REVISTA DE LA FACULTAD DE INGENIERIA
www.revistaingenieria.uda.cl

34 (2018) 46-51

ALGORITMO EM PARA LA DISTRIBUCIÓN SLASH MODIFICADA

Francisco Segovia Godoy
Departamento de Matemática
Universidad de Atacama
Copiapó-Chile

RESUMEN

En este artículo implementaremos los estimadores de máxima verosimilitud de la distribución slash modificada introducida por Reyes et al. (2013), mediante el uso del algoritmo "expectation-conditional maximization either" (ECME).

Palabras Claves: Distribución slash modificada, algoritmo ECME, máxima verosimilitud

ABSTRACT

In this article we implement the procedure to obtain the maximum likelihood estimators for the modified slash distribution introduce by Reyes et al. (2013), through the use of the expectation-conditional maximization either (ECME) algorithm.

Keywords: Modified slash distribution, ECME algorithm maximum likelihood.

1. Introducción

La distribución slash canónica, es la distribución del cociente entre dos variables aleatorias independientes, la distribución normal estándar (denotada por $N(0{,}1)$) y la uniforme en el intervalo unitario (denotamos U(0,1)). Su función se densidad es dada por

$$f(x) = \begin{cases} \dfrac{\phi(0) - \phi(x)}{x^2}, & x \neq 0 \\ \dfrac{1}{2}\phi(0), & x = 0 \end{cases}$$

donde $\phi(\cdot)$ denota la función de densidad de la distribución normal estándar.

Esta distribución tiene la característica de presentar colas más pesadas que la distribución normal estándar, permitiéndole obtener una mayor curtosis que el modelo normal. Por otro lado, Genton et al. (2014) introduce la siguiente extensión. Decimos que S tiene distribución slash si su representación estocástica es

$$S = \frac{Z}{U^{1/q}}, \qquad (\mathbf{1})$$

en que $Z \sim N(0{,}1)$, $U \sim U(0{,}1)$, y $q > 0$ es el parámetro de curtosis.

Como se puede observar, una de las principales motivaciones (si es que no es la única) de introducir este tipo de modelos, es el de obtener distribuciones con curtosis altas, para así modelar de mejor manera datos atípicos.

Una extensión del modelo slash fue introducida por Reyes et. al (2013), modificando la distribución de la variable en el cociente en ($\mathbf{1}$). Dicha distribución se presenta en la siguiente definición

Lema 1. Sea $V \sim \exp(2)$ (i.e., la distribución exponencial de radio 2), entonces $T = V^{1/q} \sim Weibull(q, 2)$ con función de densidad

$$f(t; q) = 2qt^{q-1}exp\{-2t^q\} \ ; \ t \geq 0, q > 0$$

Definición 1. Se dice que Y tiene distribución slash modificada, denotamos por $Y \sim MS(\mu, \sigma, q)$, si:

$$Y = \sigma\frac{W}{V^{1/q}} + \mu = \sigma\frac{W}{T} + \mu,$$

donde $W \sim N(0{,}1)$, $V \sim \text{Exp}(2) \Rightarrow T = V^{1/q} \sim Weibull(q, 2)$ y W es independiente de V. $\mu > 0$ es un parámetro de localización, $\sigma > 0$ es un parámetro de escala y $q > 0$ es un parámetro de curtosis.

Algunas extensiones basadas en la distribución slash modificada pueden ser encontradas en la literatura.

La distribución slash modificada Birnbaum-Saunders (MSBS) [4], es una extensión del modelo Birnbaum-Saunders (Saunders et al., 1969) basada en la distribución slash modificada. Se obtiene considerando la siguiente representación estocástica

$$T = \beta\left(\frac{\alpha}{2}X + \sqrt{\left(\frac{\alpha}{2}X\right)^2 + 1}\right)^2,$$

en donde $X \sim MS(0{,}1, q)$

Por otro lado, Reyes et al. (2019) introdujo una generalización llamada distribución slash modificada generalizada (GMS) [5], en donde se consideró la distribución gamma como denominador en lugar de la distribución exponencial.

