L-MOMENTOS:
Definición, estimación y aplicaciones

Trabajo de Investigación en
Matemáticas de la Empresa, Economía y Finanzas

Autor: David Casado de Lucas

Director: Santiago Carrillo Menéndez

Universidad Autónoma de Madrid
Doctorado en Matemáticas 2002 – 2003

Índice General

1 Introducción

En estadística es una práctica habitual describir una distribución o una población mediante el valor de un conjunto finito de cantidades, como son la media, la dispersión, la asimetría, la kurtosis, la presencia de "picos", etc.

En finanzas, por ejemplo, estas medidas de la distribución aportan soluciones simples y eficaces a gran cantidad de problemas. Los estadísticos basados en la kurtosis o en la asimetría aparecen constantemente en esta materia para estudiar la normalidad de la distribución correspondiente a unos datos conocidos. Y el momento muestral de cuarto orden estandarizado constituye la regla más utilizada actualmente en el estudio del peso de las colas en series financieras.

Hasta ahora los métodos que más se utilizaban sobre estas medidas se basaban en el estudio de los Momentos de la distribución. Más precisamente, se basaban en los momentos muestrales, puesto que lo usual es disponer sólo de una serie de datos de la distribución. El método de los momentos está aún bastante extendido en el estudio de tres o cuatro familias de distribuciones paramétricas. Esto se debe a que proporciona estimadores consistentes y viables computacionalmente, mientras que otros métodos más avanzados no se muestran eficaces para estos casos concretos. Un ejemplo lo constituye la familia de las mixturas de distribuciones normales. Pero en general estos métodos basados en los momentos proporcionan resultados insatisfactorios para otras distribuciones, debido a que son poco eficaces y a que se ven muy influidos por datos anómalos.

Recientemente se han desarrollado otros métodos que no presentan los problemas anteriores. Estas nuevas técnicas se basan en los L-momentos (en los L-momentos muestrales cuando no tengamos más información que una serie de datos). Las medidas descriptivas basadas en los L-momentos fueron introducidas en 1990 por Hosking, junto con una exposición de sus significados intuitivos. Durante la última década muchos ejercicios de simulación han mostrado la validez de los L-momentos y su mejoría frente a los momentos. Los L-momentos están asociados a una variable aleatoria o, equivalentemente, a una distribución de probabilidad, y son capaces de describir mayor número de distribuciones que los momentos convencionales.

En el presente trabajo se van a estudiar principalmente su aparición histórica, su relación con los momentos, y sus propiedades. A modo de recordatorio, las secciones 2 y 3 incluyen los momentos de una distribución de probabilidad y sus respectivos estimadores. Es en la sección 4 en la que se definen los L-momentos, y en la sección 5 se presentan algunos de los estimadores que se han propuesto para ellos: los L-momentos muestrales, los estimadores Plotting-position y las L-estimaciones. La sección 6 está dedicada a la comparación de los L-momentos, los momentos y sus respectivos estimadores. En la sección 7 se presenta el *Método de los L-momentos* como principal aplicación de estos nuevos conceptos introducidos. En el apéndice se incluyen algunas generalizaciones más, como las que nos proporcionan los L-momentos, de los conceptos de media, desviación típica, asimetría y kurtosis. Por último, la addenda final incluye el código informático (de MATLAB) que se ha desarrollado para los dos programas.

2 Momentos

A lo largo de todo el trabajo supondremos que X es una v.a. real con función de distribución

$$F(x) = Pr(X \leq x)$$

Definición 1 *Se define entonces la* función de cuantiles *de* X *como su inversa; esto es, la función* $x(u)$ *es el único valor que satisface* $F(x(u)) = u$ *para* $0 \lneq u \lneq 1$. *En el caso en que la función de distribución no es inyectiva, se define* $x(u) = \acute{\imath}nf\,\{a \in \Re / F(a) = u\}$.

En muchos casos nos interesa expresar la notación utilizando la función de cuantiles, en vez de con la función de distribución. Para ello tendremos principalmente en cuenta la expresión $u = F(x)$.

