Estimación vía método de momentos de los parámetros de la Distribución Lambda Generalizada implementado en MatLab.

Method of moments estimation of the parameters of the Generalized Lambda Distribution implemented in MatLab.

Autor 1 [1] Autor 2 [2]

Recibido: fecha recepción. Revisado: fecha revisado. Aprobado: fecha aprobado.

Resumen

Muchos de los problemas en la actualidad en diferentes áreas del conocimiento tienen solución por medio de la construcción de modelos estadísticos. Dichos modelos se soportan comúnmente en distribuciones de probabilidad y más exactamente en como se distribuyen los datos que se toman de base para su construcción. El interés de ajustar distribuciones a conjuntos de datos pretende describir el comportamiento de ellos mediante la distribución encontrada. El documento presenta una familia de estas distribuciones, la *Distribución Lambda Generalizada* (DLG) y un programa en MatLab implementando el método de momentos para estimar los parámetros de la distribución que se ajuste a conjuntos de datos o aproxime algunas distribuciones conocidas.

Palabras clave: Distribución Lambda Generalizada, Matlab, Método de Momentos.

Abstract

Distributions that ...

Keywords: Keywords1, Keywords2, Keywords3.

[1] Afiliación 1

[2] Afiliación 2

1. Introducción

La Distribución Lambda Generalizada tiene sus orígenes en la distribución lambda de Tukey[1], la cual está definida por la función cuantíl:

$$Q(P) = \begin{cases} \frac{p^{\lambda}-(1-p)^{\lambda}}{\lambda} & si \quad \lambda \neq 0 \\ \frac{log(p)}{1-p} & si \quad \lambda = 0 \quad y \quad p \neq 1 \end{cases} \tag{1}$$

donde $0 \leq p \leq 1$.

Ramberg y Schmeiser en Ramberg & Schmeiser (1974) generalizaron (1) a una familia de distribuciones de cuatro parámetros para generar variables aleatorias utilizando el método de Montecarlo. Además algunas aplicaciones de la DLG se encuentra en el modelamiento de fenómenos biológicos y físicos, Silver (1977), pruebas de bondad de ajuste en regresiones lógicas, Pregibon (1980).

Al igual que cualquier distribución clásica, para estimar los parámetros de la DLG se acude a métodos clásicos de la teoría, el Método de Máxima Verosimilitud (MV), Percentiles (MP), Método de Mínimos Cuadrados (MMC) y el Método de Momentos (MM), entre otros. Ejemplos de cómo proceder pueden consultarse en textos como Wackerly et al. (2010), Casella & Berger (2002), Mood (1950).

En el momento de utilizar algún método de los ya nombrados para la estimación de parámetros en la DLG y en particular obtener por métodos analíticos un estimador que permita calcular dichos parámetros conlleva una labor compleja, por lo tanto los métodos numéricos y los paquetes informáticos son herramientas útiles para aproximarse a ellos. Un software versátil para la implementación de métodos numéricos es MatLab, en particular el GUIDE permite construir programas que implementan estos métodos, es por esta razón que se utilizará para programar el método de momentos para la estimación de parámetros de la DLG.

2. Restricciones para el espacio de parámetros de la DLG

La familia de distribución lambda generalizada con parámetros λ_1, λ_2, λ_3, λ_4, o $DLG(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ es más sencillo especificarla en términos de la función percentil

$$Q(y) = Q(y; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2} \tag{2}$$

donde $0 \leq y \leq 1$.

Los parámetros λ_1 y λ_2 son respectivamente, parámetros de localización y escala, λ_3 y λ_4 determina sesgo y curtosis.

Teorema 2.1. *Para la $DLG(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, la función de densidad de probabilidad es*

$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1}}, \qquad x = Q(y). \tag{3}$$

[1] Nombrada así en honor a Jhon Tukey en el año 1947.

La expresión (2) no siempre especifica una f.d.p. válida. Para que una función $f(x)$ especifique una f.d.p. debe cumplir:

$$f(x) \geq 0 \quad , \quad \int_{-\infty}^{\infty} f(x)dx = 1. \tag{4}$$

La función presentada en (3) junto a las condiciones en (4) implican una DLG válida siempre que

$$\frac{\lambda_2}{\lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1}} \geq 0 \quad , \quad \int_{-\infty}^{\infty} f(Q(y))dQ(y) = 1. \tag{5}$$

La integral en (5) converge a 1 para cualquier valor de λ_1, λ_2, λ_3, λ_4. Ahora la primera condición presenta restricciones sobre el espacio de parámetros λ_1, λ_2, λ_3, λ_4 lo que conlleva al siguiente resultado

Teorema 2.2. *La DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) especifica una distribución válida si y solo si*

$$\frac{\lambda_2}{\lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1}} \geq 0 \tag{6}$$

para todo $y \in [0, 1]$.

En la Figura 1 se presentan algunas gráficas de f.d.p. que ejemplifican el hecho de ser λ_1 un parámetro de localización.

