Modelos Asimétricos (Skew)

Asignatura: Seminario

Bladimir Morales Torrez

Enero 2022

1 Introducción

Las variables aleatorias son elementos fundamentales en la estadística y en muchas ocasiones posterior al estudio de su comportamiento se adaptan muy bien a funciones de probabilidad ya conocidas o en su defecto se empieza a construir nuevas funciones de probabilidad con sus correspondientes parámetros.

En el contexto actual existe bastante teoría consolidada sobre el modelo de distribución Normal y su implicancia en la teoría de la probabilidad, inferencia estadística (clásica y bayesiana), y así también en modelos de regresión, series de tiempo y otros. Mencionar también que muchas familias de funciones de probabilidad se aproximan a la Normal a medida que un determinado parámetro tiene a un valor apropiado.

Es así que (Azzalini, 1985) define la función de densidad Skew-Normal, propiedades y estadísticos principales, generando posteriormente bastantes investigaciones y artículos referente a esta distribución. En el prefacio del libro (Azzalini, 2013), menciona que la construcción de esta teoría que parte de una función de densidad simétrica y mediante modificaciones adecuadas de la misma conduce a generar un conjunto de distribuciones no simétricas, donde el objetivo es estudiar familias paramétricas flexibles de distribuciones continuas, permitir una posible desviación de la simetría las cuales generen familias de distribuciones más flexibles y realistas.

En el presente documento se mostrará por un lado la formulación del modelo Normal, luego su modificación de parámetros para generar el modelo asimétrica Skew-Normal (Azzalini, 1985) y por otro lado la formulación del modelo t-Stundent y su correspondiente asimetría con el modelo Skew-t (Azzalini & Capitanio, 2002) y (Azzalini & Salehi, 2020), para finalizar mencionando algunas aplicaciones de estos modelos en las diferentes áreas de la teoría estadística.

2 Marco Teórico

2.1 Modelo Normal

El modelo normal es una de las más importantes y conocidos distribuciones de probabilidad continua en el campo estadístico.

2.1.1 Definición

Sea X una variable aleatoria con distribución normal con media μ y varianza σ^2, que se denota como $X \sim N(\mu, \sigma^2)$, con función de densidad dada por:

$$f(x) = (2\pi\sigma^2)^{-\frac{1}{2}} exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\} \qquad ,\forall x \in \mathbb{R} \tag{1}$$

donde los parámetros: $-\infty < \mu < \infty$ y $\sigma^2 > 0$.

2.1.2 Función generadora de momentos

Sea X una variable aleatoria con distribución $N(\mu, \sigma^2)$, entonces la función generadora de momentos es:

$$M_X(t) = \mathbb{E}(e^{tx}) = exp(\mu t + \frac{1}{2}\sigma^2 t^2) \tag{2}$$

2.1.3 Distribución normal estandar

La distribución normal con $\mu = 0$ y $\sigma^2 = 1$ se denomina como la distribución normal estándar. Así la función de densidad es:

$$\varphi(z) = (2\pi)^{-\frac{1}{2}} exp\{-\frac{z^2}{2}\} \qquad , \forall z \in \mathbb{R} \tag{3}$$

y la función de distribución será:

$$\Phi(z) = \mathbb{P}(Z < z) = \int_{-\infty}^{z} \varphi(u)du \qquad , \forall z \in \mathbb{R} \tag{4}$$

Observación. Hacer notar que cuando la tenemos una v.a. $X \sim N(\mu, \sigma^2)$, entonces Z es una variable aleatoria estandarizada $Z = \frac{X-\mu}{\sigma}$ donde $Z \sim N(0,1)$.

2.2 Modelo Skew-Normal

En principio para desarrollar el modelo Skew-Normal se debe tener en cuenta la siguiente proposición.

Proposición Sea f_0 una función de densidad de probabilidad en $\mathbb{R}^d$, sea $G_0(\cdot)$ una función de distribución continua en la recta real y sea $w(\cdot)$ una función de valor real en $\mathbb{R}^d$, tal que:

$$f_0(-x) = f_0(x), \qquad w(-x) = -w(x), \qquad G_0(-y) = 1 - G_0(y)$$

para todo $x \in \mathbb{R}^d,\ y \in \mathbb{R}$. Entonces

$$f(x) = 2f_0(x)G_0\{w(x)\} \tag{5}$$

es una función de densidad en $\mathbb{R}^d$.

2.2.1 Definición

Si por (5) se selecciona a $f_0 = \varphi$ y $G_0 = \Phi$ que representan la función de densidad (3) y la función de distribución (4) respectivamente de la $N(0,1)$ y $w(x) = \alpha x$, para algún valor real α, produce la función de densidad de la distribución Skew-Normal básica (Figura 1):

$$\varphi(x;\alpha) = 2\varphi(x)\Phi(\alpha x) \qquad (-\infty < x < \infty) \tag{6}$$

Ahora se introducirá un parámetro de escala ω y otro de localización ξ. Si Z es una variable aleatoria continua con función de densidad (6), entonces la variable Y queda expresada como:

$$Y = \xi + \omega Z \qquad (\xi \in \mathbb{R}, \omega \in \mathbb{R}^+) \tag{7}$$

Así, Y es llamada variable Skew-Normal(SN) con parámetro de localización ξ, parámetro de escala ω y parámetro de inclinación α. La función de densidad en $x \in \mathbb{R}$ de la forma (6) es:

$$\frac{1}{\omega}\varphi\left(\frac{x-\xi}{\omega};\alpha\right) \equiv \frac{2}{\omega}\varphi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\frac{x-\xi}{\omega}\right) \tag{8}$$

y se escribe:

$$Y \sim SN(\xi, \omega^2, \alpha)$$

donde el cuadrado de ω por analogía corresponde a la notación de $N(\mu, \sigma^2)$ (Figura 2).