Definición 2 *Llamamos* momentos centrados *de una distribución de probabilidad a la media,* $\mu = E(X)$, *y a los demás momentos de orden mayor, que son:*

$$\mu_r = E(X - \mu)^r \quad , \quad r = 2, 3, ...$$

<u>Observación</u>: También se pueden definir los momentos no centrados como $E(X^r)$, para $r = 1, 2, 3, ...$

Los momentos caracterizan a la función de probabilidad. La media, por ejemplo, indica el centro de "localización" de la distribución. También es de interés la *desviación estándar*, o la *varianza* $\sigma^2 = var(X)$, ya que es una medida de la dispersión de la distribución respecto de su centro

$$\sigma = \mu_2^{1/2} = [E(X - \mu)^2]^{1/2}$$

Definición 3 *El* Coeficiente de Variación *se define como* $C_v = \sigma/\mu$, *y expresa la dispersión de la distribución en relación con la media.*

Una forma de construir momentos adimensionales es normalizar de la siguiente manera:

$$\mu_r / \mu_2^{r/2}$$

De entre estos momentos son especialmente interesantes los siguientes:

Definición 4 *La* asimetría *se define como*

$$\gamma = \mu_3 / \mu_2^{3/2}$$

La kurtosis *se define como*

$$k = \mu_4 / \mu_2^2$$

Observación: Estos dos conceptos se utilizan en finanzas para estudiar si una distribución es gaussiana o no.

Como su nombre indica, la asimetría es una medida de lo asimétrica que es la distribución. La kurtosis se puede interpretar como una medida del peso que tienen las colas en la distribución o como la existencia de "picos" en esa distribución.

3 Estimadores de los Momentos

Las conceptos que hemos definido en el apartado anterior se corresponden con unos conceptos análogos (estimadores) para el caso discreto, es decir, para el caso en el que nuestro conocimiento es una muestra de datos $x_1, x_2, ..., x_n$.

Definición 5 *La* media muestral *se define como* $\overline{x} = n^{-1} \sum_{i=1}^{n} x_i$*, y es un estimador de* μ*. Los demás momentos muestrales pueden ser estimados con los momentos muestrales dados por* $m_r = n^{-1} \sum_{i=1}^{n} (x_i - \overline{x})^r$*, aunque no son estimadores centrados.*

También se conocen los estimadores insesgados siguientes:

$s^2 = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$ es insesgado de σ^2

$\widetilde{m}_3 = \frac{n^2}{(n-1)(n-2)} m_3$ es insesgado de μ_3

$\widetilde{k}_4 = \frac{n^2}{(n-2)(n-3)} \{(\frac{n+1}{n-1}) m_4 - 3m_2^2\}$ es insesgado de $k_4 = \mu_4 - 3\mu_2^2$

Observación: Aunque s^2 es un estimador insesgado de σ^2, $s = \sqrt{s^2}$ no es un estimador insesgado de σ.

Ahora podemos definir estimadores muestrales del CV, de la asimetría y de la kurtosis:

$$\widehat{C}_v = s/\overline{x} \quad , \qquad g = \widetilde{m}_3/s^3 \qquad y \qquad k = \widetilde{k}_4/s^4 + 3$$

Lo malo de estos estimadores es que tienen algunas propiedades poco deseables. Por ejemplo, g y k pueden tener un gran sesgo. Por otro lado, tienen propiedades algebraicas que dependen del tamaño de la muestra n, como son:

$$|g| \leq n^{1/2} \qquad y \qquad k \leq n+3$$

4 L-Momentos

Uno de los inconvenientes de los momentos muestrales es que no tienen buenas propiedades de comportamiento para el caso de distribuciones asimétricas. De hecho, algo que siempre debe ser tenido en cuenta a la hora de hacer inferencia es que si una distribución es suficientemente asimétrica, es imposible que toda esta asimetría se refleje en una muestra de tamaño finito. Este problema queda solucionado con la introducción de los L-momentos.

Los L-momentos, al igual que los momentos, caracterizan la distribución a la que pertenecen. Históricamente aparecieron como una variación de los momentos de probabilidad ponderados.

Definición 6 *Los* momentos de probabilidad ponderados *de una variable aleatoria X vienen definidos como $M_{p,r,s} = E[X^p\{F(X)\}^r\{1-F(X)\}^s]$, donde F() es la función de distribución de la variable.*

Observación: fueron introducidos por Greenwood et al. (1979).

Son especialmente interesantes $\alpha_r = M_{1,0,r}$ y $\beta_r = M_{1,r,0}$, que pueden expresarse como

$$\alpha_r = \int_0^1 x(u)(1-u)^r du \qquad ; \qquad \beta_r = \int_0^1 x(u)u^r du \tag{1}$$

donde $x(u)$ es la función de cuantiles.

Se descubrió que ciertas combinaciones lineales de los momentos ponderados de probabilidad contenían información de la distribución. Por ejemplo, algunos estimadores de parámetros de escala son múltiplos de $\alpha_0 - 2\alpha_1$ ó $2\beta 1 - \beta_0$; y la asimetría puede ser medida por $6\beta_2 - 6\beta_1 + \beta_0$.