2. Algoritmo "the expectation-conditional maximization either" (ECME)

En esta sección detallaremos la implementación del algoritmo ECME, quien es una extensión del algoritmo EM (Dempster et al., 1977) para la distribución slash modificada (MS).

2.1. Detalles del algoritmo ECME

Primeramente, consideraremos q conocido. Sobre esta restricción, el modelo jerárquico es dado por

$$Y|T \sim N\left(\mu, \frac{\sigma^2}{T^2}\right) \quad \text{y} \quad T \sim Weibull(q, 2),$$

donde Y es la variable observada y T es la variable latente. Entonces, la función de densidad completa es dada por

$$f(y,t \mid \Psi) = f(y|T,\Psi)\cdot f(t) = \frac{\sqrt{2}}{\sqrt{\pi}} \cdot \frac{1}{\sigma}\exp\left\{-\frac{t^2}{2\sigma^2}(y-\mu)^2\right\} q t^q exp\{-2t^q\}\ ,\ \ y,t>0\,,$$

en que $\Psi = (\mu,\sigma)$ es el vector de parámetros.
Luego, la función de log-verosimilitud completa es dada por

$$l(\Psi\,;\,\boldsymbol{y}\,,\boldsymbol{t}) = -n\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} t_i^2 (y_i-\mu)^2 + d,$$

donde d es una constante que no depende de Ψ.
Por lo que, el valor esperado de $l(\Psi\,;\,\boldsymbol{y}\,,\boldsymbol{t})$ dada la información observada, digamos $Q(\Psi,\Psi^{(k)})$, es dada por

$$Q(\Psi,\Psi^{(k)}) = E[\,(l(\Psi)\,;y_i\,,t_i) \mid y_i,\Psi = \Psi^{(k)}] \propto -nlog(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}(y_i-\mu)^2, \qquad (\mathbf{2})$$

en donde $\widehat{t_i^2}^{(k)} = E[\,t_i^2 \mid y_i,\Psi = \Psi^{(k)}]$.
Ahora, notemos que la función de densidad de $T_i \mid y_i$ es

$$f(t_i \mid y_i\,,\Psi) = \frac{f(t_i,y_i\,;\Psi)}{f(y_i\,;\Psi)} \propto f(t_i\,,y_i\,;\,\Psi)$$

$\propto t_i^q \exp\left\{-\frac{1}{2}\left[\frac{(y_I-\mu)^2}{\sigma^2}t_i^2 + 4t_i^q\right]\right\} = h(t_i\,;\,y_i\,,\Psi)\,,\ t_i > 0, y_i > 0\,,$

por ende, la distribución de $T_i \mid y_i$ tiene función de densidade

$$f(t_i \mid y_i\,,\Psi) = \frac{h(t_i\,;\,y_i\,,\Psi)}{C_i(\Psi)}\ \ ;\ t>0$$

en que $C_i(\Psi) = \int_0^\infty h(t_i\,;\,y_i\,,\Psi)\,dt_i$.

2.2. Paso E

La esperanza requerida en el paso E puede ser calculada como

$$\widehat{t_i^2}^{(k)} = \int_0^\infty t_I^2\, f(t_i \mid y_i\,,\Psi = \Psi^{(k)})\,dt_i\,.$$

2.3. Paso CME-I

A continuación, maximizamos la ecuación ($\mathbf{2}$) en relación a μ, obteniéndose

$$\hat{\mu}^{(k+1)} = \frac{\sum_{i=1}^n \widehat{t_i^2}^{(k)} y_i}{\sum_{i=1}^n \widehat{t_i^2}^{(k)}}$$

2.4. Paso CME-II

Ahora maximizando la ecuación ($\mathbf{2}$) en relación a σ, obteniéndose

$$\hat{\sigma}^{(k+1)} = \left[\frac{1}{n}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}\left(y_i-\mu^{(k+1)}\right)^2\right]^{1/2}.$$

