Investigación N° 32, 2025
ISSN i: 1684-0089 / ISSN v: 2709-8583

Modelos de regresión con errores de medida y datos censurados usando la distribución Slash multivariada

Regression models with measurement errors and censored data using the multivariate Slash distribution

Alejandro Guillermo Monzón Montoya[1]
Universidad Nacional de San Cristóbal de Huamanga, Ayacucho
https://orcid.org/0000-0002-1057-5647
alejandro.monzon@unsch.edu.pe

Instituto de investigación: Universidad Nacional de San Cristóbal de Huamanga, Área de investigación: Ingenierías y ciencias básicas, Línea de Investigación: Análisis multivariante

Resumen

En este trabajo se estudia los modelos de regresión con error de medida, que son útiles para describir diferentes fenómenos en diversas áreas del conocimiento. A este modelo se adiciona la posibilidad de considerar observaciones censuradas para la variable respuesta. Esto es relevante, una vez que en varios estudios la respuesta observada está sujeta a límites de detección máximos o mínimos. En este contexto, estudiamos la forma de desarrollar la estimación de los parámetros del modelo con error de medida multivariado y observaciones censuradas considerando la distribución Slash; es decir, estudiamos la forma de desarrollar los procedimientos de estimación e inferencia robusta, en el sentido de utilizar una distribución que acomode de forma más eficiente observaciones outliers de lo que la distribución normal. Con la implementación del algoritmo MCECM, que es una extensión del algoritmo EM, determinamos los estimadores por máxima verosimilitud de los parámetros del modelo de regresión lineal multivariado con errores de medida y respuestas censuradas basado en la distribución Slash multivariada.

Palabras claves: Algoritmo EM, datos censurados, distribución normal independiente, modelos con error de medida, distribución Slash.

Abstract

This paper studies regression models with measurement error, which are useful for describing different phenomena in various areas of knowledge. The possibility of considering censored observations for the response variable is added to this model. This is relevant, since in several studies the observed response is subject to maximum or minimum detection limits. In this context, we studied how to develop the estimation of the model parameters with multivariate measurement error and censored observations considering the Slash distribution; that is, we studied how to develop the estimation and robust inference procedures, in the sense of using a distribution that accommodates more efficiently outliers observations than the normal distribution. With the implementation of the MCECM algorithm, which is an extension of the EM algorithm, we determine the maximum likelihood estimators of the parameters of the multivariate linear regression model with measurement errors and censored responses based on

1 Doctor, área académica de Estadística, Departamento de Matemática y Física.

the multivariate Slash distribution.

Keywords: EM algorithm, censored data, independent normal distribution, measurement error models, Slash distribution.

I. Introducción

Los modelos con error de medida (MEM) son útiles para describir diferentes fenómenos en diversas áreas del conocimiento. Son utilizados para comparar dispositivos de medición que varían en costo, tiempo y eficiencia. Aun cuando haya modelos que consideran la existencia de covariables mal medidas, muchos de ellos no consideran observaciones censuradas para la variable respuesta. Esto es relevante, una vez que en varios estudios la respuesta observada está sujeta a límites de detección máximos o mínimos. En este contexto, estudiamos la forma de desarrollar la estimación de los parámetros del modelo con error de medida multivariado considerando la distribución Slash con observaciones censuradas; es decir, estudiaremos la forma de desarrollar los procedimientos de estimación e inferencia robusta, en el sentido de utilizar una distribución que acomode de forma más eficiente observaciones outliers de lo que la distribución normal. El objetivo fijado es el siguiente: Determinar los estimadores por máxima verosimilitud, a través de la implementación de alguna de las extensiones del algoritmo EM, en el modelo de regresión lineal multivariado con errores de medida y respuestas censuradas basado en la distribución Slash multivariada.

Matos et al. (2018), desarrollaron una técnica para la estimación de parámetros de un modelo multivariante de error de medida y observaciones censuradas considerando la distribución t-Student. Sin embargo, la distribución t-Student forma parte de una familia de distribuciones a la que también pertenece la distribución Slash y sería interesante extender esta idea a esta otra distribución.

En este trabajo investigamos el siguiente problema: ¿Se puede obtener procedimientos de estimación e inferencia robusta en modelos de regresión con errores de medida y datos censurados usando la distribución Slash multivariada?

II. Materiales y métodos

Los modelos de regresión multivariado son herramientas estadísticas utilizadas para analizar la relación entre múltiples variables independientes y una variable dependiente, lo que los hace aplicables en una amplia variedad de campos y situaciones. Entre los usos comunes de los modelos de regresión multivariado tenemos modelos de predicción, de análisis de impacto, de control de calidad, de econometría, de optimización de procesos, entre otros.

Equivalentemente, los modelos de regresión con errores de medida son utilizados cuando las variables independientes o dependientes en un modelo de regresión contienen errores de medición. Es decir, cuando las observaciones de las variables no son perfectamente precisas y contienen algún grado de error aleatorio. Estos modelos se desarrollan para abordar la presencia de errores de medición, lo que puede afectar la validez y la precisión de las estimaciones de los parámetros en un modelo de regresión estándar. Con estos modelos se busca proporcionar estimaciones más precisas y válidas en presencia de estas limitaciones.

Por otra parte, los modelos de regresión con errores de medición y datos censurados son aplicados cuando se enfrenta a situaciones en las cuales las variables de interés están sujetas a errores de medición y, además, algunas observaciones están censuradas. La censura ocurre cuando no todas las observaciones tienen información completa, ya sea porque están truncadas (limitadas en su rango observado) o porque solo se conoce que la variable está por debajo o por encima de cierto umbral.

Entre algunos puntos a considerar en los modelos de regresión con errores de

Investigación N° 32, 2025
ISSN i: 1684-0089 / ISSN v: 2709-8583

medición y datos censurados tenemos los siguientes:

1. Errores de medición: Al igual que en los modelos de regresión con errores de medición estándar, estos modelos buscan corregir el sesgo introducido por los errores de medición en las variables independientes o dependientes.
2. Datos censurados: La censura puede complicar el análisis de datos, ya que algunas observaciones no proporcionan información completa sobre la variable de interés. Los modelos de regresión en este contexto tratan de tener en cuenta la censura al estimar los parámetros del modelo.
3. Métodos de estimación: Para abordar tanto los errores de medición como la censura, se pueden utilizar métodos de estimación robustos. Los métodos de máxima verosimilitud y métodos de estimación por ecuaciones de estimación generalizada (GEE) son comúnmente empleados.
4. Modelos de supervivencia con errores de medición: En el caso de datos censurados, se pueden emplear modelos de supervivencia o modelos de tiempo hasta el evento. Estos modelos se utilizan para analizar el tiempo hasta que ocurre un evento particular, y pueden integrarse con consideraciones de errores de medición.
5. Técnicas de imputación: En algunos casos, se pueden utilizar técnicas de imputación para manejar datos censurados. Estas técnicas imputan valores para las observaciones censuradas basándose en la información disponible y en el modelo de regresión.

Desde que en estos modelos de regresión con errores de medición y datos censurados se tienen situaciones donde se enfrenta a la presencia simultánea de errores de medición y datos censurados, se requieren enfoques estadísticos específicos para manejar ambos desafíos y obtener estimaciones confiables de los parámetros del modelo.