Si definimos los polinomios de Legendre trasladados, $P_r^*(u)$; $r = 0, 1, 2, ...$, como:

(i) $P_r^*(u)$ es un polinomio de grado r en u.

(ii) $P_r^*(1) = 1$

(iii) $\int_{(0,1)} P_r^*(u)P_s^*(u)du = 0$ si $r \neq s$,

entonces se pueden expresar como

$$P_r^*(u) = \sum_{k=0}^{r} p_{r,k}^* u^k$$

con $p_{r,k}^* = (-1)^{r-k}\binom{r}{k}\binom{r+k}{k}$.

Definición 7 *Si $x(u)$ es la función de cuantiles de la v.a. X, se definen los* L-momentos *de X como*

$$\lambda_r = \int_0^1 x(u) P^*_{r-1}(u) du \tag{2}$$

En función de los momentos de probabilidad muestrales, los L-momentos se pueden escribir como:

$$\lambda_1 = \alpha_0 = \beta_0$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$$

y en general

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p^*_{r,k}\alpha_k = \sum_{k=0}^{r} p^*_{r,k}\beta_k \tag{3}$$

Definición 8 *Se definen los siguientes* L-momentos adimensionales $\tau_r = \lambda_r/\lambda_2$ *para* $r = 3, 4, ...$*; y el coeficiente de L-variación como* $\tau = \lambda_2/\lambda_1$.

Una manera equivalente de introducir los L-momentos hace uso de los estadísticos de orden. Si tenemos una muestra aleatoria y denotamos por $X_{k:n}$ el k-ésimo dato más pequeño de nuestra muestra de tamaño n, la muestra queda ordenada como: $X_{1:n}, X_{2:n}, ..., X_{n:n}$.

Parece lógico pensar que en una muestra de tamaño uno, el dato $X_{1:1}$ habrá tendido a tomar el valor de la media; es decir, contiene información sobre ella. En una muestra de tamaño dos podemos pensar que $X_{2:2} - X_{1:2}$ contiene información de la dispersión de la distribución. En una muestra de tamaño tres $X_{3:3} - 2X_{2:3} + X_{1:3}$ nos da una idea de la asimetría. Y, como último ejemplo, en una muestra de tamaño cuatro, $X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}$ nos sirve para estimar el peso que tienen las colas de la distribución, esto es, la kurtosis. Se llega entonces a la siguiente definición:

Definición 9 *Se definen los* L-momentos de una distribución de probabilidad *como:*

$$\lambda_1 = E(X_{1:1})$$

$$\lambda_2 = \tfrac{1}{2}E(X_{2:2} - X_{1:2})$$

$$\lambda_3 = \tfrac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$$

$$\lambda_4 = \tfrac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$$

y en general

$$\lambda_r = r^{-1} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E(X_{r-j:r}) \tag{4}$$

Observación: Esta definición es consistente con la de L-momentos de una variable aleatoria. La "L" del nombre "L-momento" procede de "Linear", por ser los L-momentos combinaciones lineales de los estadísticos de orden.

Los términos más utilizados para estudiar las distribuciones de probabilidad son λ_1, λ_2, τ, τ_3 y τ_4. Entre sus propiedades más importantes están:

- *Existencia.* Si la variable aleatoria es real, como estamos suponiendo en todo este trabajo, los L-momentos existen si y sólo si existe la media. Esta propiedad es especialmente interesante para la modelización de datos financieros, para los que no se sabe si existen la varianza y algunos momentos de orden más alto.
- *Unicidad.* Si la media de una distribución existe, los L-momentos caracterizan totalmente esa distribución. Es decir, dos distribuciones distintas tienen L-momentos distintos.
- λ_1 puede tomar cualquier valor.
- $\lambda_2 \geq 0$
- Si una distribución toma sólo valores positivos, entonces $0 \leq \tau < 1$ y $2\tau - 1 \leq \tau_3 < 1$
- $|\tau_r| < 1$ para $r \geq 3$
- $\frac{1}{4}(5\tau_3^2 - 1) \leq \tau_4 < 1$
- *Transformación lineal.* Si X e Y son v.a. con L-momentos λ_r y λ_r^*, respectivamente, y se tiene que $Y = aX + b$, entonces $\lambda_1^* = a\lambda_1 + b$, $\lambda_2^* = |a|\,\lambda_2$ y $\tau_r^* = signo(a)^r \tau_r$ si $r \geq 3$
- *Simetría.* Si una v.a. es simétrica, $\tau_r = 0$ si $r = 3, 5, 7, ...$
- λ_1 es una medida de la media, μ; λ_2 es una medida de la dispersión σ; $\tau = \lambda_2/\lambda_1$ es una medida del CV, $C_v = \sigma/\mu$; $\tau_3 = \lambda_3/\lambda_2$ es una medida de la asimetría, $\gamma = \mu_3/\mu_2^{3/2}$; y $\tau_4 = \lambda_4/\lambda_2$ es una medida de la kurtosis, $\kappa = \mu_4/\mu_2^2$. Pero estos conceptos no son los únicos que se pueden utilizar como medidas de esas cantidades; se pueden definir medidas más generales (véase el apéndice).

