Distribución para una suma de v. a. uniformes

Patricia

27 de marzo de 2007

1. Introducción

Se trata de probar que si $U_1, U_2, \ldots, U_n, \ldots$, son v.a. uniformemente distribuidas en el $(0,1)$, y N es la variable aleatoria dada por

$$N = \text{mín}\{n \mid \sum_{i=1}^{n} U_i > 1\},$$

entonces $E[N] = e$.

Para ver esto, notemos en primer lugar que N es una variable aleatoria discreta, que toma valores enteros no negativos. Más aún, $N \geq 2$, puesto que cada U_j es un número positivo menor que 1.

Así, podemos escribir usando la definición de valor esperado, que

$$E[N] = \sum_{n=2}^{\infty} n \cdot P\{N = n\}.$$

2. Cálculo de $P\{N = n\}$

Para calcular la probabilidad $P\{N = n\}$, notemos que $N = n$ si y sólo si la suma $\sum_{i=1}^{n} U_i > 1$ y $\sum_{i=1}^{n-1} U_i \leq 1$. Para simplificar la notación, llamaremos S_n a la suma de las primeras n variables aleatorias: $S_n = \sum_{i=1}^{n} U_i$. Ahora bien,

$$S_n > 1 \qquad \text{si y sólo si} \qquad S_{n-1} > 1 \ \text{ o } (S_{n-1} \leq 1 \text{ y } S_n > 1),$$

siendo estas dos últimas posibilidades disjuntas. Por lo tanto,

$$P\{S_n > 1\} = P\{S_{n-1} > 1\} + P\{N = n\},$$

de donde obtenemos que

$$P\{N = n\} = P\{S_n > 1\} - P\{S_{n-1} > 1\}. \tag{1}$$

Así, será preciso calcular las densidades para S_n, con $n \geq 2$.

3. Cálculo de la densidad de S_n

Para este problema, es suficiente encontrar la densidad para S_n, para valores menores que 1.

Afirmación: Sea f_n la función de densidad de la variable aleatoria S_n, $n \geq 1$. Entonces

$$f_n(x) = \frac{x^{n-1}}{(n-1)!} \qquad \text{para todo } x,\ 0 < x < 1.$$

Probamos esto por inducción.

Para $n = 1$, tenemos que $S_1 = U_1$, y sabemos que su densidad es $\mathbb{I}_{(0,1)}(x)$, por lo que la afirmación es cierta en este caso.

Para $n = 2$, tenemos que $S_2 = U_1 + U_2$. Así, S_2 es una v.a. que toma valores en el intervalo $(0, 2)$. Sea $f_{U_i}(t)$ la densidad de la v.a. U_i, con $i = 1, 2$. Tenemos que

$$f_2(x) = \int_{-\infty}^{\infty} f_{U_1}(t)\, f_{U_2}(x-t)\, dt = \int_{-\infty}^{\infty} \mathbb{I}_{(0,1)}(t)\mathbb{I}_{(0,1)}(x-t).$$

Para $x < 1$, esto es igual a

$$\int_0^x 1\, dt = x.$$

Para $1 < x < 2$, esto es igual a

$$\int_{x-1}^{1} dt = 1 - (x-1) = 2 - x.$$

Por lo tanto, la hipótesis inductiva se cumple para $x < 1$.

Supongamos ahora que $f_n(x) = \frac{x^{n-1}}{(n-1)!}$, para un cierto n, y para todo x, $0 < x < 1$. Veamos que la afirmación se cumple para $n + 1$.

Tenemos que $f_{n+1}(x) = f_{S_n+U_{n+1}}(x)$, y utilizando la fórmula de la densidad para suma de variables aleatorias, obtenemos

$$\begin{aligned}
f_{n+1}(x) &= \int_{-\infty}^{\infty} f_n(t)\, f_{U_n}(x-t) \\
&= \int_{-\infty}^{\infty} f_n(t)\, \mathbb{I}_{(0,1)}(x-t)\, dt \\
&= \int_0^x \frac{t^{n-1}}{(n-1)!}(t)\, dt \\
&= \frac{x^n}{n!}.
\end{aligned}$$

Aquí hemos usado que $S_n(t)$ toma valores en el intervalo $(0, n)$, mientras que $U_n(t)$ los toma en el $(0, 1)$. Puesto que $(x - t) \in (0, 1)$ es equivalente a $x - 1 < t < x$, y dado que $x - 1 < 0$, el integrando resulta no nulo sólo en el intervalo $(0, x)$.

4. Cálculo de $E[N]$

Ahora podemos calcular $E[N]$. Tenemos que

$$P\{S_n > 1\} = 1 - P\{S_n \leq 1\} = 1 - \int_0^1 f_n(t)\,dt = 1 - \frac{1}{n!}.$$

Por lo tanto, y usando la fórmula (1), tenemos que

$$P\{N = n\} = (1 - \frac{1}{n!}) - (1 - \frac{1}{(n-1)!}) = \frac{n-1}{n!}.$$

Luego

$$\begin{aligned} E[N] &= \sum_{n=2}^{\infty} n \cdot \frac{n-1}{n!} \\ &= \sum_{n=2}^{\infty} \frac{1}{(n-2)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \\ &= e. \end{aligned}$$