

GENERACION DE NUMEROS ALEATORIOS CON DISTRIBUCIONES DE LEVY

Miguel A. Ré

Facultad de Matemática, Astronomía y Física - Universidad Nacional de Córdoba
Ciudad Universitaria - (5010) - Córdoba - Argentina
e-mail: re@roble.fis.uncor.edu

Los procesos estocásticos con leyes estables de Lévy están recibiendo creciente atención. Estas distribuciones están caracterizadas básicamente por la ley de potencias que determinan su decaimiento asintótico: $L(x, \gamma) \sim x^{-(1+\gamma)}$ para x suficientemente grande, donde $0 < \gamma < 2$. El método de Monte Carlo, de amplia difusión para la modelación de diversos procesos físicos o químicos, resulta de difícil aplicación para variables aleatorias con distribuciones de Lévy debido a que sólo se conocen expresiones analíticas de las funciones de Lévy para unos pocos valores del exponente γ . En esta comunicación se presenta un algoritmo simple para la generación de secuencias de números pseudo aleatorios con distribuciones de Lévy de índice arbitrario. El método se basa en el uso de la generalización de la distribución q-normal de la estadística generalizada de Tsallis y el teorema generalizado de Lévy-Gnedenko.

Lévy stable law processes are receiving increasing attention. These distributions are essentially characterized by the power law decay in the asymptotic behaviour: $L(x, \gamma) \sim x^{-(1+\gamma)}$ for large enough x , with $0 < \gamma < 2$. Monte Carlo method, of great diffusion in modeling physical or chemical processes, is difficult to use with Lévy distributed stochastic variables since an analytical expression for Lévy functions is not known except for a few values of the parameter γ . In this communication it is proposed a simple algorithm for generating pseudo random sequences with Lévy distribution for any value of γ . This algorithm is based on the generalized q-normal distribution of Tsallis's generalized statistics and the Lévy Gnedenko central limit theorem.

Pacs N^o 02.70.-c; 02.50.-r; 05.40.j

1 Introducción

La distribución Gaussiana, íntimamente relacionada con el proceso de difusión normal (el movimiento browniano), exhibe una gran ubicuidad en los procesos naturales, atribuible al teorema central del límite. Sin embargo resulta un hecho bien establecido en el presente⁽¹⁾ que una gran variedad de fenómenos estocásticos en el área de la física, la biología y otras disciplinas caen fuera del paradigma gaussiano, respondiendo a distribuciones de Lévy. Una característica común a estos procesos es que corresponden a procesos estocásticos con leyes estables: una variable aleatoria z definida como una adecuada combinación lineal de varias variables aleatorias x , independientes e idénticamente distribuidas tiene una densidad de probabilidad de igual forma que la exhibida por las x . El paradigma gaussiano resulta el ejemplo más conocido de un proceso de ley estable: la suma de n variables gaussianas tiene también distribución gaussiana. Los procesos con leyes estables resultan además atractores estables en el espacio funcional de las variables aleatorias: el teorema central del límite establece que la suma de n variables aleatorias con varianza finita tiene una densidad de probabilidad que converge a una

distribución de Gauss cuando $n \rightarrow \infty$.

Las distribuciones de Lévy, a su vez, resultan ser los atractores para las densidades de probabilidad de variables aleatorias con un decaimiento asintótico de ley de potencias: $P(x) \simeq |x|^{-(1+\gamma)}$, para $x > x_m$ con x_m suficientemente grande y $0 < \gamma < 2$, de manera tal que no existe el segundo momento para estas distribuciones. Denotamos a las distribuciones de Lévy por $L(x; \gamma)$, siendo γ el valor de la potencia que exhibe el decaimiento asintótico.

Los procesos de Lévy han recibido una creciente atención y estudios teóricos, numéricos o experimentales se han llevado a cabo en diferentes áreas. Sin embargo estos procesos no son simples de tratar dado que no se conoce una expresión analítica para las funciones de Lévy, a excepción de algunos valores especiales del parámetro γ : por ejemplo $L(x; 1)$ corresponde a la distribución de Cauchy. De esta manera un algoritmo que pueda generar una secuencia pseudo aleatoria con valor arbitrario del parámetro γ resultará de interés.