5 Estimadores de los L-momentos

Como suele suceder en Estadística, no podemos conocer la expresión exacta de muchos de nuestros conceptos, y tenemos que recurrir a estimarlos a partir de una muestra de datos. Veamos cómo podemos construir nuestros estimadores para los L-momentos.

5.1 L-momentos muestrales ó U-estimaciones

Sea $X_{1:n}, X_{2:n}, ..., X_{n:n}$ la muestra ordenada de menor a mayor, un estimador insesgado del momento ponderado de probabilidad β_r es:

$$b_r = n^{-1}\binom{n-1}{r}^{-1}\sum_{j=r+1}^{n}\binom{j-1}{r}x_{j:n} \tag{5}$$

Definición 10 *Entonces, y según las ecuaciones que relacionaban los L-momentos con los momentos ponderados de probabilidad, los* L-momentos muestrales *ó* U-estimaciones *se definen como*

$l_1 = b_0$

$l_2 = 2b_1 - b_0$

$l_3 = 6b_2 - 6b_1 + b_0$

$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$

y en general

$$l_{r+1} = \sum_{k=0}^{r} p^*_{r,k} b_k \tag{6}$$

para $r = 0, 1, 2, ..., n-1$

Observación: Los L-momentos muestrales son estimadores insesgados de los L-momentos.

Definición 11 *Se definen también los siguientes estimadores muestrales:* $t = l_2/l_1$ *es estimador del L-CV;* $t_r = l_r/l_2$ *es estimador de* τ_r*.*

Observación: Aunque estos estimadores anteriores no son insesgados, su sesgo es pequeño para muestras grandes.

5.2 Estimadores Plotting-position

Definición 12 *Landwehr et al. (1979) propusieron como estimador de* α_r *a* $\widetilde{\alpha}_r = n^{-1}\sum_{j=1}^{n}(1-p_{j:n})^r x_{j:n}$. *Y análogamente* $\widetilde{\lambda}_r = n^{-1}\sum_{j=1}^{n} P^*_{r-1}(p_{j:n})x_{j:n}$ *se pueden definir como estimadores de* λ_r, *y* $\widetilde{\tau}_r = \widetilde{\lambda}_r/\widetilde{\lambda}_2$ *para* τ_r.

Observación: Aunque $\widetilde{\lambda}_r$ no son insesgados, su sesgo tiende a cero para muestras grandes

Elecciones razonables para $p_{j:n}$ son $p_{j:n} = (j+\gamma)/(n+\delta)$ para $\delta > \gamma > -1$.

5.3 L-estimaciones

Un ejemplo de L-estimación lo constituye el siguiente funcional

Definición 13 *El funcional empírico se define como:*

$$L_r(F_n) = \int_0^1 P_{r-1}(u)F_n^{-1}(u)du = \sum_{i=1}^{n}\left(\int_{(i-1)/n}^{i/n} P_{r-1}(u)du\right)X_{i:n} \tag{7}$$

Estos estimadores son también consistentes cuando hacemos algunas hipótesis suaves sobre la distribución F.

6 Comparación de Momentos, L-Momentos y sus estimadores

Resulta de gran interés comparar los momentos con los L-momentos, así como sus correspondientes conceptos muestrales, puesto que ambos se pueden utilizar para medir las mismas propiedades de las distribuciones.

- El primer L-momento mide la localización, y coincide con el primer momento μ.
- Como $\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$ y $\sigma^2 = \frac{1}{2}E(X_{2:2} - X_{1:2})^2$, entonces podemos decir que σ^2 da más peso a las diferencias entre los dos valores de la muestra de tamaño dos. Se satisface $\sigma \geq \sqrt{3}\lambda_2$ (la igualdad se alcanza sólo para la distribución uniforme). Para distribuciones con menor asimetría se tiene que $\sigma \approx 2\lambda_2$; la igualdad se alcanza para la distribución exponencial ($\tau_3 = 0,3333$) y para la generalized extreme-value con $\tau_3 = 0,2628$.
- Para CV y L-CV se verifica que $\widehat{C}_v = (\frac{3n}{n+1})^{1/2}t$. Esta cota se alcanza para algunas muestras de distribuciones simétricas o casi simétricas. Para distribuciones poco asimétricas, frecuentemente se toma $\widehat{C}_v$ como el doble que t; pero puede se mayor aún.