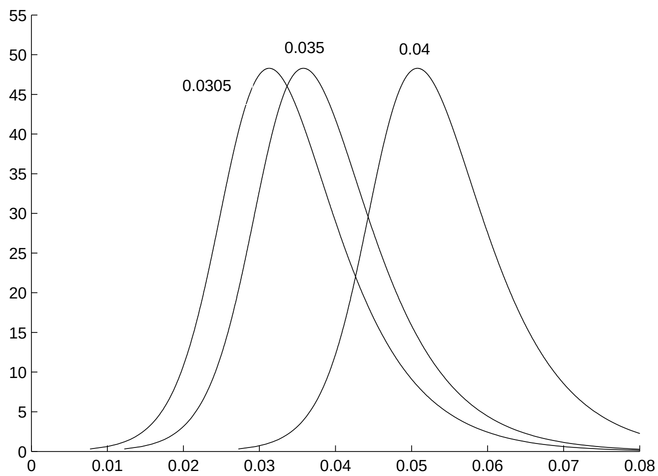


Figura 1: DLG con diferentes λ_1.

En el teorema 2.2 la fracción en (6) es no negativa cuando tanto numerador como denominador tienen el mismo signo, para $0 \leq y \leq 1$, resultado presente en

Corolario 2.3. *La DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) especifica una distribución válida si y solo si*

$$g(y, \lambda_3, \lambda_4) \equiv \lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1} \tag{7}$$

tiene el mismo signo para todo $y \in [0,1]$, siempre que λ_2 tome ese signo también. En particular, la DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) especifica una distribución válida si λ_2, λ_3, λ_4 tienen todos el mismo signo.

La Figura 2 presenta el espacio (λ_3, λ_4) donde la DLG es una función de densidad válida, donde la región designada por X no cumple que (2.3) tenga el mismo signo para todo $y \in [0,1]$.

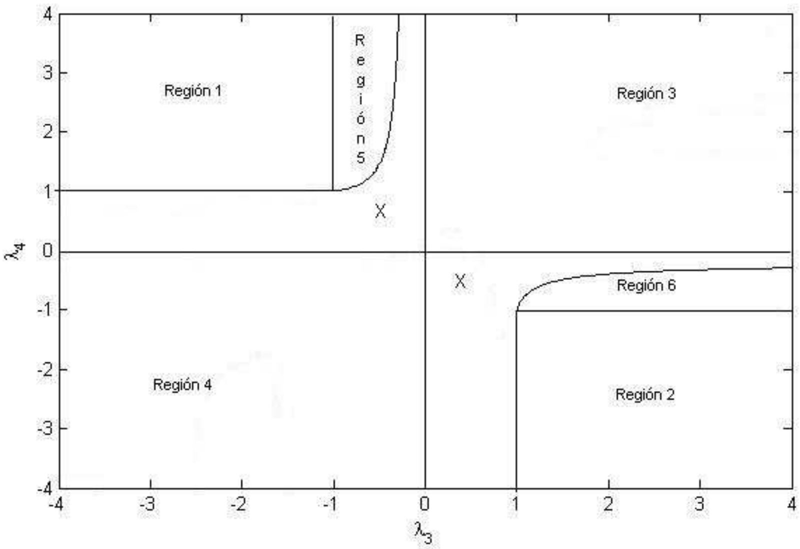


Figura 2: Espacio (λ_3, λ_4) válido.

3. Momentos de la DLG

En Karian (2010) el autor propone especificar los cuatro primeros momentos teóricos, es decir los momentos correspondientes a la distribución teórica, calcula los cuatro primeros momentos muestrales si se trata de un conjunto de datos, o los momentos de la variable aleatoria de interés, para finalmente igualar los momentos teóricos a los momentos muestrales correspondientes, y así halla los valores de los parámetros de la función teórica.

Se procederá entonces a especificar los momentos no centrados de la DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$); y finalmente, se derivaran los momentos centrales de la DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$). Se comenzará por ajustar $\lambda_1 = 0$, para simplificar los análisis siguientes.

Teorema 3.1. *Si X es una variable aleatoria DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$), entonces $Z = X - \lambda_1$ es DLG($0, \lambda_2, \lambda_3, \lambda_4$).*

El teorema anterior muestra que λ_1 es un parámetro de localización; se determinarán ahora los momentos no centrados de la DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$).

Teorema 3.2. *Si Z es DLG($0, \lambda_2, \lambda_3, \lambda_4$), entonces $E(Z^k)$, está dada por*

$$E(Z^k) = \frac{1}{\lambda_2^k} \sum_{i=0}^{k} \left[\binom{k}{i} (-1)^i \beta(\lambda_3(k-i)+1, \lambda_4 i + 1) \right] \tag{8}$$

donde $\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$.

Expresión de la función generadora de momentos de la DLG, puesto que en la función generadora de momentos interviene la función beta $\beta(a,b)$ y esta converge bajo ciertas restricciones en sus argumentos, es así que la expresión (8) presenta restricciones para su convergencia, el siguiente corolario presenta las condiciones para esta.

