r) * (π1*νg1[φ1]/Ξg12) * (1 - μ1 * Exp[-φ1]) * Exp[-φ1 - μ1 * Exp[-φ1]] + (ng - r) * (π1*Dνg1[φ1]/Ξg12) * Exp[-φ1 - μ1 * Exp[-φ1]];
qg14[μ1_, φ1_] := 0; qg15[μ1_, φ1_] := 0; qg16[μ1_, φ1_] := 0;
qg21[μ1_, φ1_] := 0;
qg22[μ2_, φ2_] := Exp[-2 * φ2] * sumgn2[μ2, φ2] + (ng - r) * (π2*νg2[φ2]/Ξg12) * Exp[-2 * φ2 - μ2 * Exp[-φ2]];
qg23[μ2_, φ2_] := 0;
qg24[μ2_, φ2_] := -r2 * Exp[-φ2] + Exp[-2 * φ2] * sumgk2[μ2, φ2] + Exp[-φ2] * sumgn2[μ2, φ2] - (ng - r) * Exp[-φ2] * (π2*FF2/Ξg12) + (ng - r) * (π2 * νg2[φ2]/Ξg12) * (1 - μ2 * Exp[-φ2]) * Exp[-φ2 - μ2 * Exp[-φ2]] + (ng - r) * (π2*Dνg2[φ2]/Ξg12) * Exp[-φ2 - μ2 * Exp[-φ2]];
qg25[μ2_, φ2_] := 0; qg26[μ2_, φ2_] := 0;
qg31[μ1_, φ1_] := -r1 * Exp[-φ1] + Exp[-2 * φ1] * sumgk1[μ1, φ1] + Exp[-φ1] * sumgn1[μ1, φ1] - (ng - r) * Exp[-φ1] * (π1*FF1/Ξg12) + (ng - r) * (π1*νg1[φ1]/Ξg12) * (1 - μ1 * Exp[-φ1]) * Exp[-φ1 - μ1 * Exp[-φ1]] + (ng - r) * (π1*Dνg1[φ1]/Ξg12) * Exp[-φ1 - μ1 * Exp[-φ1]];
qg32[μ1_, φ1_] := 0;

```

$qg33[\mu1_ , \varphi1_] := -1 * \text{sumgl1}[\mu1] * \text{Exp}[-\varphi1] + \text{Exp}[-\varphi1] * \text{sumgk1}[\mu1, \varphi1] + \text{Exp}[-2 * \varphi1] * \text{sumgo1}[\mu1, \varphi1] - (ng - r) * \left(\frac{\text{pi1} * (\text{mg1} - \mu1 * \text{FF1})}{\Xi g12}\right) * \text{Exp}[-\varphi1] + (ng - r) * \left(\frac{\text{pi1}}{\Xi g12}\right) * ((\mu1 - \mu1) * (\nu g1[\varphi1] * (\mu1 * \text{Exp}[-\varphi1] - 1) * \text{Exp}[-\varphi1 - \mu1 * \text{Exp}[-\varphi1]]) - \text{Exp}[-\varphi1 - \mu1 * \text{Exp}[-\varphi1]]) * \text{D}\nu g1[\varphi1] + \text{Exp}[\text{Exp}[-\varphi1] * (\mu1 - \mu1) + (\text{lgsig1} - \varphi1)] * \text{Exp}[\text{lgsig1} - \varphi1] * \Gamma g1[\varphi1] + (1 + (\mu1 - \mu1) * \text{Exp}[-\varphi1]) * \text{Exp}[\text{Exp}[-\varphi1] * (\mu1 - \mu1) + (\text{lgsig1} - \varphi1)] * \Lambda g1[\varphi1];$

$qg34[\mu1_ , \varphi1_] := 0; \quad qg35[\mu1_ , \varphi1_] := 0; \quad qg36[\mu1_ , \varphi1_] := 0;$

$qg41[\mu2_ , \varphi2_] := 0;$

$qg42[\mu2_ , \varphi2_] := -r2 * \text{Exp}[-\varphi2] + \text{Exp}[-2 * \varphi2] * \text{sumgk2}[\mu2, \varphi2] + \text{Exp}[-\varphi2] * \text{sumgn2}[\mu2, \varphi2] - (ng - r) * \text{Exp}[-\varphi2] * \left(\frac{\text{pi2} * \text{FF2}}{\Xi g12}\right) + (ng - r) * \left(\frac{\text{pi2} * \nu g2[\varphi2]}{\Xi g12}\right) * (1 - \mu2 * \text{Exp}[-\varphi2]) * \text{Exp}[-\varphi2 - \mu2 * \text{Exp}[-\varphi2]] + (ng - r) * \left(\frac{\text{pi2} * \text{D}\nu g2[\varphi2]}{\Xi g12}\right) * \text{Exp}[-\varphi2 - \mu2 * \text{Exp}[-\varphi2]];$

$qg43[\mu2_ , \varphi2_] := 0;$

$qg44[\mu2_ , \varphi2_] := -1 * \text{sumgl2}[\mu2] * \text{Exp}[-\varphi2] + \text{Exp}[-\varphi2] * \text{sumgk2}[\mu2, \varphi2] + \text{Exp}[-2 * \varphi2] * \text{sumgo2}[\mu2, \varphi2] - (ng - r) * \left(\frac{\text{pi2} * (\text{mg2} - \mu2 * \text{FF2})}{\Xi g12}\right) * \text{Exp}[-\varphi2] + (ng - r) * \left(\frac{\text{pi2}}{\Xi g12}\right) * ((\mu2 - \mu2) * (\nu g2[\varphi2] * (\mu2 * \text{Exp}[-\varphi2] - 1) * \text{Exp}[-\varphi2 - \mu2 * \text{Exp}[-\varphi2]]) - \text{Exp}[-\varphi2 - \mu2 * \text{Exp}[-\varphi2]]) * \text{D}\nu g2[\varphi2] + \text{Exp}[\text{Exp}[-\varphi2] * (\mu2 - \mu2) + (\text{lgsig2} - \varphi2)] * \text{Exp}[\text{lgsig2} - \varphi2] * \Gamma g2[\varphi2] + (1 + (\mu2 - \mu2) * \text{Exp}[-\varphi2]) * \text{Exp}[\text{Exp}[-\varphi2] * (\mu2 - \mu2) + (\text{lgsig2} - \varphi2)] * \Lambda g2[\varphi2];$

$qg45[\mu2_ , \varphi2_] := 0; \quad qg46[\mu2_ , \varphi2_] := 0; \quad qg51[\pi1_] := 0;$

$qg52[\pi1_] := 0; \quad qg53[\pi1_] := 0; \quad qg54[\pi1_] := 0;$

$qg55[\pi1_] := (-ug1 * \left(\frac{1}{(\pi1)^2}\right));$

$qg56[\pi1_] := 0; \quad qg61[\pi2_] := 0; \quad qg62[\pi2_] := 0;$

$qg63[\pi2_] := 0; \quad qg64[\pi2_] := 0; \quad qg65[\pi2_] := 0;$

$qg66[\pi2_] := -ug2 * \left(\frac{1}{(\pi2)^2}\right);$

$q[1, 1] = qg11[\mu1, \text{lgsig1}]; \quad q[1, 2] = qg12[\mu1, \text{lgsig1}];$

$q[1, 3] = qg13[\mu1, \text{lgsig1}]; \quad q[1, 4] = qg14[\mu1, \text{lgsig1}];$

$q[1, 5] = qg15[\mu1, \text{lgsig1}]; \quad q[1, 6] = qg16[\mu1, \text{lgsig1}];$

$q[2, 1] = qg21[\mu2, \text{lgsig2}]; \quad q[2, 2] = qg22[\mu2, \text{lgsig2}];$

$q[2, 3] = qg23[\mu2, \text{lgsig2}]; \quad q[2, 4] = qg24[\mu2, \text{lgsig2}];$

$q[2, 5] = qg25[\mu2, \text{lgsig2}]; \quad q[2, 6] = qg26[\mu2, \text{lgsig2}];$

$q[3, 1] = qg31[\mu1, \text{lgsig1}]; \quad q[3, 2] = qg32[\mu1, \text{lgsig1}];$

$q[3, 3] = qg33[\mu1, \text{lgsig1}]; \quad q[3, 4] = qg34[\mu1, \text{lgsig1}];$

$q[3, 5] = qg35[\mu1, \text{lgsig1}]; \quad q[3, 6] = qg36[\mu1, \text{lgsig1}];$

$q[4, 1] = qg41[\mu2, \text{lgsig2}]; \quad q[4, 2] = qg42[\mu2, \text{lgsig2}];$

$q[4, 3] = qg43[\mu2, \text{lgsig2}]; \quad q[4, 4] = qg44[\mu2, \text{lgsig2}];$

$q[4, 5] = qg45[\mu2, \text{lgsig2}]; \quad q[4, 6] = qg46[\mu2, \text{lgsig2}];$

```

q[5, 1] = qg51[pi1];   q[5, 2] = qg52[pi1];
q[5, 3] = qg53[pi1];   q[5, 4] = qg54[pi1];
q[5, 5] = qg55[pi1];   q[5, 6] = qg56[pi1];
q[6, 1] = qg61[pi2];   q[6, 2] = qg62[pi2];
q[6, 3] = qg63[pi2];   q[6, 4] = qg64[pi2];
q[6, 5] = qg65[pi2];   q[6, 6] = qg66[pi2]; Qg ]

```

El tercer programa lee los datos simulados por los programas hechos en R, y manda a llamar las funciones $Mf[]$ treinta veces para encontrar un estimador puntual de Ξ^n , y después manda a llamar a la función $DG[]$ para construir los intervalos de confianza.