- Aunque para distribuciones simétricas τ_3 y γ son cero, y para muchas distribuciones casi simétricas se tiene que $\gamma \approx 6\tau_3$, en general no hay una relación entre ambos conceptos. Lo que sí es cierto es que γ se ve más afectada por un peso de las colas grande. De hecho, para distribuciones con colas pesadas, γ se toma como infinito, mientras que τ_3 tiene aún un valor pequeño (por ejemplo, 0,33 en el caso de una distribución generalized logistic).

- Tanto la L-asimetría muestral, t_3, como la L-kurtosis muestral, t_4, son mucho menos sesgadas que la asimetría y la kurtosis ordinarias. La distribución conjunta de ambas es asintóticamente normal, pero para tamaños pequeños de muestra la aproximación es mala, sobre todo si la distribución que subyace es medianamente sesgada.

- Los L-momentos de órdenes altos pueden existir aunque los correspondientes momentos del mismo orden no existan. Esto es una ventaja más que tienen sobre éstos.

- Mientras que γ y κ están sin acotar, τ_3 y τ_4 tienen la cota $|\tau_r| < 1$; esto es una ventaja. Intuitivamente es más fácil interpretar la medida de algo que está contenido en un intervalo que si no lo está.

- Quizá la principal diferencia entre los momentos y los L-momentos es que los momentos dan más importancia a las colas de las distribuciones. Los momentos muestrales también están más afectados por los valores extremos que los equivalentes L-momentos muestrales.

- Los momentos muestrales no son adecuados para el estudio de distribuciones asimétricas, porque los resultados pueden estar bastante equivocados.

- Los estimadores Plotting-position fueron introducidos para estimar los parámetros de la distribución Wakeby. En concreto se vio que $p_{j:n} = (j - 0,35)/n$ daba buenos resultados para las distribuciones Wakeby, generalized extreme-value y generalized Pareto; pero no son adecuados para otras aplicaciones. Por ejemplo, con la elección anterior de $p_{j:n}$ los estimadores $\widetilde{\lambda}_r$ no son invariantes si sumamos a los datos de la muestra una cantidad fija. En casos extremos los estimadores Plotting-position toman valores que no pueden tomar los L-momentos (incluso $\widetilde{\lambda}_2$ puede tomar valores negativos).

- Los estimadores $\widetilde{\tau}_r$ (Plotting-position) tienen, en general, un sesgo mayor que los análogos estimadores muestrales. En particular, el estimador de la L-kurtosis tiene un gran sesgo positivo. Cuando queremos estimar los L-momentos o sus cocientes, son preferibles los estimadores muestrales a los Plotting-position.

- Aun con todo lo anterior, el método de los momentos sigue siendo utilizado frecuentemente para el estudio de algunas distribuciones con tres o cuatro parámetros. Los momentos se muestran especialmente consistentes cuando las distribuciones que se estudian son mixtura de distribuciones normales, caso en el que otros métodos más elaborados fallan.

- Los L-momentos muestrales (ó U-estimaciones) y las L-estimaciones son asintóticamente equivalentes.

7 Aplicaciones

Como hemos indicado ya, una de las principales aplicaciones de los L-momentos es su utilización para la identificación de distribuciones a partir de muestras de datos. Aunque los momentos también caracterizan la distribución a la que pertenecen, el hecho de que los L-momentos existan en casos en que los momentos no existen, y la mayor robustez frente a datos anómalos, hace que los L-momentos sean más adecuados para esta identificación, y especialmente para distribuciones asimétricas.

En finanzas, por ejemplo, resulta útil medir la forma de las distribuciones, porque es una solución simple y eficaz para muchos problemas.

A la hora de trabajar con los datos de una muestra, parte de la utilidad de los L-momentos muestrales radica en el hecho de que el sesgo de sus estimadores es significativamente menor que las diferencias que hay entre los L-momentos de las distintas distribuciones. De ahí que podamos utilizarlos al hacer inferencia insesgada para identificar la distribución de probabilidad.

Pero una de las limitaciones de los L-momentos aparece en distribuciones como la uniforme, ya que los L-momentos de orden superior a dos son nulos, por lo que no existen algunos conceptos, como el de varianza estandarizada, por ejemplo. Esta dificultad se supera parcialmente restringiéndonos a los funcionales que existen, y se supera totalmente añadiendo una condición más.