En esta investigación implementamos una de las extensiones del algoritmo EM para obtener los estimadores de máxima verosimilitud, utilizando el software estadístico R (versión 4.4.1 o posterior). de los parámetros del modelo que nos permiten obtener procedimientos de estimación e inferencia robusta en modelos de regresión con errores de medida y datos censurados usando la distribución Slash multivariada.

III. Resultados y discusión

3.1 Distribución normal independiente

La familia de distribuciones normal independente ha sido investigada por varios autores, entre ellos Andrews y Mallows (1974) y Lange y Sinsheimer (1993).

Una distribución normal independente (Lange y Sinsheimer, 1993) o simplemente distribución NI es definida como el vector aleatorio p−dimensional

$$Y = \mu + U^{-1/2} Z, \qquad (1)$$

donde $\mu \in R^p$ es un vector de locación (constante), Z es un vector aleatorio normal con vector de medias **0**, matriz de covarianzas Σ y U es una variable aleatoria positiva con función de distribución acumulada (fda) $H(u,\nu)$ y función de densidad de probabilidad (fdp) $h(u,\nu)$, indexado por el parámetro ν, independiente de Z. Dado U, Y sigue una distribución normal multivariada con vector de medias μ y matriz de covarianzas $u^{-1}\Sigma$, o sea, $Y \mid U = u \sim N_p\left(\mu, u^{-1}\Sigma\right)$. En consecuencia, la fdp de Y es dada por

$$f(y) = \int_0^\infty \phi_p\left(y; \mu, u^{-1}\Sigma\right) dH(u,\nu), \qquad (2)$$

donde $\phi_p(.;\mu,\Sigma)$ representa la fdp de la distribución normal p–variada con vector de medias μ y matriz de covarianzas Σ. Un caso particular de esta distribución es la distribución normal, para el cual $U = 1$.

La familia de distribuciones normal independente incluye modelos tales como las distribuciones t-Student, Slash, normal contaminada, entre otros. Todas estas distribuciones tienen colas más pesadas que de una normal y pueden ser usadas para inferencia robusta.

3.2 Especificación del modelo

Sea $Y_i = \left(Y_{i1},\ldots,Y_{ir}\right)^T$ el vector de respuestas para la i-ésima unidad experimental, donde Y_{ij} es la j-ésima respuesta observada de la unidad $i\left(i=1,...,n, j=1,...,r\right)$. Siguiendo a Barnett (1969), el modelo con error de medida multivariado es formulado como

$$X_i = x_i + \varepsilon_i \quad y \quad Y_i = \alpha + \beta x_i + e_i \tag{3}$$

donde

X_i es el i-ésimo valor observado de la covariable para la unidad i;
x_i es el valor no observado (verdadero);
$e_i = \left(e_{i1},...,e_{ir}\right)^T$ es un vector de errores de medición;
$\alpha = \left(\alpha_1,...,\alpha_r\right)^T$ y $\beta = \left(\beta_1,\ldots,\beta_r\right)^T$ son vectores con los parámetros de regresión.
Sea $\in_i = \left(\varepsilon_i, e_i^T\right)^T$ y $Z_i = \left(X_i, Y_i^T\right)^T = \left(Z_{i1},\ldots,Z_{ip}\right)^T$. Entonces, (3) implica en

$$Z_i = a + bx_i + \varepsilon_i = a + Br_i, \qquad i = 1,\ldots,n, \tag{4}$$

en que $a = \left(0,\alpha^T\right)^T$ y $b = \left(1,\beta^T\right)^T$ son vectores $p\times 1$, con $p = r+1$, $B = \left[b; I_p\right]$ es una matriz $p\times\left(p+1\right)$ donde I_p es la matriz identidad de orden p y $r_i = \left(x_i, \varepsilon_i^T\right)^T$.

De (4), la distribución de Z_i se torna especificada una vez que la distribución de r_i es especificada. Usualmente es utilizada una suposición de normalidad, tal que

$$r_i = \begin{bmatrix} x_i \\ \varepsilon_i \end{bmatrix} \sim N_{1+p}\left(\begin{bmatrix} \mu_x \\ 0_p \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & 0_p^T \\ 0_p & \Omega \end{bmatrix}\right), i = 1,\ldots,n, \tag{5}$$

$0_p = \left(0,\ldots,0\right)^T$ es un vector $p\times 1$, $\Omega = diag\left(\phi_1^2,\ldots,\phi_p^2\right)$.
Bajo normalidad, marginalmente $x_i \sim N\left(\mu_x,\sigma^2\right)$ y $\varepsilon_i \sim N_p\left(0,\Omega\right)$ son independientes $\forall i = 1,...,n$.

Con el fin de obtener una estimación robusta de los parámetros en el modelo, substituimos la normalidad de la suposición (5) por la distribución Slash, quedando

$$r_i = \begin{bmatrix} x_i \\ \varepsilon_i \end{bmatrix} \sim Sl_{1+p}\left(\begin{bmatrix} \mu_x \\ 0_p \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & 0_p^T \\ 0_p & \Omega \end{bmatrix}, \nu\right), i = 1,\ldots,n \tag{6}$$

Utilizando la Ecuación (1), esta formulación implica que

$$\begin{bmatrix} x_i \\ \varepsilon_i \end{bmatrix} | U_i = u_i \sim N_{1+p}\left(\begin{bmatrix} \mu_x \\ 0_p \end{bmatrix}, u_i^{-1}\begin{bmatrix} \sigma_x^2 & 0_p^T \\ 0_p & \Omega \end{bmatrix}\right), U_i \sim Beta\left(\nu,1\right).$$

para $i = 1,\ldots,n$. Como consecuencia,

$$x_i \,|\, U_i = u_i \sim N\left(\mu_x, u_i^{-1}\sigma_x^2\right) y \tag{7}$$

$$\varepsilon_i \,|\, U_i = u_i \sim N_p\left(0_p, u_i^{-1}\Omega\right). \tag{8}$$

Además de eso, $\varepsilon_i \sim Sl_p\left(0,\Omega,\nu\right)$ y $x_i \sim Sl\left(\mu_x,\sigma_x^2,\nu\right)$.
Por la Ecuación (4), $Z_i \sim Sl_p\left(\mu_z,\Sigma_z,\nu\right), i = 1,\ldots,n$, siendo $\mu_z = a + b\mu_x$ y $\Sigma_z = \sigma_x^2 bb^T + \Omega$.

Considerando ahora el modelo con observaciones censuradas, en este caso tenemos

Investigación N° 32, 2025
ISSN i: 1684-0089 / ISSN v: 2709-8583

que la respuesta Z_{ij} no es totalmente observada para todo i, j. Lo que observamos, para cada $i = 1,...,n$, es el vector aleatorio $V_i = \left(V_{i1},...,V_{ip}\right)^T$, tal que si κ_{ij} es un nivel de censura,

$$V_{ij} = \begin{cases} Z_{ij}, \; si \, Z_{ij} > k_{ij} \\ k_{ij}, \; si \, Z_{ij} \leq k_{ij} \end{cases} \qquad (9)$$

El modelo definido por (3), (6) y (9) es denominado Modelo con error de medida estructural y respuestas censuradas basados en la distribución Slash (MEMC-Sl). Por conveniencia, escogimos trabajar con el caso de censura a la izquierda, pero los resultados son fácilmente extendidos para otros tipos de censura.

