

PRESENTACIONES ESTADISTICAS

Distribuciones de frecuencia:

Son tablas en las que se agrupan los valores posibles de una variable y se registra el número de valores observados que corresponden a cada clase.

Como ejemplo se verá la distribución de frecuencias de los salarios semanales para 100 trabajadores:

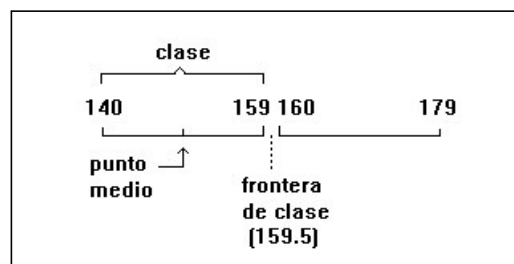
Salario semanal (en pesos)	Número de Trabajadores (frecuencia)
140 – 159	07
160 – 179	20
180 – 199	33
200 – 219	25
220 – 239	11
240 – 259	04
	Total=100 (tamaño de la muestra)

Intervalos de Clase:

Un ejemplo de ellos lo constituye 140 - 159, los límites de clase superior e inferior son: 159 y 140 respectivamente. Estos indican los valores incluidos dentro de la clase. Las fronteras de clase o límites exactos son puntos específicos de la escala que sirven para "separar clases adyacentes". Se pueden determinar identificando los puntos que están en la mitad entre los límites superior e inferior de las clases adyacentes.

Por ejemplo, para 140 - 159 y 160 - 179 la frontera de clase está en 159.50.

El intervalo de clase indica el rango de valores incluidos dentro de una clase y se puede determinar restando la frontera de clase inferior de la frontera de clase superior. Puede ser útil el punto medio de clase, que se determina sumando la mitad del intervalo de clase a la frontera inferior de clase.



Resumiendo:

Salario semanal (límite de clases)	Fronteras de clase	Punto medio de clase	Nro. de Trabajadores
140 - 159	139.5 - 159.5	149.5	7
160 - 179	159.5 - 179.5	169.5	20
180 - 199	179.5 - 199.5	189.5	33
200 - 219	199.5 - 219.5	209.5	25
220 - 239	219.5 - 239.5	229.5	11
240 - 259	239.5 - 259.5	249.5	4
		Total:	100

Conviene que los intervalos de clase sean iguales. Se puede determinar el intervalo de clase en forma aproximada del siguiente modo:

$$\text{Intervalo aproximado} = \frac{\left(\begin{array}{c} \text{Valor mayor en} \\ \text{datos no agrup.} \end{array} \right) - \left(\begin{array}{c} \text{Valor menor en} \\ \text{datos no agrup.} \end{array} \right)}{\text{número deseado de clases}}$$

Para el caso anterior, supuesto que los valores menores y mayores de los datos no agrupados fueran 142 y 258 respectivamente, y se desean 6 clases:

$$\text{Int}_{\text{aprox}} := \frac{258 - 142}{6} \quad \text{Int}_{\text{aprox}} = 19.333$$

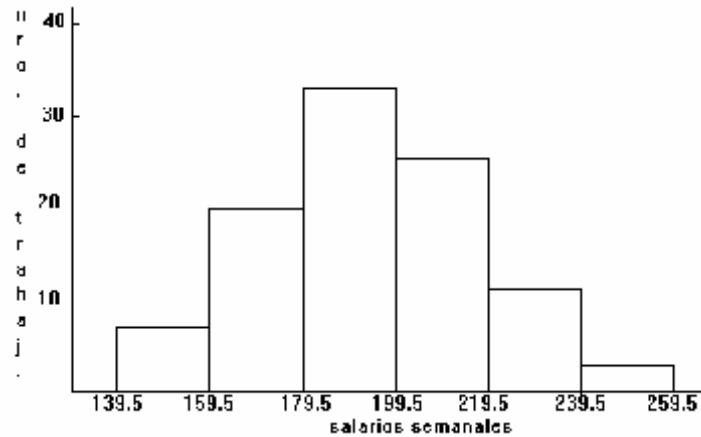
El más próximo será 20.

Para datos distribuidos "irregularmente" pueden ser convenientes intervalos irregulares de clase. Se utilizan intervalos mayores para los rangos de valores en los que hay relativamente pocas observaciones.

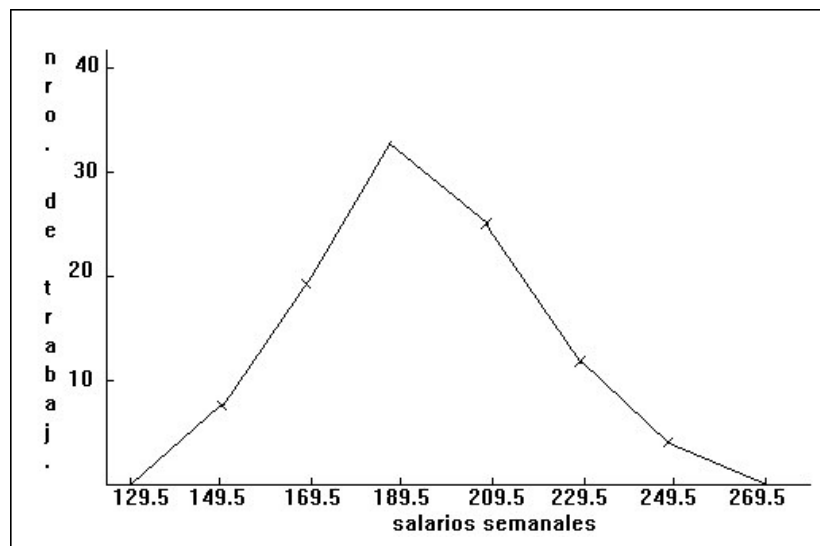
HISTOGRAMAS Y POLIGONOS DE FRECUENCIA

Un histograma es un diagrama de barras de una distribución de frecuencias, donde en ordenadas se representan las observaciones y en abscisas las fronteras de clase. En general es conveniente que el número de observaciones sea superior a las 30.

Para el caso del ejemplo considerado:



Un polígono de frecuencias es un gráfico de distribución de frecuencias de tipo lineal.



Es posible trabajar con estas figuras de resumen en forma más dinámica a través de la utilización de Matlab.

Para la realización del histograma, se deben definir el vector que comprenda a los datos no agrupados y el vector de los intervalos. Luego, mediante una operación que se verá más adelante, se podrá observar cómo se calcula un tercer vector correspondiente a la distribución de frecuencia en cada intervalo.