Observación. Cuando $\xi = 0$ y $\omega = 1$ se vuelve a la definición básica de la Skew-Normal (6).

2.2.2 Función generadora de momentos

La función generadora de momentos de Y es obtenida por:

$$\begin{aligned}
M_Y(t) &= \mathbb{E}(e^{yt}) = \mathbb{E}(exp\{(\xi + wz)t\}) \\
&= 2exp\{\xi t + \frac{1}{2}\omega^2 t^2\} \int_{-\infty}^{\infty} \varphi(z - \omega t)\Phi(\alpha z) dz \\
&= 2exp\{\xi t + \frac{1}{2}\omega^2 t^2\}\Phi(\delta\omega t)
\end{aligned} \tag{9}$$

donde $\delta = \delta(\alpha) = \frac{\alpha}{\sqrt{1+\alpha^2}}$

Luego se utiliza la función generadora cumulante:

$$K_Y(t) = log\ M_Y(t) = \xi t + \frac{1}{2}\omega^2 t^2 + \zeta_0(\delta\omega t) \tag{10}$$

donde

$$\zeta_0(x) = log\{2\Phi(x)\} \tag{11}$$

y sus derivada $\zeta_r(x) = \frac{d^r}{dx^r}\zeta_0(x)$ con $r = 1, 2, ...$

así, se tiene:

$$\begin{aligned}
\zeta_1(x) &= \frac{d}{dx}\zeta_0(x) = \frac{\varphi(x)}{\Phi(x)} \\
\zeta_2(x) &= \frac{d^2}{dx^2}\zeta_0(x) = \frac{d}{dx}\zeta_1(x) \\
&= \frac{\varphi'(x)\Phi(x) - \varphi(x)^2}{\Phi(x)^2} = \frac{-x\varphi(x)}{\Phi(x)} - \frac{\varphi(x)^2}{\Phi(x)^2} = -x\zeta_1(x) - \zeta_1(x)^2 \\
\zeta_3(x) &= \frac{d^3}{dx^3}\zeta_0(x) = \frac{d}{dx}\zeta_2(x) \\
&= -\zeta_1(x) - x\zeta_1(x)' - 2\zeta_1(x)'\zeta_1(x) \\
&= -\zeta_1(x) - x[-x\zeta_1(x) - \zeta_1(x)^2] - 2[-x\zeta_1(x) - \zeta_1(x)^2]\zeta_1(x) \\
&= -\zeta_1(x) + x^2\zeta_1(x) + x\zeta_1(x)^2 + 2x\zeta_1(x)^2 + 2\zeta_1(x)^3 \\
&= 2\zeta_1(x)^3 + 3x\zeta_1(x)^2 + x^2\zeta_1(x) - \zeta_1(x) \\
\zeta_4(x) &= \frac{d^4}{dx^4}\zeta_0(x) = \frac{d}{dx}\zeta_3(x) \\
&= -6\zeta_1(x)^2\zeta_1(x)' + 3[\zeta_1(x)^2 + 2x\zeta_1(x)\zeta_1(x)'] + [2x\zeta_1(x) + x^2\zeta_1(x)'] - \zeta_1(x)' \\
&= -6\zeta_1(x)^2[-x\zeta_1(x) - \zeta_1(x)^2] + 3\{\zeta_1(x)^2 + 2x\zeta_1(x)[-x\zeta_1(x) - \zeta_1(x)^2]\} + \\
&\qquad \{2x\zeta_1(x) + x^2[-x\zeta_1(x) - \zeta_1(x)^2]\} - [-x\zeta_1(x) - \zeta_1(x)^2] \\
&= -6x\zeta_1(x)^3 - 6\zeta_1(x)^4 + 3\zeta_1(x)^2 - 6x^2\zeta_1(x)^2 - 6x\zeta_1(x)^3 + 2x\zeta_1(x) - x^3\zeta_1(x) - x^2\zeta_1(x)^2 + x\zeta_1(x) + \zeta_1(x)^2 \\
&= -6\zeta_1(x)^4 - 12x\zeta_1(x)^3 - 7x^2\zeta_1(x)^2 + 4\zeta_1(x)^2 - x^3\zeta_1(x) + 3x\zeta_1(x)
\end{aligned} \tag{12}$$

Para todo $\zeta_r(x)$ con $r > 1$ pueden escribirse como funciones de $\zeta_1(x)$ y pontencias de x. Notar que $\zeta_1(x) > 0$, $x + \zeta_1(x) > 0$ y $\zeta_2(x) < 0$.

Utilizando (10), las derivadas de $K_Y(t)$ hasta el cuarto orden son:

$$\mathbb{E}\{Y\} = \xi + \omega\mu_z \tag{13}$$
$$Var\{Y\} = (\omega\sigma_z^2)^2 \tag{14}$$
$$\mathbb{E}\{(Y - \mathbb{E}(Y))^3\} = \frac{1}{2}(4-\pi)(\omega\mu_z)^3 \tag{15}$$
$$\mathbb{E}\{(Y - \mathbb{E}(Y))^4\} = 2(\pi - 3)(\omega\mu_z)^4 \tag{16}$$

Donde $\mu_z = \mathbb{E}(Z) = b\delta$, $\sigma_z^2 = Var(Z) = 1 - \mu_z^2 = 1 - b^2\delta^2$ y $b = \zeta_1(0) = \sqrt{(\frac{2}{\pi})}$

2.2.3 Asimetría

Para la asimetría estandarizando se tiene:

$$\gamma_1\{Y\} = \gamma_1\{Z\} = \frac{\mathbb{E}\{(Y - \mathbb{E}(Y))^3\}}{\omega^3\sigma_z^3} = \frac{4-\pi}{2}\frac{\mu_z^3}{\sigma_z^3}$$

2.2.4 Kurtosis

Para la Kurtosis estandarizando se tiene:

$$\gamma_2\{Y\} = \gamma_2\{Z\} = \frac{\mathbb{E}\{(Y - \mathbb{E}(Y))^4\}}{\omega^4\sigma_z^4} = 2(\pi - 3)\frac{\mu_z^4}{\sigma_z^4}$$

2.3 Modelo t

La familia t de Student es también simétrica a 0 y se emplea habitualmente cuando se necesita regular el grosor de las colas.