El $MEMC-Sl$ puede ser formulado em una representación jerárquica flexible que es útil para la obtención de las derivadas. Es obtenida a través de las ecuaciones (4), (7) y (8) y es dado por

$$Z_i \,|\, x_i, U_i = u_i \sim N_p\left(a + bx_i, u_i^{-1}\Omega\right), \qquad (10)$$

$$x_i \,|\, U_i = u_i \sim N\left(\mu_x, u_i^{-1}\sigma_x^2\right), \qquad (11)$$

$$U_i \sim Beta\left(\nu, 1\right), \; i = 1,\ldots,n, \qquad (12)$$

Proposición 1. Considere la representación jerárquica del $MEMC-Sl$ dado en (10)-(12). Entonces,

$$x_i \,|\, U_i = u_i, Z_i = z_i \; N\left(\frac{\mu_x + \sigma_x^2 b^T \Omega^{-1}\left(z_i - a\right)}{1 + \sigma_x^2 b^T \Omega^{-1} b}, \frac{\sigma_x^2}{u_i\left(1 + \sigma_x^2 b^T \Omega^{-1} b\right)}\right).$$

La prueba sigue de la relación $f\left(x_i \,|\, u_i, z_i\right) \propto f\left(z_i \,|\, x_i, u_i\right) f\left(x_i \,|\, u_i\right)$, en que $f\left(\cdot\right)$ denota una fdp genérica.

3.2.1 Distribución Slash multivariada

La distribución Slash multivariada, denotada por $Sl_p\left(\mu, \Sigma, \nu\right)$, puede ser derivada a partir del modelo de mixtura (1) cuando la distribución de U es $Beta\left(\nu, 1\right)$, con $0 < u < 1$ y $\nu > 0$. Su fdp es dada por

$$f\left(y_i\right) = Sl_p\left(y_i; \mu, \Sigma, \nu\right) = \nu \int_0^1 u^{\nu-1} \phi_p\left(y_i; \mu, u^{-1}\Sigma\right) du = \frac{\nu}{\left(2\pi\right)^{\frac{p}{2}} \left|\Sigma\right|^{\frac{1}{2}}} \int_0^1 u^{\frac{p}{2}+\nu-1} e^{\frac{-u\delta_i}{2}} du$$

donde ν es el parámetro de forma y $\delta_i = \left(y_i - \mu\right)^T \Sigma^{-1}\left(y_i - \mu\right)$ es la distancia de Mahalanobis. La fda de Y es denotada por $SL_p\left(. \,|\, \mu, \Sigma, \nu\right)$ y cuando $\nu \uparrow \infty$, la distribución Slash se reduce a la distribución normal.

El vector aleatorio Y admite la representación estocástica

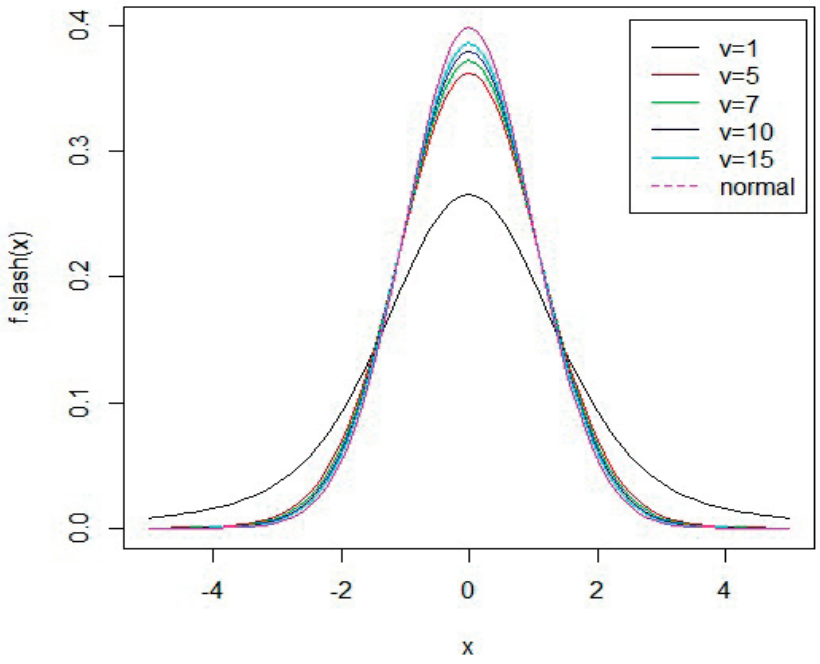
$$Y = \mu + U^{-1/2} Z, \; Z \sim N_p\left(0, \Sigma\right), \; U \sim Beta\left(\nu, 1\right), \qquad (13)$$

donde Z y U son independientes, y $Beta\left(a, b\right)$ denota la distribución beta.

En la Figura 1, podemos observar la variación de las curvas de la distribución Slash para diferentes valores de ν y vemos como ellas se aproximan a la normal cuando el valor de ν crece.

Figura 1

Distribución Slash para diferentes grados de libertad, comparados con la curva normal estándar



A continuación, revisamos las siguientes proposiciones que son importantes en la implementación del algoritmo EM, utilizado en el cálculo de las EMV de los parámetros en este modelo.

Proposición 2. Los momentos recíprocos

$$E\left[U^{-m}\right]=\frac{\nu}{\nu-m}$$

existen para $\nu > m$.

Proposición 3. Suponga que $Y \sim Sl_p\left(\mu,\Sigma,\nu\right)$. Entonces,

i) $E\left[Y\right]=\mu,\nu>\frac{1}{2}$;

ii) $Cov\left[Y\right]=\frac{\nu}{\nu-1}\Sigma,\nu>1.$

Proposición 4. Sea $X \sim Sl_p\left(\mu,\Sigma,\nu\right)$. Si $a \in R^q$ y B es una matriz $q \times p$ con $r\left(B\right)=q$, entonces,

$Y=a+BX \sim Sl_q\left(a+B\mu,B\Sigma B^T,\nu\right)$.

La prueba puede ser encontrada en Fang et al. (1990) para la clase de distribuciones elípticas, clase más general que contiene a las distribuciones normal independientes.

Proposición 5. Sea $YSl_p\left(\mu,\Sigma,\nu\right)$. Considere la partición de Y , μ y Σ como

$$Y=\begin{bmatrix}Y_1\\Y_2\end{bmatrix},\ \mu=\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix}\quad y\quad \Sigma=\begin{bmatrix}\Sigma_{11}&\Sigma_{12}\\\Sigma_{21}&\Sigma_{22}\end{bmatrix}$$

Investigación N° 32, 2025
ISSN i: 1684-0089 / ISSN v: 2709-8583

en que Y_1 y μ_1 son vectores $p_1 \times 1$ y Σ_{11} es una matriz $p_1 \times p_1$. Entonces

i) $Y_1 \sim Sl_{p_1}(\mu_1, \Sigma_{11}, \nu)$;

ii) $Y_2 \mid Y_1 = y_1 \sim Sl_{p_2}(\mu_{2.1}, \Sigma_{22.1}, \nu + p_1)$, donde

$$\mu_{2.1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(y_1 - \mu_1)$$

$$\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

La siguiente definición es importante en el cálculo de la función de verosimilitud.