7.1 Método de los L-momentos

Este método es para los L-momentos lo que para los momentos es el *Método de los Momentos*. Y como éste, permite estimar p parámetros de una distribución de probabilidad. Para ello se igualan los p primeros L-momentos muestrales con sus correspondientes L-momentos teóricos. Por tanto, hay que obtener expresiones que relacionen los parámetros con los L-momentos.

Las distribuciones exactas de los estimadores de los parámetros obtenidos por este método de los L-momentos, son difíciles de obtener; pero para muestras con muchos datos se puede disponer de aproximaciones teóricas.

Hosking et al. (1985) y Hosking y Wallis (1987) demostraron que para muestras pequeñas y medianas, este método es normalmente mejor que el de máxima verosimilitud.

Como parte de este trabajo teórico, se han realizado dos programas con el lenguaje de programación del programa MATLAB. El código de estos programas se incluye en la Adenda. El lector debe tener en cuenta que gran parte del código que se presenta tiene como única misión la de conseguir una presentación agradable para el usuario que ejecuta los programas.

7.2 Programa 1

Este primer programa nos permite estimar los valores de los L-momentos a partir de una muestra de datos de la distribución. De los estimadores presentados en la sección 5 se utilizan los L-momentos muestrales (ó U-estimaciones) y las L-estimaciones. No se utilizan los estimadores Plotting-position porque no tenemos definidos valores adecuados del parámetro $p_{j:n}$ para todas las distribuciones del programa.

Básicamente la teoría que se utiliza en este programa son las ecuaciones 6 y 7.

Una observación que se puede hacer —por otra parte previsible— tras ejecutar este programa es que computacionalmente es bastante más costoso calcular las L-estimaciones que los L-momentos muestrales.

Los resultados obtenidos para tres distribuciones concretas se muestran en las siguientes tablas:

En la cabecera de cada tabla se indica la distribución de probabilidad y el tamaño de la muestra utilizada.

TABLA 1.1

UNIFORME(0,1) (5000 datos)

	λ_1	λ_2	τ_3	τ_4
Valor teórico	0,5000	0,1667	0,0000	0,0000
L-momento muestral	0,5045	0,1669	-0,0065	0,0026
L-estimación	0,5045	0,1669	-0,0065	0,0026

TABLA 1.2

EXPONENCIAL(2) (5000 datos)

	λ_1	λ_2	τ_3	τ_4
Valor teórico	2,0000	1,0000	0,3333	0,1667
L-momento muestral	1,9999	1,0232	0,3519	0,1786
L-estimación	1,9999	1,0229	0,3518	0,1784

TABLA 1.3

NORMAL(1,1) (5000 datos)

	λ_1	λ_2	τ_3	τ_4
Valor teórico	1,0000	0,5642	0,0000	0,1226
L-momento muestral	1,0087	0,5656	0,0026	0,1279
L-estimación	1,0087	0,5555	0,0026	0,1277

Observación:

λ_r es el L-momento r-ésimo

$\tau_r = \lambda_r/\lambda_2$ es el L-momento normalizado r-ésimo

A la vista de las tablas, podemos afirmar que los dos métodos de estimación son bastante aproximados entre sí y al valor real.

7.3 Programa 2

El segundo programa hace lo contrario que el primero, por lo que es algo más complicado. Permite construir una muestra de datos independientes idénticamente distribuidos de la distribución que elijamos (de la lista dada), a partir del valor de los L-momentos.

La teoría que se utiliza para ello es el Método de los L-momentos. Como hemos descrito anteriormente, al igualar los primeros L-momentos teóricos con los primeros L-momentos muestrales, se obtienen ecuaciones que relacionan los parámetros con los L-momentos. Es decir, que conociendo los L-momentos tenemos una estimación de los parámetros. Después el programa genera las variables aleatorias utilizando la función de cuantiles, es decir, invirtiendo la función de distribución.