2.3.1 Definición

Sea X una variable aleatoria con distribución t con ν grados de libertad, y se denota como $X \sim t_{(\nu)}$, con función de densidad dada por:

$$f(x) = \frac{\Gamma(\frac{1}{2}(\nu+1))}{\sqrt{(\pi\nu)}\Gamma(\frac{\nu}{2})}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \qquad ,\forall x \in \mathbb{R},\ \nu > 0 \tag{17}$$

2.3.2 Relación Modelo t

Se puede igualar una distribución t de Student a una razón de una distribuión $V \sim \frac{\chi^2_\nu}{\nu}$ y una $Z_0 \sim N(0,1)$, siempre y cuando sean ambos independientes. Así se tiene la siguiente igualdad.

$$Z = \frac{Z_o}{\sqrt{V}} \sim t_{(\nu)} \tag{18}$$

2.4 Modelo Skew-t

2.4.1 Definición

Las colas de esta densidad son siempre más pesadas que las de la normal, para ν finito. Por lo tanto, utilizando el mecanismo de la (Proposición 5), se puede producir distribuciones a lo sumo una cola más ligera que la normal, mientras que la otra cola siempre será más pesada para cualquier ν finito. Así se tiene

$$2t(x;\nu)T(\alpha x;\nu) \tag{19}$$

donde $T(\alpha x;\nu)$ denota la función de distribución de (17).

Pero por otro lado también se puede definir de la forma (18), pero en este caso $Z_0 \sim SN(0,1,\alpha)$ y $h(\cdot)$ denotará la función de densidad de V, asi se tiene:

$$\begin{aligned} t(x;\alpha,\nu) &= \int_0^\infty 2\varphi(x\sqrt{t})\Phi(\alpha x\sqrt{t})\sqrt{t}h(t)dt \\ &= \frac{2}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\int_0^\infty e^{-u}u^{\frac{(\nu-1)}{2}}\Phi\left(\frac{\alpha x\sqrt{2u}}{\sqrt{x^2+\nu}}\right)du \\ &= 2t(x;\nu)T\left(\alpha x\sqrt{\frac{\nu+1}{\nu+x^2}};\nu+1\right) \end{aligned} \tag{20}$$

Observación 1. Si $\alpha = 0$ se reduce a la densidad t de Student habitual.

Observación 2. Si $\nu \to \infty$ converge a la $SN(0,1,\alpha)$.

Ahora se introduce un parametro de lozalización y escala, considerando:

$$Y = \xi + \omega Z$$

la función de densidad estará dada por:

$$\frac{1}{\omega}t\left(\frac{x-\xi}{\omega};\alpha,\nu\right) \tag{21}$$

Así se tiene:

$$Y \sim ST(\xi,\omega^2,\alpha,\nu)$$

2.4.2 Momentos

Si se toma la forma de (18), el $m-\acute{e}simo$ momento de Z se expresa como:

$$\mathbb{E}(Z^m) = \mathbb{E}(V^{-\frac{m}{2}})\mathbb{E}(Z_0^m)$$

Los componentes se proporcionana del resultado estándar:

$$\mathbb{E}(V^{-\frac{m}{2}}) = \frac{(\frac{v}{2}^{\frac{m}{2}})\Gamma(\frac{v-m}{2})}{\Gamma(\frac{v}{2})} \qquad ,\text{si } m < v$$

Ahora si se toma las expresiones de $\mathbb{E}(Z_0^m)$ denotadas en la sección 2.2.2, donde ahora se toma:

$$b_\nu = \frac{\sqrt{v}\Gamma(\frac{\nu-1}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})} \qquad ,\nu > 1$$

Luego se tiene:

$$\mathbb{E}(Y) = \xi + \omega b_\nu \delta \qquad ,\nu > 1$$

$$Var(Y) = \omega^2 \left[\frac{\nu}{\nu-2} - (b_\nu\delta)^2\right] = \omega^2\sigma_z^2 \qquad ,\nu > 2$$

$$\gamma_1 = \frac{b_\nu\delta}{\sigma_z^{3/2}}\left[\frac{\nu(3-\delta^2)}{\nu-3} - \frac{3\nu}{\nu-2} + 2(b_\nu\delta)^2\right] \qquad ,\nu > 3$$

$$\gamma_2 = \frac{1}{\sigma_z^4}\left[\frac{3\nu^2}{(\nu-2)(\nu-4)} - \frac{4(b_\nu\delta)^2\nu(3-\delta^2)}{\nu-3} + \frac{6(b_\nu\delta)^2\nu}{\nu-2} - 3(b_\nu\delta)^4\right] - 3 \qquad ,\nu > 4$$

donde $\delta = \delta(\alpha) = \frac{\alpha}{\sqrt{1+\alpha^2}}$, γ_1 corresponde a la asimetría y γ_2 la Kurtosis.