Definición 1. Sea $Y \sim Sl_p(\mu, \Sigma, \nu)$ y D un conjunto de Borel en R^p. Decimos que el vector aleatorio Z tiene una distribución Slash truncada en D cuando Z tiene la misma distribución que $Y \mid (Y \in D)$. En este caso, la fdp de Z es dada por

$$TSl_p(z \mid \mu, \Sigma, \nu; D) = \frac{Sl_p(z \mid \mu, \Sigma, \nu)}{P(Y \in D)} I_D(z),$$

donde $I_D(.)$ es la función indicadora de D, o sea, $I_D(z) = 1$ si $z \in D$ y $I_D(z) = 0$ en otro caso. Aquí usamos la notación $ZTSl_p(\mu, \Sigma, \nu; D)$. Si D tiene la forma

$$D = \{(x_1, \ldots, x_p) \in R^p; x_1 \le d_1, \ldots, x_p \le d_p\}, \qquad (14)$$

entonces usamos la notación $(Y \in D) = (Y \le d)$, donde $d = (d_1, \ldots, d_p)^T$. En este caso, $P(Y \le d) = Sl_p(d \mid \mu, \Sigma, \nu)$. Note que podemos tener $d_i = +\infty, i = 1, \ldots, p$.

3.3 Función de verosimilitud:

Como en Matos et al. (2018), particionamos Z_i en las componentes observadas y censuradas, $Z_i = vec(Z_i^o, Z_i^c)$, donde $Z_i^o \in R^{p_o}$ está formado por componentes observadas, $Z_i^c \in R^{p_c}$ está formado por componentes censuradas y vec denota la función que apila vectores. De la misma forma, consideramos $V_i = vec(V_i^o, V_i^c)$, $Z_i \sim Sl_p(\mu_z, \Sigma_z, \nu)$, $\mu_z = vec(\mu_z^o, \mu_z^c)$ y $\Sigma_z = \begin{bmatrix} \Sigma_z^{oo} & \Sigma_z^{oc} \\ \Sigma_z^{co} & \Sigma_z^{cc} \end{bmatrix}$, donde κ_i^c es el vector con los correspondientes niveles de censura para Z_i^c. Por la proposición 5, tenemos que

$$Z_i^o \sim Sl_{p_o}(\mu_z^o, \Sigma_z^{oo}, \nu) \; y \; Z_i^c \mid Z_i^o = z_i^o \sim Sl_{p_c}(\mu_z^{co}, \Sigma_z^{cc.o}, \nu + p_o), \qquad (15)$$

donde

$$\mu_z^{co} = \mu_z^c + \Sigma_z^{co}(\Sigma_z^{oo})^{-1}(z_i^o - \mu_z^o), \; y \qquad (16)$$

$$\Sigma_z^{cc.o} = \Sigma_z^{cc} - \Sigma_z^{co}(\Sigma_z^{oo})^{-1}\Sigma_z^{oc}. \qquad (17)$$

La muestra observada para la i-ésima unidad experimental es $\{z_i^o, k_i^c\}$ y la verosimilitud asociada es

$$L_i(\theta) = P(V_i^c = k_i^c \mid Z_i^o = z_i^o) f(z_i^o),$$

en que $f(.)$ es la densidad marginal de z_i^o. Pero $V_i^c = k_i^c$ sí y solamente sí $Z_i^c \le k_i^c$. De esta

forma, por la Ecuación (15) obtenemos que la función de verosimilitud de la i-ésima unidad experimental es:

$$L_i(\theta) = SL_{p_c}(k_i^c \mid \mu_z^{co}, \Sigma_z^{cc.o}, \nu + p_o) Sl_{p_o}(z_i^o \mid \mu_z^o, \Sigma_z^{oo}, \nu),$$

y que puede obtenerse considerando tres casos diferentes:

i) Si el *i*-ésimo individuo no tiene componentes censurados:

$$L_i(\theta) = Sl_p(z_i \mid \mu_z, \Sigma_z, \nu) = \int_0^1 v u^{\nu-1} \phi_p(z_i \mid \mu_z, u^{-1}\Sigma_z) du = \frac{\nu\Gamma\left(\frac{p}{2}+\nu\right) P_1\left(\frac{p}{2}+\nu, \frac{\delta_i}{2}\right)}{(2\pi)^{\frac{p}{2}} |\Sigma_z|^{\frac{1}{2}} \left(\frac{\delta_i}{2}\right)^{\frac{p}{2}+\nu}},$$

que es la fdp de la distribución Slash multivariada en el punto z_i, δ_i es la distancia de Mahalanobis, y $P_x(a,b)$ es la fda de la distribución $Gama(a,b)$, con media $\frac{a}{b}$, evaluada en *x*.

ii) Si el *i*-ésimo individuo tiene solo componentes censurados:

$$L_i(\theta) = SL_p(k_i \mid \mu_z, \Sigma_z, \nu) = \int_{-\infty}^{k_{i1}} \ldots \int_{-\infty}^{k_{ip}} \int_0^1 v u^{\nu-1} \phi_p(z_i \mid \mu_z, u^{-1}\Sigma_z) du dz_{ip} \ldots dz_{i1}.$$

iii) Si el *i*-ésimo individuo tiene componentes censurados y no censurados:

$$L_i(\theta) = SL_{p_c}(k_i^c \mid \mu_z^{co}, \Sigma_z^{cc.o}, \nu + p_o) Sl_{p_o}(z_i^o \mid \mu_z^o, \Sigma_z^{oo}, \nu) =$$

$$\int_{-\infty}^{k_{i1}^c} \ldots \int_{-\infty}^{k_{ip_c}^c} \int_0^1 (v + p_o) u^{\nu+p_o-1} \phi_{p_c}(z_i^c \mid \mu_z^{co}, u^{-1}\Sigma_z^{cc.o}) du dz_{ip_c}^c \ldots dz_{i1}^c \frac{\nu\Gamma\left(\frac{p_o}{2}+\nu\right) P_1\left(\frac{p_o}{2}+\nu, \frac{\delta_i^o}{2}\right)}{(2\pi)^{\frac{p_o}{2}} |\Sigma_z^{oo}|^{\frac{1}{2}} \left(\frac{\delta_i^o}{2}\right)^{\frac{p_o}{2}+\nu}}$$

donde $\delta_i^o = (z_i^o - \mu_z^o)^{\cdot} (\Sigma_z^{oo})^{-1} (z_i^o - \mu_z^o)$ y μ_z^{co} y $\Sigma_z^{cc.o}$ son dados por las ecuaciones (16) y (17), respectivamente.