La forma más sencilla de representar el histograma es cuando se tienen los datos agrupados como en el ejemplo. Para esto basta representar "la función" en diagrama de barras", mediante el siguiente segmento de programa que utiliza la función `bar`:

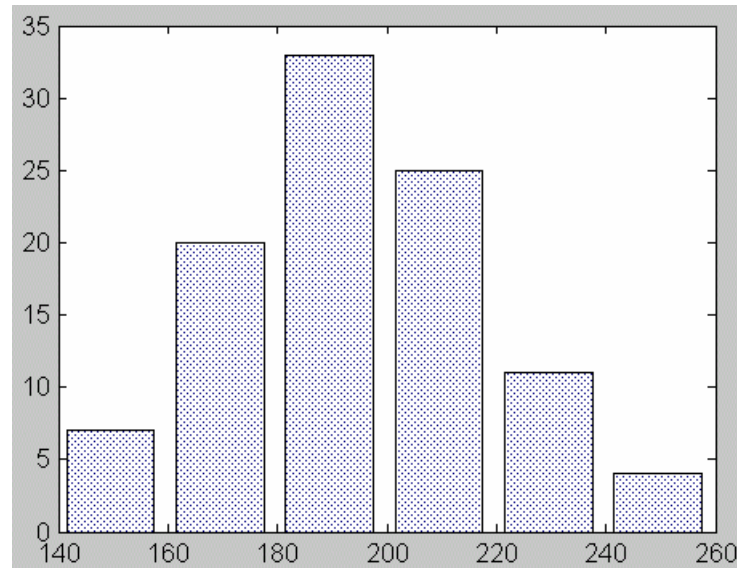
```
function barras
% Permite representar graficamente en un diagrama de barras el vector
% de puntos medios de clase y el de frecuencias
% Entradas: puntos: vector de puntos medios de clase
```

```
%      frecuencias: vector con las correspondientes frecuencias
% Salida: Histograma

% Vectores de entrada
puntos=[149.5000 169.5000 189.5000 209.5000 229.5000 249.5000];
frecuencias=[7 20 33 25 11 4];

% Grafico de barras
bar(puntos,frecuencias)
```

que da el siguiente resultado:



Si se poseen los datos no agrupados, como podrían ser 1000 números aleatorios comprendidos entre 0 y 10, se procede del siguiente modo utilizando la función [hist](#):

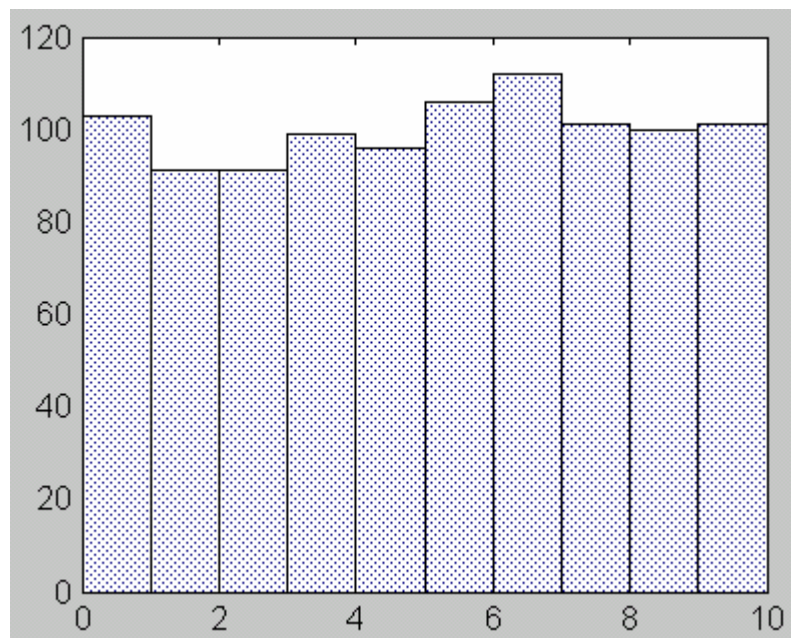
```
function aleatorio
% Permite ver el histograma de un conjunto de datos no agrupados, como podrían
% ser 1000 números aleatorios comprendidos entre 0 y 10
% Entrada: datos, vector de datos
%
% Salida: Histograma

% Vector de datos
for i=1:1000, datos(i)=rand*10;end
% Vector con los centros de intervalos
x=0.5:1:9.5;
% Histograma
m=hist(datos,x);
hist(datos,x)
m      %vector de frecuencias en cada intervalo
```

Para esta corrida, el vector de frecuencias resultante es:

[103 91 91 99 96 106 112 101 100 101]

y el histograma correspondiente:



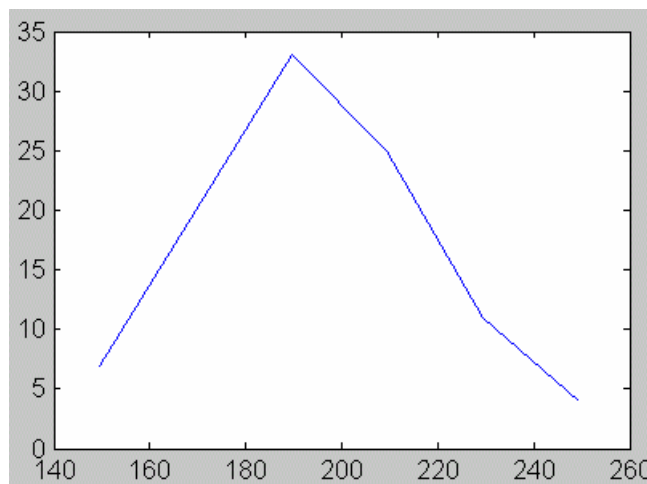
Si se quisiera obtener el "polígono de frecuencia" basta con cambiar las características de la presentación gráfica, pasando a representar en el formato "línea". Para ello se recurre a:

```
function poligono
% Permite representar graficamente el poligono de frecuencias
% Entradas: puntos: vector de puntos medios de clase
%           frecuencias: vector con las correspondientes frecuencias
% Salida: Poligono de frecuencias

% Vectores de entrada
puntos=[149.5000 169.5000 189.5000 209.5000 229.5000 249.5000];
frecuencias=[7 20 33 25 11 4];

% Grafico
plot(puntos,frecuencias)
```

que entrega por resultado:



Es posible hallar la "Curva de Frecuencia" haciendo un polígono suavizado. Para ello se puede utilizar la función "spline cúbica" incorporada en Matlab.

SPLINE Interpolación de datos con spline Cúbico

YY = SPLINE(X,Y,XX) usa la interpolación spline cúbica para encontrar YY, los valores de la función subyacente Y en los puntos en los cuales están dados los datos Y.