```

x1 = ReadList["C:/tesis/cap tres/k2/y1_80.txt"];
x2 = ReadList["C:/tesis/cap tres/k2/y2_80.txt"];
tamr1 = Length[x1];
tamr2 = Length[x2];
xbarr1 =  $\sum_{i=1}^{tamr1} \frac{x1[[i]]}{tamr1}$ ;
xbarr2 =  $\sum_{i=1}^{tamr2} \frac{x2[[i]]}{tamr2}$ ;
fi1 = fi1 /. FindRoot[N[  $\frac{1}{tamr1} (\sum_{i=1}^{tamr1} x1[[i]] * Exp[x1[[i]] - Exp[fi1]] * Log[\sum_{i=1}^{tamr1} \frac{Exp[x1[[i]] * Exp[-fi1]]}{tamr1}] Exp[-fi1] ) - Exp[-fi1] == xbarr1$  , { fi1, -8, 8 }];
φ1 = fi1 + 0.5;
φ01 = fi1 - 0.5;
μ1 = Exp[ fi1 ] * Log[  $\sum_{i=1}^{tamr1} \frac{Exp[x1[[i]] * Exp[-fi1]]}{tamr1}$  ];
fi2 = fi2 /. FindRoot[N[  $\frac{1}{tamr2} (\sum_{i=1}^{tamr2} x2[[i]] * Exp[x2[[i]] - Exp[fi2]] * Log[\sum_{i=1}^{tamr2} \frac{Exp[x2[[i]] * Exp[-fi2]]}{tamr2}] Exp[-fi2] ) - Exp[-fi2] == xbarr2$  , { fi2, -8, 8 }];
φ2 = fi2 + 0.5;
φ02 = fi2 - 0.5;
μ2 = Exp[ fi2 ] * Log[  $\sum_{i=1}^{tamr2} \frac{Exp[x2[[i]] * Exp[-fi2]]}{tamr2}$  ];
π1 =  $\frac{tamr1}{tamr1+tamr2}$ ;
π2 = 1 - π1;
n=369;
Ξstar=Array[matΞ,30,6];
Do[Ξstar[[kk]]=Mf[μ1, μ2, φ1, φ2, π1, π2, x1, x2, n, φ01, φ02]; φ01=φ1; φ02=φ2; μ1 = Ξstar[[kk, 1]]; μ2=Ξstar[[kk, 2]]; φ1=Ξstar[[kk, 3]]; φ2=Ξstar[[kk, 4]]; π1=Ξstar[[kk, 5]]; π2=Ξstar[[kk, 6]]; Print[Ξstar[[kk]]; , {kk, 30} ]

```

```

Ξstar[[30]]
μ1star=Ξstar[[30,1]];
μ2star=Ξstar[[30,2]];
φ1star=Ξstar[[30,3]];
φ2star=Ξstar[[30,4]];
π1star=Ξstar[[30,5]];
π2star=Ξstar[[30,6]];
Q=Array[q,6,6];
Q = DG[μ1star, μ2star, φ1star, φ2star, π1star, π2star, x1, x2, n];
ERRE=Array[matr,6,6];
θd=Array[θ,6];
θtilda=Array[θt,6];
Do[Do[θd = Ξstar[[30]]; θd[[ii]] = Ξstar[[tt, ii]];
θtilda = Mf[θd[[1]], θd[[2]], θd[[3]], θd[[4]], θd[[5]], θd[[6]], x1, x2, n, Ξstar[[21, 3]],
Ξstar[[21, 4]]];
Do[ matr[[ii, jj]] =  $\frac{\theta_{tilda}[[jj]] - \Xi_{star}[[30, jj]]}{\Xi_{star}[[tt, ii]] - \Xi_{star}[[30, ii]]}$  ; , {jj, 6}];, {ii, 6}]; Print[tt];, {tt, 12, 19}]
MatrixForm[ERRE]
II = DiagonalMatrix[{1,1,1,1,1,1}];
VMAT=Inverse[Q]+Inverse[Q].ERRE.Inverse[II-ERRE];
MatrixForm[VMAT]
LIlu1 =Ξstar[[30,1]]-1.96*Sqrt[VMAT[[1,1]]]
LSmu1 =Ξstar[[30,1]]+1.96*Sqrt[VMAT[[1,1]]]
LIlu2 =Ξstar[[30,2]]-1.96*Sqrt[VMAT[[2,2]]]
LSmu2 =Ξstar[[30,2]]+1.96*Sqrt[VMAT[[2,2]]]
LIphi1 =Ξstar[[30,3]]-1.96*Sqrt[VMAT[[3,3]]]
LSphi1 =Ξstar[[30,3]]+1.96*Sqrt[VMAT[[3,3]]]
LIphi2 =Ξstar[[30,4]]-1.96*Sqrt[VMAT[[4,4]]]
LSphi2 =Ξstar[[30,4]]+1.96*Sqrt[VMAT[[4,4]]]
LIp1 =Ξstar[[30,5]]-1.96*Sqrt[VMAT[[5,5]]]
LSp1 =Ξstar[[30,5]]+1.96*Sqrt[VMAT[[5,5]]]
LIp2 =Ξstar[[30,6]]-1.96*Sqrt[VMAT[[6,6]]]
LSp2 =Ξstar[[30,6]]+1.96*Sqrt[VMAT[[6,6]]]

```

Ahora se muestran dos códigos hechos en el lenguaje R el primero (como en el caso $k = 2$) se uso para simular una mezcla de tres poblaciones con

distribución Weibull y el otro programa simula una muestra para ser usada por el algoritmo EM.

```
simula <- function(N, mu1, mu2, mu3, sig1, sig2, sig3, p1, p2) {  
  p3 <- 1-p1-p2  
  Te <- 600  
  
  a1 <- exp(mu1/sig1)*Te ^ {1/sig1}  
  b1 <- 1/sig1  
  
  a2 <- exp(mu2/sig2)*Te ^ {1/sig2}  
  b2 <- 1/sig2  
  
  a3 <- exp(mu3/sig3)*Te ^ {1/sig3}  
  b3 <- 1/sig3  
  
  u <- runif(N)  
  uaux <- runif(N)  
  indexu <- {1:length(u)}  
  
  ind1 <- indexu[ u < p1 ]  
  u1 <- uaux[ind1]  
  
  ind2 <- indexu[ u >= p1 & u < p2 ]  
  u2 <- uaux[ind2]  
  
  ind3 <- indexu[ u >= p2 ]  
  u3 <- uaux[ind3]  
  
  t1 <- -a1*log(1-u1)  
  t1 <- t1 ^ {1/b1}  
  len1 <- length(t1)  
  
  t2 <- -a2*log(1-u2)  
  t2 <- t2 ^ {1/b2}  
  len2 <- length(t2)  
  
  t3 <- -a3*log(1-u3)
```

```
t3 <- t3 ^ {1/b3}
len3 <- length(t3)

t1 <- t1[ t1 <= Te]
len1 <- length(t1)

t2 <- t2[ t2 <= Te]
len2 <- length(t2)

t3 <- t3[ t3 <= Te]
len3 <- length(t3)

t1 <- - log(t1/Te)
t2 <- - log(t2/Te)
t3 <- - log(t3/Te)

tsamples <- list(pop1 = t1, pop2 = t2, pop3 = t3, leng1 = len1, leng2 = len2, leng3 =
len3)
tsamples
}
```

```
sampler <- function(k){
n <- 500
Tsigma1 <- 0.8214
Tsigma2 <- 0.6882
Tsigma3 <- 0.6755
Ttheta1 <- -0.8680
Ttheta2 <- -0.6495
Ttheta3 <- -0.7496
Tp1 <- 0.3229
Tp2 <- 0.2433 + 0.3229
set.seed(k)
```

```

tsample <- simula(n, Ttheta1, Ttheta2, Ttheta3, Tsigma1, Tsigma2, Tsigma3, Tp1, Tp2)
y1 <- tsample$pop1
y2 <- tsample$pop2
y3 <- tsample$pop3
r1 <- tsample$leng1
r2 <- tsample$leng2
r3 <- tsample$leng3
write(y1, file = "c:/tesis/k tres/y1.txt", ncol = 1)
write(y2, file = "c:/tesis/k tres/y2.txt", ncol = 1)
write(y3, file = "c:/tesis/k tres/y3.txt", ncol = 1)
vect <- c(k, r1, r2, r3, r1+r2+r3, n, n-(r1+r2+r3))
paste(c("k = ", "r1 = ", "r2 = ", "r3 = ", "r = ", "N = ", "N-r = "), vect)
}

```

Por último se muestra los tres programas hechos en *Mathematica*, el primero se uso para encontrar un estimador puntual para el vector de parámetros de la distribución posterior ($Mf3[]$), el segundo para evaluar el Hessiano $DG3[]$ y el último es el que lee la muestra simulada para encontrarle el estimador puntual después de treinta iteraciones y hacer los intervalos de confianza.