Corolario 3.3. *Los k-ésimos momentos de la DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) existen si y solamente si $\lambda_3 > -1/k$ y $\lambda_4 > -1/k$.*

Condiciones válidas si $\lambda_3 > -1/k$ y $\lambda_4 > -1/k$, que dependen de la cantidad de momentos, y como el interés es la estimación de los cuatro parámetros de la DLG, la cantidad mínima de ecuaciones que se requieren para utilizar el método de momentos serán cuatro. Es así que se impondra la condición $\lambda_3 > -1/4$ y $\lambda_4 > -1/4$. El siguiete teorema presenta una formula explicita para los primeros cuatro momentos centrales de la DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$).

Teorema 3.4. *Si X es DLG($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) con $\lambda_3 > -1/4$ y $\lambda_4 > -1/4$, entonces los primeros cuatro momentos, α_1, α_2, α_3, α_4 (media, varianza, sesgo, y curtosis, respectivamente), estan dados por*

$$\alpha_1 = \mu = E(X) = \lambda_1 + \frac{A}{\lambda_2}, \tag{9}$$

$$\alpha_2 = \sigma^2 = E\left[(X-\mu)^2\right] = \frac{B - A^2}{\lambda_2^2}, \tag{10}$$

$$\alpha_3 = E(X - E(X))^3/\sigma^3 = \frac{C - 3AB + 2A^3}{\lambda_2^3 \sigma^3}, \tag{11}$$

$$\alpha_4 = E(X - E(X))^4/\sigma^4 = \frac{D - 4AC + 6A^2B - 3A^4}{\lambda_2^4 \sigma^4} \tag{12}$$

donde

$$
\begin{aligned}
A &= \frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4}, && (13)\\
B &= \frac{1}{1+2\lambda_3} + \frac{1}{1+2\lambda_4} - 2\beta(1+\lambda_3, 1+2\lambda_4), && (14)\\
C &= \frac{1}{1+3\lambda_3} - \frac{1}{1+3\lambda_4} - 3\beta(1+2\lambda_3, 1+\lambda_4) + 3\beta(1+\lambda_3, 1+2\lambda_4), && (15)\\
D &= \frac{1}{1+4\lambda_4} + \frac{1}{1+4\lambda_4} - 4\beta(1+3\lambda_3, 1+\lambda_4) \\
&\quad +6\beta(1+2\lambda_3, 1+2\lambda_4) - 4\beta(1+\lambda_3, 1+3\lambda_4). && (16)
\end{aligned}
$$

4. Ajustando la DLG vía método de momentos

Para un conjunto de datos $X_1, X_2, \ldots, X_n$ cualquiera, se pueden calcular algunos valores que describen dicho conjunto, estos valores reciben el nombre de estadísticos y son valores numéricos que dependen solo del conjunto de datos. Se conocen algunos estadísticos importantes como la media o la varianza, conocidos también como primer y segundo momento muestral, para utilizar el método de momentos son necesarios los estadísticos $\hat{\alpha_1}$, $\hat{\alpha_2}$, $\hat{\alpha_3}$ y $\hat{\alpha_4}$, que se hallan mediante

$$
\begin{aligned}
\hat{\alpha_1} &= \bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} && (17)\\
\hat{\alpha_2} &= \hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n} && (18)\\
\hat{\alpha_3} &= \sum_{i=1}^{n} \frac{(X_i - \bar{X})^3}{n\hat{\sigma}^3} && (19)\\
\hat{\alpha_4} &= \sum_{i=1}^{n} \frac{(X_i - \bar{X})^4}{n\hat{\sigma}^4} && (20)
\end{aligned}
$$

y así, estimar los parámetros de la DLG mediante el método de momentos que implica la solución del sistema de ecuaciones:

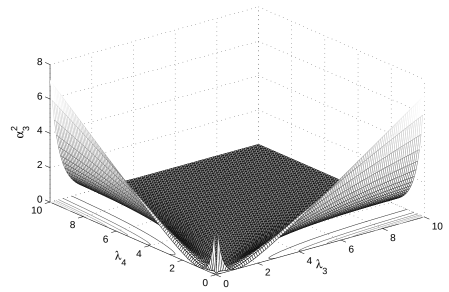
$$\alpha_i = \hat{\alpha}_i \qquad \text{con} \qquad i = 1, 2, 3, 4 \tag{21}$$

para λ_1, λ_2, λ_3 y λ_4.