2.5 La función log-verosimilitud

Si y es un valor muestreado de una variable aleatoria $Y \sim SN(\xi, \omega^2, \alpha)$ su función log-verosimilitud es:

$$\ell(\theta^{DP}; y) = constante - log\omega - \frac{(y-\xi)^2}{2\omega^2} + \zeta_0\left(\alpha\frac{y-\xi}{\omega}\right) \tag{22}$$

donde $\theta^{DP} = (\xi, \omega, \alpha)^t$ y $\zeta_0(\cdot)$ esta definida como (11). El superindice DP significa *parametros directos (direct parameters).*

Si $z = \left(\frac{y-\xi}{\omega}\right)$ y $\zeta_1(\cdot)$ es definida como (12) el componente del vector SCORE son:

$$\begin{aligned}
\frac{\partial \ell_1}{\partial \xi} &= \frac{z}{\omega} - \frac{\alpha}{\omega}\zeta_1(\alpha z)\\
\frac{\partial \ell_1}{\partial \omega} &= -\frac{1}{\omega} + \frac{z^2}{\omega} - \frac{\alpha}{\omega}\zeta_1(\alpha z)z\\
\frac{\partial \ell_1}{\partial \alpha} &= \zeta_1(\alpha z)z
\end{aligned} \tag{23}$$

Si se dispone de una muestra aleatoria $y_1, ..., y_n$ de $Y \sim SN(\xi, \omega^2, \alpha)$, la probabilidad logarítmica $\ell(\theta)^{DP}$ se obtiene mediante la suma de n términos del tipo (22) y la correspondiente suma de términos (23) conduce a las ecuaciones de verosimilitud.

$$\begin{gathered}
\sum_i z_i - \alpha\sum_i \zeta_1(\alpha z_i) = 0\\
\sum_i z_i^2 - \alpha\sum_i z_i\zeta_1(\alpha z_i) = n\\
\sum_i z_i\zeta_1(\alpha z_i) = 0
\end{gathered} \tag{24}$$

La presencia de la función no lineal ζ_1 impide la solución de estas ecuaciones, por lo que se emplea métodos numéricos. De la tercera ecuación, la segunda requiere que ξ y *omega* satisfagan

$$\hat{\omega}^2 = \frac{1}{n} \sum_i (y_i - \hat{\xi})^2$$

3 Desarrollo

3.1 Avance teórico en las distribuciones Skew

Las distribuciones mostradas en la sección anterior son bastante utilizadas en diferentes áreas de la estadística, es así que luego de realizar una búsqueda se pudo evidenciar que se esta profundizando la teoría en las siguientes áreas y así también se nombra los nombres de los artículos:

- Distribuciones univariadas-multivariables Skew-Normal.
  - A New Class of Skew Normal Distribution : Tanh-Skew Normal Distribution and its Properties
  - On the bivariate skew-normal and log-skew-normal distributions
  - An Extension of the Truncated-Exponential Skew- Normal Distribution
  - A new class of bimodal skewed distributions
  - The Exponential-Centred Skew-Normal Distribution
  - Generalized alpha skew normal distribution
  - Alpha-skew-normal distribution
  - A Generalized-Alpha–Beta-Skew Normal Distribution with Applications
  - The Balakrishnan skew–normal density
  - The Kumaraswamy skew-normal distribution
- Distribuciones Skew-Elipticas (colas pesadas).
  - A New Robust Class of Skew Elliptical Distributions
  - Bivariate Power-Skew-Elliptical Distribution
  - Multivariate unified skew-elliptical distributions
  - The F statistic of the skew elliptical variables
  - The Skewed-Elliptical Log-Linear Birnbaum–Saunders Alpha-Power Model
- Inferencia estadística
  - Mixtures of skew normal distributions and an efficient EM algorithm to robustly estimate them
  - A Generalization of the Skew-Normal Distribution: The Beta Skew-Normal
  - A new class of bimodal skewed distributions
  - Statistical inference on skew-normal distribution based on order statistics
  - Bayesian Estimation of skewness parametric of skew normal
- Modelos de regresión lineal y no lineal.
  - Regularized Extended Skew-Normal Regression
  - A new Balakrishnan type Skew-Normal distribution
  - Nonlinear mixed-effects model with skew-scale of skew normal distributions
  - Scale Mixture of Skew-Normal Linear Mixed Models
  - A Skew-Normal Mixture Regression Model
  - Bayesian Analysis of Skew Normal Mixture Regression
  - A skew-normal dynamic linear model and Bayesian forecasting
- Series temporales
  - AR(1) model with skew-normal innovations
  - Skew Normal and Skew Student-t distributiones on GARCH(1,1) model

– A Skew-Normal Spatial Simultaneous Autoregressive Model and its Implementation

- Modelo lineal generalizado .
  - GLM bernoulli distribution
  - A Skew-normal copula-driven GLMM
- Geoestadística
- Non-Gaussian geostatistical modeling using (skew) t processes
- On Spatial (Skew) t Processes and Applications

3.2 Aplicación

Para poder mostrar de manera práctica lo visto en secciones anteriores se realizará algunas simulaciones con el software R de las distribuciones (Azzallini, 2017) y así también un modelo de regresión lineal para errores skew-normal, con la librería `sn` que muestra bastantes aplicaciones (Azzallini, 2021).