En esta comunicación se presenta un algoritmo simple para la simulación de procesos de Lévy simétricos y valor de índice γ arbitrario en el intervalo de interés. Este algoritmo está basado en la generación de una secuencia con distribución q-normal⁽²⁾ definida

en el contexto de la mecánica estadística generalizada propuesta por C. Tsallis⁽³⁾. Estas distribuciones presentan un decaimiento asintótico según una ley de potencias, por lo que eligiendo adecuadamente el valor del parámetro q de la estadística no extensiva podemos obtener el comportamiento asintótico buscado para el parámetro γ .

Las secuencias de Lévy se obtienen de las anteriores efectuando la correspondiente combinación lineal determinada por el teorema de Lévy-Gnedenko. El grado de aproximación mejora con el número de variables sumadas en la combinación lineal. Dado que para $\gamma = 1$ la función de Lévy y la q -normal son coincidentes en este caso sólo es necesario considerar $n = 1$. El número de valores necesarios aumenta a medida que nos alejamos con γ del valor 1.

A continuación presentamos las distribuciones de Lévy y sus principales propiedades de interés para este trabajo definiéndolas a partir de la representación de Fourier. También presentamos las distribuciones q -normal normalizadas⁽⁴⁾ de la estadística no extensiva y comparamos los comportamientos asintóticos de ambas distribuciones a fin de conseguir su coincidencia en este límite. En la tercera sección, presentamos el algoritmo para generar la secuencia con distribución q -normal. Comparamos aquí los resultados obtenidos a partir de la suma de variables con los valores analíticos de las funciones de Lévy obtenidos por inversión numérica y generamos un estimador de la bondad de la aproximación. Finalmente, en la cuarta sección, se discuten los resultados obtenidos y se proponen líneas de acción futuras para mejorar la aproximación.

2 Distribuciones de Lévy simétricas y distribuciones q -normales

Definimos la distribución de Lévy simétrica a partir de la representación integral

$$L(x; \gamma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \exp[-ikx - |\alpha k|^\gamma] \quad (1)$$

siendo γ el exponente en el decaimiento asintótico que determina las propiedades de escala de la distribución. El factor α que aparece en la expresión está relacionado con la escala espacial elegida y será tomado igual a 1 en el resto del trabajo. Sólo se conoce una expresión analítica para las distribuciones de Lévy para algunos valores particulares del parámetro γ ($\gamma = 2$ corresponde a la distribución de Gauss; $\gamma = 1$ que corresponde a la distribución de Cauchy y también $\gamma = 2/3$ y $\gamma = 1/2$). Estas funciones presentan momentos divergentes: $\langle |x|^\eta \rangle$ es divergente para todo $\eta > \gamma$. Sin embargo debe notarse que aún cuando algunos momentos sean divergentes el proceso estocástico está completamente definido desde el punto de vista matemático.

De particular interés para este trabajo será el de-

sarrollo asintótico para valores grandes de x ($x \gg 1$)

$$L(x; \gamma) \simeq -\frac{1}{\pi} \sum_{k=1}^n \frac{(-1)^k \Gamma(\gamma k + 1)}{k! x^{\gamma k + 1}} \operatorname{sen} \left(\frac{k\pi\gamma}{2} \right) + \mathcal{O}(x^{-\gamma(n+1)-1}) \quad (2)$$

con $\Gamma(z)$ la función factorial. En particular consideraremos el término dominante en el desarrollo

$$L(x; \gamma) \simeq \frac{\Gamma(1 + \gamma)}{\pi} \frac{\operatorname{sen} \left(\frac{\pi\gamma}{2} \right)}{x^{1+\gamma}} \quad (3)$$

evidenciando el comportamiento asintótico con un decaimiento según una ley de potencias. Podemos ver también que la potencia γ determina el valor de la función de Lévy en el origen. En efecto, a partir de la representación integral se muestra de manera directa

$$L(0; \gamma) = \frac{1}{\pi} \Gamma \left(1 + \frac{1}{\gamma} \right) \quad (4)$$

Las funciones de Lévy han sido relativamente bien estudiadas y otras propiedades pueden encontrarse en la literatura^(5,6).

