La Distribución de Irwin–Hall

Lic. Raúl Delgado Álvarez
dea_5@hotmail.com

1. Introducción

Una manera de entender, por qué el teorema de límite central tiene una importancia gravitante en el estudio de la Estadística, es observar el comportamiento de la suma de variables aleatorias independientes.

Es así que una distribución nada acampanada como es la distribución uniforme nos permite observar la rápida convergencia de la suma de variables aleatorias hacia la distribución Normal en forma muy didáctica en base a los gráficos de las sumas de variables aleatorias independientes con distribución Uniforme.

En este trabajo primero observamos la distribución de la suma de dos variables aleatorias distribuidas uniformemente y se construye su gráfico, posteriormente se estudia la distribución de la suma de tres variables aleatorias independientes con distribución uniforme, de esta manera la distribución de Irwin-Hall se constituye en una generalización del problema citado y por último se construye un programa para la suma de un número finito de variables aleatorias con distribución uniforme.

Problema. Sean x y y variables aleatorias independientes, cada una con distribución uniforme en el intervalo (0,1). Hallar la distribución de $z = x + y$.

Solución. Si

$$x \sim U[0,1) \Rightarrow f(x) = 1, \quad 0 \le x \le 1 \qquad y \sim U[0,1) \Rightarrow f(y) = 1, \quad 0 \le y \le 1$$

Siendo x e y independientes la densidad conjunta será:

$$f(x,y) = f(x)f(y) = 1, \quad 0 \le x \le 1, 0 \le y \le 1$$

definida en el cuadrado.

Sea $z = x + y$ y la variable aleatoria auxiliar $w = y$

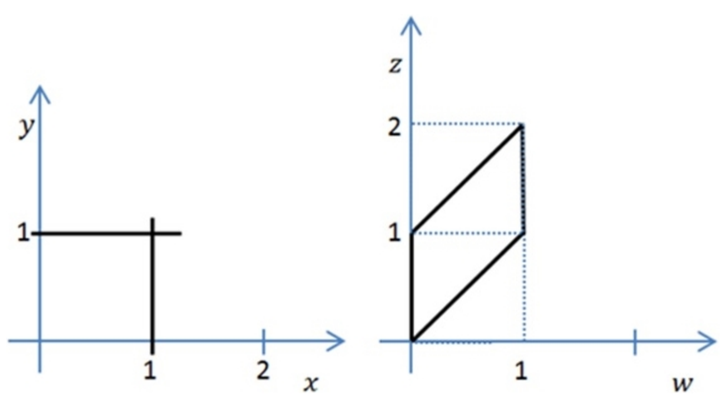

Gráfico 1 Región inicial y región transformada

Se tiene $\begin{matrix} x = z - w \\ y = w \end{matrix}$ luego el Jacobiano de la transformación es $J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$. Luego la función conjunta de z y w será:

$$K(z,w) = 1: \quad 0 \le z - w \le 1, 0 \le w \le 1 \qquad K(z,w) = 1: \quad w \le z \le w + 1, 0 \le w \le 1$$

Es decir $k(z,w)$ se define en el paralelogramo.

Para hallar la distribución de la suma, se halla la distribución de la marginal

$$g_Z(z) = \int_{Rw} k(z,w)dw = \begin{cases} \int_0^z dw; & 0 \leq z \leq 1 \\ \int_z^1 dw; & 1 \leq z \leq 2 \end{cases} = \begin{cases} z; & 0 \leq z \leq 1 \\ 2-z; & 1 \leq z \leq 2 \end{cases}$$

Sea $z = x_1 + x_2 + x_3$, donde x_1, x_2 y x_3 son variables aleatorias independientes cada una con distribución uniforme en el intervalo (0,1). Hallar la función de densidad de z.