```

Mf3[mu1_, mu2_, mu3_, lgsig1_, lgsig2_, lgsig3_, pi1_, pi2_, pi3_, xm1_, xm2_, xm3_, nm_,
lgsig01_, lgsig02_, lgsig03_] := Module[{r1, r2, r3, r, mu, phi, mu1, mu2, mu3, phi1, phi2, phi3, pi1, pi2,
pi3, fi1, fi2, fi3 },
r1 = Length[xm1]; r2 = Length[xm2]; r3 = Length[xm3];
r = r1 + r2 + r3;
summn1[phi_] := Sum[Exp[xm1*Exp[-phi]] [ [i] ],
i=1, r1];
summn2[phi_] := Sum[Exp[xm2*Exp[-phi]] [ [i] ],
i=1, r2];

```

$$\text{summn3}[\varphi_-] := \sum_{i=1}^{r3} (\text{Exp}[\text{xm3} * \text{Exp}[-\varphi]] [[i]]);$$

$$\text{summl1}[\mu_-] := \sum_{i=1}^{r1} ((\text{xm1}[[i]] - \mu));$$

$$\text{summl2}[\mu_-] := \sum_{i=1}^{r2} ((\text{xm2}[[i]] - \mu));$$

$$\text{summl3}[\mu_-] := \sum_{i=1}^{r3} ((\text{xm3}[[i]] - \mu));$$

$$\text{summk1}[\varphi_-, \mu_-] := \sum_{i=1}^{r1} ((\text{xm1}[[i]] - \mu) * \text{Exp}[\text{xm1} * \text{Exp}[-\varphi]] [[i]]);$$

$$\text{summk2}[\varphi_-, \mu_-] := \sum_{i=1}^{r2} ((\text{xm2}[[i]] - \mu) * \text{Exp}[\text{xm2} * \text{Exp}[-\varphi]] [[i]]);$$

$$\text{summk3}[\varphi_-, \mu_-] := \sum_{i=1}^{r3} ((\text{xm3}[[i]] - \mu) * \text{Exp}[\text{xm3} * \text{Exp}[-\varphi]] [[i]]);$$

$$\alpha3 = \text{Exp}[-\text{mu3} * \text{Exp}[-\text{lgsig3}]] ;$$

$$\alpha2 = \text{Exp}[-\text{mu2} * \text{Exp}[-\text{lgsig2}]] ;$$

$$\alpha1 = \text{Exp}[-\text{mu1} * \text{Exp}[-\text{lgsig1}]] ;$$

$$\text{FFF1} = \text{Exp}[-\alpha1] ;$$

$$\text{FFF2} = \text{Exp}[-\alpha2] ;$$

$$\text{FFF3} = \text{Exp}[-\alpha3] ;$$

$$\Xi_{m123} = \text{pi1} * \text{FFF1} + \text{pi2} * \text{FFF2} + \text{pi3} * \text{FFF3};$$

$$\text{um1} = -r1 - (\text{nm} - r) * \left(\frac{\text{pi1} * \text{FFF1}}{\Xi_{m123}} \right) ;$$

$$\text{um2} = -r2 - (\text{nm} - r) * \left(\frac{\text{pi2} * \text{FFF2}}{\Xi_{m123}} \right) ;$$

$$\text{um3} = -r3 - (\text{nm} - r) * \left(\frac{\text{pi3} * \text{FFF3}}{\Xi_{m123}} \right) ;$$

$$\text{uum1} = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{ w, \alpha1, \text{Infinity} \}] ;$$

$$\text{uum2} = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{ w, \alpha2, \text{Infinity} \}] ;$$

$$\text{uum3} = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{ w, \alpha3, \text{Infinity} \}] ;$$

$$\text{vvm1} = \text{Integrate}[\text{Exp}[-w], \{ w, \alpha1, \text{Infinity} \}] ;$$

$$\text{vvm2} = \text{Integrate}[\text{Exp}[-w], \{ w, \alpha2, \text{Infinity} \}] ;$$

$$\text{vvm3} = \text{Integrate}[\text{Exp}[-w], \{ w, \alpha3, \text{Infinity} \}] ;$$

$$\text{mm3} = \text{Exp}[\text{lgsig3}] * \text{uum3} + \text{mu3} * \text{vvm3};$$

$mm2 = \text{Exp}[\text{lgsig}2]*uum2 + mu2*vvm2;$
 $mm1 = \text{Exp}[\text{lgsig}1]*uum1 + mu1*vvm1;$
 $fm1[t_ , \varphi_] := t^{\frac{e^{\text{lgsig}1}}{e^{\varphi}}} * \text{Exp}[-t];$
 $fm2[t_ , \varphi_] := t^{\frac{e^{\text{lgsig}2}}{e^{\varphi}}} * \text{Exp}[-t];$
 $fm3[t_ , \varphi_] := t^{\frac{e^{\text{lgsig}3}}{e^{\varphi}}} * \text{Exp}[-t];$
 $gm1[\varphi_] := \text{Integrate}[fm1[t, \varphi], \{t, \alpha1, \text{Infinity}\}];$
 $gm2[\varphi_] := \text{Integrate}[fm2[t, \varphi], \{t, \alpha2, \text{Infinity}\}];$
 $gm3[\varphi_] := \text{Integrate}[fm3[t, \varphi], \{t, \alpha3, \text{Infinity}\}];$
 $\nu m1[\varphi_] := \text{Exp}[mu1*\text{Exp}[-\varphi]]*gm1[\varphi];$
 $\nu m2[\varphi_] := \text{Exp}[mu2*\text{Exp}[-\varphi]]*gm2[\varphi];$
 $\nu m3[\varphi_] := \text{Exp}[mu3*\text{Exp}[-\varphi]]*gm3[\varphi];$
 $hm1[t_ , \varphi_] := t^{\frac{e^{\text{lgsig}1}}{e^{\varphi}}} * \text{Log}[t]*\text{Exp}[-t];$
 $hm2[t_ , \varphi_] := t^{\frac{e^{\text{lgsig}2}}{e^{\varphi}}} * \text{Log}[t]*\text{Exp}[-t];$
 $hm3[t_ , \varphi_] := t^{\frac{e^{\text{lgsig}3}}{e^{\varphi}}} * \text{Log}[t]*\text{Exp}[-t];$
 $\omega m1[\varphi_] := \text{NIntegrate}[hm1[t, \varphi], \{t, \alpha1, \text{Infinity}\}];$
 $\omega m2[\varphi_] := \text{NIntegrate}[hm2[t, \varphi], \{t, \alpha2, \text{Infinity}\}];$
 $\omega m3[\varphi_] := \text{NIntegrate}[hm3[t, \varphi], \{t, \alpha3, \text{Infinity}\}];$
 $\Phi m1[\varphi_] := mu1*\nu m1[\varphi] + (\text{Exp}[\text{lgsig}1]*\text{Exp}[mu1*\text{Exp}[-\varphi]]*\omega m1[\varphi]);$
 $\Phi m2[\varphi_] := mu2*\nu m2[\varphi] + (\text{Exp}[\text{lgsig}2]*\text{Exp}[mu2*\text{Exp}[-\varphi]]*\omega m2[\varphi]);$
 $\Phi m3[\varphi_] := mu3*\nu m3[\varphi] + (\text{Exp}[\text{lgsig}3]*\text{Exp}[mu3*\text{Exp}[-\varphi]]*\omega m3[\varphi]);$
 $fi1 = \frac{\text{lgsig}01 + \text{lgsig}1}{2};$
 $\mu1 = -\text{Exp}[fi1]* \text{Log}\left[\frac{-um1}{\text{summn}1[fi1] + (nm-r)*\left(\frac{\text{pi}1*\nu m1[fi1]}{\Xi m123}\right)}\right];$
 $Bma1[\varphi_] := um1 * \text{Exp}[\varphi] - \text{summl}1[\mu1] + \text{Exp}[-\mu1 * \text{Exp}[-\varphi]] * \text{summk}1[\varphi, \mu1] - \frac{(nm-r)*\text{pi}1}{\Xi m123}$
 $* \text{Exp}[-\mu1 * \text{Exp}[-\varphi]] * ((\nu m1[\varphi] * \mu1) - \Phi m1[\varphi]) - (nm - r) * \left(\frac{\text{pi}1*mm1 - \mu1*\text{pi}1*FFF1}{\Xi m123}\right);$
 $\varphi1 = \varphi1 /. \text{FindRoot}[Bma1[\varphi1] == 0, \{\varphi1, \text{lgsig}01, \text{lgsig}1\}];$