Por otra parte consideramos las funciones q -normales, definidas en el contexto de la mecánica estadística no aditiva de Constantino Tsallis

$$N(x; q) = \sqrt{\frac{q-1}{3-q}} \frac{\left[B \left(\frac{1}{2}, \frac{1}{2} \frac{3-q}{q-1} \right) \right]^{-1}}{\left[1 + \frac{q-1}{3-q} x^2 \right]^{\frac{1}{q-1}}} \quad (5)$$

donde $B(x, y)$ simboliza la función beta (integral de Euler de primera clase)⁽⁷⁾.

Notamos en particular que en el límite $q \rightarrow 1$ esta función converge a una distribución normal. Por otra parte si consideramos el comportamiento asintótico de estas funciones encontramos que para valores grandes de x ($(q-1)/(3-q)x^2 \gg 1$)

$$N(x; q) \simeq \frac{\left[B \left(\frac{1}{2}, \frac{1}{2} \frac{3-q}{q-1} \right) \right]^{-1}}{\left[\frac{q-1}{3-q} \right]^{\frac{3-q}{q-1}}} \frac{1}{x^{\frac{2}{q-1}}} \quad (6)$$

que corresponde a un decaimiento según una ley de potencias. Esta ley de potencias para el decaimiento resulta coincidente con el comportamiento de las funciones de Lévy si hacemos la identificación

$$\gamma = \frac{3-q}{q-1} \quad (7)$$

Un caso particularmente interesante corresponde a $\gamma = 1$ ($q = 2$), valor para el cual la función de Lévy y la q -normal coinciden, resultando la distribución de Cauchy

$$L(x; 1) = \frac{1}{\pi} \frac{1}{1+x^2} \quad (8)$$

La coincidencia del comportamiento asintótico entre ambos conjuntos de funciones, tomando en cuenta la relación establecida entre q y γ , nos permite afirmar, por el teorema central del límite de Lévy-Gnedenko⁽⁸⁾, que la variable aleatoria

$$Z_n = \frac{1}{n^{1/\gamma}} \sum_{k=1}^n X_k \quad (9)$$

tendrá una densidad de probabilidad que converge a una distribución de Lévy en el límite $n \rightarrow \infty$. En particular hemos visto que para $\gamma = 1$ ambas funciones N y L coinciden, en cuyo caso $n = 1$ basta para asegurar la coincidencia con la distribución de Lévy.

3 Algoritmo

El algoritmo propuesto en esta comunicación para generar secuencias de números pseudo-aleatorios con distribución q-normal es una extensión de un conocido algoritmo para generar una secuencia con distribución de Cauchy⁽⁹⁾. Recordamos que en última instancia la distribución de Cauchy es una q-normal con $q = 2$.

Consideremos dos variables aleatorias U y V con distribución conjunta uniforme en la región del plano (u, v) definida por la condición

$$u^2 + |v|^{2/\gamma} \leq \frac{1}{4} \quad (10)$$

Para el caso $\gamma = 1$ la región corresponde a un círculo de radio $1/2$. A partir de estas variables definimos una nueva variable aleatoria $Y = U / |V|^{1/\gamma}$ que, de acuerdo con el teorema de transformación de variables aleatorias^(10,11), tiene una densidad de probabilidad

$$f(y, q) = \frac{\left[B\left(\frac{1}{2}, \frac{\gamma}{2}\right) \right]^{-1}}{(1 + y^2)^{\frac{1+\gamma}{2}}} \quad (11)$$

que, recordando la relación establecida entre q y γ , corresponde a una distribución q-normal para la variable reescalada $y = \sqrt{\gamma}x$. Es decir que, salvo por un factor de escala, el algoritmo propuesto nos permite generar una secuencia con distribución q-normal en forma exacta. Por supuesto este factor de escala puede elegirse arbitrariamente con una simple modificación del algoritmo a emplear. La implementación del algoritmo se incluye en el apéndice.