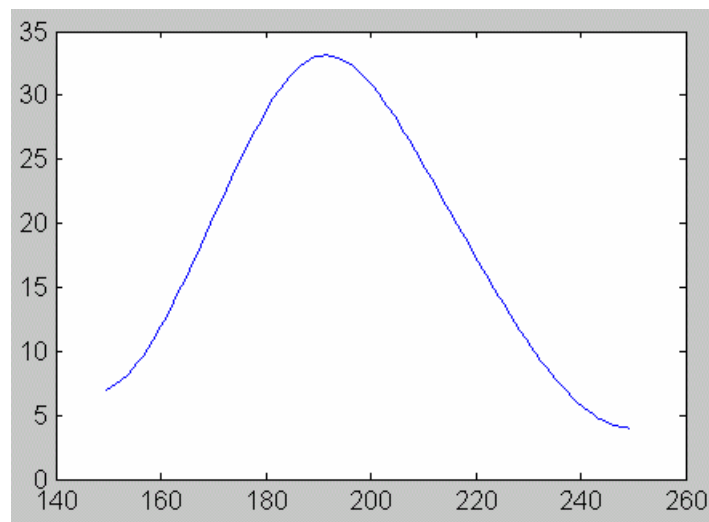
Esto se puede observar corriendo el siguiente segmento de programa:

```
function suave
% Permite representar graficamente el poligono de frecuencias en forma suavizada
% Entradas: puntos: vector de puntos medios de clase
%           frecuencias: vector con las correspondientes frecuencias
% Salida: Poligono de frecuencias suavizado

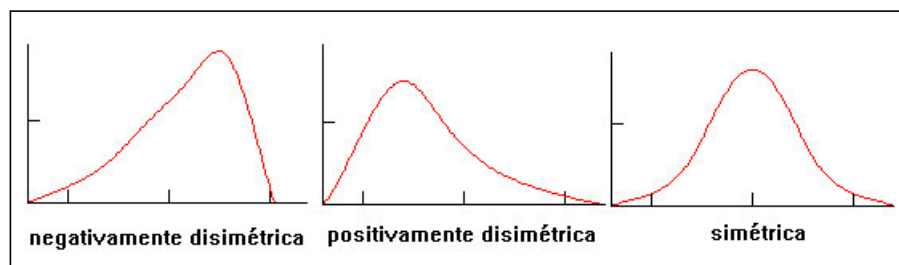
% Vectores de entrada
puntos =[149.5000 169.5000 189.5000 209.5000 229.5000 249.5000];
frecuencias =[7 20 33 25 11 4];

% Calculo del spline
x=puntos(1):puntos(length(puntos)); % malla para representar el suavizado
PP = spline(puntos,frecuencias,x); % vector resultante de la interpolacion
% Grafico
plot(x,PP)
```

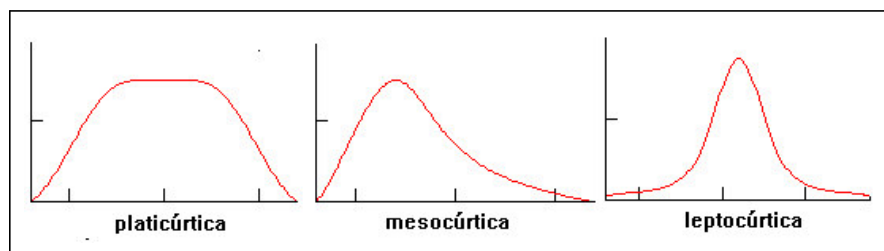
que da como salida:



En términos de **disimetría** se las puede clasificar en:



En términos de **curtosis** se las puede clasificar en:



DISTRIBUCION DE FRECUENCIA ACUMULADA

Salario semanal (límite de clases)	Fronteras Superior de clase	Nro. de Trabajadores	Frecuencia Acumulada
140 - 159	159.5	7	7
160 - 179	179.5	20	20+7=27
180 - 199	199.5	33	27+33=60
200 - 219	219.5	25	60+25=85
220 - 239	239.5	11	85+11=96
240 - 259	259.5	4	96+4=100
	Total:	100	

La tabla de arriba explicita este concepto:

El gráfico de una distribución de frecuencia acumulada se llama "Ojiva". Para el tipo "**menor que**", el gráfico indica la frecuencia acumulada debajo de cada frontera de clase, obtenido mediante el siguiente segmento de programa:

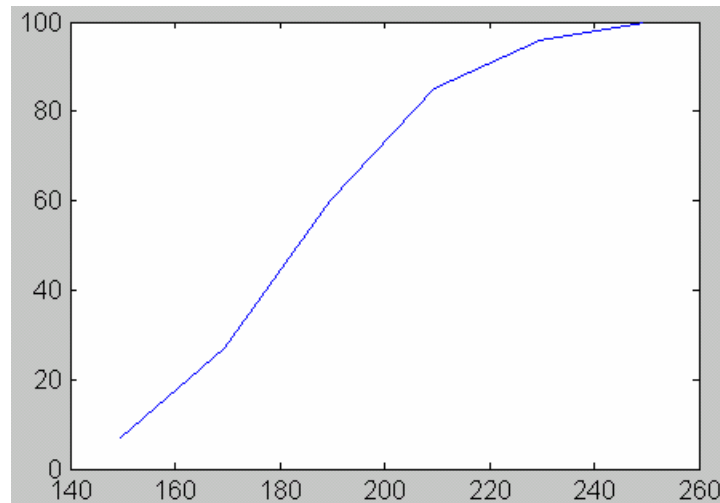
```
function ojiva
% Permite representar graficamente la distribucion de frecuencia acumulada
% Entradas: puntos: vector de puntos medios de clase
%           frecuencias: vector con las correspondientes frecuencias
% Salida: distribucion de frecuencia acumulada (ojiva)

% Vectores de entrada
puntos =[149.5000 169.5000 189.5000 209.5000 229.5000 249.5000];
frecuencias =[7 20 33 25 11 4];

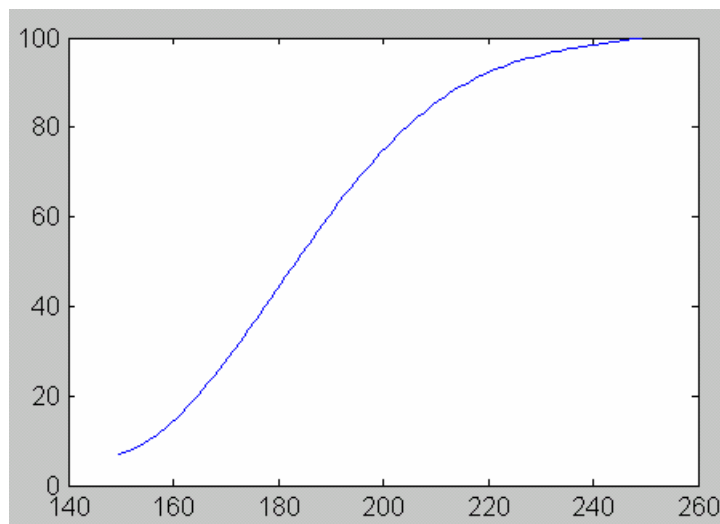
% Vector de frecuencias acumuladas
f_acumulada(1)=frecuencias(1);
for i=2:length(puntos), f_acumulada(i)=f_acumulada(i-1)+frecuencias(i);
end

% Grafico
plot(puntos,f_acumulada)
```

que gráficamente entrega el siguiente resultado:



Si se suaviza se obtiene la "curva ojiva suavizada".



DISTRIBUCION DE FRECUENCIA RELATIVA

Es aquella en que el número de observaciones de cada clase se convierte en una frecuencia relativa dividiéndola por el número total de observaciones en la distribución.