Desde que el cálculo de la verosimilitud no tiene expresiones de forma cerrada, es aproximada utilizando métodos computacionales (en este trabajo, por su simplicidad, utilizamos la regla del trapecio), basados en el hecho que dado que la fdp de la distribución Slash es $f(z_i) = \int_0^1 v u^{\nu-1} \phi_p(z_i \mid \mu_z, u^{-1}\Sigma_z) du$, con $z_i = (z_{i1}, \ldots, z_{ip})^T$, tenemos que

$$\begin{aligned} P(Z_i \le z_i) &= \int_{-\infty}^{z_{i1}} \ldots \int_{-\infty}^{z_{ip}} \int_0^1 v u^{\nu-1} \phi_p(z_i \mid \mu_z, u^{-1}\Sigma_z) du dz_{ip} \ldots dz_{i1} \\ &= \int_0^1 v u^{\nu-1} \int_{-\infty}^{z_{i1}} \ldots \int_{-\infty}^{z_{ip}} \phi_p(z_i \mid \mu_z, u^{-1}\Sigma_z) dz_{ip} \ldots dz_{i1} du \\ &= \int_0^1 v u^{\nu-1} \phi_p(z_i \mid \mu_z, u^{-1}\Sigma_z) du \end{aligned}$$

donde $\phi_p(.\mid \mu, \Sigma)$ es la fda de la distribución normal p-variada. En este trabajo, la regla del trapecio fue usada con $m = 1000$ particiones del intervalo (0, 1), y para calcular la fda de la distribución normal *p*-variada fue utilizado la librería *mvtnorm* del R, disponible en el CRAN.

La log-verosimilitud asociada a la muestra completa es

$$l(\theta) = \sum_{i=1}^{n} log L_i(\theta)$$

3.4 Algoritmo MCECM:

El algoritmo EM (Dempster et al., 1977) es un popular algoritmo iterativo para calcular estimaciones de parámetros vía máxima verosimilitud en modelos con datos faltantes o en modelos que pueden ser formulados como tal. En circunstancias como las que prevalecen aquí, la maximización de la función de log-verosimilitud con base en los datos observados es difícil de ejecutar debido a la presencia de integrales sin solución analítica.

Además, algunas veces la maximización tiene que ser realizada por bloques del parámetro θ utilizando el algoritmo ECM (Meng y Rubin, 1993). Sin embargo, en algunas aplicaciones del algoritmo EM (o ECM), el paso E no puede ser obtenido analíticamente y debe ser calculado por simulación. Wei y Tanner (1990) propusieron el algoritmo EM Monte Carlo (MCEM), en el cual el paso E es sustituido por una aproximación de Monte Carlo basado en un gran número de simulaciones independientes de los datos faltantes. El algoritmo MCEM puede ser resumido em los siguientes pasos:

Paso E: Calcule la esperanza condicional (en la log-verosimilitud) de los datos faltantes condicionado a las variables observadas y a la estimación $\hat{\theta}^{(k)}$ en la *k*-ésima etapa del algoritmo, vía integración por Monte Carlo, o sea, simule *M* conjuntos de valores para $x, U \vee Z_0, \hat{\theta}^{(k)}$ y calcule

$$Q(\theta \mid \hat{\theta}^{(k)}) = E\left[l_c(\theta \mid Z_c) \mid Z_0, \hat{\theta}^{(k)} \right] \approx \frac{1}{M} \sum_{i=1}^{M} l_c(\theta \mid Z_c).$$

Paso M: Maximizar $Q(. \mid \hat{\theta}^{(k)})$ con relación a θ**,** o sea, obtener

$$\hat{\theta}^{(k+1)} = argmax_{\theta} Q(\theta \mid \hat{\theta}^{(k)}).$$

La proposición siguiente es útil para obtener las esperanzas en el paso E del algoritmo MCECM, que será utilizado para calcular las estimaciones de máxima verosimilitud de los parámetros en el *MEMC–Sl*.

Proposición 6. Para el $MEMC - Sl$,

$$E\left[U_i \mid Z_i = z_i\right] = \frac{p+2v}{\delta_i} \frac{P_1\left(\frac{p}{2}+v+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+v, \frac{\delta_i}{2}\right)},$$

donde δ_i es la distancia de Mahalanobis, que fue definido anteriormente, y $P_x(a,b)$ denota la fda de la distribución $Gama(a,b)$, con media con media $\frac{a}{b}$, evaluada en *x*.

3.4.1 Paso E

Sea $Z = \left(Z_1^T, \ldots, Z_n^T\right)^T$, $x = \left(x_1, \ldots, x_n\right)^T$, $u = \left(u_1, \ldots, u_n\right)^T$ y $\theta = \left(\alpha^T, \beta^T, \mu_x, \sigma_x^2, \phi^T\right)^T$ el vector con todos os parámetros en el modelo. Además de constantes que no dependen de θ, la log-verosimilitud completa asociada a los datos completos $Z_c = (Z, x, u)$ es dada por

$$\begin{aligned} l_c(\theta \mid Z_c) \propto & -\frac{n}{2} \sum_{j=1}^{p} log\phi_j^2 - \frac{1}{2} \sum_{i=1}^{n} u_i \left(Z_i - a - bx_i\right)^T \Omega^{-1} \left(Z_i - a - bx_i\right) \\ & -\frac{n}{2} log\sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^{n} u_i \left(x_i - \mu_x\right)^2. \end{aligned} \tag{18}$$

Suponga que en la *k*-ésima etapa del algoritmo obtenemos una estimación $\hat{\theta}^{(k)}$ de θ. El paso E consiste en el cálculo de la esperanza condicional

$$Q(\theta \mid \hat{\theta}^{(k)}) = E_{\hat{\theta}^{(k)}}\left[l_c\left(\theta \mid Z_c\right) \mid V \right],$$

donde $E_{\hat{\theta}^{(k)}}$ significa que la esperanza está siendo afectada usando $\hat{\theta}^{(k)}$ como el valor verdadero del parámetro y $V = \left(V_1^T, \ldots, V_n^T\right)^T$. El paso M consiste en maximizar $Q(.\mid \hat{\theta}^{(k)})$ en θ. Para hacer eso, observe que la función $Q(.\mid \hat{\theta}^{(k)})$ puede ser descompuesta en

$$Q(\theta \mid \hat{\theta}^{(k)}) = Q_1(\alpha, \beta, \phi \mid \hat{\theta}^{(k)}) + Q_2(\mu_x, \sigma_x^2 \mid \hat{\theta}^{(k)}), \qquad (19)$$

donde $\phi = \left(\phi_1^2, \ldots, \phi_p^2\right)$, y

$$\begin{aligned} Q_1(\alpha, \beta, \phi \mid \hat{\theta}^{(k)}) &= E_{\hat{\theta}^{(k)}}\left[-\frac{n}{2}\sum_{j=1}^{p} log\phi_j^2 - \frac{1}{2}\sum_{i=1}^{n} u_i \left(Z_i - a - bx_i\right)^T \Omega^{-1}\left(Z_i - a - bx_i\right) \mid V \right] \\ Q_2(\mu_x, \sigma_x^2 \mid \hat{\theta}^{(k)}) &= E_{\hat{\theta}^{(k)}}\left[-\frac{n}{2} log\sigma_x^2 - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n} u_i \left(x_i - \mu_x\right)^2 \mid V \right]. \end{aligned} \qquad (20)$$