Si consideramos el comportamiento asintótico de las funciones de Lévy, aún cuando la ley de potencias para el decaimiento sea la misma, encontramos una diferencia dada por un factor multiplicativo. A fin de conseguir un comportamiento asintótico coincidente planteamos un cambio de escala C_γ en la variable y de manera tal que la variable aleatoria $X = Y/C_\gamma$ tiene como densidad de probabilidad^(10,11)

$$P(x; \gamma) = C_\gamma \frac{\left[B\left(\frac{1}{2}, \frac{\gamma}{2}\right) \right]^{-1}}{\left[1 + (C_\gamma x)^2 \right]^{\frac{1+\gamma}{2}}} \quad (12)$$

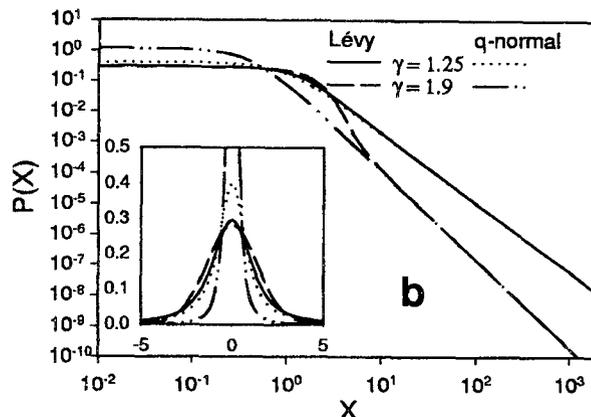
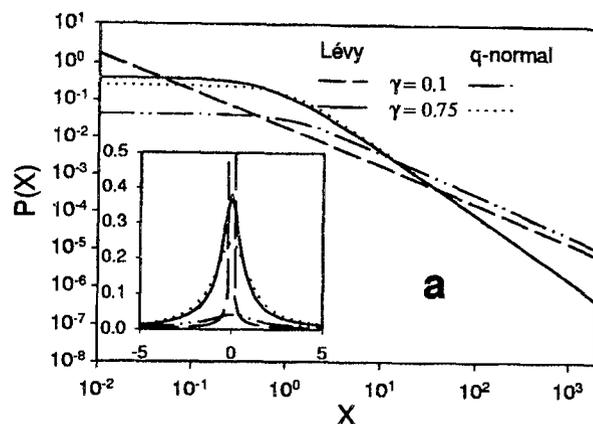


Figura 1: Comparación de funciones Lévy y q-normal para distintos valores del índice γ : a) para $\gamma < 1$; b) para $\gamma > 1$.

con un comportamiento asintótico dado por

$$P(x; \gamma) \simeq \frac{\left[B\left(\frac{1}{2}, \frac{\gamma}{2}\right) \right]^{-1}}{|C_\gamma|^\gamma |x|^{1+\gamma}} \quad (13)$$

Requiriendo ahora que el comportamiento asintótico de la función $P(x; \gamma)$ sea coincidente con el de la función $L(x; \gamma)$, encontramos que el valor apropiado para el factor de escala resulta

$$C_\gamma = 2 \left[\frac{\Gamma\left(1 + \frac{\gamma}{2}\right)}{\Gamma\left(1 - \frac{\gamma}{2}\right)} \right]^{1/\gamma} \quad (14)$$