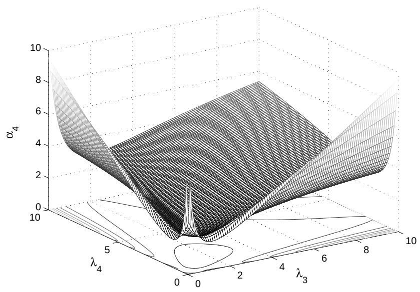
Las ecuaciones (13) a (16) no dependen de λ_1 y λ_2. Además α_3 y α_4 tienen en su denominador el término $\lambda_2\sigma$, por (14) se tiene que $\lambda_2^i\sigma^i = (B - A^2)^{i/2}$ para $i = 3$ y 4, así α_3 y α_4 solo dependen de λ_3 y λ_4. Por tanto, λ_3 y λ_4 se pueden obtener a partir del subsistema:

$$\alpha_3 = \hat{\alpha_3} \quad , \quad \alpha_4 = \hat{\alpha_4}. \tag{22}$$

Hallados los valores λ_3 y λ_4 por medio de (22) se procede a utilizar (14) y (13) para encontrar λ_2 y λ_1 respectivamente. Pero el sistema de ecuaciones (22) es difícil de solucionar por métodos analíticos, por lo que se utilizan métodos numéricos para aproximar su solución. Algunos algoritmos para encontrar una solución numérica a sistemas de ecuaciones son generalmente nombrados de "búsqueda". Se presentará un método de búsqueda para solucionar el sistema (22) a partir de las curvas de nivel de las superficies de la tabla 1.



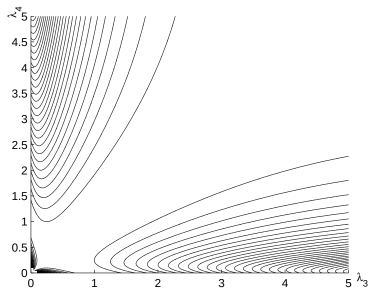
Superficie α_3^2



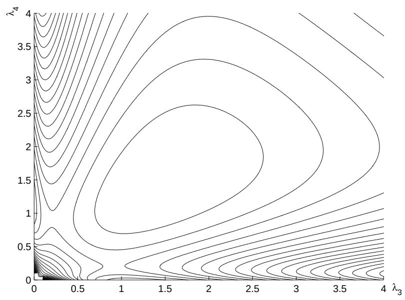
Superficie α_4

Tabla 1: Superficie α_3 y α_4

Cuando se hace $\hat{\alpha_3} = \alpha_3$ se toma la curva de nivel $\hat{\alpha_3}$ de la superficie α_3^2 teniendo así un curva en el plano (λ_3, λ_4) para cada $\hat{\alpha_3}$, de forma similar se hace para $\hat{\alpha_4} = \alpha_4$. La tabla 2 presentan diferentes curvas de nivel para cada una de las superficies α_3^2 y α_4.



Curvas de nivel para α_3^2



Curvas de nivel para α_4

Tabla 2: Curvales de Nivel de α_3^2 y α_4.

Al tomar un conjunto de datos con α_3 y α_4 las soluciones del sistema resultante están dadas por los puntos de intersección de las dos curvas de nivel. Se presentan las curvas $\alpha_3 = 0{,}03$ y $\alpha_4 = 2$, en la Figura 3 que las dos curvas tienen varios puntos de intersección, esto es que no solo un par de valores (λ_3, λ_4) resuelven el sistema y por tanto diferentes DLG se ajustan bien a un mismo conjunto de datos.

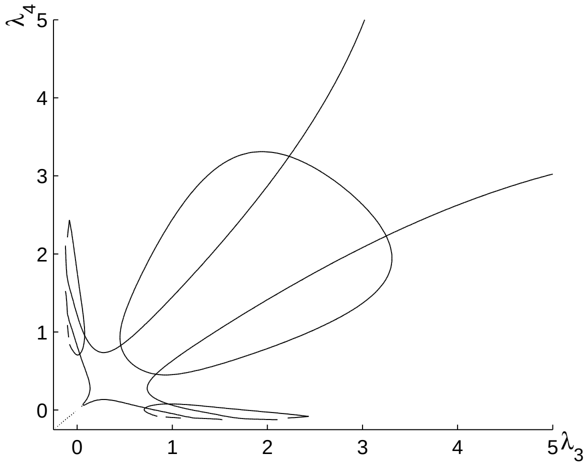


Figura 3: Curvas de nivel $\alpha_3^2 = 0{,}03$ y $\alpha_4 = 0{,}4$.

5. Programa para estimar parámetros de la DLG

5.1. Algoritmo de Busqueda

La solución del sistema (22) está dado por el punto (λ_3, λ_4) que es intersección de las curvas $\alpha_3^2 = \hat{\alpha_3}^2$ y $\alpha_4 = \hat{\alpha_4}$. Una solución numérica para calcular tal punto se obtiene teniendo en cuenta la diferencia entre $\hat{\alpha_i}$ y α_i, $i = 3, 4$, cuanto menor sea esta diferencia, la solución numérica se aproximará más a la solución. El algoritmo implementado para la solución de (22) se describe a continuación.