2.5. Paso CME-III

Hasta ahora, hemos asumido que el parámetro q es fijo. Sin embargo, en la práctica desconocemos el valor de q, por lo que también debe ser estimado. Para ello, fijamos μ y σ (utilizando los valores obtenidos en el paso CME-I) y maximizamos una versión perfilada de la función de log-verosimilitud observada en términos de q.

$$l(q\,;\,\boldsymbol{y}) = nlog(2) + nlog(q) - \frac{n}{2}\log(2\pi) - nlog(\sigma) + \sum_{i=1}^{n}\log(C_i(\Psi)).$$

Finalmente, el algoritmo queda como sigue:

<u>Paso E</u>: Dado los valores $\left(\mu^{(k)}\,,\sigma^{(k)}\,,q^{(k)}\right)$, para $i = 1,2,\dots,n$, calcular la esperanza

$$\widehat{t_i^2}^{(k)} = \int_0^\infty t_I^2\, f(t_i \mid y_i\,,\Psi = \Psi^{(k)})\,dt_i\,,$$

<u>Paso CME-I</u>: Actualizar $\mu^{(k)}$ a $\mu^{(k+1)}$, utilizando

$$\hat{\mu}^{(k+1)} = \frac{\sum_{i=1}^n \widehat{t_i^2}^{(k)} y_i}{\sum_{i=1}^n \widehat{t_i^2}^{(k)}}$$

<u>Paso CME-II</u>: Actualizar $\sigma^{(k)}$ a $\sigma^{(k+1)}$, utilizando

$$\hat{\sigma}^{(k+1)} = \left[\frac{1}{n}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}\left(y_i-\mu^{(k+1)}\right)^2\right]^{1/2}.$$

Paso CME-III: Considerando $\mu^{(k+1)}$ y $\sigma^{(k+1)}$, actualizar q maximizando la función de log-verosimilitud observada

$$l(q\,;\,\boldsymbol{y}) = nlog(2) + nlog(q) - \frac{n}{2}\log(2\pi) - nlog(\sigma) + \sum_{i=1}^{n}\log(C_i(\Psi)).$$

Repetir los pasos E, CME-I y CME-II hasta cumplir algún criterio de convergencia, por ejemplo:

$$max\left[\,\left|\,\left(\mu^{(k)}, \sigma^{(k)}, q^{(k)}\right) - \left(\mu^{(k+1)}, \sigma^{(k+1)}, q^{(k+1)}\right)\,\right|\,\right] < \epsilon.$$

En este trabajo consideramos ϵ = 0,0001

3. Algoritmo ECME para regresión

En el contexto de regresión, suele ser bastante conveniente considerar distribuciones distintas a las usuales para los residuos, para así poder adaptar mejor el modelo de regresión a los datos que se disponen.

Sea $Y_i \sim MS(\mu_i, \sigma, q)$, en donde $\mu_i = \beta_0 + \beta_1 z_i$ con z_i covariables conocidas. El desarrollo del algoritmo ECME es exactamente igual al expuesto en la subsección 2.1., pero haciendo $\mu = \mu_i$. Por lo que el valor esperado de $l(\Psi\,;\,\boldsymbol{y}, \boldsymbol{t})$ dada la información observada, digamos $Q\left(\Psi_{\mathrm{i}}, \Psi_{\mathrm{i}}^{(k)}\right)$, es dada por

$$Q\left(\Psi_{\mathrm{i}}, \Psi_{\mathrm{i}}^{(k)}\right) \propto -nlog(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}(y_i - \mu_i)^2 \qquad (\mathbf{3}).$$

Dicho esto, los pasos son del algoritmo son:

3.1. Paso E

La esperanza requerida en el paso E puede calcular como

$$\widehat{t_i^2}^{(k)} = \int_0^\infty t_i^2\, f\left(t_i \mid y_i, \Psi_{\mathrm{i}} = \Psi_{\mathrm{i}}^{(k)}\right) dt_i\,,$$

en que $\Psi_i = (\mu_i, \sigma, q)$ es el vector de parámetros.