En la figura 1 se presenta una comparación de las funciones de Lévy y las q-normales corregidas por el cambio de escala para algunos valores del índice γ . Como se esperaba encontramos un comportamiento coincidente en el límite asintótico de x grande. Sin embargo el valor de x a partir del cual se alcanza la coincidencia asintótica depende del valor de γ , en particular para $\gamma < 1$ como puede apreciarse en la figura 1a. En efecto para valores de γ aproximándose a 0 la convergencia se da a valores cada vez mayores de x . Además para valores de $\gamma < 1$ la diferencia entre $L(0; \gamma)$ y $N(0; q)$ es más marcada para valores menores de γ , como se desprende de (4) y el hecho

de que N es una función acotada en el origen. Por el contrario, para $\gamma > 1$, las diferencias en $x = 0$ no son tan pronunciadas, pero se registran diferencias relativamente importantes para valores intermedios como puede apreciarse para el caso $\gamma = 1.9$ en la figura 1b. Este comportamiento de las funciones harán que en la generación de las secuencias para las variables Z_n de (9) la densidad de probabilidad correspondiente aproxime a la función de Lévy con valores pequeños de n en el límite asintótico en tanto que para valores pequeños de Z_n los apartamientos de las distribuciones serán mayores y se requerirán más términos en la suma (valores de n mayores) para mejorar la aproximación.

4 Resultados obtenidos

Implementamos el algoritmo propuesto como se indica en el apéndice generando secuencias de 10^9 números con la distribución q-normal modificada por el factor de escala para distintos valores del parámetro $\gamma(q)$. Como ya fue mencionado, la variable aleatoria definida en (9) tiene una densidad de probabilidad que se aproxima a una función de Lévy tanto mejor cuanto mayor es n , conforme al teorema central del límite de Lévy-Gnedenko. La convergencia está dada por la elección del factor $n^{1/\gamma}$ en la combinación lineal.

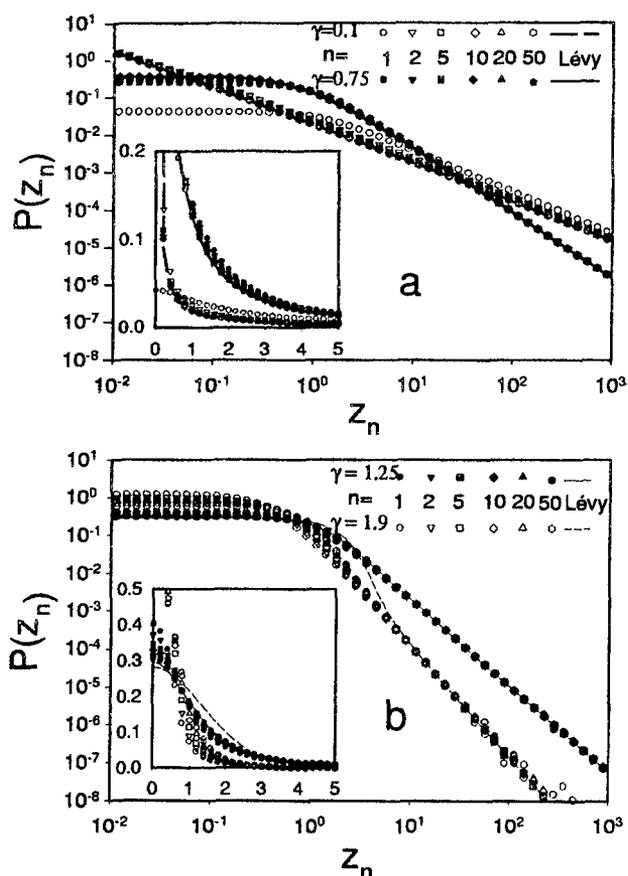


Figura 2: Densidad de probabilidad aproximada para la variable Z_n (ver texto) para distintos valores de n . Las correspondientes funciones de Lévy se han incluido como referencia. Nótese la coincidencia para valores grandes de z_n y la convergencia para valores de n crecientes.