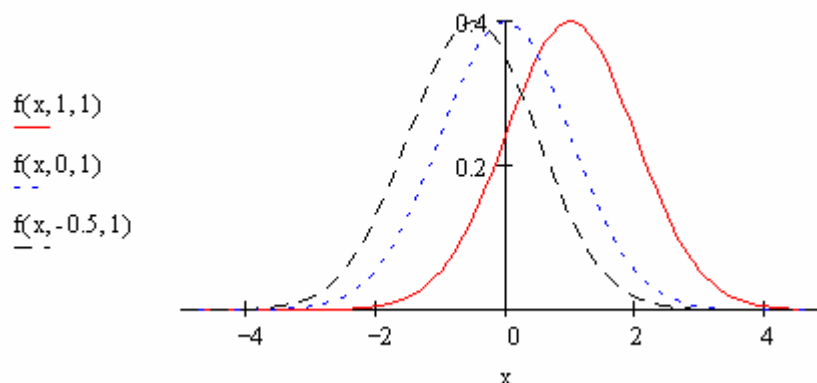
LEY NORMAL DE DISTRIBUCION (LEY DE GAUSS)

La densidad de probabilidad de una magnitud aleatoria X normalmente repartida se expresa por la fórmula:

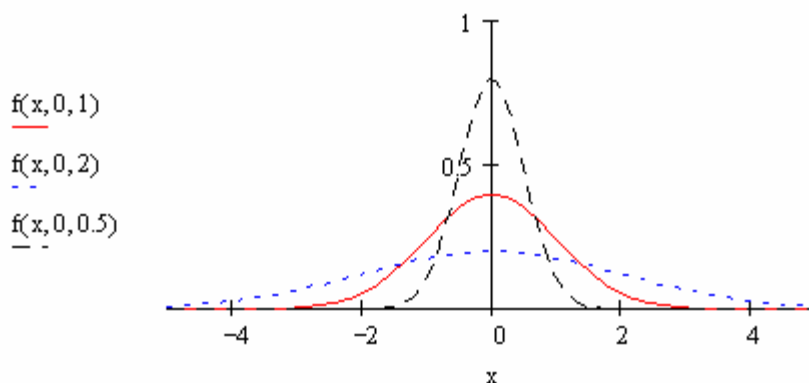
$$f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{x - \mu}{\sigma} \right)^2 \right] \quad (1)$$

En las siguientes figuras se muestra el efecto de cada uno de los parámetros σ y μ sobre la forma de la curva.

$$f(x, \mu, \sigma) := \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{x - \mu}{\sigma} \right)^2 \right] \quad x := -5, -4.9 \dots 5$$



Como se puede apreciar, la variación de μ (con σ fijo) produce el desplazamiento de la función en el sentido del eje x . Siendo la abscisa del máximo, el valor de μ .



En cambio, la variación de σ (con μ fijo) produce el "aplastamiento" (σ grande) o "aguzamiento" (σ chico) de la función.

En todos los casos, la función es simétrica respecto del punto $x = \mu$. Por eso al parámetro μ se lo llama Centro de distribución o Centro de dispersión. Cuando $x = \mu$ la ordenada de la curva de densidad normal de la probabilidad es igual a:

$$f(\mu) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}}$$

Al disminuir σ esta ordenada crece ilimitadamente y viceversa. Pero cualquiera sea la situación, el área bajo la curva a lo largo de todo el eje x es 1.

De este modo, al disminuir σ , la densidad de probabilidad en el centro de dispersión crece y en los demás puntos, a partir de cierto valor de σ , se reduce. En otras palabras, al disminuir σ disminuye la dispersión de los valores posibles de la magnitud aleatoria.

La ley normal de distribución está muy difundida en los problemas prácticos. Liapunov fue el primero en aclarar las causas de esta amplia difusión.

Liapunov demostró que si una magnitud aleatoria puede considerarse como la suma de un gran número de pequeños sumandos, entonces, al ser las condiciones lo suficientemente comunes, la ley de distribución de dicha magnitud se aproxima a la normal

sin que importe cuales sean las leyes de distribución de los sumandos tomados por separados.

Puesto que prácticamente las magnitudes aleatorias, en la mayoría de los casos, suelen ser el resultado de la acción de un gran número de diferentes causas, la ley normal resulta ser el principio de distribución más difundido. No obstante, en muchos problemas se pueden encontrar leyes de distribución distintas de la normal.

Es importante hallar los **momentos** de una magnitud normalmente distribuida. Para esto se hace uso de las siguientes fórmulas:

$$\int_{-\infty}^{\infty} t^{2k-1} \cdot \exp\left(-\frac{h^2}{t^2}\right) dt = 0 \quad \int_{-\infty}^{\infty} t^{2k} \cdot \exp\left(-\frac{h^2}{t^2}\right) dt = \frac{(2k-1)!!}{2^k \cdot h^{2k+1}} \cdot \sqrt{\pi}$$

donde con $(2k-1)!!$ se designa al producto de todos los números enteros impares a partir de 1 hasta $2k-1$.

De acuerdo con estas expresiones, la esperanza matemática de una magnitud aleatoria normalmente distribuida (momento de primer orden) se expresa por la fórmula:

$$m_x = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \int_{-\infty}^{\infty} x \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{x - \mu}{\sigma}\right)^2\right] dx \quad (2)$$

sustituyendo las variables $t = x - \mu$ y valiéndose de las expresiones anteriores, siendo:

$$k=1 \quad h = \frac{1}{\sigma \cdot \sqrt{2}}$$

$$m_x = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \int_{-\infty}^{\infty} (\mu + t) \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{t}{\sigma}\right)^2\right] dt$$

$$m_x = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \int_{-\infty}^{\infty} \mu \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{t}{\sigma}\right)^2\right] dt + \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \int_{-\infty}^{\infty} t \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{t}{\sigma}\right)^2\right] dt$$

El segundo término del segundo miembro es nulo, como se puede verificar integrando por partes y aplicando Barrow.

$$m_x = \mu \cdot \left[\frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \cdot \left(\frac{t}{\sigma} \right)^2 \right] dt \right] = \mu$$

ya que la expresión entre corchetes es 1. Este resultado podía ser previsto de antemano, teniendo en cuenta la simetría de la distribución normal con respecto al centro de distribución μ .

Para determinar los momentos centrales, se recurre a:

$$\psi_r = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \int_{-\infty}^{\infty} (x - \mu)^r \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{x - \mu}{\sigma} \right)^2 \right] dx$$

$$\psi_r = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \int_{-\infty}^{\infty} t^r \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{t}{\sigma} \right)^2 \right] dt$$

Conforme a la resolución de este tipo de integral (visto más arriba) es nula para valores impares de r . Por consiguiente, **todos los momentos centrales de los ordenes impares de una magnitud normalmente distribuida son iguales a cero.**

$$\psi_{2 \cdot k - 1} = 0 \quad \text{para } k = 1, 2, \dots$$

Si $r = 2 \cdot k$ es par:

$$h = \frac{1}{\sigma \cdot \sqrt{2}} \quad \int_{-\infty}^{\infty} t^{2 \cdot k} \cdot \exp \left(-\frac{h^2}{t^2} \right) dt = \frac{(2 \cdot k - 1)!!}{2^k \cdot h^{2 \cdot k + 1}} \cdot \sqrt{\pi}$$

$$\psi_{2 \cdot k} = (2 \cdot k - 1)!! \cdot \sigma^2 = 1 \cdot 3 \cdot 5 \dots (2 \cdot k - 1) \cdot \sigma^2$$

En particular, para la **dispersión** y la **desviación cuadrática media** de una magnitud X repartida normalmente, se obtienen las fórmulas:

$$D_x = \psi_2 = \sigma^2 \quad \sigma_x = \sqrt{D_x} = \sigma \quad (3)$$

De este modo, el parámetro σ_x de la ley normal de distribución representa la desviación cuadrática media de la magnitud aleatoria.