TABLA 2.1

Distribución	λ_1	λ_2	λ_3	λ_4	λ_4	Tamaño
U	2,9940 (3)	0,3457 (0,3333)				500
E	1,9679 (2)	0,9940 (1)				500
G	4,9325 (5)	2,0562 (2)				500
N	2,9442 (3)	0,4752 (0,5000)				500
GP	3,9440 (4)	1,9907 (2)	1,0284 (1)			1000
GEV	7,0071 (7)	4,9601 (5)	2,9779 (3)			1000
GL	9,2706 (9)	2,1447 (2)	1,0569 (1)			1000
L	7,7975 (8)	5,8269 (6)	3,8480 (4)			20000
W	5,7466 (5,7500)	1,6370 (1,6548)	0,8330 (0,8643)	0,5598 (0,5962)	0,4545) (0,4476	20000

donde: U Uniforme ; E Exponencial ; G Gumbel ; N Normal ; GP Generalized Pareto ;

GEV Generalized extreme-value ; GL Generalized logistic ; L Lognormal ; W Wakeby. Y los números que están entre paréntesis son los valores reales introducidos

Para hacer este ejercicio se han utilizado las ecuaciones que se incluyen en el apéndice de la obra [2] de la bibliografía, que son el resultado de aplicar el Método de los L-momentos a las distintas distribuciones.

En esta tabla se representan los valores de los L-momentos a priori, es decir, los conocidos, y después los estimados a partir de la muestra obtenida (vista con el Programa 1 la equivalencia de ambos estimadores y la mayor rapidez de cálculo de los L-momentos muestrales, se han utilizado estos últimos). Es decir, que aplicamos el Programa 1 a los datos de salida).

Como se aprecia, para las dos últimas distribuciones ha sido necesario utilizar tamaños muestrales bastante grandes. Esto significa que el método que hemos utilizado para generar las variables es mucho menos preciso que el utilizado para generar las variables de las demás distribuciones.

Apéndice

Mediante la definición de relaciones de orden parcial, se pueden dar definiciones generales de medidas de la localización, la dispersión, la asimetría y la kurtosis.

A.1 Localización

Sean X e Y variables con funciones de distribución F y G, respectivamente. Diremos que Y es estocásticamente mayor que X ($F \prec G$) si

$$F^{-1}(u) \leq G^{-1}(u) \quad \text{para todo} \quad 0 < u < 1.$$

Observación: La idea intuitiva es que Y está distribuida con el peso más a la derecha que X.

Definición 14 *Entendemos por una medida de la localización un funcional lineal μ definido en el subconjunto Λ de las funciones de distribución y que verifica los siguientes axiomas:*

(i) Si $X \prec Y$ entonces $\mu(X) \leq \mu(Y)$

(ii) Para todo número real c, se tiene que $\mu(X+c) = \mu(X) + c$

(iii) $-\mu(X) = \mu(-X)$

Ejemplos

1) La esperanza de F, $\mu(F) = \int_0^1 F^{-1}(u)du$

2) La mediana de F, $\mu(F) = F^{-1}(1/2)$

3) La familia de las medias truncadas $\mu_\alpha(F) = \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} F^{-1}(u)du$

4) Los ejemplos anteriores están incluidos en la clase de medidas dada por $\mu(F) = \int_0^1 F^{-1}(u)d\lambda(u)$, donde λ es una medida de probabilidad simétrica respecto al punto $1/2$.

A.2 Dispersión

Se dice que G es más dispersa que F ($F \prec G$) si

$$F^{-1}(1-\alpha) - F^{-1}(\alpha) \leq G^{-1}(1-\alpha) - G^{-1}(\alpha) \quad \text{para todo} \quad 0 < \alpha < 1/2.$$

Definición 15 *Se dice que σ es una medida de la dispersión si es un funcional positivo definido en el subconjunto Λ de las funciones de distribución continuas, y que verifica:*

(i) Si G es más dispersa que F, entonces $\sigma(F) \leq \sigma(G)$

(ii) Para todo número real c, $\sigma(X + c) = \sigma(X)$

(iii) Para todo número real a, $\sigma(a \cdot X) = |a|\, \sigma(X)$

Ejemplos

1) El ejemplo más conocido es la desviación típica $\sigma = \left[E\left(X - E(X)\right)^2\right]^{1/2}$

2) Una clase entera de estas medidas la constituyen $\tau(F) = \int_0^1 F^{-1}(u)d\lambda(u)$, donde λ es una medida en (0,1) que satisface que $\lambda(A) = -\lambda(1 - A)$ y $\lambda(\{0\}) = 0$.

A.3 Asimetría

Aunque han sido definidas varias relaciones de orden en distintos textos, nosotros diremos que F es menos asimétrica por la derecha ($F \prec G$) que G si y sólo si

$$G^{-1}(F(X)) \text{ es convexa.}$$

Definición 16 *Por una medida γ de la asimetría entendemos un funcional real definido en un conjunto dado Λ de funciones de distribución, tal que verifica:*

(i) Si $F \prec G$ entonces, $\gamma(F) \leq \gamma(G)$

(ii) Si F es una distribución simétrica, entonces $\gamma(F) = 0$

(iii) Para cualesquiera números reales a y b, $\gamma(a \cdot X + b) = \gamma(X)$

Ejemplos

1) Es frecuentemente utilizado $\gamma = \frac{E(X - E(X))^3}{[E(X-E(X))^2]^{3/2}}$, tercer momento central estandarizado.