Los resultados correspondientes a esta aproximación se ilustran en la figura 2 para distintos valores de n . Vemos que efectivamente la aproximación mejora a medida que aumentamos el valor de n . Encontramos en particular que en las "colas" de la distribución la convergencia es comparativamente más rápida que alrededor del origen con el aumento de n . Este efecto ya fue previsto cuando elegimos el factor de escala para la variable q-normal a fin de conseguir la coincidencia asintótica de las distribuciones. La convergencia de las distribuciones para valores pequeños o intermedios del argumento es más lenta y requiere un mayor número de variables q-normales sumadas para mejorar la aproximación. Notamos nuevamente aquí un comportamiento diferente según que $\gamma < 1$ o $\gamma > 1$. En el caso $\gamma < 1$ la discrepancia principal se da para $Z_n = 0$, lográndose una aproximación relativamente buena para valores intermedios como se aprecia en el gráfico 2a aún para el valor $\gamma = 0.1$. La situación es distinta para $\gamma > 1$, dado que en estos casos la diferencia no alcanza valores tan grandes como en el caso anterior, pero a valores intermedios se dan diferencias importantes. Esto se aprecia en particular en la figura 2b para el caso $\gamma = 1.9$. En ambos casos las diferencias mencionadas son más pronunciadas cuanto más nos alejamos del valor $\gamma = 1$, es decir con γ aproximándose a los valores 0 o 2.

A fin de estimar la bondad de la aproximación para distintos valores del parámetro γ y con distinto grado de aproximación n , definimos la suma de cuadrados de apartamientos

$$\epsilon^2(n) = \sum_{z_n=-10}^{z_n=10} [P(z_n) - L(z_n; \gamma)]^2 \quad (15)$$

tomando 101 términos ($\Delta Z_n = 0.2$). Nos restringimos a valores relativamente pequeños del argumento puesto que, según ya fue mencionado, los apartamientos son comparativamente más importantes en este intervalo de valores. En la figura 3 ilustramos los valores obtenidos para $\epsilon^2(n)$ vs. n para distintos valores de γ . Como ya se mencionó la calidad de la aproximación disminuye con el apartamiento del valor $\gamma = 1$ para un valor fijo n . En particular ya se mencionó que el apartamiento es muy marcado para valores de $\gamma \ll 1$ en $z = 0$. A fin de separar esta contribución definimos una nueva variable $\epsilon_0^2(n)$ obtenida de eliminar en (15) el término correspondiente a $Z_n = 0$. Graficamos los valores correspondientes en la figura 3b. Vemos efectivamente una disminución en la suma ϵ^2 , disminución más marcada para los casos $\gamma < 1$, en los que la influencia del valor $z=0$ es mayor.

5 Discusión y Conclusiones

Se ha presentado un algoritmo para la generación de secuencias pseudo-aleatorias con distribución q-normal⁽⁴⁾. El valor de γ y el correspondiente valor asociado de γ puede elegirse en forma arbitraria en el rango $0 < \gamma < 2$. Haciendo uso del teorema central del