1. Graficar las curvas para $\alpha_3^2 = \hat{\alpha_3}^2$ y $\alpha_4 = \hat{\alpha_4}$.

2. Tomar un punto $(\lambda_{3_0}, \lambda_{4_0})$ próximo a la intersección de las curvas de (1).

3. Calcular $|\hat{\alpha_i} - \alpha_i|$ $i = 3, 4$ con el punto en (2) y nombrar `Error` al mayor de los dos valores.

4. Construir intervalos $[\lambda_{3_0} - \epsilon, \lambda_{3_0} + \epsilon]$ y $[\lambda_{4_0} - \epsilon, \lambda_{4_0} + \epsilon]$, partiendo de un valor $\epsilon = 0{,}2$, realizar una partición de cada intervalo con distancia $\gamma = 4/1000$.

5. Evaluar cada punto de (4) en

$$\begin{aligned}\alpha_3 &= E(X-E(X))^3/\sigma^3 = \frac{C-3AB+2A^3}{\lambda_2^3\sigma^3}\\ \alpha_4 &= E(X-E(X))^4/\sigma^4 = \frac{D-4AC+6A^2B-3A^4}{\lambda_2^4\sigma^4}\end{aligned}$$

donde A, B, C y D se especifican en 3.4.

6. Para cada punto comparar $|\hat{\alpha_3}-\alpha_3|$ y $|\hat{\alpha_4}-\alpha_4|$ y tomar el mayor de los dos y nombrar `NuevoError`.
7. Llamar $(\lambda_{3_0}, \lambda_{4_0})$ al punto que calculo el menor de todos los `NuevoError`, $\epsilon = \gamma$ y hacer `Error` igual a `NuevoError`.
8. Repetir tres veces más pasos (3) al (7) con punto (7).
9. $(\hat{\lambda_3}, \hat{\lambda_4}) = (\lambda_{3_0}, \lambda_{4_0})$.

Al hacer una red en el punto (4) puede que los puntos sobre los que se calculan los errores no aproximen bien la solución y hacer una red fina con el fin de evitar el problema puede resultar en un largo tiempo de computación. Por ejemplo, pasar de una red 10×10 a una 25×25 aumenta el tiempo de computación por más de un factor 6 y pasar a una 250×250 aumenta el factor en un 625. Si el original 10×10 requiere 5 segundos de tiempo de computación, la red 250×250 requerirá 52 minutos de computación.

Un algoritmo más aceptables es buscar un intervalo próximo a la solución y subdividir tanto como se desee. Suponga que una búsqueda para $(\alpha.\beta)$ conduce a un intervalo $[A_1, A_2] \times [B_1, B_2]$. En lugar de construir una red 10000×10000, utilizar el algoritmo descrito implica que en la primera iteración se reduce el área de búsqueda en un 0.04 % a la cuarta iteración se tendrá un mejor efecto que la red 10000×10000 y el tiempo de computación será menor, por lo tanto se toman un total de cuatro iteraciones en el paso (9).

El punto $(\hat{\lambda_3}, \hat{\lambda_4})$ es una solución numérica para el sistema $\alpha_3^2 = \hat{\alpha_3}^2$ y $\alpha_4 = \hat{\alpha_4}$, con un *error absoluto* Mathews & Fink (1999) igual a `Error`. Los puntos finales $\hat{\lambda_1}$, $\hat{\lambda_2}$, $\hat{\lambda_3}$, $\hat{\lambda_4}$ se presentan como solución numérica del sistema (21) y son utilizados para construir la gráfica $(x, f(x))$ donde $f(x)$ es la función de densidad de probabilidad de la DLG$(\hat{\lambda_1}, \hat{\lambda_2}, \hat{\lambda_3}, \hat{\lambda_4})$.

Como se observa en la Figura 3, la solución al sistema presentado en (22) puede tener varias soluciones, se tomara la intersección más próxima al eje λ_4. Se construye un programa llamado “Seleccion” utilizando como herramienta el GUIDE de MATLAB en el que se implementó el algoritmo. A continuación se presentan algunas imágenes y comentarios del programa de tal manera que expliquen su funcionamiento.

5.2. Programa

En Matlab se construye programa llamado “Seleccion”, dicho programa comienza con la ventana

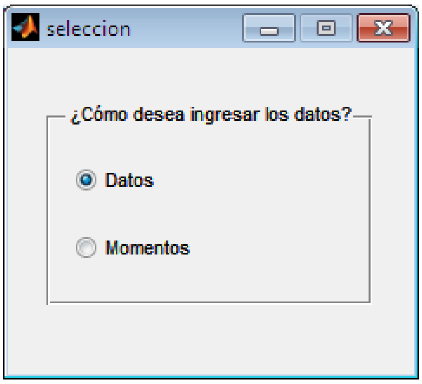


Figura 4: Programa Seleccion

En la que se elige la forma de como se ingresan los momentos muestrales $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_3$ y $\hat{\alpha_4}$, si se requiere que el programa los calcule a partir de un conjunto de datos o si el mismo usuario los ingresa. Si la elección es que el programa los calcule se selecciona “Datos” y emerge la Figura 5.