3.2.1 Programación distribución Skew-Normal

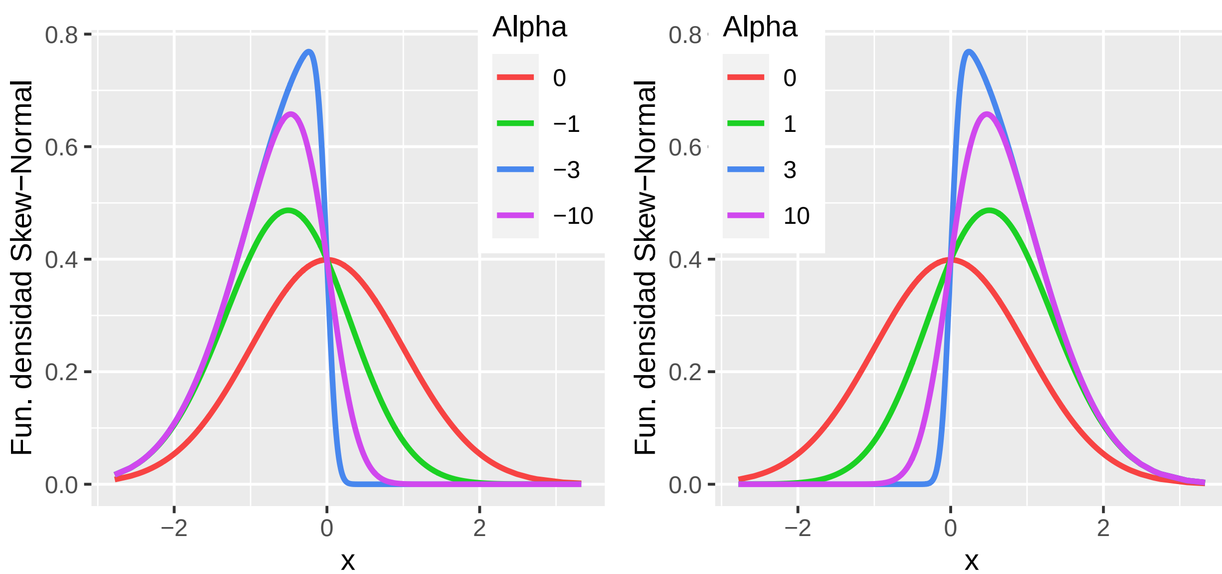


Figure 1: Distribución básica Skew-Normal

3.2.2 Libreria sn

Skew-Normal

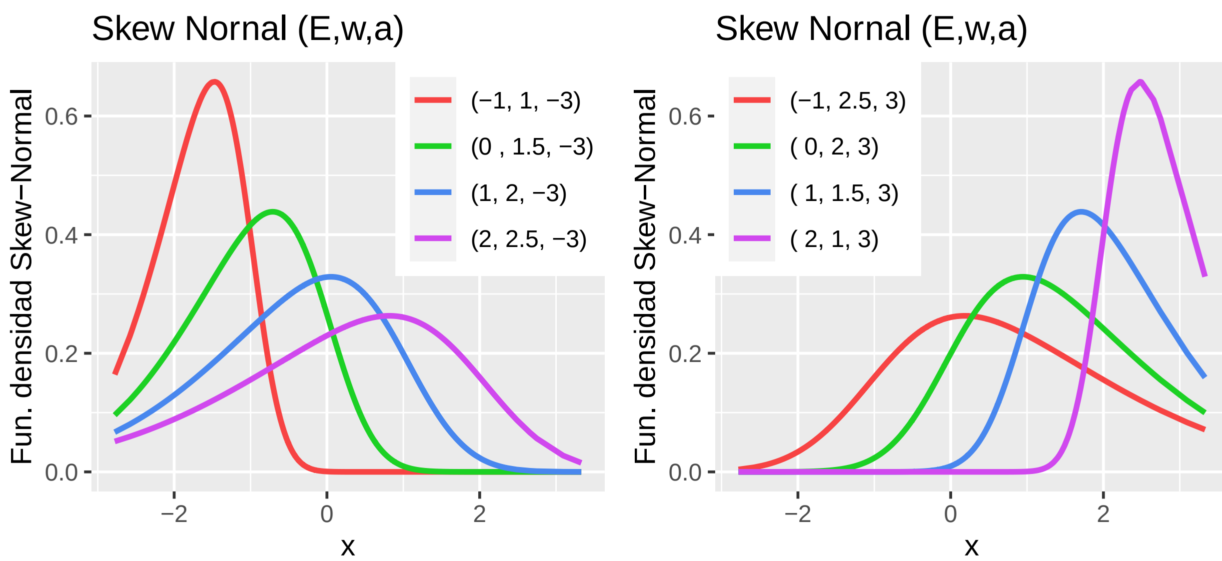
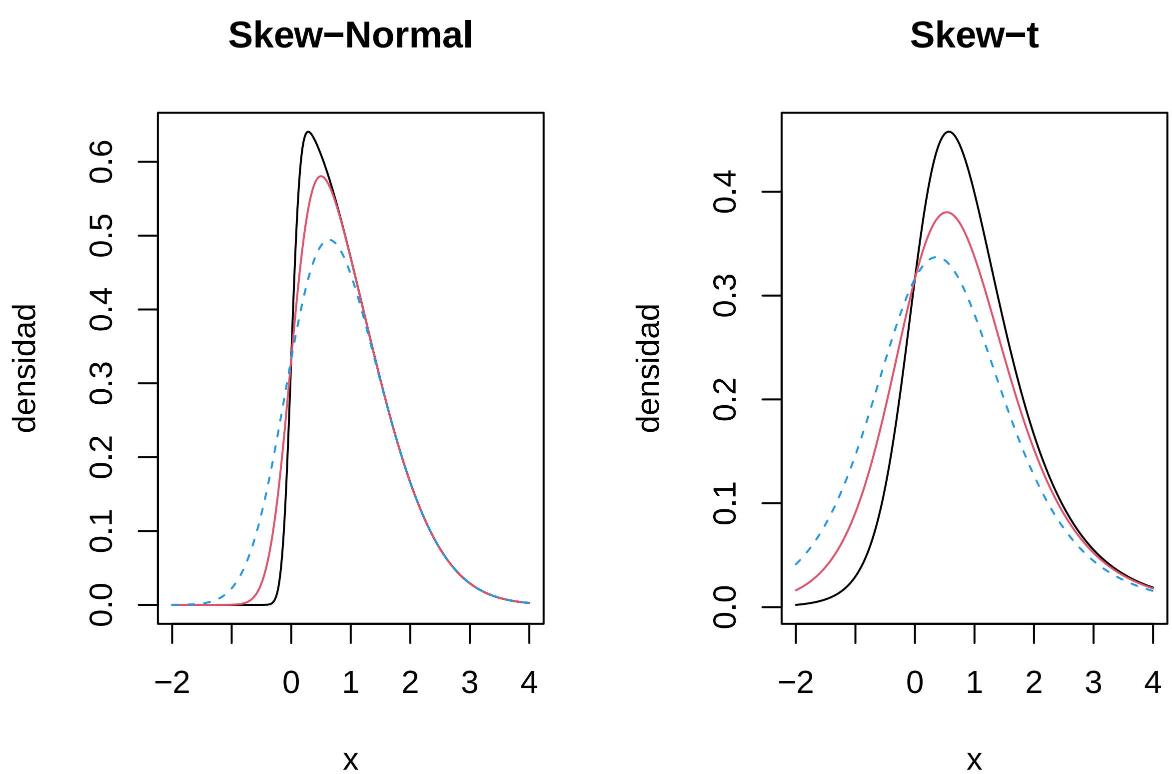