3.2. Paso ECM-I

Maximizamos la ecuación $(\mathbf{3})$ en relación a β_0, obteniéndose

$$\widehat{\beta_0}^{(k+1)} = \frac{\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} y_i - \widehat{\beta_1}^{(k)}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} z_i}{\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}}$$

3.3. Paso ECM-II

Maximizamos la ecuación $(\mathbf{3})$ en relación a β_1, obteniéndose

$$\widehat{\beta_1}^{(k+1)} = \frac{\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} z_i y_i - \widehat{\beta_0}^{(k+1)}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} z_i}{\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} z_i^2}$$

3.4. Paso ECM-III

Maximizamos la ecuación $(\mathbf{3})$ en relación a σ, obteniéndose

$$\hat{\sigma}^{(k+1)} = \left[\frac{1}{n}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}\left(y_i - \mu_i^{(k+1)}\right)^2\right]^{1/2}$$

3.5. Paso ECM-IV

Al igual que en la sección 2, para estimar q fijamos μ_i y σ (utilizando los valores obtenidos en el paso CME-I), y maximizamos una versión perfilada de la función de log-verosimilitud observada en términos de q.

$$l(q\,;\,\boldsymbol{y}) = nlog(2) + nlog(q) - \frac{n}{2}\log(2\pi) - nlog(\sigma) + \sum_{i=1}^{n}\log(C_i^*(\Psi_i)).$$

en que $C_i^*(\Psi_i) = \int_0^\infty h(t_i\,;\,y_i, \Psi_i)\,dt$

Finalmente, el algoritmo es:

Paso E: Dado los valores $\left(\beta_0^{(k)}, \beta_1^{(k)}, \sigma^{(k)}, q^{(k)}\right)$, para $i = 1,2,\ldots,n$, calcular la esperanza

$$\widehat{t_i^2}^{(k)} = \int_0^\infty t_i^2\, f\left(t_i \mid y_i, \Psi_{\mathrm{i}} = \Psi_{\mathrm{i}}^{(k)}\right) dt_i\,,$$

Paso CME-I: Actualizar $\beta_0^{(k)}$ a $\beta_0^{(k+1)}$, utilizando

$$\widehat{\beta_0}^{(k+1)} = \frac{\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} y_i - \widehat{\beta_1}^{(k)}\sum_{i=1}^{n}\widehat{t_i^2}^{(k)} z_i}{\sum_{i=1}^{n}\widehat{t_i^2}^{(k)}}$$

Paso CME-II: Actualizar $\beta_1^{(k)}$ a $\beta_1^{(k+1)}$, utilizando

$$\widehat{\beta_1}^{(k+1)} = \frac{\sum_{i=1}^{n} \widehat{t_i^2}^{(k)} z_i y_i - \widehat{\beta_0}^{(k+1)} \sum_{i=1}^{n} \widehat{t_i^2}^{(k)} z_i}{\sum_{i=1}^{n} \widehat{t_i^2}^{(k)} z_i^2}$$

Paso CME-III: Actualizar $\sigma^{(k)}$ a $\sigma^{(k+1)}$ utilizando

$$\hat{\sigma}^{(k+1)} = \left[\frac{1}{n}\sum_{i=1}^{n} \widehat{t_i^2}^{(k)}\left(y_i - \mu_i^{(k+1)}\right)^2\right]^{1/2}.$$

Paso CME-IV: Considerando $\mu_i^{(k+1)}$ y $\sigma^{(k+1)}$, actualizar q maximizando la función de log-verosimilitud observada

$$l(q\,;\,\boldsymbol{y}) = nlog(2) + nlog(q) - \frac{n}{2}\log(2\pi) - nlog(\sigma) + \sum_{i=1}^{n} \log(C_i^*(\Psi_i)).$$