Para los momentos de ordenes pares superiores:

$$\psi_{2 \cdot k} = (2 \cdot k - 1)!! \quad D_x^k = 1 \cdot 3 \cdot 5 \dots (2 \cdot k - 1) \cdot \sigma^2$$

Como se ve, todos los momentos de los órdenes pares de la magnitud normalmente repartida se expresan por medio de su dispersión.

Sustituyendo en (1) la expresión de σ en función de la dispersión de la fórmula (3) y teniendo en cuenta (2):

$$f(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot D_x}} \cdot \exp \left[-\frac{1}{2} \cdot \frac{(x - m_x)^2}{D_x} \right] \quad (4)$$

Esta expresión muestra que la ley normal de distribución se determina por completo por la esperanza matemática y la dispersión de la magnitud aleatoria. Así pues, la esperanza matemática y la dispersión caracterizan de un modo completo la magnitud aleatoria normalmente repartida.

Esto no es cierto para otros tipos de distribuciones.

Distribuciones Estadísticas

página 4

Universidad de Mendoza

Ing. Jesús Rubén Azor Montoya

Matemática

Superior con Mathcad

Para hallar la **Función Característica** de una magnitud aleatoria X normalmente distribuida, se procede del siguiente modo:

$$g(\lambda) = \frac{1}{\sqrt{2 \cdot \pi \cdot D_x}} \cdot \int_{-\infty}^{\infty} \exp \left[i \cdot \lambda \cdot x - \frac{1}{2} \cdot \frac{(x - m_x)^2}{D_x} \right] dx$$

haciendo la sustitución

$$t = x - m_x \quad dt = dx$$

$$g(\lambda) = \frac{1}{\sqrt{2 \cdot \pi \cdot D_x}} \cdot \int_{-\infty}^{\infty} \exp \left[i \cdot \lambda \cdot (t + m_x) - \frac{1}{2} \cdot \frac{t^2}{D_x} \right] dt$$

$$g(\lambda) = \frac{1}{\sqrt{2 \cdot \pi \cdot D_x}} \cdot \int_{-\infty}^{\infty} \exp(i \cdot \lambda \cdot t) \cdot \exp(i \cdot \lambda \cdot m_x) \cdot \exp \left(-\frac{1}{2} \cdot \frac{t^2}{D_x} \right) dt$$

$$g(\lambda) = \frac{\exp(i \cdot \lambda \cdot m_X)}{\sqrt{2 \cdot \pi \cdot D_X}} \cdot \int_{-\infty}^{\infty} \exp\left(i \cdot \lambda \cdot t - \frac{1}{2} \cdot \frac{t^2}{D_X}\right) dt$$

Esta última integral se puede calcular por la fórmula:

$$\int_{-\infty}^{\infty} \exp(\eta \cdot t - h^2 \cdot t^2) dt = \frac{\sqrt{\pi}}{h} \cdot \exp\left(\frac{\eta^2}{4 \cdot h^2}\right)$$

Aplicando la fórmula, se obtiene:

$$g(\lambda) = \exp\left(i \cdot \lambda \cdot m_X - \frac{1}{2} \cdot D_X \cdot \lambda^2\right)$$

Esta expresión es un pie para calcular los momentos de las magnitudes aleatorias normalmente distribuidas, a través de las derivadas sucesivas de esta función evaluadas en $\lambda = 0$.

Por ejemplo:

1)

$$g'(\lambda) = (i \cdot m_X - D_X \cdot \lambda) \cdot \exp\left(i \cdot \lambda \cdot m_X - \frac{1}{2} \cdot D_X \cdot \lambda^2\right)$$

$$g'(0) = (i \cdot m_X - D_X \cdot 0) \cdot \exp\left(i \cdot 0 \cdot m_X - \frac{1}{2} \cdot D_X \cdot 0^2\right)$$

$$|g'(0)| = m_X \quad \text{momento de primer orden}$$

2)

$$g''(\lambda) = -D_X \cdot \exp\left(i \cdot \lambda \cdot m_X - \frac{1}{2} \cdot D_X \cdot \lambda^2\right) + (i \cdot m_X - D_X \cdot \lambda)^2 \cdot \exp\left(i \cdot \lambda \cdot m_X - \frac{1}{2} \cdot D_X \cdot \lambda^2\right)$$

$$g''(0) = -D_X \cdot \exp\left(i \cdot 0 \cdot m_X - \frac{1}{2} \cdot D_X \cdot 0^2\right) + (i \cdot m_X - D_X \cdot 0)^2 \cdot \exp\left(i \cdot 0 \cdot m_X - \frac{1}{2} \cdot D_X \cdot 0^2\right)$$

$$|g''(0)| = D_X + m_X^2 \quad \text{momento de segundo orden}$$

Es sumamente importante deducir la fórmula para la probabilidad de que los valores de una magnitud aleatoria X normalmente repartida se encuentren en el intervalo dado (α, β) .

Sustituyendo en la expresión (4) D_X por σ_X^2 :

$$f(x) = \frac{1}{\sigma_X \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left[-\frac{1}{2} \cdot \frac{(x - m_X)^2}{\sigma_X^2} \right]$$

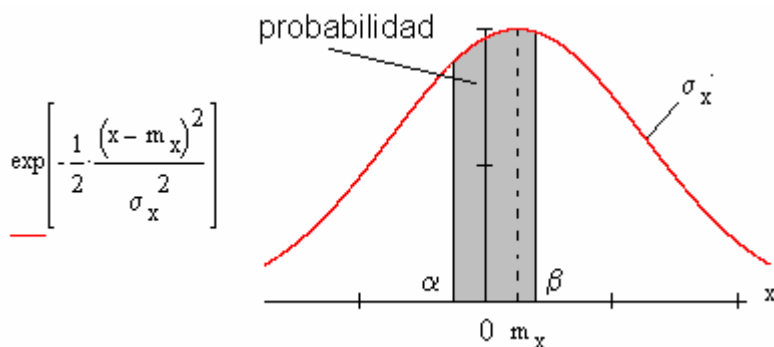
$$P(\alpha < X < \beta) = \frac{1}{\sigma_X \cdot \sqrt{2 \cdot \pi}} \cdot \int_{\alpha}^{\beta} \exp \left[-\frac{1}{2} \cdot \frac{(x - m_X)^2}{\sigma_X^2} \right] dx$$

Gráfica y numéricamente:

Se quiere encontrar la probabilidad de que la magnitud aleatoria X , con distribución normal de media m_X y desviación standard σ_X , se encuentre en el intervalo (α, β) .

$$m_X := 0.5 \quad \sigma_X := 2 \quad \alpha := -0.5 \quad \beta := 0.8 \quad \text{datos}$$

$$\frac{1}{\sigma_X \cdot \sqrt{2 \cdot \pi}} \cdot \int_{\alpha}^{\beta} \exp \left[-\frac{1}{2} \cdot \frac{(x - m_X)^2}{\sigma_X^2} \right] dx = 0.251 \quad \text{Probabilidad}$$



Haciendo la sustitución:

$$z = \frac{x - m_X}{\sigma_X} \quad dz = \frac{1}{\sigma_X} \cdot dx$$

queda

$$P(\alpha < X < \beta) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_{\frac{\alpha - m_x}{\sigma_x}}^{\frac{\beta - m_x}{\sigma_x}} \exp\left(-\frac{1}{2} \cdot z^2\right) dz \quad (5)$$

Esta integral no se expresa por medio de las funciones elementales. Sin embargo, puede hallarse una aproximación por serie que mecaniza el cálculo para computadora (ver ALGORITMO PARA EL CALCULO DE AREAS BAJO LA DISTRIBUCION GAUSSIANA, Aparecida en Revista de la Universidad de Mendoza. Mendoza, N° 14, año 1995, págs. 225/230).