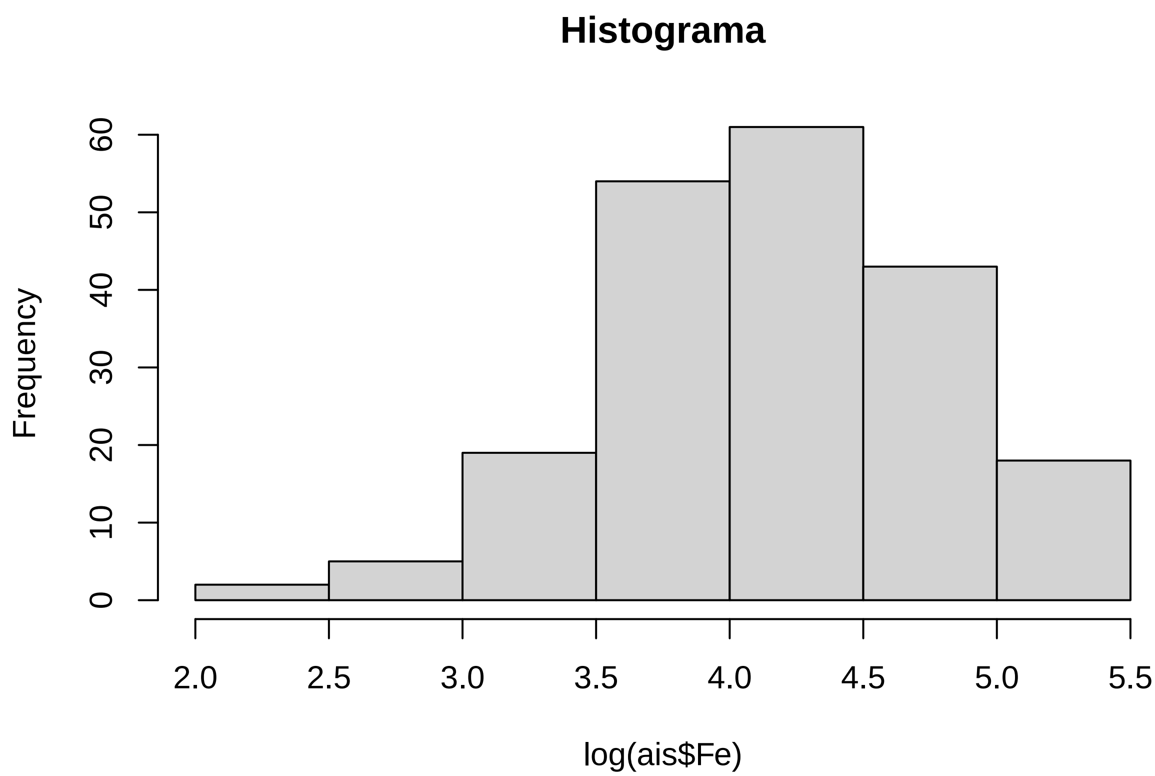
Figure 2: Distribución Skew-Normal con parámetro de localización y escala



3.2.3 Regresión lineal

Datos.

```
data(ais)
hist(log(ais$Fe), main="Histograma")
```



Modelo lineal Skew-Normal

```
mod_sn <- selm(log(Fe) ~ BMI + LBM, family="SN", data=ais)
summary(mod_sn,param.type="DP")
```

```
## Call: selm(formula = log(Fe) ~ BMI + LBM, family = "SN", data = ais)
## Number of observations: 202
## Family: SN
## Estimation method: MLE
## Log-likelihood: -175.9094
## Parameter type: DP
##
## DP residuals:
##      Min       1Q   Median       3Q      Max
## -2.44776 -0.82832 -0.39582 -0.02169  0.90086
##
## Regression coefficients
##                 estimate  std.err  z-ratio Pr{>|z|}
## (Intercept.DP) 3.165958 0.396889 7.976945    0.000
## BMI            0.035270 0.020760 1.698934    0.089
## LBM            0.009465 0.004467 2.118964    0.034
##
```

```
## Parameters of the SEC random component
##       estimate std.err
## omega   0.7212    0.10
## alpha  -1.1195    0.56
```