Solución. Sea $z = (x_1 + x_2) + x_3$ y $y = x_1 + x_2$ del problema anterior $w = x_3$

$$f(y) = \begin{cases} y; & 0 \leq y \leq 1 \\ 2-y; & 1 \leq y \leq 2 \end{cases}$$

$f(w) = 1; 0 \leq w \leq 1$, la densidad conjunta de y y w será:

$$f(y,w) = \begin{cases} y; & 0 \leq y \leq 1; 0 \leq w \leq 1 \\ 2-y; & 1 \leq y \leq 2; 0 \leq w \leq 1 \end{cases}$$

Luego

$$z = y + w \Rightarrow y = z - h \qquad H = w \Rightarrow w = h$$

Cuyo $J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$

Luego la densidad conjunta de z,H será:

$$k(z,H) = \begin{cases} -h; & 0 \leq z-h \leq 1; 0 \leq h \leq 1 \\ 2-z+h; & 1 \leq z-h \leq 2; 0 \leq h \leq 1 \\ z-h; & h \leq z \leq h+1; 0 \leq h \leq 1 \\ (2+h)-z; & h+1 \leq z \leq h+2; 0 \leq h \leq 1 \end{cases}$$

Definida en el paralelogramo:

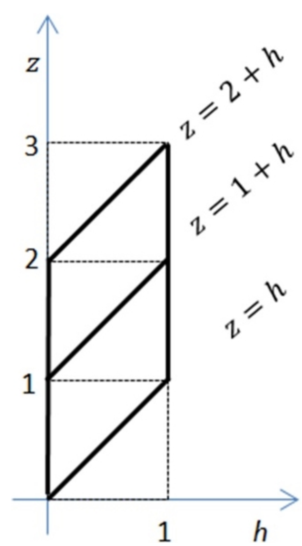

Gráfico 2 Región transformada para tres variables

Para hallar la distribución de z, hallamos la marginal:

$$g(z) = \int_{Rw} K(z,h)dh = \begin{cases} \frac{z^2}{2}; & 0 \leq z \leq 1 \\ -z^2 + 3z - \frac{3}{2}; & 1 \leq z \leq 2 \\ \frac{z^2}{2} - 3z + \frac{9}{2}; & 2 \leq z \leq 3 \end{cases}$$

2. La distribución de Irwin–Hall

La *distribución de Irwin-Hall*, llamada así por Joseph Irwin y Phillip Hall, es la distribución que rige el comportamiento la suma de variables aleatorias independientes, cada una con la distribución uniforme estándar. También se conoce como la *distribución uniforme suma*. También sirve como un buen ejemplo del teorema del límite central, observando la convergencia de esta suma como veremos a continuación.

Sea $Z = \sum_{i=1}^{N} x_i$, donde x_i se distribuye uniformemente en el intervalo (0,1), entonces se distribuye de acuerdo a:

$$f(z) = \frac{1}{(n-1)!} \sum_{j=0}^{k} \binom{N}{j} (-1)^j (z-j)^{N-1}; \quad k \leq z < k+1, \; k = 0,1,2,\ldots,N-1.$$