 $fi2 = \frac{\text{lgsig}02 + \text{lgsig}2}{2};$
 $\mu2 = -\text{Exp}[fi2]* \text{Log}\left[\frac{-um2}{\text{summn}1[fi2] + (nm-r)*\left(\frac{\text{pi}2*\nu m2[fi2]}{\Xi m123}\right)}\right];$
 $Bma2[\varphi_] := um2 * \text{Exp}[\varphi] - \text{summl}2[\mu2] + \text{Exp}[-\mu2 * \text{Exp}[-\varphi]] * \text{summk}2[\varphi, \mu2] - \frac{(nm-r)*\text{pi}2}{\Xi m123}$
 $* \text{Exp}[-\mu2 * \text{Exp}[-\varphi]] * ((\nu m2[\varphi] * \mu2) - \Phi m2[\varphi]) - (nm - r) * \left(\frac{\text{pi}2*mm2 - \mu2*\text{pi}2*FFF2}{\Xi m123}\right);$
 $\varphi2 = \varphi2 /. \text{FindRoot}[Bma1[\varphi2] == 0, \{\varphi2, \text{lgsig}02, \text{lgsig}2\}];$
 $fi3 = \frac{\text{lgsig}03 + \text{lgsig}3}{2};$

$$\mu_3 = -\text{Exp}[f_3] * \text{Log}\left[\frac{-\text{um}_3}{\text{summn}_1[f_3] + (nm-r) * \left(\frac{\text{pi}_3 * \nu \text{m}_3[f_3]}{\Xi_{m123}}\right)}\right];$$

$$\text{Bma}_3[\varphi_-] := \text{um}_3 * \text{Exp}[\varphi] - \text{summl}_3[\mu_3] + \text{Exp}[-\mu_3 * \text{Exp}[-\varphi]] * \text{summk}_3[\varphi, \mu_3] - \frac{(nm-r) * \text{pi}_3}{\Xi_{m123}} * \text{Exp}[-\mu_3 * \text{Exp}[-\varphi]] * ((\nu \text{m}_3[\varphi] * \mu_3) - \Phi \text{m}_3[\varphi]) - (nm-r) * \left(\frac{\text{pi}_3 * \text{mm}_3 - \mu_3 * \text{pi}_3 * \text{FFF}_3}{\Xi_{m123}}\right);$$

$$\varphi_3 = \varphi_3 /. \text{FindRoot}[\text{Bma}_1[\varphi_3] == 0, \{\varphi_3, \text{lgsig03}, \text{lgsig3}\}];$$

$$\pi_1 = -\frac{\text{um}_1}{nm};$$

$$\pi_2 = -\frac{\text{um}_2}{nm};$$

$$\pi_3 = -\frac{\text{um}_3}{nm};$$

$$\{\mu_1, \mu_2, \mu_3, \varphi_1, \varphi_2, \varphi_3, \pi_1, \pi_2, \pi_3\}$$

DG3[mu1_, mu2_, mu3_, lgsig1_, lgsig2_, lgsig3_, pi1_, pi2_, pi3_, xg1_, xg2_, xg3_, ng_] :=
Module[{\r1, r2, r3, r, \mu1, \mu2, \mu3, \varphi1, \varphi2, \varphi3, \pi1, \pi2, \pi3, a1, a2, a3, FF1, FF2, FF3, ug1,
ug2, ug3, uug1, uug2, uug3, vvg1, vvg2, vvg3, mg1, mg2, mg3, Qg = Array[q, {9, 9}]},

r1 = Length[xg1]; r2 = Length[xg2]; r3 = Length[xg3];

r = r1 + r2 + r3;

$$\text{sumgl}_1[\mu_{1-}] := \sum_{i=1}^{r_1} ((xg1[[i]] - \mu_1);$$

$$\text{sumgl}_2[\mu_{2-}] := \sum_{i=1}^{r_2} ((xg2[[i]] - \mu_2);$$

$$\text{sumgl}_3[\mu_{3-}] := \sum_{i=1}^{r_3} (xg3[[i]] - \mu_3);$$

$$\text{sumgk}_1[\mu_{1-}, \varphi_{1-}] := \sum_{i=1}^{r_1} ((xg1[[i]] - \mu_1) * \text{Exp}[(xg1[[i]] - \mu_1) * \text{Exp}[-\varphi_1]);$$

$$\text{sumgk}_2[\mu_{2-}, \varphi_{2-}] := \sum_{i=1}^{r_2} ((xg2[[i]] - \mu_2) * \text{Exp}[(xg2[[i]] - \mu_2) * \text{Exp}[-\varphi_2]);$$

$$\text{sumgk}_3[\mu_{3-}, \varphi_{3-}] := \sum_{i=1}^{r_3} (xg3[[i]] - \mu_3) * \text{Exp}[(xg3[[i]] - \mu_3) * \text{Exp}[-\varphi_3];$$

$$\text{sumgo}_1[\mu_{1-}, \varphi_{1-}] := \sum_{i=1}^{r_1} ((xg1[[i]] - \mu_1)^2 * \text{Exp}[(xg1[[i]] - \mu_1) * \text{Exp}[-\varphi_1]);$$

$$\text{sumgo2}[\mu2_ , \varphi2_] := \sum_{i=1}^{r2} ((\text{xg2}[[i]] - \mu2)\hat{2} * \text{Exp}[(\text{xg2}[[i]] - \mu2) * \text{Exp}[-\varphi2]]);$$

$$\text{sumgo3}[\mu3_ , \varphi3_] := \sum_{i=1}^{r3} (\text{xg3}[[i]] - \mu3)\hat{2} * \text{Exp}[(\text{xg3}[[i]] - \mu3) * \text{Exp}[-\varphi3]];$$

$$\text{sumgn1}[\mu1_ , \varphi1_] := \sum_{i=1}^{r1} \text{Exp}[(\text{xg1}[[i]] - \mu1) * \text{Exp}[-\varphi1]];$$

$$\text{sumgn2}[\mu2_ , \varphi2_] := \sum_{i=1}^{r2} \text{Exp}[(\text{xg2}[[i]] - \mu2) * \text{Exp}[-\varphi2]];$$

$$\text{sumgn3}[\mu3_ , \varphi3_] := \sum_{i=1}^{r3} \text{Exp}[(\text{xg3}[[i]] - \mu3) * \text{Exp}[-\varphi3]];$$

$$\text{a1} = \text{Exp}[-\mu1 * \text{Exp}[-\text{lgsig1}]];$$

$$\text{a2} = \text{Exp}[-\mu2 * \text{Exp}[-\text{lgsig2}]];$$

$$\text{a3} = \text{Exp}[-\mu3 * \text{Exp}[-\text{lgsig3}]];$$

$$\text{FF1} = \text{Exp}[-\text{a1}];$$

$$\text{FF2} = \text{Exp}[-\text{a2}];$$

$$\text{FF3} = \text{Exp}[-\text{a3}];$$

$$\Xi\text{g123} = \text{pi1} * \text{FF1} + \text{pi2} * \text{FF2} + \text{pi3} * \text{FF3};$$

$$\text{ug1} = -r1 - (\text{ng} - r) * (\text{pi1} * \text{FF1} / \Xi\text{g123});$$

$$\text{ug2} = -r2 - (\text{ng} - r) * (\text{pi2} * \text{FF2} / \Xi\text{g123});$$

$$\text{ug3} = -r3 - (\text{ng} - r) * (\text{pi3} * \text{FF3} / \Xi\text{g123});$$

$$\text{uug1} = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{w, \text{a1}, \text{Infinity}\}] ;$$

$$\text{uug2} = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{w, \text{a2}, \text{Infinity}\}] ;$$

$$\text{uug3} = \text{Integrate}[\text{Log}[w] * \text{Exp}[-w], \{w, \text{a3}, \text{Infinity}\}] ;$$

$$\text{vvg1} = \text{Integrate}[\text{Exp}[-w], \{w, \text{a1}, \text{Infinity}\}];$$

$$\text{vvg2} = \text{Integrate}[\text{Exp}[-w], \{w, \text{a2}, \text{Infinity}\}];$$

$$\text{vvg3} = \text{Integrate}[\text{Exp}[-w], \{w, \text{a3}, \text{Infinity}\}];$$

$$\text{mg3} = \text{Exp}[\text{lgsig3}] * \text{uug3} + \mu3 * \text{vvg3};$$

$$\text{mg2} = \text{Exp}[\text{lgsig2}] * \text{uug2} + \mu2 * \text{vvg2};$$

$$\text{mg1} = \text{Exp}[\text{lgsig1}] * \text{uug1} + \mu1 * \text{vvg1};$$

$$\text{fg1}[t_ , \varphi1_] := t^{\frac{e^{\text{lgsig1}}}{e^{\varphi1}}} * \text{Exp}[-t];$$