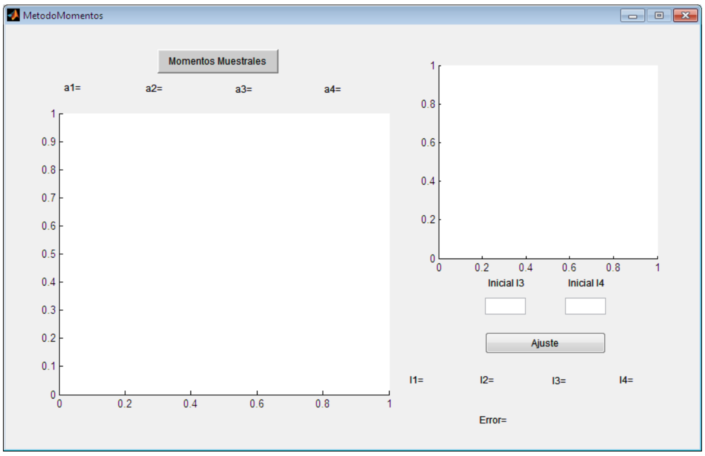


Figura 5: Método Momentos

En ella se da click en el botón “Momentos Muestrales”, se busca el archivo .txt con los datos a describir mediante la DLG y abrir. La Figura 6 presenta lo anterior para el archivo datosnormales.txt que contiene un conjunto de 1000 datos con distribución Normal Estándar. Ver Anexo B.

En los espacios “Inicial l3” e “Inicial l4” se ingresan los valores de un punto de la forma (λ_3, λ_4) cercano a una intersección. En la Figura 7 se observa que los valores iniciales para comenzar el algoritmo son $\lambda_{3_0} = 0{,}2$ y $\lambda_{4_0} = 0{,}25$, obteniendo los valores $\hat{\lambda_1} = -0{,}111667$, $\hat{\lambda_2} = 0{,}284952$, $\hat{\lambda_3} = 0{,}197614$, $\hat{\lambda_4} = 0{,}256231$, el error absoluto de

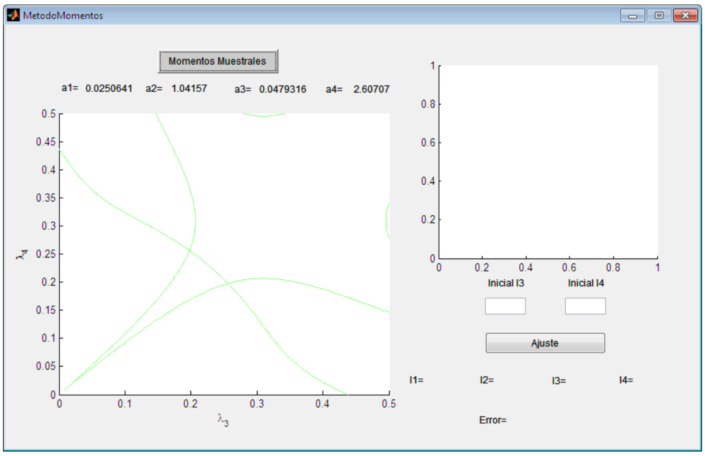


Figura 6: Curvas $\alpha_3 = 0{,}0479316$ y $\alpha_4 = 2{,}60707$.

$8{,}12248e^{-10}$ y la gráfica de la función de densidad de probabilidad de una DLG(-0.111667,0.284952,0.197614,0.256231).

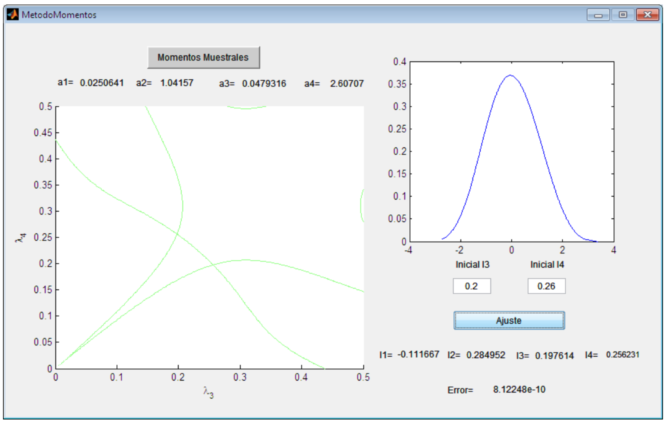


Figura 7: Estimaciones del conjunto de datos.

Si el usuario decide ingresar los momentos directamente, se selecciona "Momentos" y la ventana emergente es la Figura 8.

Se ingresan los momentos en los espacios $a1$, $a2$, $a3$ y $a4$ y se da click en el botón "Graficar", obteniendo las curvas $\alpha_3 = a3$ y $\alpha_4 = a4$. En la Figura 9 se observan las curvas para $a3 = 0{,}01$ y $a4 = 1{,}8$.

Se ingresa el valor "Inicial l3" e "Inicial l4" próximos a la intersección de las curvas, y se da click en Ajuste para obtener los parámetros de la DLG, el error de la solución y la gráfica de $(x, f(x))$ que es la función de densidad de probabilidad con los parámetros estimados mostrado en la Figura 10.