En general, para calcularla, se introduce una nueva función.

$$\Phi(u) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_0^u \exp\left(-\frac{1}{2} \cdot z^2\right) dz$$

Esta función se denomina **función de Laplace** o **integral de las probabilidades**.

Los valores de esta función están tabulados y aparecen en toda bibliografía referente al tema.

La expresión (5) se puede desdoblar en:

$$\begin{aligned} \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_{\frac{\alpha - m_x}{\sigma_x}}^{\frac{\beta - m_x}{\sigma_x}} \exp\left(-\frac{1}{2} \cdot z^2\right) dz &= \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_0^{\frac{\beta - m_x}{\sigma_x}} \exp\left(-\frac{1}{2} \cdot z^2\right) dz \dots \\ &\quad + \left[- \int_0^{\frac{\alpha - m_x}{\sigma_x}} \exp\left(-\frac{1}{2} \cdot z^2\right) dz \right] \\ P(\alpha < X < \beta) &= \Phi\left(\frac{\beta - m_x}{\sigma_x}\right) - \Phi\left(\frac{\alpha - m_x}{\sigma_x}\right) \end{aligned}$$

Mathcad tiene incorporadas funciones que trabajan directamente sobre la distribución normal y que son de gran aplicación. Ellas son:

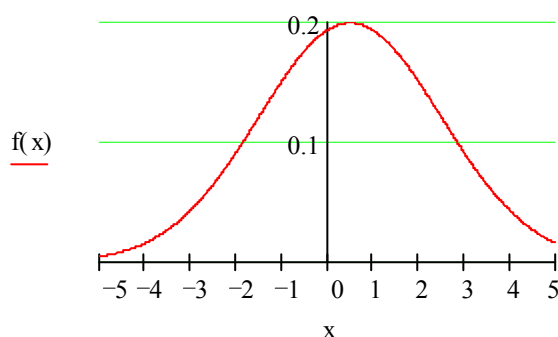
1) dnorm(x,m_x,σ_x): Retorna la densidad de probabilidad para la distribución normal

$$m_x = 0.5 \quad \sigma_x = 2$$

$$\text{dnorm}(1, m_x, \sigma_x) = 0.193 \quad \text{ordenada de la función densidad en la abscisa } x$$

$$f(x) := \frac{1}{\sigma_x \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{x - m_x}{\sigma_x} \right)^2 \right] \quad f(1) = 0.193$$

$$x := -5, -4.99 \dots 5$$



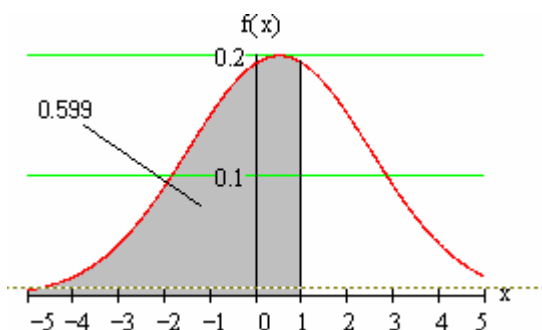
2) pnorm(x,m_x,σ_x): Retorna la distribución de probabilidad acumulativa, área bajo la curva entre - infinito y el valor x.

Las Tablas standards entregan los valores de esta función (entrando por las abscisas).

$$\text{pnorm}(1, m_x, \sigma_x) = 0.598706$$

$$\frac{1}{\sigma_x \cdot \sqrt{2 \cdot \pi}} \cdot \int_{-10}^1 \exp \left[-\frac{1}{2} \cdot \frac{(x - m_x)^2}{\sigma_x^2} \right] dx = 0.598706$$

Se usa, en este caso, en lugar de -infinito el valor de -10, ya que para los valores de media y desviación standard del ejemplo, la función prácticamente se anula. Si se usara -infinito se produce un desborde (overflow).



3) `pnorm(x)`: Idem que la anterior pero para una distribución con media 0 y desviación standard 1.

$$\text{pnorm}(1, 0, 1) = 0.841 \quad \text{cnorm}(1) = 0.841$$

4) `qnorm(p, mx, σx)`: Retorna la inversa de la distribución de probabilidad acumulativa.

Actúa de modo inverso respecto de **`pnorm`**, es decir dada la probabilidad **`p`** calcula la abscisa correspondiente **`x`**.

Las Tablas standards entregan los valores de esta función (entrando por las probabilidades).

$$\text{qnorm}(0.598706, m_x, \sigma_x) = 1$$

5) `rnorm(m, mx, σx)`: Retorna un vector de **`m`** números aleatorios que tienen la distribución normal con media **`mx`** y desviación standard **`sx`**.

$$\text{rnorm}(6, m_x, \sigma_x)^T = (-0.378 \ -0.859 \ -0.447 \ -1.403 \ -2.871 \ 0.587)$$

EJEMPLO:

Verificar que el siguiente vector tiene elementos que constituyen una muestra proveniente de población normalmente distribuida.

`n := 1000` Tamaño de la muestra

`v := rnorm(n, mx, σx)`

`n1 := 10` número de intervalos

`j := 0 .. n1` índice para intervalos

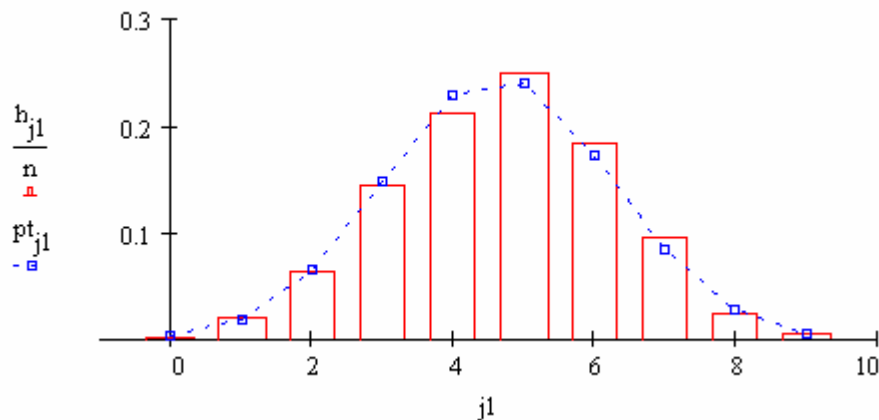
$$I_j := \min(v) + \frac{\max(v) - \min(v)}{n1} \cdot j \quad \text{vector de intervalos}$$

$$j1 := 0..n1 - 1 \quad \text{índice auxiliar}$$

$$h := \text{hist}(I, v) \quad \text{histograma}$$

Para establecer comparaciones con los valores esperados en cada intervalo de la distribución, se define la probabilidad teórica en cada intervalo:

$$pt_{j1} := \int_{I_{j1}}^{I_{j1+1}} \frac{1}{\sigma_x \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{x - m_x}{\sigma_x} \right)^2 \right] dx$$



Para generar distintos vectores aleatorios, basta poner el cursor sobre la función **rnorm** y apretar la tecla F9. En el gráfico se apreciará permanentemente el buen ajuste de la distribución obtenida respecto de la teórica usada como referencia.