$$\text{fg2}[t_ , \varphi2_] := t^{\frac{e^{\text{lgsig2}}}{e^{\varphi2}}} * \text{Exp}[-t];$$

$$\text{fg3}[t_ , \varphi3_] := t^{\frac{e^{\text{lgsig3}}}{e^{\varphi3}}} * \text{Exp}[-t];$$

$gg1[\varphi1_] := NIntegrate[fg1[t, \varphi1], \{t, a1, Infinity \}];$
 $gg2[\varphi2_] := NIntegrate[fg2[t, \varphi2], \{t, a2, Infinity \}];$
 $gg3[\varphi3_] := NIntegrate[fg3[t, \varphi3], \{t, a3, Infinity \}];$
 $\nu g1[\varphi1_] := Exp[\mu1 * Exp[-\varphi1]] * gg1[\varphi1];$
 $\nu g2[\varphi2_] := Exp[\mu2 * Exp[-\varphi2]] * gg2[\varphi2];$
 $\nu g3[\varphi3_] := Exp[\mu3 * Exp[-\varphi3]] * gg3[\varphi3];$
 $hg1[t_ , \varphi1_] := t^{\frac{e^{lgsig1}}{e^{\varphi1}}} * Exp[-t] * Log[t];$
 $hg2[t_ , \varphi2_] := t^{\frac{e^{lgsig2}}{e^{\varphi2}}} * Exp[-t] * Log[t];$
 $hg3[t_ , \varphi3_] := t^{\frac{e^{lgsig3}}{e^{\varphi3}}} * Exp[-t] * Log[t];$
 $\Lambda g1[\varphi1_] := NIntegrate[hg1[t, \varphi1], \{t, a1, Infinity \}];$
 $\Lambda g2[\varphi2_] := NIntegrate[hg2[t, \varphi2], \{t, a2, Infinity \}];$
 $\Lambda g3[\varphi3_] := NIntegrate[hg3[t, \varphi3], \{t, a3, Infinity \}];$
 $\gamma g1[t_ , \varphi1_] := t^{\frac{e^{lgsig1}}{e^{\varphi1}}} * Exp[-t] * (Log[t])^2;$
 $\gamma g2[t_ , \varphi2_] := t^{\frac{e^{lgsig2}}{e^{\varphi2}}} * Exp[-t] * (Log[t])^2;$
 $\gamma g3[t_ , \varphi3_] := t^{\frac{e^{lgsig3}}{e^{\varphi3}}} * Exp[-t] * (Log[t])^2;$
 $\Gamma g1[\varphi1_] := NIntegrate[\gamma g1[t, \varphi1], \{t, a1, Infinity \}];$
 $\Gamma g2[\varphi2_] := NIntegrate[\gamma g2[t, \varphi2], \{t, a2, Infinity \}];$
 $\Gamma g3[\varphi3_] := NIntegrate[\gamma g3[t, \varphi3], \{t, a3, Infinity \}];$
 $D\nu g1[\varphi1_] := Exp[\mu1 * Exp[-\varphi1]] * Exp[lgsig1 - \varphi1] * \Lambda g1[\varphi1] + Exp[-\varphi1] * \mu1 * \nu g1[\varphi1];$
 $D\nu g2[\varphi2_] := Exp[\mu2 * Exp[-\varphi2]] * Exp[lgsig2 - \varphi2] * \Lambda g2[\varphi2] + Exp[-\varphi2] * \mu2 * \nu g2[\varphi2];$
 $D\nu g3[\varphi3_] := Exp[\mu3 * Exp[-\varphi3]] * Exp[lgsig3 - \varphi3] * \Lambda g3[\varphi3] + Exp[-\varphi3] * \mu3 * \nu g3[\varphi3];$
 $qg11[\mu1_ , \varphi1_] := Exp[-2 * \varphi1] * sumgn1[\mu1, \varphi1] + (ng - r) * (\frac{\pi1 * \nu g1[\varphi1]}{\Xi g123}) * (Exp[-2 * \varphi1 - \mu1 * Exp[-\varphi1]]);$
 $qg12[\mu1_ , \mu2_] := 0; \quad qg13[\mu1_ , \mu3_] := 0;$
 $qg14[\mu1_ , \varphi1_] := -r1 * Exp[-\varphi1] + Exp[-2 * \varphi1] * sumgk1[\mu1, \varphi1] + Exp[-\varphi1] * sumgn1[\mu1, \varphi1]$
 $- (ng - r) * Exp[-\varphi1] * (\frac{\pi1 * FF1}{\Xi g123}) + (ng - r) * (\frac{\pi1 * \nu g1[\varphi1]}{\Xi g123}) * (1 - \mu1 * Exp[-\varphi1]) * Exp[-\varphi1 - \mu1 * Exp[-\varphi1]] + (ng - r) * (\frac{\pi1 * D\nu g1[\varphi1]}{\Xi g123}) * Exp[-\varphi1 - \mu1 * Exp[-\varphi1]];$
 $qg15[\mu1_ , \varphi2_] := 0; \quad qg16[\mu1_ , \varphi3_] := 0;$
 $qg17[\mu1_ , \pi1_] := 0; \quad qg18[\mu1_ , \pi2_] := 0;$
 $qg19[\mu1_ , \pi3_] := 0; \quad qg21[\mu1_ , \mu2_] := 0;$

$$\begin{aligned}
\text{qg22}[\mu 2_{-}, \varphi 2_{-}] &:= \text{Exp}[-2 * \varphi 2] * \text{sumgn2}[\mu 2, \varphi 2] + (\text{ng} - r) * \left(\frac{\text{pi} 2 * \nu \text{g} 2[\varphi 2]}{\Xi \text{g} 123} \right) * \text{Exp}[-2 * \varphi 2 - \mu 2 * \text{Exp}[-\varphi 2]]; \\
\text{qg23}[\mu 2_{-}, \mu 3_{-}] &:= 0; \quad \text{qg24}[\mu 2_{-}, \varphi 1_{-}] := 0; \\
\text{qg25}[\mu 2_{-}, \varphi 2_{-}] &:= -r 2 * \text{Exp}[-\varphi 2] + \text{Exp}[-2 * \varphi 2] * \text{sumgk2}[\mu 2, \varphi 2] + \text{Exp}[-\varphi 2] * \text{sumgn2}[\mu 2, \varphi 2] \\
&- (\text{ng} - r) * \text{Exp}[-\varphi 2] * \left(\frac{\text{pi} 2 * \text{FF} 2}{\Xi \text{g} 123} \right) + (\text{ng} - r) * \left(\frac{\text{pi} 2 * \nu \text{g} 2[\varphi 2]}{\Xi \text{g} 123} \right) * (1 - \mu 2 * \text{Exp}[-\varphi 2]) * \text{Exp}[-\varphi 2 - \mu 2 * \text{Exp}[-\varphi 2]] \\
&+ (\text{ng} - r) * \left(\frac{\text{pi} 2 * \text{D} \nu \text{g} 2[\varphi 2]}{\Xi \text{g} 123} \right) * \text{Exp}[-\varphi 2 - \mu 2 * \text{Exp}[-\varphi 2]]; \\
\text{qg26}[\mu 2_{-}, \varphi 3_{-}] &:= 0; \quad \text{qg27}[\mu 2_{-}, \pi 1_{-}] := 0; \\
\text{qg28}[\mu 2_{-}, \pi 2_{-}] &:= 0; \quad \text{qg29}[\mu 2_{-}, \pi 3_{-}] := 0; \\
\text{qg31}[\mu 1_{-}, \mu 3_{-}] &:= 0; \quad \text{qg32}[\mu 2_{-}, \mu 3_{-}] := 0; \\
\text{qg33}[\mu 3_{-}, \varphi 3_{-}] &:= \text{Exp}[-2 * \varphi 3] * \text{sumgn3}[\mu 3, \varphi 3] + (\text{ng} - r) * \left(\frac{\text{pi} 3 * \nu \text{g} 3[\varphi 3]}{\Xi \text{g} 123} \right) * (\text{Exp}[-2 * \varphi 3 - \mu 3 * \text{Exp}[-\varphi 3]]); \\
\text{qg34}[\mu 3_{-}, \varphi 1_{-}] &:= 0; \quad \text{qg35}[\mu 3_{-}, \varphi 2_{-}] := 0; \\
\text{qg36}[\mu 3_{-}, \varphi 3_{-}] &:= -r 3 * \text{Exp}[-\varphi 3] + \text{Exp}[-2 * \varphi 3] * \text{sumgk3}[\mu 3, \varphi 3] + \text{Exp}[-\varphi 3] * \text{sumgn3}[\mu 3, \varphi 3] \\
&- (\text{ng} - r) * \text{Exp}[-\varphi 3] * \left(\frac{\text{pi} 3 * \text{FF} 3}{\Xi \text{g} 123} \right) + (\text{ng} - r) * \left(\frac{\text{pi} 3 * \nu \text{g} 3[\varphi 3]}{\Xi \text{g} 123} \right) * (1 - \mu 3 * \text{Exp}[-\varphi 3]) * \text{Exp}[-\varphi 3 - \mu 3 * \text{Exp}[-\varphi 3]] \\