2) Una medida más robusta que la anterior es la siguiente $\gamma = \frac{E(X)-m}{E|X-m|}$, con $m = F^{-1}(1/2)$

3) La medida cuartil es $\gamma = \frac{F^{-1}(0{,}75)+F^{-1}(0{,}25)-F^{-1}(1/2)}{F^{-1}(0{,}75)-F^{-1}(0{,}25)}$, y su generalización a la medida cuartil es $\gamma_\alpha = \frac{F^{-1}(1-\alpha)+F^{-1}(\alpha)-F^{-1}(1/2)}{F^{-1}(1-\alpha)-F^{-1}(\alpha)}$

A.4 Kurtosis

Aunque hay definiciones de esta relación de orden para distribuciones asimétricas, en la mayoría de los libros se define sólo para distribuciones simétricas. Sean F y G distribuciones simétricas, diremos que F tiene menos kurtosis que G $(F \prec G)$ si y sólo si

$$\begin{gathered} G^{-1}(F(X)) \text{ es convexa para } x > m_F \\ \text{y} \\ G^{-1}(F(X)) \text{ es cóncava para } x < m_F \end{gathered}$$

(m_F es una constante que depende de F).

Definición 17 *Entendemos por una medida κ de la kurtosis un funcional real definido en un subconjunto dado funciones de cuantiles simétricos y que satisface:*

(i) Si $F \prec G$ entonces $\kappa(F) \leq \kappa(G)$

(ii) Para cualesquier valores reales a y b, $\kappa(a \cdot X + b) = \kappa(X)$

Ejemplos

1) El ejemplo más conocido es $\kappa_4 = \frac{E(X-E(X))^4}{E^2(X-E(X))^2}$, cuarto momento central estandarizado.

2) Una clase importante de ejemplos se puede dar como $\kappa(F) = \frac{\tau_1(F^{-1})}{\tau_2(F^{-1})}$, donde τ_1 y τ_2 son funcionales que satisfacen que $\tau_i(a \cdot F^{-1} + b) = |a| \cdot \tau_i(F^{-1})$ para unos números reales a y b cualesquiera.

3) $\tau_p(F) = \frac{F^{-1}(\frac{1}{2}+p)-F^{-1}(\frac{1}{2}-p)}{F^{-1}(0{,}75)-F^{-1}(0{,}25)}$ para $0 \leq p \leq 1/2$

4) Para cada $0 \lneqq p \lneqq 1/4$, $\kappa_p(F) = \frac{F^{-1}(0{,}75+p)+F^{-1}(0{,}75-p)-2F^{-1}(0{,}75)}{F^{-1}(0{,}75+p)-F^{-1}(0{,}75-p)}$

5) Por último $\kappa_\delta(F) = \frac{U_\delta - L_\delta}{U_{0,5}-L_{0,5}}$, donde $U_\delta = E(X \mid X \gneqq F^{-1}(1-\delta))$ y $L_\delta = E(X \mid X \lneqq F^{-1}(\delta))$

Lista de obras citadas

En este apartado se incluyen las obras que no se han utilizado directamente en la elaboración de este trabajo, pero que son mencionadas en él.

- Greenwood, J.A., Landwehr, J.M., Matalas, N.C. y Wallis, J.R. (1979). *Probability weigthed moments: Definition and relation to parameters of several distributions expressable in inverse form.* Water Resources Research, 15, 1049-54.

- Hosking, J.R.M. y Wallis, J.R. (1987). *Parameter and quantile estimation for the generalized Pareto distribution.* Technometrics, 29, 339-49.

- Hosking, J.R.M., Wallis, J.R. y Wood, E.F.(1985). *Estimation of the generalized extreme-value distribution by the method of probability-weigthed moments.* Technometrics, 27, 251-61.

- Landwehr, J.M., Matalas, N.C. y Walis, J.R. (1979). *Estimation of parameters and quantiles of Wakeby distributions.* Water Resources Research, 15, 1361-79 (Correction: Water Resources Research, 15 (1979), 1672).

Referencias

[1] Hernández Pérez, Nicolás (2002) *Applications of Descriptive Measures in Risk Management*, Tesis Doctoral, Universidad de Toronto.

[2] Hosking, J.R.M. y Walis, J.R. (1997) *Regional Frequency Analysis: An Approach Based on L-moments*, New York, Cambridge University Press.