También se puede calcular la media y la desviación standard del vector de datos que no diferirá apreciablemente de los parámetros elegidos.

$$\text{mean}(v) = 0.599 \quad m_x = 0.5 \quad \sigma_x = 2 \quad \text{stdev}(v) = 1.953$$

LEY DE POISSON

Supóngase que ciertos acontecimientos sucedan en los mementos aleatorios continuamente repartidos en el eje numérico. Tales acontecimientos forman una secuencia de sucesos llamada habitualmente flujo de acontecimientos. En calidad de ejemplo de un flujo de acontecimiento pueden servir las llamadas de los abonados en una central telefónica, paso de los medios de transporte por un cruce, fallo de los elementos de algún sistema técnico.

En muchos problemas de la práctica se puede considerar que el flujo de acontecimientos satisface las condiciones siguientes:

- 1) si (t_1, t_2) y (t_3, t_4) son cualesquiera intervalos de tiempo no superpuestos, entonces, la probabilidad de que se produzca cualquier número de acontecimientos en el transcurso de uno de dichos intervalos no depende de cuántos acontecimientos ocurren en el transcurso de otro.
- 2) la probabilidad de que un acontecimiento se produzca en el transcurso de un intervalo de tiempo infinitesimal $(t, t+\Delta t)$ es un infinitesimal de orden Δt .
- 3) la probabilidad de que se produzcan más de un acontecimiento en el transcurso del intervalo de tiempo $(t, t+\Delta t)$ es un infinitesimal de orden superior en comparación con Δt .

Es evidente que el número de acontecimientos que sucedan durante cualquier intervalo de tiempo (t_0, t) representa una magnitud aleatoria discreta cuyos valores posibles son todos los números enteros no negativos 0, 1, 2, ...

Si se designa esta magnitud aleatoria por $X(t)$ y la probabilidad de que esta tome el valor de m se designa por $p_m(t_0, t)$ o, más brevemente, $p_m(t)$:

$$P(X=m)=p_m(t_0, t)=p_m(t) \quad (m = 0, 1, 2, \dots)$$

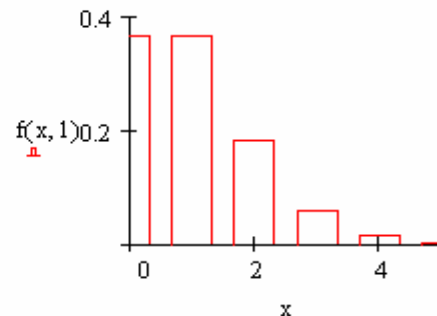
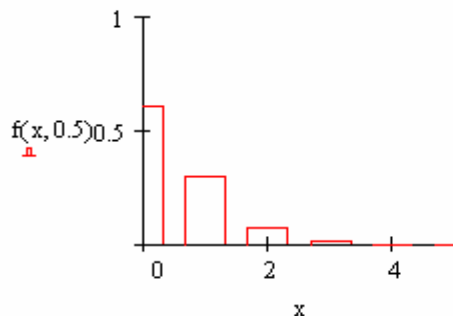
La ley de distribución de esta magnitud aleatoria tiene la siguiente forma:

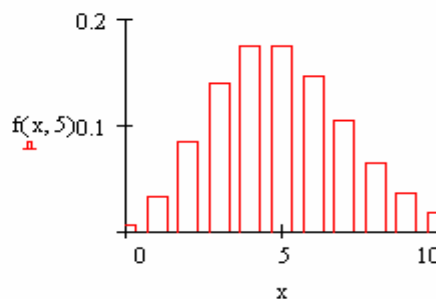
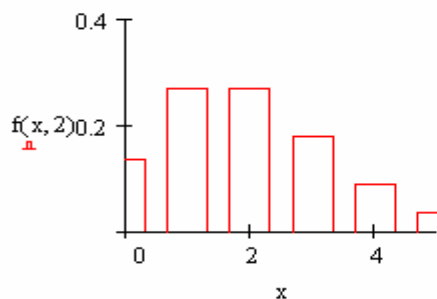
$$f(x)=\frac{\mu^x}{x!} \cdot e^{-\mu} \quad \text{Función densidad de la Distribución de Poisson}$$

A continuación se muestra a $f(x)$ para varios μ .

$$f(x, \mu) := \frac{\mu^x}{x!} \cdot e^{-\mu} \quad \text{función}$$

$x := 0..10 \quad \text{rango}$





El parámetro μ es la media de la distribución, como se demostrará más adelante.

De la expresión de la función densidad se deduce la función distribución correspondiente.

$$F(x) = e^{-\mu} \cdot \sum_{s=0}^x \frac{\mu^s}{s!} \quad \text{Función distribución de la Distribución de Poisson}$$

y $F(x) = 0$ cuando $x < 0$.

APLICACIONES DE LA DISTRIBUCION DE POISSON

La distribución de Poisson tiene aplicaciones en diferentes campos.

Ejemplo 1: de la Física de Partículas, la variable x es el número de partículas Alfa en intervalos de tiempo de longitud 7.5 seg, conforme a la siguiente tabla:

x	Número de Intervalos de tiempo (longitud 7.5 seg) con x partículas por intervalo	
	Observadas	Teóricas (redondeadas)
0	57	54
1	203	210
2	383	407
3	525	526
4	532	509
5	408	394
6	273	254
7	139	140
8	45	68
9	27	29
10	10	11
11	4	4
12	2	1
13	0	1

Las dos primeras columnas surgen de valores experimentales y se representarán por los vectores:

$i := 0..13$ $x_i := i$ vector x
 $y := (57 \ 203 \ 383 \ 525 \ 532 \ 408 \ 273 \ 139 \ 45 \ 27 \ 10 \ 4 \ 2 \ 0)$
 $y := y^T$ vector de frecuencias observadas

La media ponderada es:

$$\mu := \frac{\sum_i x_i \cdot y_i}{\sum y} \quad \mu = 3.87$$

$\mu = 3.87$

Las probabilidades normalizada se obtienen aplicando la función densidad de probabilidad de Poisson:

$$p_i := \frac{\mu^{x_i} \cdot \exp(-\mu)}{x_i!} \quad \text{vector de probabilidades normalizadas}$$

Desnormalizando y redondeando a entero, para el ejemplo:

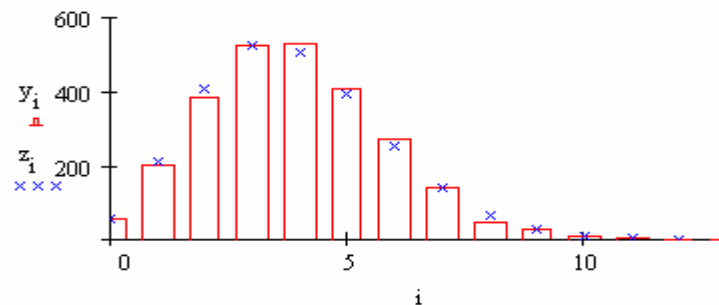
$$z_i := p_i \cdot \sum y$$

$$z_i := p_i \cdot \sum y$$

$$z^T =$$

	0	1	2	3	4	5	6	7	8	9
	54	210	407	525	508	394	254	140	68	29

que es la tercera columna de la Tabla, valores teóricos de probabilidad para cada intervalo. Si se observan gráficamente los valores experimentales y teóricos se aprecia que la distribución de este fenómeno obedece a la distribución de Poisson.