Dada esta descomposición, podemos reducir el problema a la maximización de dos funciones independentes, buscando puntos críticos de $Q_1(.\mid \hat{\theta}^{(k)})$ y $Q_2(.\mid \hat{\theta}^{(k)})$ separadamente. Expandiendo las expresiones de $Q_1(.\mid \hat{\theta}^{(k)})$ y $Q_2(.\mid \hat{\theta}^{(k)})$ y tomando esperanzas, sigue que

$$\begin{aligned} Q_1(\alpha, \beta, \phi \mid \hat{\theta}^{(k)}) = &-\frac{n}{2}\sum_{j=1}^{p} log\phi_j^2 - \frac{1}{2}\sum_{i=1}^{n} [tr\left(\grave{U}^{-1}\widehat{uz_i^2}\right) - 2a^T\grave{U}^{-1}\widehat{uz_i} - 2\widehat{uxz_i}\grave{U}^{-1}b \\ &+ a^T\grave{U}^{-1}a\widehat{u_i} + 2a^T\grave{U}^{-1}b\widehat{ux_i} + b^T\grave{U}^{-1}b\widehat{ux_i^2}], \end{aligned} \qquad (21)$$

$$Q_2(\mu_x, \sigma_x^2 \mid \hat{\theta}^{(k)}) = -\frac{n}{2} log\sigma_x^2 - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}\left(\widehat{ux_i^2} - 2\mu_x\widehat{ux_i} + \mu_x^2\widehat{u_i}\right),$$

donde $tr(\cdot)$ denota la traza de una matriz,

$$\begin{aligned} \widehat{uz_i^2} &= E[U_i Z_i Z_i^T \mid V_i], & \widehat{uz_i} &= E[U_i Z_i \mid V_i], \\ \widehat{u_i} &= E[U_i \mid V_i], & \widehat{uxz_i} &= E[U_i x_i Z_i^T \mid V_i], \\ \widehat{ux_i} &= E[U_i x_i \mid V_i], & \widehat{ux_i^2} &= E[U_i x_i^2 \mid V_i], \end{aligned}$$

omitiéndose $\hat{\theta}^{(k)}$ para simplificar la notación.

Para obtener expresiones para estas esperanzas, usaremos una propiedad de la esperanza condicional: Se X y Y son vectores aleatorios arbitrarios y $f(\cdot)$ es una función medible, entonces

$$E[E(X \mid Y)\big| f(Y)] = E[X \big| f(Y)]. \qquad (22)$$

Para una prueba, ver Ash (2000, Teorema 5.5.10). Ahora, observemos que por (9), V_i es una función de Z_i. Entonces, por la propiedad de la Ecuación (22), podemos escribir

$$\begin{aligned}\widehat{uz_i^2} &= E[U_i Z_i Z_i^T \mid V_i] = E[E[U_i Z_i Z_i^T \mid Z_i] \mid V_i],\\ \widehat{uz_i} &= E[U_i Z_i \mid V_i] = E[E[U_i Z_i \mid Z_i] \mid V_i] \quad y \\ \widehat{u_i} &= E[U_i \mid V_i] = E[E[U_i \mid Z_i] \mid V_i].\end{aligned} \qquad (23)$$

Utilizando la Proposición 6 sobre esperanza condicional $E[U_i \mid Z_i]$ obtenemos los resultados de las siguientes expresiones para $\widehat{u_i}$, $\widehat{uz_i}$, y $\widehat{uz_i^2}$, considerando tres diferentes casos:

i) El i-ésimo individuo no tiene componentes censurados. Aqui, $V_i = Z_i$, entonces

$$\widehat{u_i} = E[U_i \mid V_i] = E[U_i \mid Z_i] = \frac{p+2v}{\delta_i} \frac{P_1\left(\frac{p}{2}+v+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+v, \frac{\delta_i}{2}\right)},$$

$$\widehat{uz_i} = E[U_i Z_i \mid V_i] = E[U_i Z_i \mid Z_i] = E[U_i \mid Z_i] z_i = \widehat{u_i} z_i \text{, y}$$

$$\widehat{uz_i^2} = E[U_i Z_i Z_i^T \mid V_i] = E[U_i Z_i Z_i^T \mid Z_i] = E[U_i \mid Z_i] z_i = \widehat{u_i} z_i z_i^T,$$

con δ_i siendo la distancia de Mahalanobis y $P_x(a,b)$ es la fda de la distribución $Gama(a,b)$, con media $\frac{a}{b}$, evaluada en el punto x.

ii) El i-ésimo individuo tiene solo componentes censurados.

Por (9), $Z_i \le K_i$, donde K_i es el vector con los niveles de censura para el individuo i. Así, por la definición de la distribución Slash truncada, tenemos que $Z_i \mid (Z_i \le K_i) \sim TSl_p(\mu_z, \Sigma_z, v; D_i)$, donde D_i es como en (14) con $d = \kappa_i$. Entonces tenemos

$$\widehat{u_i} = E\left[E[U_i \mid Z_i] \mid V_i\right] = E\left[E[U_i \mid Z_i] \mid Z_i \le K_i\right] = E\left[\frac{p+2v}{\delta_i} \frac{P_1\left(\frac{p}{2}+v+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+v, \frac{\delta_i}{2}\right)} \mid Z_i \le K_i\right],$$

$$\widehat{uz_i} = E[E[U_i Z_i \mid Z_i] \mid V_i] = E\left[\frac{p+2v}{\delta_i} \frac{P_1\left(\frac{p}{2}+v+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+v, \frac{\delta_i}{2}\right)} Z_i \mid Z_i \le K_i\right],$$

$$\widehat{uz_i^2} = E[E[U_i Z_i Z_i^T \mid Z_i] \mid V_i] = E\left[\frac{p+2v}{\delta_i} \frac{P_1\left(\frac{p}{2}+v+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+v, \frac{\delta_i}{2}\right)} Z_i Z_i^T \mid Z_i \le K_i\right].$$

iii) El i-ésimo individuo tiene componentes censurados y no censurados.

Descomponemos el vector V_i en dos subvectores, Z_i^0 y k_i^c, correspondientes a las observaciones no censuradas y a los niveles de censura, respectivamente. Particionamos el vector Z_i como $Z_i = vec(Z_i^0, Z_i^c)$. Los componentes son censurados si y solamente si $Z_i^c \le k_i^c$.

También tenemos que $Z_i \mid \left(Z_i^c \le k_i^c\right) \sim TSl_p\left(\mu_z, \Sigma_z, \nu; D_i^c\right)$, con

$$D_i^c = \left\{\left(x_1, \ldots, x_p\right) \in R^p;\ x_i \le k_i^c,\ i \in C\right\}, \qquad (24)$$

donde $\mathbb{C}$ es el conjunto de índices para los componentes censurados; por consiguiente, hacemos $d_i = +\infty$ para $i \notin \mathbb{C}$ en (14).