&+ (\text{ng} - r) * \left(\frac{\text{pi} 3 * \text{D} \nu \text{g} 3[\varphi 3]}{\Xi \text{g} 123} \right) * \text{Exp}[-\varphi 3 - \mu 3 * \text{Exp}[-\varphi 3]]; \\
\text{qg37}[\mu 3_{-}, \pi 1_{-}] &:= 0; \quad \text{qg38}[\mu 3_{-}, \pi 2_{-}] := 0; \\
\text{qg39}[\mu 3_{-}, \pi 3_{-}] &:= 0; \\
\text{qg41}[\mu 1_{-}, \varphi 1_{-}] &:= -r 1 * \text{Exp}[-\varphi 1] + \text{Exp}[-2 * \varphi 1] * \text{sumgk1}[\mu 1, \varphi 1] + \text{Exp}[-\varphi 1] * \text{sumgn1}[\mu 1, \varphi 1] \\
&- (\text{ng} - r) * \text{Exp}[-\varphi 1] * \left(\frac{\text{pi} 1 * \text{FF} 1}{\Xi \text{g} 123} \right) + (\text{ng} - r) * \left(\frac{\text{pi} 1 * \nu \text{g} 1[\varphi 1]}{\Xi \text{g} 123} \right) * (1 - \mu 1 * \text{Exp}[-\varphi 1]) * \text{Exp}[-\varphi 1 - \mu 1 * \text{Exp}[-\varphi 1]] \\
&+ (\text{ng} - r) * \left(\frac{\text{pi} 1 * \text{D} \nu \text{g} 1[\varphi 1]}{\Xi \text{g} 123} \right) * \text{Exp}[-\varphi 1 - \mu 1 * \text{Exp}[-\varphi 1]]; \\
\text{qg42}[\mu 2_{-}, \varphi 1_{-}] &:= 0; \quad \text{qg43}[\mu 3_{-}, \varphi 1_{-}] := 0; \\
\text{qg44}[\mu 1_{-}, \varphi 1_{-}] &:= -1 * \text{sumgl1}[\mu 1] * \text{Exp}[-\varphi 1] + \text{Exp}[-\varphi 1] * \text{sumgk1}[\mu 1, \varphi 1] + \text{Exp}[-2 * \varphi 1] * \text{sumgo1}[\mu 1, \varphi 1] \\
&- (\text{ng} - r) * \left(\frac{\text{pi} 1 * (\text{mg} 1 - \mu 1 * \text{FF} 1)}{\Xi \text{g} 123} \right) * \text{Exp}[-\varphi 1] + (\text{ng} - r) * \left(\frac{\text{pi} 1}{\Xi \text{g} 123} \right) * ((\mu 1 - \mu 1) * (\nu \text{g} 1[\varphi 1] * (\mu 1 * \text{Exp}[-\varphi 1] - 1) * \text{Exp}[-\varphi 1 - \mu 1 * \text{Exp}[-\varphi 1]] - \text{Exp}[-\varphi 1 - \mu 1 * \text{Exp}[-\varphi 1]] * \text{D} \nu \text{g} 1[\varphi 1]) \\
&+ \text{Exp}[\text{Exp}[-\varphi 1] * (\mu 1 - \mu 1) + (\text{lgsig} 1 - \varphi 1)] * \text{Exp}[\text{lgsig} 1 - \varphi 1] * \Gamma \text{g} 1[\varphi 1] + (1 + (\mu 1 - \mu 1) * \text{Exp}[-\varphi 1]) * \text{Exp}[\text{Exp}[-\varphi 1] * (\mu 1 - \mu 1) + (\text{lgsig} 1 - \varphi 1)] * \Lambda \text{g} 1[\varphi 1]); \\
\text{qg45}[\varphi 1_{-}, \varphi 2_{-}] &:= 0; \quad \text{qg46}[\varphi 1_{-}, \varphi 3_{-}] := 0; \\
\text{qg47}[\varphi 1_{-}, \pi 1_{-}] &:= 0; \quad \text{qg48}[\varphi 1_{-}, \pi 2_{-}] := 0; \\
\text{qg49}[\varphi 1_{-}, \pi 3_{-}] &:= 0; \quad \text{qg51}[\mu 1_{-}, \varphi 2_{-}] := 0; \\
\text{qg52}[\mu 2_{-}, \varphi 2_{-}] &:= -r 2 * \text{Exp}[-\varphi 2] + \text{Exp}[-2 * \varphi 2] * \text{sumgk2}[\mu 2, \varphi 2] + \text{Exp}[-\varphi 2] * \text{sumgn2}[\mu 2, \varphi 2] \\
&- (\text{ng} - r) * \text{Exp}[-\varphi 2] * \left(\frac{\text{pi} 2 * \text{FF} 2}{\Xi \text{g} 123} \right) + (\text{ng} - r) * \left(\frac{\text{pi} 2 * \nu \text{g} 2[\varphi 2]}{\Xi \text{g} 123} \right) * (1 - \mu 2 * \text{Exp}[-\varphi 2]) * \text{Exp}[-\varphi 2 - \mu 2 * \text{Exp}[-\varphi 2]];
\end{aligned}$$

$$\mu_2 * \text{Exp}[-\varphi_2] + (ng-r) * \left(\frac{\text{pi}_2 * D\nu g_2[\varphi_2]}{\Xi g_{123}} \right) * \text{Exp}[-\varphi_2 - \mu_2 * \text{Exp}[-\varphi_2]];$$

$$\text{qg53}[\mu_{3-}, \varphi_{2-}] := 0; \quad \text{qg54}[\varphi_{1-}, \varphi_{2-}] := 0;$$

$$\text{qg55}[\mu_{2-}, \varphi_{2-}] := -1 * \text{sumgl}_2[\mu_2] * \text{Exp}[-\varphi_2] + \text{Exp}[-\varphi_2] * \text{sumgk}_2[\mu_2, \varphi_2] + \text{Exp}[-2 * \varphi_2] * \text{sumgo}_2[\mu_2, \varphi_2] - (ng-r) * \left(\frac{\text{pi}_2 * (mg_2 - \mu_2 * FF_2)}{\Xi g_{123}} \right) * \text{Exp}[-\varphi_2] + (ng-r) * \left(\frac{\text{pi}_2}{\Xi g_{123}} \right) * ((\mu_2 - \mu_2) * (\nu g_2[\varphi_2] * (\mu_2 * \text{Exp}[-\varphi_2] - 1) * \text{Exp}[-\varphi_2 - \mu_2 * \text{Exp}[-\varphi_2]] - \text{Exp}[-\varphi_2 - \mu_2 * \text{Exp}[-\varphi_2]]) * D\nu g_2[\varphi_2] + \text{Exp}[\text{Exp}[-\varphi_2] * (\mu_2 - \mu_2) + (lgsig_2 - \varphi_2)] * \text{Exp}[lgsig_2 - \varphi_2] * \Gamma g_2[\varphi_2] + (1 + (\mu_2 - \mu_2) * \text{Exp}[-\varphi_2]) * \text{Exp}[\text{Exp}[-\varphi_2] * (\mu_2 - \mu_2) + (lgsig_2 - \varphi_2)] * \Lambda g_2[\varphi_2]);$$

$$\text{qg56}[\varphi_{3-}, \varphi_{2-}] := 0; \quad \text{qg57}[\pi_{1-}, \varphi_{2-}] := 0;$$

$$\text{qg58}[\pi_{2-}, \varphi_{2-}] := 0; \quad \text{qg59}[\pi_{3-}, \varphi_{2-}] := 0;$$

$$\text{qg61}[\mu_{1-}, \varphi_{3-}] := 0; \quad \text{qg62}[\mu_{2-}, \varphi_{3-}] := 0;$$

$$\text{qg63}[\mu_{3-}, \varphi_{3-}] := -r_3 * \text{Exp}[-\varphi_3] + \text{Exp}[-2 * \varphi_3] * \text{sumgk}_3[\mu_3, \varphi_3] + \text{Exp}[-\varphi_3] * \text{sumgn}_3[\mu_3, \varphi_3] - (ng-r) * \text{Exp}[-\varphi_3] * \left(\frac{\text{pi}_3 * FF_3}{\Xi g_{123}} \right) + (ng-r) * \left(\frac{\text{pi}_3 * \nu g_3[\varphi_3]}{\Xi g_{123}} \right) * (1 - \mu_3 * \text{Exp}[-\varphi_3]) * \text{Exp}[-\varphi_3 - \mu_3 * \text{Exp}[-\varphi_3]] + (ng-r) * \left(\frac{\text{pi}_3 * D\nu g_3[\varphi_3]}{\Xi g_{123}} \right) * \text{Exp}[-\varphi_3 - \mu_3 * \text{Exp}[-\varphi_3]];$$