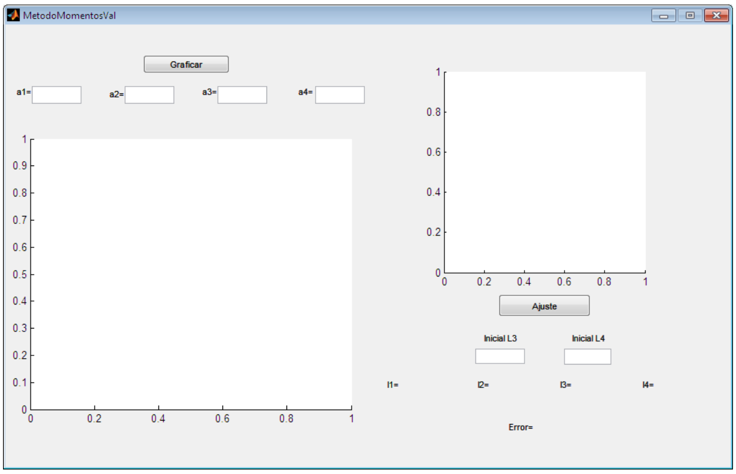


Figura 8: Momentos

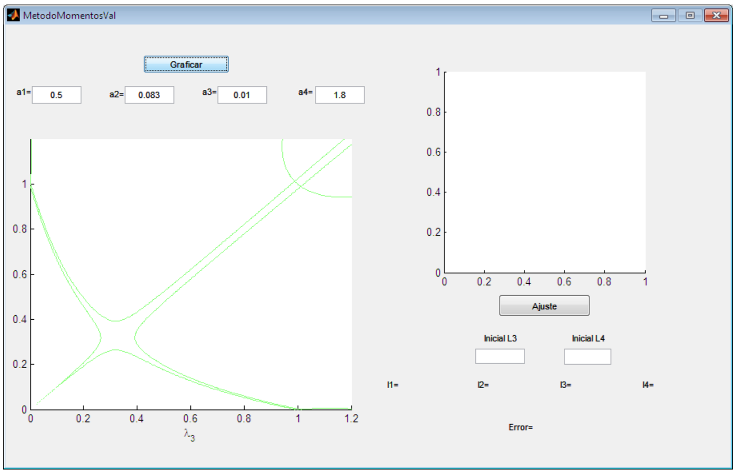


Figura 9: Curvas α_3^2 y α_4 por Momentos

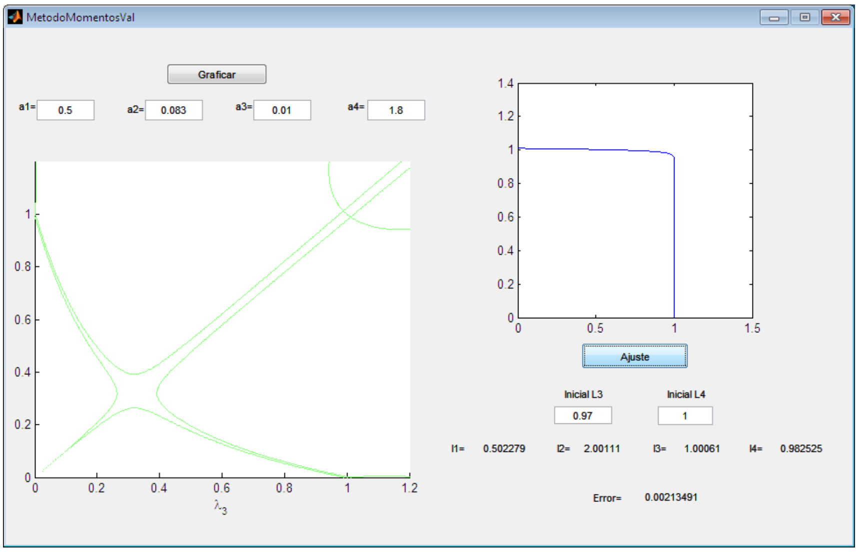


Figura 10: Ajuste a partir de conjunto de datos

6. Resultados

Se tomaron tres ejemplos para variables aleatorias con diferente distribución. El primer ejemplo se tomó con un conjunto de 1000 datos generados con Excel cuya distribución es normal estándar, la Figura 7 presenta las dos curvas $\alpha_3 = \hat{\alpha_3}$ y $\alpha_4 = \hat{\alpha_4}$ y la gráfica de la funciòn de densidad de probabilidad de una DLG para este conjunto de datos. La Figura 11 muestra la función de densidad de probabilidad de la DLG(-0.111667,0.284952,0.197614,0.256231) y el histograma de frecuencias relativas del conjunto de datos.