Modelo lineal Skew-Normal

```r
mod_st <- selm(log(Fe) ~ BMI + LBM, family="ST", fixed.param=list(nu=8), data=ais)
summary(mod_st,param.type="DP")
```

```
## Call: selm(formula = log(Fe) ~ BMI + LBM, family = "ST", data = ais,
##     fixed.param = list(nu = 8))
## Number of observations: 202
## Family: ST
## Fixed parameters: nu = 8
## Estimation method: MLE
## Log-likelihood: -177.661
## Parameter type: DP
##
## DP residuals:
##     Min      1Q  Median      3Q     Max
## -2.2733 -0.6328 -0.2083  0.1607  1.0883
##
## Regression coefficients
##                  estimate  std.err  z-ratio Pr{>|z|}
## (Intercept.DP) 3.018494 0.438440 6.884618    0.000
## BMI            0.031800 0.020511 1.550351    0.121
## LBM            0.010061 0.004391 2.290969    0.022
##
## Parameters of the SEC random component
##       estimate std.err
## omega   0.5633   0.090
## alpha  -0.5542   0.636
```

Modelo Lineal general

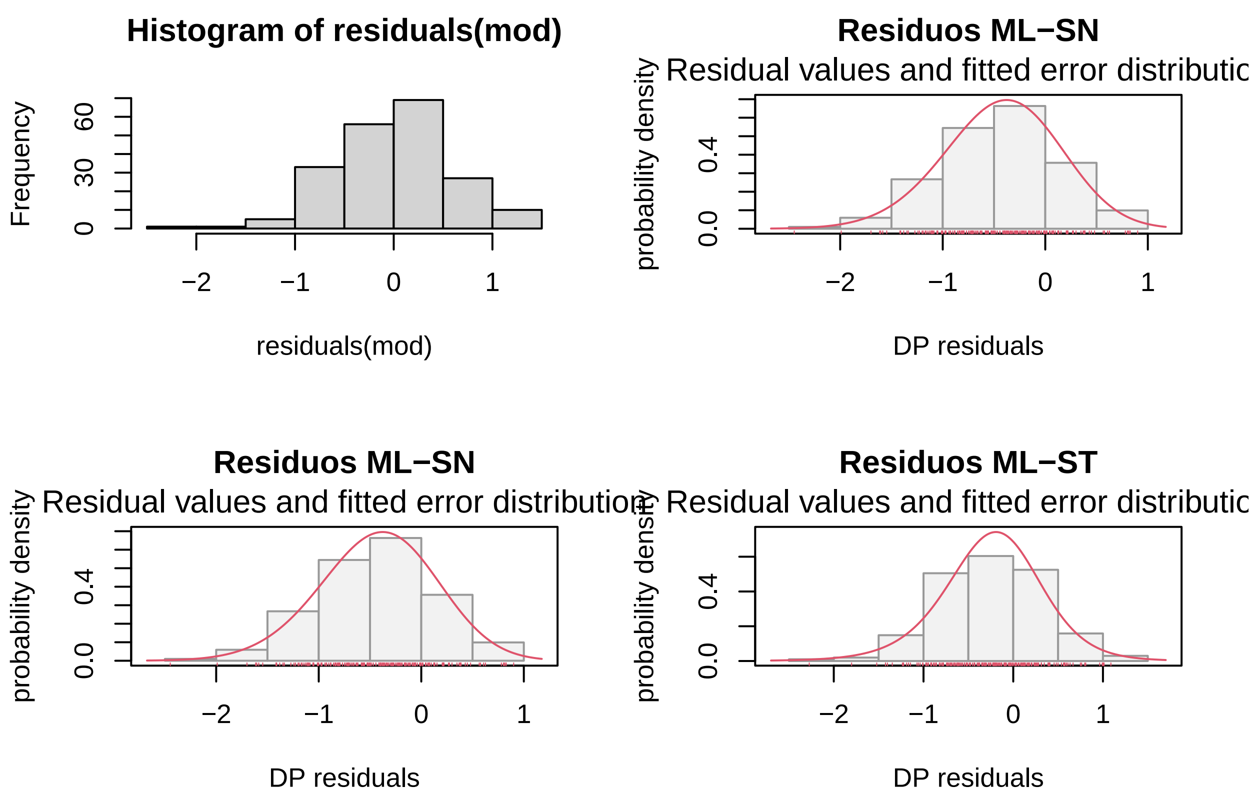
```r
mod <- lm(log(Fe) ~ BMI + LBM, data=ais)
summary(mod)
```

```
##
## Call:
## lm(formula = log(Fe) ~ BMI + LBM, data = ais)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.00158 -0.39884  0.03481  0.41113  1.33198
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.659418   0.335641   7.923 1.61e-13 ***
## BMI         0.039945   0.020535   1.945   0.0532 .
## LBM         0.009004   0.004500   2.001   0.0467 *
```