Utilizando la Proposición 5, con Z_i^0 y Z_i^c desempeñando el papel de Y_1 y Y_2, respectivamente, y la Definición 1, tenemos que

$$Z_i^c \mid Z_i^o = z_i^o, Z_i^c \le k_i^c \sim TSl_{p_c}\left(\mu_z^{co}, \Sigma_z^{cc.o}, \nu + p_o; D_i^c\right),$$

donde p_o y p_c son las dimensiones de los vectores Z_i^0 y Z_i^c, respectivamente, y μ_z^{co} y $\Sigma_z^{cc.o}$ son dados en (16) y (17), respectivamente. Así,

$$\widehat{u_i} = E[U_i \mid V_i] = E[E[U_i \mid Z_i] \mid V_i] = E\left[\frac{p+2\nu}{\delta_i} \frac{P_1\left(\frac{p}{2}+\nu+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+\nu, \frac{\delta_i}{2}\right)} \mid Z_i^o = z_i^o, Z_i^c \le k_i^c\right],$$

$$\widehat{uz_i} = E[E[U_i Z_i \mid Z_i] \mid V_i] = E\left[vec(E[U_i \mid Z_i] Z_i^o, E[U_i \mid Z_i] Z_i^c) \mid Z_i^o = z_i^o, Z_i^c \le k_i^c\right] = vec\left(\widehat{u_i} z_i^o, \widehat{uz_i^c}\right),$$

$$\widehat{uz_i^2} = E[U_i Z_i Z_i^T \mid V_i] = E\left[E[U_i \mid Z_i] \begin{bmatrix} Z_i^o Z_i^{oT} & Z_i^o Z_i^{cT} \\ Z_i^c Z_i^{oT} & Z_i^c Z_i^{cT} \end{bmatrix} \mid Z_i^o = z_i^o, Z_i^c \le k_i^c\right] = \begin{bmatrix} \widehat{u_i} z_i^o z_i^{oT} & z_i^o \widehat{uz_i^c}^T \\ \widehat{uz_i^c} z_i^{oT} & \widehat{uz_i^c z_i^{cT}} \end{bmatrix},$$

donde

$$\widehat{uz_i^c} = E\left[\frac{p+2\nu}{\delta_i} \frac{P_1\left(\frac{p}{2}+\nu+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+\nu, \frac{\delta_i}{2}\right)} Z_i^c \mid Z_i^o = z_i^o, Z_i^c \le k_i^c\right] y$$

$$\widehat{uz_i^c z_i^{cT}} = E\left[\frac{p+2\nu}{\delta_i} \frac{P_1\left(\frac{p}{2}+\nu+1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2}+\nu, \frac{\delta_i}{2}\right)} Z_i^c Z_i^{cT} \mid Z_i^o = z_i^o, Z_i^c \le k_i^c\right].$$

Con relación a los restantes valores esperados, tenemos

$$E[x_i U_i \mid V_i = v_i] = \iint x_i u_i \pi(x_i, u_i \mid v_i) dx_i du_i = E[x_i \mid U_i = u_i, V_i = v_i] E[U_i \mid V_i = v_i] \qquad (25)$$

donde π(·) denota una función de densidad de probabilidad genérica.

Por la propiedad de esperanza condicional dada en la ecuación (22), tenemos

$$E\left[x_i \mid U_i, V_i\right] = E\left[E\left[x_i \mid U_i, Z_i\right] \mid U_i, V_i\right].$$

En consecuencia,

$$\widehat{ux_i} = E[\mathrm{E}\left[x_i \mid U_i, \mathrm{Z}_i\right] \mid U_i, V_i] E[U_i \mid V_i] = E\left[\frac{\mu_x + \sigma_x^2 b^T \Omega^{-1}\left(\mathrm{Z}_i - a\right)}{1 + \sigma_x^2 b^T \Omega^{-1} b} \mid U_i, V_i\right] \mathrm{E}[U_i \mid V_i] =$$

$$= \frac{\mu_x \mathrm{E}[U_i \mid V_i] + \sigma_x^2 b^T \Omega^{-1} \mathrm{E}\left[Z_i \mid U_i, V_i\right] \mathrm{E}[U_i \mid V_i] - \sigma_x^2 b^T \grave{\mathrm{U}}^{-1} a \mathrm{E}[U_i \mid V_i]}{1 + \sigma_x^2 b^T \Omega^{-1} b}$$

Pero

$$E\left[\mathrm{Z}_i \mid U_i, V_i\right] E[U_i \mid V_i] = E[U_i \mathrm{Z}_i \mid V_i] \equiv \widehat{uz_i}.$$

Por (2.9) tenemos que $\mu_z = a + b\mu_x$; entonces, $a = \mu_z - b\mu_x$. Así

$$\widehat{ux_i} = \frac{\mu_x \widehat{u_i} + \sigma_x^2 b^T \Omega^{-1} \widehat{uz_i} - \sigma_x^2 b^T \Omega^{-1} (\grave{\imath}_z - b\mu_x) \widehat{u_i}}{1 + \sigma_x^2 b^T \Omega^{-1} b} = \mu_x \widehat{u_i} + \varphi\left(\widehat{uz_i} - \mu_z \widehat{u_i}\right) \qquad (26)$$

donde

$$\varphi = \frac{\sigma_x^2 b^T \grave{\mathrm{U}}^{-1}}{1 + \sigma_x^2 b^T \grave{\mathrm{U}}^{-1} b} \qquad (27)$$

De forma semejante, obtenemos

$$\widehat{ux_i^2} = \Lambda + \mu_x^2 \widehat{u_i} + 2\mu_x \varphi\left(\widehat{uz_i} - \grave{\imath}_z \widehat{u_i}\right) + \varphi\left(\widehat{uz_i^2} - \widehat{uz_i} \grave{\imath}_z^T - \grave{\imath}_z \widehat{uz_i}^T + \grave{\imath}_z \grave{\imath}_z^T \widehat{u_i}\right) \varphi^T. \qquad (28)$$

$$\widehat{uxz_i} = \mu_x \widehat{uz_i}^T + \varphi\left(\widehat{uz_i^2} - \grave{\imath}_z \widehat{uz_i}^T\right) \qquad (29)$$

donde

$$\Lambda = \frac{\sigma_x^2}{1 + \sigma_x^2 b^T \grave{\mathrm{U}}^{-1} b}. \qquad (30)$$

3.4.2. Paso *CM*

Cuando el paso M del algoritmo EM es complicado, este puede ser simplificado realizando el proceso de maximización condicional a alguna función de los parámetros que están siendo estimados. Dada la estimación actual $\theta = \hat{\theta}^{(k)}$ en la *k*-ésima etapa, el paso *CM* del algoritmo *ECM* (Meng y Rubin, 1993) consiste en la maximización condicional de la función *Q* dada en (19). El *ECM* substituye cada paso *M* del algoritmo *EM* de Dempster et al. (1977) por una secuencia de S pasos de maximización condicional, llamados pasos CM, cada uno de los cuales maximiza la función *Q* sobre θ pero con alguna función vectorial de θ , como $\left(g_1(\theta), \ldots, g_s(\theta)\right)$, fijado en su valor anterior. En nuestro caso, por ejemplo, primero maximizamos condicionalmente la función $Q_1(\alpha, \beta, \phi \mid \hat{\theta}^{(k)})$ en (21) sobre α fijando los valores $\beta = \hat{\beta}^{(k)}$ y $\phi = \hat{\phi}^{(k)}$. Entonces maximizamos $Q_1(\alpha, \beta, \phi \mid \hat{\theta}^{(k)})$ sobre β fijando los valores $\alpha = \hat{\alpha}^{(k+1)}$ y $\phi = \hat{\phi}^{(k)}$ y así por delante. Obtenemos las siguientes expresiones cerradas:

$$\hat{\alpha}^{(k+1)} = \overline{z}_u^{(k)} - \overline{x}_u^{(k)} \hat{\beta}^{(k)},$$

$$\hat{\beta}^{(k+1)} = \frac{n\overline{\hat{u}^{(k)}} \sum_{i=1}^n \widehat{uxz_i}^{*\top(k)} - \sum_{i=1}^n \widehat{uz_i}^{*(k)} \sum_{i=1}^n \widehat{ux_i}^{(k)}}{n\overline{\hat{u}^{(k)}} \sum_{i=1}^n \widehat{ux_i^2}^{(k)} - \left(\sum_{i=1}^n \widehat{ux_i}^{(k)}\right)^2},$$