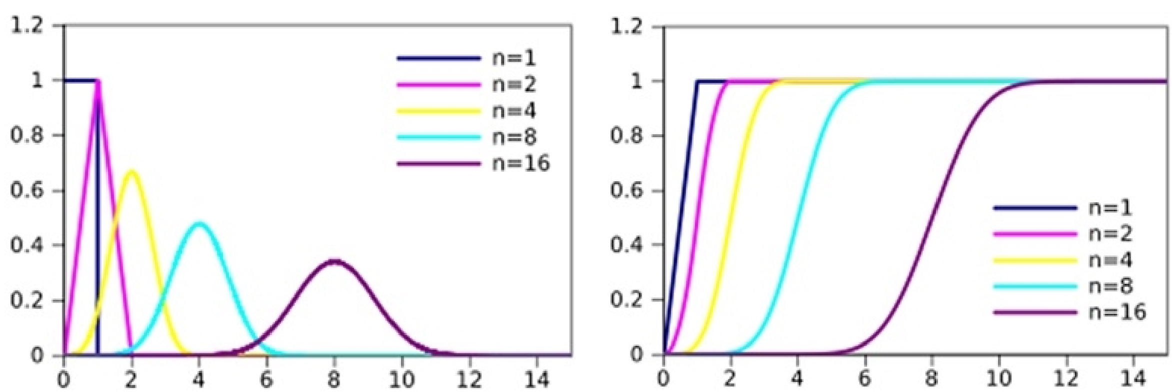

Gráfico 3 Evolución de la función de densidad y función acumulada. Extraído de Wikipedia

Para evidenciar el comportamiento de Z, se realiza un programa que permite generalizar La distribución de Z, además de mostrar el gráfico resultante para distintos valores de N, para $N = 8$, se muestra la indicada distribución y su gráfico, obsérvese que prácticamente se comporta como una normal, lo mismo se puede hacer para distintos valores de N.

	f(z)	k
1	z^7/5040	0<= k <1
2	z^6/90 - z^7/720 - z^5/30 + z^4/18 - z^3/18 + z^2/30 - z/90 ...	1<= k <2
3	z^7/240 - z^6/15 + (13*z^5)/30 - (3*z^4)/2 + (55*z^3)/18 - (...	2<= k <3
4	z^6/6 - z^7/144 - (5*z^5)/3 + 9*z^4 - (256*z^3)/9 + 53*z^2 - ...	3<= k <4
5	z^7/144 - (2*z^6)/9 + 3*z^5 - (199*z^4)/9 + 96*z^3 - (737*z...	4<= k <5
6	z^6/6 - z^7/240 - (17*z^5)/6 + (53*z^4)/2 - (2647*z^3)/18 + ...	5<= k <6
7	z^7/720 - z^6/15 + (41*z^5)/30 - (31*z^4)/2 + (1889*z^3)/18...	6<= k <7
8	z^6/90 - z^7/5040 - (4*z^5)/15 + (32*z^4)/9 - (256*z^3)/9 + ...	7<= k <8

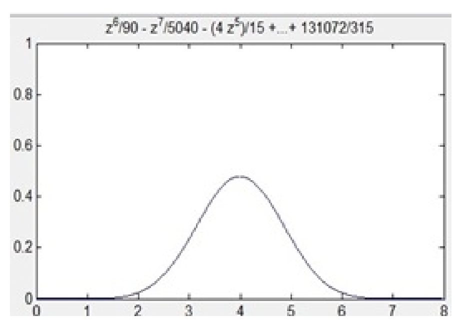

Gráfico 4 Resultado del programa

Programa

```
%programa que genera la funcion de Irwin-Hall en Matlab
%definimos variables simbolicas
syms z c;
%preguntamos si n es mayor a 1
n=input('Ingrese el valor de N=');
% inicializamos el vector donde se guardaran las funciones
A=zeros(n,1);
%creamos un vector simbolico a partir de A
B=sym(A);
%inicializamos las variables auxiliares
f=0;
%definimos la constante simbolica
c=1/factorial(n-1);
% comenzamos el calculo de las variables
for j=0:n-1
   %guardamos en el vector f la suma de los
    f=f+nchoosek(n,j)*(-1)^j*(z-j)^(n-1);
   %comenzamos el calculo de las variables
    f=collect(f);
   %desplegamos el calculo de las variables multiplicada por la constante definida antes
    disp(c*f);
   %almacenamos cada valor impreso en el vector simbolico
    B(j+1,1)=c*f;
End
%graficamos las variables simbolicas halladas en el vector
for i=1:n
  ezplot(B(i,1),[i-1,i]);
%mientras no terminemos el vector cada variable debe ser trazada en el mismo lienzo
if i<n
hold on
end
end
hold off
%delimitamos el eje de x entre 0 y n y y entre 0 y 1
axis([0 n 0 1]);
```