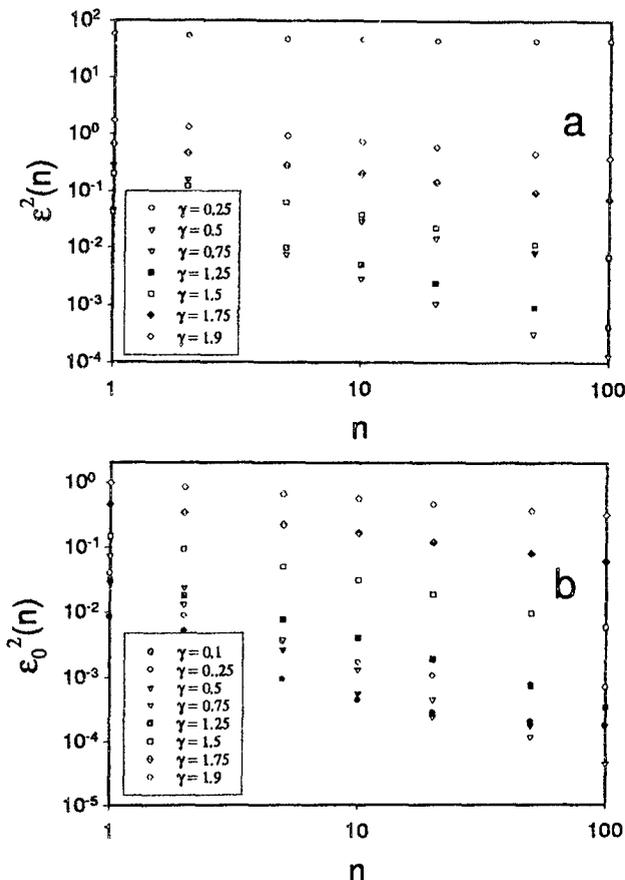


Figura 3: Suma de cuadrados de diferencias entre la densidad de probabilidad $P(z_n)$ y la correspondiente función de Lévy: a) incluyendo la diferencia en $z_n = 0$; b) sin considerar la diferencia en $z_n = 0$. Nótese como disminuye el valor en el caso b), consecuencia de que la mayor contribución proviene de la diferencia en $z_n = 0$.

límite de Lévy-Gnedenko encontramos que la variable Z_n definida en (9) tiene una densidad de probabilidad que aproxima una función de Lévy. La aproximación, como en el caso del teorema central del límite, mejora con el aumento del valor de n . Dado el factor de escala elegido (con valores pequeños de n) en las "colas" de la distribución. La convergencia para valores chicos o intermedios de Z_n resulta más "lenta" (requiriéndose valores mayores de n para la aproximación). Queda así planteado para la continuidad de este trabajo mejorar el algoritmo de aproximación a fin de lograr una convergencia más rápida en estas regiones. Esta línea de trabajo está desarrollándose.

A Apéndice: función para generación de secuencia con distribución q-normal

Presentamos en este apéndice la función Pascal implementada para la generación de un número aleatorio de la distribución $N(x; \gamma)$ definida en (5) y con la relación dada en (7) entre q y γ .

El argumento de la función: $gami$ corresponde al valor de γ y el valor de salida vuelve con el nombre de

la función $genx$.

```
function genx(gami:real): extended;
var
  dx,x,y,yt,expy : extended;

begin
  Repeat
    x:=2*Random-1;
    y:=2*Random-1;
    if (y=0) then expy:= -5000
      else expy:= ln(abs(y))/gami;
    if (expy>=-4900) then yt := exp(expy)
      else yt := 0;
  until(x*x+yt*yt<=0.25);
  if (yt>0) then if (y<0) then dx:=-x/yt
    else dx:= x/yt
    else dx:= 1e2500;
  genx:= dx;
end;
```

Agradecimientos: El autor agradece el financiamiento de CONICOR y SeCyT-UNC para este proyecto.

Referencias

- [1] M. F. Shlessinger, G. M. Zaslavsky y U. Frisch, Editores, "Lévy flights and related topics in Physics", Lecture Notes in Physics, Springer-Verlag, Berlín (1995).
- [2] E. P. Borges, J. Phys. A **31**, 5281 (1998).
- [3] C. Tsallis, S.V.F. Levy, A.M.C. Souza y R. Maynard. Phys. Rev. Lett. **75**, 3598 (1995). [Erratum: Phys. Rev. Lett. **27**, 5442 (1996)].
- [4] D. Prato and C. Tsallis. Phys. Rev. E. **60**, 2398 (1999).
- [5] E. W. Montroll y J. T. Bendler, J. Stat. Phys. **34**, 129 (1984).
- [6] J. P. Bouchard y A. Georges, Phys. Rep. **195**, 128 (1990).
- [7] I. S. Gradshteyn e I. M. Ryzhik, "Table of Integrals, Series and Products", 5^{ta} edición, editor A. Jeffrey, Academic Press, Londres (1965).
- [8] A. Araujo y E. Gine, "The Central Limit Theorem for Real and Banach Valued Random Variables", J. Wiley, N. York (1980).
- [9] G. S. Fishman, "Monte Carlo: concepts, algorithms and applications", Springer Series in Operations Research, Editor P. Glynn, Springer, N. York (1996).
- [10] D. T. Gillespie, Am. J. Phys. **51**, 520 (1983).
- [11] A. Papoulis, "Probability, Random Variables and Stochastic Processes", 3^{ra} edición, Mc Graw-Hill, N. York (1991).