$$\text{qg64}[\varphi_{1-}, \varphi_{3-}] := 0; \quad \text{qg65}[\varphi_{2-}, \varphi_{3-}] := 0;$$

$$\text{qg66}[\mu_{3-}, \varphi_{3-}] := -1 * \text{sumgl}_3[\mu_3] * \text{Exp}[-\varphi_3] + \text{Exp}[-\varphi_3] * \text{sumgk}_3[\mu_3, \varphi_3] + \text{Exp}[-3 * \varphi_3] * \text{sumgo}_3[\mu_3, \varphi_3] - (ng-r) * \left(\frac{\text{pi}_3 * (mg_3 - \mu_3 * FF_3)}{\Xi g_{123}} \right) * \text{Exp}[-\varphi_3] + (ng-r) * \left(\frac{\text{pi}_3}{\Xi g_{123}} \right) * ((\mu_3 - \mu_3) * (\nu g_3[\varphi_3] * (\mu_3 * \text{Exp}[-\varphi_3] - 1) * \text{Exp}[-\varphi_3 - \mu_3 * \text{Exp}[-\varphi_3]] - \text{Exp}[-\varphi_3 - \mu_3 * \text{Exp}[-\varphi_3]]) * D\nu g_3[\varphi_3] + \text{Exp}[\text{Exp}[-\varphi_3] * (\mu_3 - \mu_3) + (lgsig_3 - \varphi_3)] * \text{Exp}[lgsig_3 - \varphi_3] * \Gamma g_3[\varphi_3] + (1 + (\mu_3 - \mu_3) * \text{Exp}[-\varphi_3]) * \text{Exp}[\text{Exp}[-\varphi_3] * (\mu_3 - \mu_3) + (lgsig_3 - \varphi_3)] * \Lambda g_3[\varphi_3]);$$

$$\text{qg67}[\pi_{1-}, \varphi_{3-}] := 0; \quad \text{qg68}[\pi_{2-}, \varphi_{3-}] := 0;$$

$$\text{qg69}[\pi_{3-}, \varphi_{3-}] := 0; \quad \text{qg71}[\pi_{1-}] := 0;$$

$$\text{qg72}[\pi_{1-}] := 0; \quad \text{qg73}[\pi_{1-}] := 0;$$

$$\text{qg74}[\pi_{1-}] := 0; \quad \text{qg75}[\pi_{1-}] := 0;$$

$$\text{qg76}[\pi_{1-}] := 0;$$

$$\text{qg77}[\pi_{1-}] := -ug_1 * \left(\frac{1}{(\pi_1)^2} \right);$$

$$\text{qg78}[\pi_{1-}] := 0; \quad \text{qg79}[\pi_{1-}] := 0;$$

$$\text{qg81}[\pi_{2-}] := 0; \quad \text{qg82}[\pi_{2-}] := 0;$$

$$\text{qg83}[\pi_{2-}] := 0; \quad \text{qg84}[\pi_{2-}] := 0;$$

$$\text{qg85}[\pi_{2-}] := 0; \quad \text{qg86}[\pi_{2-}] := 0;$$

$$\text{qg87}[\pi_{2-}] := 0;$$

$$\text{qg88}[\pi_{2-}] := -ug_2 * \left(\frac{1}{(\pi_2)^2} \right);$$

$qg89[\pi2.] := 0;$ $qg91[\pi3.] := 0;$
 $qg92[\pi3.] := 0;$ $qg93[\pi3.] := 0;$
 $qg94[\pi3.] := 0;$ $qg95[\pi3.] := 0;$
 $qg96[\pi3.] := 0;$ $qg97[\pi3.] := 0;$
 $qg98[\pi3.] := 0;$
 $qg99[\pi3.] := -ug3*(\frac{1}{(\pi3)^2});$
 $q[1, 1] = qg11[\mu1, lgsig1];$ $q[1, 2] = qg12[\mu1, \mu2];$
 $q[1, 3] = qg13[\mu1, \mu3];$ $q[1, 4] = qg14[\mu1, lgsig1];$
 $q[1, 5] = qg15[\mu1, lgsig2];$ $q[1, 6] = qg16[\mu1, lgsig3];$
 $q[1, 7] = qg17[\mu1, \pi1];$ $q[1, 8] = qg18[\mu1, \pi2];$
 $q[1, 9] = qg19[\mu1, \pi3];$ $q[2, 1] = qg21[\mu1, \mu2];$
 $q[2, 2] = qg22[\mu2, lgsig2];$ $q[2, 3] = qg23[\mu2, \mu3];$
 $q[2, 4] = qg24[\mu2, lgsig1];$ $q[2, 5] = qg25[\mu2, lgsig2];$
 $q[2, 6] = qg26[\mu2, lgsig3];$ $q[2, 7] = qg27[\mu2, \pi1];$
 $q[2, 8] = qg28[\mu2, \pi2];$ $q[2, 9] = qg29[\mu2, \pi3];$
 $q[3, 1] = qg31[\mu1, \mu3];$ $q[3, 2] = qg32[\mu1, lgsig1];$
 $q[3, 3] = qg33[\mu3, lgsig3];$ $q[3, 4] = qg34[\mu3, lgsig1];$
 $q[3, 5] = qg35[\mu3, lgsig2];$ $q[3, 6] = qg36[\mu3, lgsig3];$
 $q[3, 7] = qg37[\mu3, \pi1];$ $q[3, 8] = qg38[\mu3, \pi2];$
 $q[3, 9] = qg39[\mu3, \pi3];$ $q[4, 1] = qg41[\mu1, lgsig1];$
 $q[4, 2] = qg42[\mu2, lgsig1];$ $q[4, 3] = qg43[\mu3, lgsig1];$
 $q[4, 4] = qg44[\mu1, lgsig1];$ $q[4, 5] = qg45[lgsig1, lgsig2];$
 $q[4, 6] = qg46[lgsig1, lgsig3];$ $q[4, 7] = qg47[lgsig1, \pi1];$
 $q[4, 8] = qg48[lgsig1, \pi2];$ $q[4, 9] = qg49[lgsig1, \pi3];$
 $q[5, 1] = qg51[\mu1, lgsig2];$ $q[5, 2] = qg52[\mu2, lgsig2];$
 $q[5, 3] = qg53[\mu3, lgsig2];$ $q[5, 4] = qg54[lgsig1, lgsig2];$
 $q[5, 5] = qg55[\mu2, lgsig2];$ $q[5, 6] = qg56[lgsig3, lgsig2];$
 $q[5, 7] = qg57[\pi1, lgsig2];$ $q[5, 8] = qg58[\pi2, lgsig2];$
 $q[5, 9] = qg59[\pi3, lgsig2];$ $q[6, 1] = qg61[\mu1, lgsig3];$
 $q[6, 2] = qg62[\mu2, lgsig3];$ $q[6, 3] = qg63[\mu3, lgsig3];$
 $q[6, 4] = qg64[lgsig1, lgsig3];$ $q[6, 5] = qg65[lgsig2, lgsig3];$
 $q[6, 6] = qg66[\mu3, lgsig3];$ $q[6, 7] = qg67[\pi1, lgsig3];$
 $q[6, 8] = qg68[\pi2, lgsig3];$ $q[6, 9] = qg69[\pi3, lgsig3];$

$q[7, 1] = qg71[pi1]; \quad q[7, 2] = qg72[pi1];$
 $q[7, 3] = qg73[pi1]; \quad q[7, 4] = qg74[pi1];$
 $q[7, 5] = qg75[pi1]; \quad q[7, 6] = qg76[pi1];$
 $q[7, 7] = qg77[pi1]; \quad q[7, 8] = qg78[pi1];$
 $q[7, 9] = qg79[pi1]; \quad q[8, 1] = qg81[pi2];$
 $q[8, 2] = qg82[pi2]; \quad q[8, 3] = qg83[pi2];$
 $q[8, 4] = qg84[pi2]; \quad q[8, 5] = qg85[pi2];$
 $q[8, 6] = qg86[pi2]; \quad q[8, 7] = qg87[pi2];$
 $q[8, 8] = qg88[pi2]; \quad q[8, 9] = qg89[pi2];$
 $q[9, 1] = qg91[pi3]; \quad q[9, 2] = qg92[pi3];$
 $q[9, 3] = qg93[pi3]; \quad q[9, 4] = qg94[pi3];$
 $q[9, 5] = qg95[pi3]; \quad q[9, 6] = qg96[pi3];$
 $q[9, 7] = qg97[pi3]; \quad q[9, 8] = qg98[pi3];$
 $q[9, 9] = qg99[pi3]; \text{ Qg}$

$x1 = \text{ReadList}["C:/tesis/cap tres/k3/y1.txt"];$
 $x2 = \text{ReadList}["C:/tesis/cap tres/k3/y2.txt"];$
 $x3 = \text{ReadList}["C:/tesis/cap tres/k3/y3.txt"];$
 $tamr1 = \text{Length}[x1];$
 $tamr2 = \text{Length}[x2];$
 $tamr3 = \text{Length}[x3];$

$$xbarr1 = \sum_{i=1}^{tamr1} \frac{x1[[i]]}{tamr1};$$

$$xbarr2 = \sum_{i=1}^{tamr2} \frac{x2[[i]]}{tamr2};$$

$$xbarr3 = \sum_{i=1}^{tamr3} \frac{x3[[i]]}{tamr3};$$

$fi1 =$

$$fi1/.FindRoot[N[\frac{1}{tamr1} \sum_{i=1}^{tamr1} x1[[i]] * \text{Exp}[x1[[i]] - \text{Exp}[fi1] * \text{Log}[\sum_{i=1}^{tamr1} \frac{\text{Exp}[x1[[i]] * \text{Exp}[-fi1]]}{tamr1}] \text{Exp}[-fi1]] -$$