$$\widehat{\phi_1^2}^{(k+1)} = \frac{1}{\mathrm{n}} \sum_{i=1}^n \left(\widehat{uz_{i\,11}^2}^{(k)} - 2\widehat{uxz_{i1}}^{(k)} + \widehat{ux_i^2}^{(k)}\right),$$

$$\widehat{\phi_{j+1}^2}^{(k+1)} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n}(\widehat{uz_i^2}^{(k)}_{(j+1)(j+1)} + \widehat{u_i}^{(k)}\widehat{\alpha_j}^{2(k+1)} + \widehat{ux_i^2}^{(k)}\widehat{\beta_j}^{2(k+1)} +$$
$$2\widehat{ux_i}^{(k)}\widehat{\alpha_j}^{(k+1)}\widehat{\beta_j}^{(k+1)} - 2\widehat{uxz_i}^{(k)}_{(j+1)}\widehat{\beta_j}^{(k+1)} - 2\widehat{uz_i}^{(k)}_{(j+1)}\widehat{\alpha_j}^{(k+1)}),\ j = 1,...,r$$

$$\widehat{\mu_x}^{(k+1)} = \overline{x}_u^{(k)},$$

$$\widehat{\sigma_x^2}^{(k+1)} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n}\left(\widehat{ux_i^2}^{(k)} - 2\widehat{ux_i}^{(k)}\widehat{\mu_x}^{(k+1)} + \widehat{u_i}^{(k)}\widehat{\mu_x}^{2(k+1)}\right),$$

donde $\overline{z}_u^{(k)} = \dfrac{\sum_{i=1}^{n}\widehat{uz_i}^{*(k)}}{\sum_{i=1}^{n}\widehat{u_i}^{(k)}}$, $\overline{x}_u^{(k)} = \dfrac{\sum_{i=1}^{n}\widehat{ux_i}^{(k)}}{\sum_{i=1}^{n}\widehat{u_i}^{(k)}}$ y $\overline{\hat{u}^{(k)}} = \dfrac{1}{\mathrm{n}}\sum_{i=1}^{n}\widehat{u_i}^{(k)}$ con $\widehat{uz_i}^{*(k)} = \left(\widehat{uz}_{i2},\ldots,\widehat{uz}_{ip}\right)^T$ y $\widehat{uxz_i}^{*(k)} = \left(\widehat{uxz}_{i2},\ldots,\widehat{uxz}_{ip}\right)^T$.

3.5 Matriz de información observada

Con el fin de obtener las estimaciones de los errores estándar para el vector de parámetros θ, calculamos antes la matriz de información observada. Bajo algunas condiciones de regularidad, seguimos Lin (2010) para encontrar un método basado en información para obtener la covarianza asintótica de los estimadores de máxima verosimilitud de los parámetros del *MEMC–Sl*. Como fue definido por Meilijson (1989), la matriz de información empírica puede ser calculada como

$$I_c(\theta \mid Z) = \sum_{i=1}^{n} s(Z_i \mid \theta)s^T(Z_i \mid \theta) - \frac{1}{n}S(Z \mid \theta)S^T(Z \mid \theta)$$

IV. Conclusiones

En este trabajo fue estudiado el modelo con error de medida estructural y respuestas censuradas basados en la distribución Slash multivariada.

Para el modelo presentado desarrollamos estimación por máxima verosimilitud e implementamos una extensión del algoritmo EM, denominado algoritmo MCECM, para estimar los parámetros de regresión del modelo estudiado.

Se recomienda estudiar la influencia local en este modelo, haciendo uso de los tipos de perturbación más usuales y también desarrollar metodologías de inferencia bajo el enfoque Bayesiano para el modelo propuesto en este trabajo.

V. Agradecimiento

Agradezco a la Universidad Nacional de San Cristóbal de Huamanga por el apoyo financiero durante el desarrollo de la presente investigación.

VI. Contribuciones de los autores

El autor participó en todas las fases de la investigación.

VII. Referencias

Andrews, D. F. y Mallows, S. L. (1974). Scale mixtures of normal distributions. Journal of the Royal Statistical Society, Series B, 36, 99–102.

Arellano-Valle, R. B., Ozan, S., Bolfarine, H. y Lachos, V. H. (2005). Skew normal measurement error models. Journal of Multivariate Analysis, 98, 265–281.

Ash, R. B. y Doleans-Dade, C. A. (2000). Probability and Measure Theory. Academic Press, San Diego.

Barnett, V. D. (1969). Simultaneous pairwise linear structural relationships. Biometrics, **25**, 129–142.

Bolfarine, H. y Galea-Rojas, M. (1996). On structural comparative calibration under

a t-model. Computational Statistics, 11, 63–85.

Buonaccorsi, J. (2010). Measurement Error: Models, Methods, and Applications. Chapman and Hall/CRC, Boca Raton.

Carroll, R. J., Ruppert, D., Stefanski, L. A. y Crainiceanu, C. M. (2006). Measurement Error in Nonlinear Models. Chapman & Hall/CRC, Boca Raton, second edition.

Cheng, C. L. y Van-Ness, J. W. (1999). Statistical regression with measurement error. Arnold, London.

Chipkevitch, E., Nishimura, R. T., Tu, D. G. S. y Galea-Rojas, M. (1996). Clinical measurement of testicular volume in adolescents: Comparison of the reliability of methods. The Journal of Urology, 156, 2050–2053.

Dempster, A. P., Laird, N. M., y Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B, 39, 1–22.

Dunn, G. (1992). Design and Analysis of Reliability: The statistical evaluation of measurement errors. Edward Arnold, New York.

Fang, K. T., Kotz, S., y Ng, K. W. (1990). Symmetric multivariate and related distribution. Chapman and Hall, London.

Fuller, W. A. (1987). Measurement Error Models. John Wiley and Sons, New York.

Galea-Rojas, M., Bolfarine, H. y Vilca, L. F. (2005). Local influence in Comparative calibration models under elliptical t-distribution. Biometrical Journal, 47, 691–706.

Kelly, G. (1984). The influence function in the errors in variables problem. Annals of Statistics, 12, 87–100.

Lange, K. L. y Sinsheimer, J. S. (1993). Normal/independent distributions and their applications in robust regression. J. Comput. Graph. Stat, 2, 175–198.

Lin, T. I. (2010). Robust mixture modeling using multivariate skew t distributions. Statistics and Computing, 20, 343–356.

Lu, Y., Ye, K., Mathur, A., Hui, S., Fuerst, T. y Genant, H. (1997). Comparative calibration without a gold standard. Statistics in Medicine, 16, 1889–1905.

Matos L. A., Castro L. M., Cabral C. R. B. y Lachos, V. H. (2018) Multivariate measurement error models based on student-t distribution under censored responses. Statistics 52(6):1395–1416. https://doi.org/10.1080/02331888.2018.1527841.

Meilijson, I. (1989). A fast improvement to the EM algorithm on its own terms. Journal of the Royal Statistical Society, Series B, 51, 127–138.

Meng, X. L. y Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika, 80, 267–278.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Wei, G. y Tanner, M. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. J. Am. Stat. Assoc., 85, 699–704.