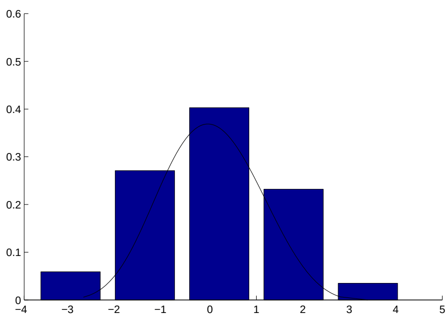


Figura 11: Histograma

El segundo ejemplo es tomado de un conjunto de datos extraídos de Karian (2010), se presenta a continuación los parámetros estimados por el programa “Seleccion”.

$$DLG(41{,}7897, 0{,}0113449, 0{,}0985255, 0{,}36058)$$

$$DLG(41{,}78965, 0{,}01134493, 0{,}09852547, 0{,}3605804)$$

El último ejemplo se utiliza “momentos” para ajustar la DLG a una distribución Weibull (1,5), la Figura 10 presenta los estimadores para la distribución Weibull con parámetros $\alpha = 1$, $\beta = 5$ cuyos momentos alrededor de la media son $\hat{\alpha}_1 = 0{,}91816$, $\hat{\alpha}_2 = 0{,}04423$, $\hat{\alpha}_3 = -0{,}2541$, y $\hat{\alpha}_4 = 2{,}8802$, y la función de densidad de probabilidad de la DLG(0.993495,1.04918,0.212114,0.106139), A continuación se presentan las dos funciones de densidad de probabilidad.

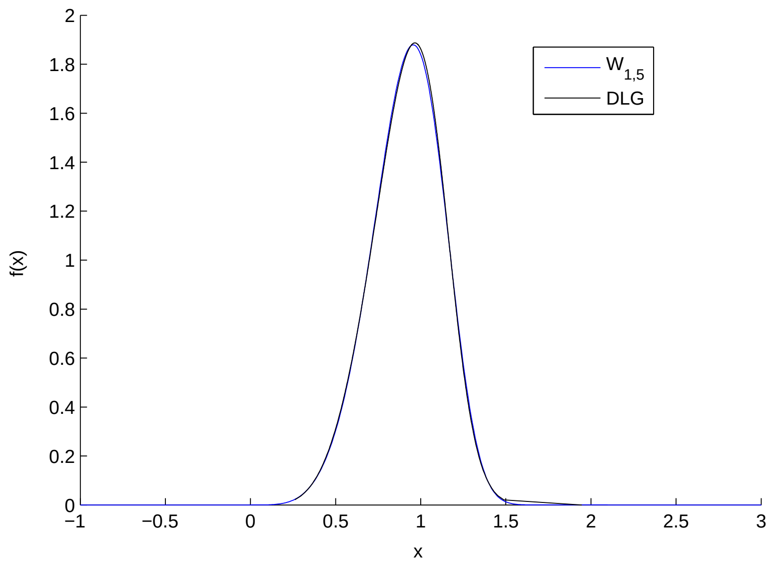


Figura 12: Weibull

Como se observa el programa “seleccion” que implementa la estimación de parámetros de la familia DLG vía método de momentos se ajusta a diferentes conjuntos de datos o distribuciones conocidas, como se observó con los tres ejemplos anteriores.

7. Conclusiones

Acudir al método de momentos para la estimación de los parámetros de la DLG, recurre a obtener la solución de un sistema de ecuaciones que por métodos analíticos es demasiado difícil. Es por ello que el procedimiento que se implementa para obtener las estimaciones se basa en métodos numéricos y en su respectiva programación.

En particular se logró programar el método de manera eficiente e implementarlo a diferentes conjuntos de datos.

Los ajustes a conjuntos de datos utilizando el programa `Seleccion` que estima los parámetros de la DLG$(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ vía método de momentos de forma numérica presenta una buena aproximación tanto a dichos conjuntos, como a distribuciones conocidas, haciendo del programa `Seleccion` una herramienta útil en el modelamiento de conjunto de datos.

El conocimiento de la DLG, las restricciones por su forma funcional y las restricciones del método de momentos, permitieron construir un programa en Matlab llamado `seleccion` conveniente para el modelamiento de datos.

Bibliografía

G. Casella & R. L. Berger (2002). *Statistical inference*, vol. 2. Duxbury Pacific Grove, CA.

Z. A. Karian (2010). *Handbook of fitting statistical distributions with R.* CRC Press.

J. H. Mathews & K. D. Fink (1999). *Numerical methods using MATLAB*, vol. 31. Prentice hall Upper Saddle River, NJ.

A. M. Mood (1950). *Introduction to the Theory of Statistics.* McGraw-hill.

D. Pregibon (1980). 'Goodness of link tests for generalized linear models'. *Applied statistics* pp. 15–14.

J. S. Ramberg & B. W. Schmeiser (1974). 'An approximate method for generating asymmetric random variables'. *Communications of the ACM* **17**(2):78–82.

E. A. Silver (1977). 'A safety factor approximation based upon TukeyŠs Lambda distribution.'. *Operational Research Quarterly* **28**:743–746.

D. D. Wackerly, et al. (2010). *Estadística matemática con aplicaciones.* Cengage Learning Editores.