```

Exp[-fi1]==
xbarr1,{fi1,-8,8};
φ1 = fi1+.5;
φ01 = fi1-0.5;
μ1 = Exp[fi1] * Log[ Sum_{i=1}^{tamr1} \frac{Exp[x1[[i]] * Exp[-fi1]}{tamr1} ];
fi2 =
fi2/.FindRoot[N[\frac{1}{tamr2} \sum_{i=1}^{tamr2} x2[[i]] * Exp[x2[[i]] - Exp[fi2] * Log[ \sum_{i=1}^{tamr2} \frac{Exp[x2[[i]] * Exp[-fi2]]}{tamr2} ]Exp[-fi2]-
Exp[-fi2]]==
xbarr2,{fi2,-8,8};
φ2 = fi2 + 0.5;
φ02 = fi2 - 0.5;
μ2 = Exp[fi2] * Log[ Sum_{i=1}^{tamr2} \frac{Exp[x2[[i]] * Exp[-fi2]}{tamr2} ];
fi3 =
fi3/.FindRoot[N[\frac{1}{tamr3} \sum_{i=1}^{tamr3} x3[[i]] * Exp[x3[[i]] - Exp[fi3] * Log[ \sum_{i=1}^{tamr3} \frac{Exp[x2[[i]] * Exp[-fi3]]}{tamr3} ]Exp[-fi3]-
Exp[-fi3]]==
xbarr3,{fi3,-8,8};
φ3 = fi3 + 0.5;
φ03 = fi3 - 0.5;
μ3 = Exp[fi3] * Log[ Sum_{i=1}^{tamr3} \frac{Exp[x3[[i]] * Exp[-fi3]}{tamr3} ];
π1 = \frac{tamr1}{tamr1 + tamr2 + tamr3};
π2 = \frac{tamr2}{tamr1 + tamr2 + tamr3};
π3 = 1 - π1 - π2;
n = 500;
Ξstar = Array[matΞ, {30, 9}];
Do[Ξstar[[kk]] = Mf3[μ1, μ2, μ3, φ1, φ2, φ3, π1, π2, π3, x1, x2, x3, n, φ01, φ02, φ03];
φ01 = φ1; φ02 = φ2; φ03 = φ3; μ1 = Ξstar[[kk, 1]]; μ2 = Ξstar[[kk, 2]]; μ3 = Ξstar[[kk,
3]]; φ1 = Ξstar[[kk, 4]]; φ2 = Ξstar[[kk, 5]]; φ3 = Ξstar[[kk, 6]]; π1 = Ξstar[[kk, 7]]; π2 =
Ξstar[[kk, 8]]; π3 = Ξstar[[kk, 9]]; Print[Ξstar[[kk]]];, {kk, 30}
Ξstar[[30]]
μ1star = Ξstar[[30, 1]];

```

```

μ2star = Ξstar[[30, 2]];
μ3star = Ξstar[[30, 3]];
φ1star = Ξstar[[30, 4]];
φ2star = Ξstar[[30, 5]];
φ3star = Ξstar[[30, 6]];
π1star = Ξstar[[30, 7]];
π2star = Ξstar[[30, 8]];
π3star = Ξstar[[30, 9]];
Q = Array[q, {9, 9}];
Q = DG3[μ1star, μ2star, μ3star, φ1star, φ2star, φ3star, π1star, π2star, π3star, x1, x2, x3,
n];
ERRE = Array[matr, {9, 9}];
θd = Array[θ, {9}];
θtilda = Array[θt, {9}];
Do[ Do[θd = Ξstar[[30]]; θd[[ii]] = Ξstar[[tt, ii]];
θtilda = Mf3[θd[[1]], θd[[2]], θd[[3]], θd[[4]], θd[[5]], θd[[6]], θd[[7]], θd[[8]], θd[[9]], x1, x2,
x3, n, Ξstar[[21, 4]], Ξstar[[21, 5]], Ξstar[[21, 6]]];
Do[
matr[[ii, jj]] =  $\frac{\theta_{tilda}[[jj]] - \Xi_{star}[[30, jj]]}{\Xi_{star}[[tt, ii]] - \Xi_{star}[[30, ii]]}$  ;, {jj, 9}];, {ii, 9}]; Print[tt];, {tt, 12, 19}]
MatrixForm[ERRE]
II = DiagonalMatrix[1, 1, 1, 1, 1, 1, 1, 1, 1];
VMAT = Inverse[Q]+Inverse[Q].ERRE.Inverse[II-ERRE];
MatrixForm[VMAT]
Llmu1 = Ξstar[[30, 1]]-1.96*Sqrt[VMAT[[1, 1]]]
LSmu1 = Ξstar[[30, 1]]+1.96*Sqrt[VMAT[[1, 1]]]
Llmu2 = Ξstar[[30, 2]]-1.96*Sqrt[VMAT[[2, 2]]]
LSmu2 = Ξstar[[30, 2]]+1.96*Sqrt[VMAT[[2, 2]]]
Llmu3 = Ξstar[[30, 3]]-1.96*Sqrt[VMAT[[3, 3]]]
LSmu3 = Ξstar[[30, 3]]+1.96*Sqrt[VMAT[[3, 3]]]
Llphi1 = Ξstar[[30, 4]]-1.96*Sqrt[VMAT[[4, 4]]]
LSphi1 = Ξstar[[30, 4]]+1.96*Sqrt[VMAT[[4, 4]]]
Llphi2 = Ξstar[[30, 5]]-1.96*Sqrt[VMAT[[5, 5]]]
LSphi2 = Ξstar[[30, 5]]+1.96*Sqrt[VMAT[[5, 5]]]
Llphi3 = Ξstar[[30, 6]]-1.96*Sqrt[VMAT[[6, 6]]]

```

$$LSphi3 = \Xi star[[30, 6]] + 1.96 * Sqrt[VMAT[[6, 6]]]$$

$$LIp1 = \Xi star[[30, 7]] - 1.96 * Sqrt[VMAT[[7, 7]]]$$

$$LSp1 = \Xi star[[30, 7]] + 1.96 * Sqrt[VMAT[[7, 7]]]$$

$$LIp2 = \Xi star[[30, 8]] - 1.96 * Sqrt[VMAT[[8, 8]]]$$

$$LSp2 = \Xi star[[30, 8]] + 1.96 * Sqrt[VMAT[[8, 8]]]$$

$$LIp3 = \Xi star[[30, 9]] - 1.96 * Sqrt[VMAT[[9, 9]]]$$

$$LSp3 = \Xi star[[30, 9]] + 1.96 * Sqrt[VMAT[[9, 9]]]$$

Bibliografía

- [1] Bernardo, J. M. (1981). *Bioestadística*. 1ra ed. España: Vincens-Vives.
- [2] Bernardo, J. M. y Smith, A.F.M. (1994). *Bayesian Theory*. Ed. Wiley.
- [3] Contreras-Cristán, A., Gutiérrez-Peña, E. & O'Reilly, F. (2003). Statistical Inference for mixtures of distributions for censored data which partial identification. *Communication in Statistics*, **32**, issues 4, Theory and Methods, 749-774.
- [4] Cox, D.R. (1958) The analysis of exponentially distributed lifetimes with two types of failure. *J. R. Statist. Soc. B*, *21*, 411-421.
- [5] Dempster, A. P., Laird, N. & Rubin, D.B. (1977). Maximun likelihood from incomplete data via the EM algorithm. *J. R. Statist. Soc. B*, **39**, 1-38.
- [6] Díaz Francés, E. (1998) *Scientific application of maximum likelihood in multi-parametric problems*, Ph.D. Dissertation, CIMAT, Mexico.
- [7] Gelman, Carlin, Stern y Rubin (1995). *Bayesian Data Analysis*. Ed. Chapman & Hall.
- [8] Gutiérrez-Peña, E. (1997). *Métodos computacionales en la inferencia Bayesiana*. Monográficas. Vol. 6, No.15. I.I.M.A.S., México.
- [9] Lehmann, E. L. (1983). *Theory of point estimation*. Ed. Wiley.
- [10] McLachlan, G. J. & Krishnan, T. (1997). *The EM algorithm and expresions*.1ra ed. New York: John Wiley and Sons.

- [11] Mendenhall, W. and Hader, R.J. (1958). Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data. *Biometrika*, **45**, 504-520.
- [12] Meng, X.L. and Rubin, D.B. (1991). Using EM to obtain asymptotic variance-covariance matrices: the SEM algorithm. *Journal of the American Statistical Association*, **86** (416) 899-909.
- [13] Ruíz-Maya-Pérez, L., Martín, J. (1995). *Estadística I* .2da ed. Madrid: Universidad Autónoma de Madrid.
- [14] Ruíz-Maya-Pérez, L., Martín, J. (1996). *Estadística II* .2da ed. Madrid: Universidad Autónoma de Madrid.
- [15] Tanner, M. A. (1996). *Tools for Statistical Inference*.3rd ed. New York: Springer Verlag.