Repetir los pasos hasta cumplir algún criterio de convergencia, por ejemplo:

$$max\left[\left|\,\left(\beta_0^{(k)}, \beta_1^{(k)}, \sigma^{(k)}, q^{(k)}\right) - \left(\beta_0^{(k+1)}, \beta_1^{(k+1)}, \sigma^{(k+1)}, q^{(k+1)}\right)\,\right|\right] < \epsilon.$$

En este trabajo consideramos ϵ = 0,0001

4. Aplicación

Utilizaremos los mismos datos expuestos por Reyes et al. (2013). Las siguientes tablas presentan los estimadores de máxima verosimilitud con sus respectivas desviaciones estándar en paréntesis, y el criterio AIC de ambos sets de estimadores.

4.1. Ilustración 1

Se consideró un conjunto de datos que proviene de un experimento sobre la entomología de ciertas hormigas. Un total de n = 300 hormigas fueron colocadas en el centro de una pista. La medición corresponde a la dirección inicial en la cual decidieron moverse, en relación con un estímulo visual. Los datos fueron inicialmente presentados por Jander [2].

Parámetros	MS (Reyes et al. 2013)	MS^* (EM algorithm)
μ	181.67 (1.22)	181.61 (1.22)
σ	16.7 (1.27)	16.66 (1.24)
q	1.50 (0.09)	1.49 ()
AIC	7921.28	7921.55

4.2. Ilustración 2

Para el conjunto de datos proveniente de Daniel [3], se estudió la regresión lineal entre la altura (cms) y la latencia espinal máxima (CV) de 155 pacientes utilizando el siguiente modelo:

$$ALTURA_i = \beta_0 + \beta_1 * CV_i + \epsilon_i\ , \qquad i = 1,2,\ldots,155$$

asumiendo $\epsilon_i \sim MS(0,\sigma,q)$.

Parámetros	MS (Reyes et al. 2013)	MS^* (EM algorithm)
β_0	70.27(1.45)	69.36(2.11)
β_1	6.188 (0.08)	6.27 (0.16)
σ	5.535 (0.56)	6.13 (0.23)
q	6.01 (0.05)	50.34(3.75)
AIC	1024.09	1016.47

5. Agradecimientos

El autor agradece el financiamiento parcial de la Beca de Magíster en Estadística de la Universidad de Atacama.

6. Referencias

[1] Reyes, J., Gómez, H. W., Bolfarine, H., Gómez, H. (2013). Modified slash distribution, Statistics: A journal of theoretical and applied Statistics, 47:5, 929-941, DOI: 10.1080/02331888.2012.694441

[2] Jander, R. (1957) Die optische Richtungsorientierung der Roten Waldameise (Formica rufa L.) Zeitschrift tiir vergMehende Physiologie Bd, vol 40.

[3] W. Daniel (2004), Biostatistics: A Foundation for Analysis in the Health

Sciences, 8th ed., Johhn Wiley & Sons, Hoboken, NJ.

[4] Vilca, F., Gallardo, D., Gómez G., Gómes. Héctor, Reyes, J. (2016). Modified slash Birnbaum-Saunders distribution, Hacettepe Journal of Mathematics and Statistics, vol 46.

[5] Reyes. J, Inmaculada, B-C, Gómez, H. (2019), Generalized modified slash distribution, Communications in Statistics - Theory and Methods.

[6] Geoffrey, J., Thriyambakam, K. (2008), The EM Algorithm and Extensions, 2th ed, John Wiley & Sons

[7] Meng, X., Rubin, D. (1993). Maximum Likelihood Estimation via the ECM Algorithm: A General Framework. Biometrika, vol. 80

[8] Liu, C., Donald, R. (1994), The ECME Algorithm: A Simple Extension of EM and ECM with Faster Monotone Convergence, Biometrika, vol. 81