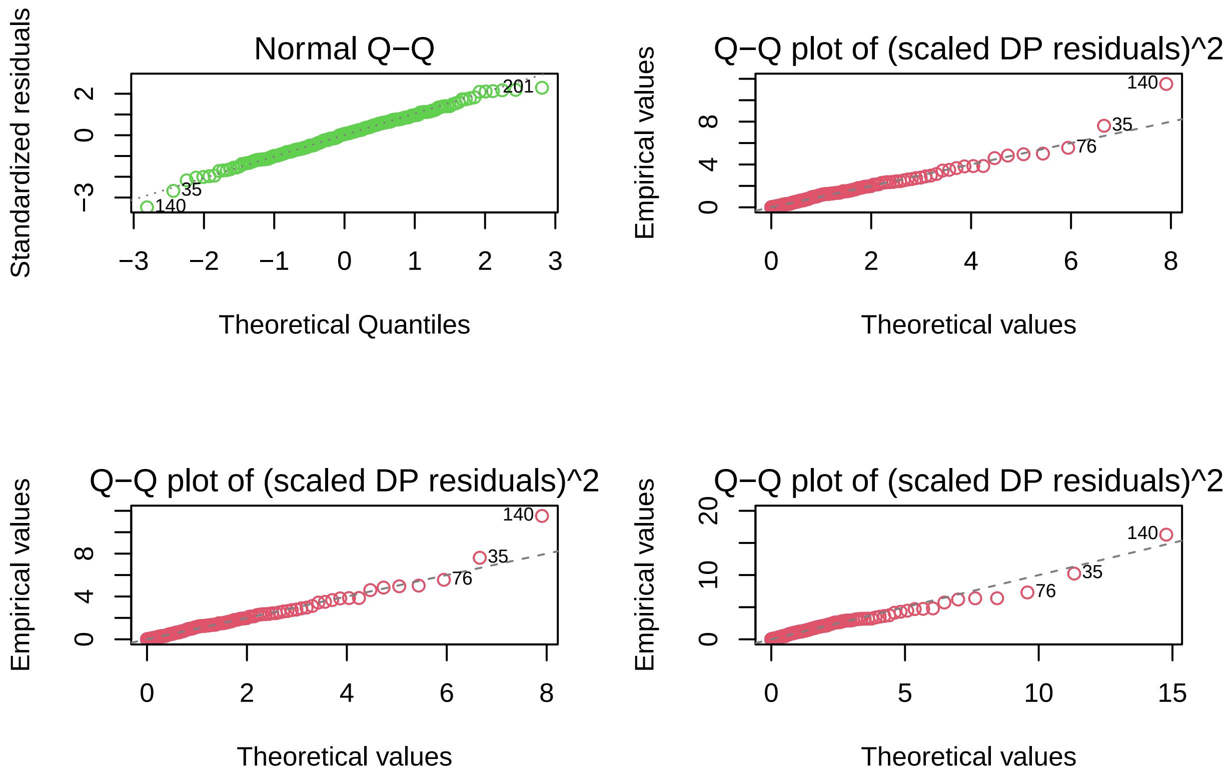
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5839 on 199 degrees of freedom
## Multiple R-squared:  0.1203, Adjusted R-squared:  0.1115
## F-statistic: 13.61 on 2 and 199 DF,  p-value: 2.892e-06
```

Gráficos

```
par(mfrow=c(2,2))
hist(residuals(mod))
plot(mod_sn,col=2,param.type="DP",which=2,main="Residuos ML-SN")
plot(mod_sn,col=2,param.type="DP",which=2,main="Residuos ML-SN")
plot(mod_st,col=2,param.type="DP",which=2,main="Residuos ML-ST")
```




```
par(mfrow=c(2,2))
plot(mod,col=3,which=2)
plot.selm(mod_sn,col=2,which=3,param.type="DP")
plot.selm(mod_sn,col=2,which=3,param.type="DP")
plot.selm(mod_st,col=2,which=3,param.type="DP")
```



4 Conclusiones

Se hizo una revisión bibliográfica bastante extensa de la distribuciones asimétricas (Skew), encontrando así diferentes investigaciones de diferentes autores, donde el principal precursor y creador es Azzalini, el cual desde 1985 presentó esta nueva familia de distribuciones asimétricas. Inicio con la distribución normal estandár, encontrando sus momentos, luego realizando inferencia encontró sus estimadores y tuvo aproximaciones bayesianas. De la misma forma trabajo con la distribución t de Student tomando en cuenta para ambos casos parámetros de centralidad. Pasó luego a la parte multivariable de las distribuciones e investigó sobre las distribuciones elipticas (colas pesadas).

En este trabajo se contextualizó las distribuciones Skew-Normal y Skew-t, mostrando su formulación con parámetros establecidos y que son una generalización de las distribuciones Normal y t de Student respectivamente. Se mostró los momentos de cada una de las distribuciones para finalmente realizar la función log de verosimilitud de la Skew-Normal.

Se pudo también conocer respecto a las áreas de investigación que generó este tipo de distribuciones siendo bastante extensa la literatura, nombrando algunas de las investigaciones ya generadas referente al tema de estudio.

Al finalizar se programo la simulación de la distribución skew-normal con el software R y se indago respecto a la libreria `sn` la cual cuenta con bastantes funciones relacionadas a las distribuciones asimétricas simulando como ejemplos las distribuciones skew-normal y skew-t. También se pudo comparar tres modelos de regresión lineal, el primero modelo lineal general, luego un modelo lineal skew-normal y finalmente un modelo lineal skew-t, de los cuales se pudo verificar de manera visual con el QQ-plot que el que tiene mejor ajuste a los datos según los residuos es el modelo Skew-t.

Bibliografia

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*(2), 171–178.

Azzalini, A. (2013). *The skew-normal and related families*. Cambridge University Press. https://doi.org/10.1017/CBO9781139248891

Azzalini, A., & Capitanio, A. (2002). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society*, *65*(2), 367–389.

Azzalini, A., & Salehi, M. (2020). *Some computational aspects of maximum likelihood estimation of the skew-t distribution*. Springer Nature Switzerland. https://doi.org/10.1007/978-3-030-42196-0_1

Azzallini, A. (2017). *An introduction to the r package sn*.

Azzallini, A. (2021). Package sn, the skew-normal and related distributions such as the skew-t and the SUN. *CRAN*.