Ejemplo 2: Distribución del tráfico en una calle, según la siguiente tabla:

x	Número de Intervalos de tiempo (longitud 30 seg) con x autos por intervalo	
	Observadas	Teóricas (redondeadas)
0	6	6
1	18	17
2	21	24
3	26	22
4	16	15
5	8	9
6	2	4
7	1	2
8	2	1
9	0	0

Las dos primeras columnas surgen de valores experimentales y se representarán por los vectores:

$i := 0..9$ $x_i := i$ vector x
 $y := (6 \ 18 \ 21 \ 26 \ 16 \ 8 \ 2 \ 1 \ 2 \ 0)$
 $y := y^T$ vector de frecuencias observadas

La media ponderada es:

$$\mu := \frac{\sum_i x_i \cdot y_i}{\sum y} \quad \mu = 2.77$$

Las probabilidades normalizada se obtienen aplicando la función densidad de probabilidad de Poisson:

$$p_i := \frac{\mu^{x_i} \cdot \exp(-\mu)}{x_i!} \quad \text{vector de probabilidades normalizadas}$$

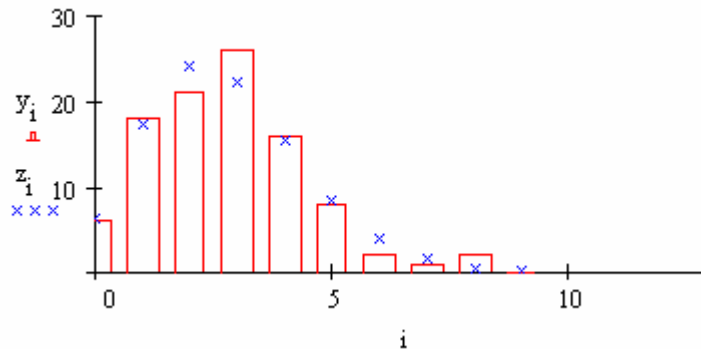
Desnormalizando y redondeando a entero, para el ejemplo:

$$z_i := p_i \cdot \sum y$$

$$z^T =$$

	0	1	2	3	4	5	6	7	8	9	10	11
0	6	17	24	22	15	9	4	2	1	0	11	4

que es la tercera columna de la Tabla, valores teóricos de probabilidad para cada intervalo. Si se observan gráficamente los valores experimentales y teóricos se aprecia que la distribución de este fenómeno obedece a la distribución de Poisson.



VARIANZA Y SESGO DE LA DISTRIBUCION DE POISSON

En una distribución discreta, la función generadora de momentos [$G(t)$] se expresa:

$$G(t) = \sum_j e^{t \cdot x_j} \cdot f(x_j)$$

donde $f(.)$ es la función densidad en consideración. Para el caso de Poisson:

$$G(t) = \sum_j e^{t \cdot x_j} \cdot \frac{\mu^{x_j} \cdot e^{-\mu}}{x_j!} = e^{-\mu} \cdot \sum_j e^{t \cdot x_j} \cdot \frac{\mu^{x_j}}{x_j!}$$

Dado que los x_j toman los valores 0, 1, 2, ...

$$\sum_j e^{t \cdot x_j} \cdot \frac{\mu^{x_j}}{x_j!} = 1 + e^{t \cdot \mu} + e^{t \cdot 2} \cdot \frac{\mu^2}{2!} + e^{t \cdot 3} \cdot \frac{\mu^3}{3!} \dots$$

$$\sum_j e^{t \cdot x_j} \cdot \frac{\mu^{x_j}}{x_j!} = 1 + (e^t \cdot \mu) + \frac{(e^t \cdot \mu)^2}{2!} + \frac{(e^t \cdot \mu)^3}{3!} \dots$$

Distribuciones Estadísticas

página 16

Universidad de Mendoza

Ing. Jesús Rubén Azor Montoya

Matemática

Superior con Mathcad

$$G(t) = e^{-\mu} \cdot \left(\sum_j e^{t \cdot x_j} \cdot \frac{\mu^{x_j}}{x_j!} \right) = e^{-\mu} \cdot e^{\mu \cdot e^t} \quad \text{Función generalizada de momentos}$$

Si se quiere hallar la media de la distribución, basta con evaluar la derivada de la función $G(t)$ en $t=0$:

$$\begin{aligned} \frac{d}{dt} G(t) &= e^{-\mu} \cdot \mu \cdot e^t \cdot e^{\mu \cdot e^t} \\ \frac{d}{dt} G(0) &= e^{-\mu} \cdot \mu \cdot e^0 \cdot e^{\mu \cdot e^0} = e^{-\mu} \cdot \mu \cdot 1 \cdot e^{\mu \cdot 1} = \mu \end{aligned}$$

El resultado es el esperado. Para hallar el momento de segundo orden no centrado, se recurre a calcular la derivada segunda de la función generadora de momentos y evaluarla en 0.

$$\begin{aligned} \frac{d^2}{dt^2} G(t) &= e^{-\mu} \cdot \mu \cdot e^t \cdot e^{\mu \cdot e^t} + e^{-\mu} \cdot \mu^2 \cdot (e^t)^2 \cdot e^{\mu \cdot e^t} \\ \frac{d^2}{dt^2} G(0) &= e^{-\mu} \cdot \mu \cdot e^0 \cdot e^{\mu \cdot e^0} + e^{-\mu} \cdot \mu^2 \cdot (e^0)^2 \cdot e^{\mu \cdot e^0} = e^{-\mu} \cdot \mu \cdot 1 \cdot e^{\mu \cdot 1} + e^{-\mu} \cdot \mu^2 \cdot 1^2 \cdot e^{\mu \cdot 1} \\ \frac{d^2}{dt^2} G(0) &= \mu + \mu^2 \quad E(X^2) = \mu + \mu^2 \quad \text{Momento no centrado} \end{aligned}$$

Para calcular la varianza (o dispersión) que es el momento de segundo orden centrado, se recurre a:

$$\sigma^2 = E[(X - \mu)^2] = E(X^2) - 2 \cdot E(X) \cdot \mu + \mu^2 = E(X^2) - 2 \cdot \mu \cdot \mu + \mu^2 = E(X^2) - \mu^2$$

Reemplazando por lo obtenido en el paso anterior:

$$\sigma^2 = (\mu + \mu^2) - \mu^2 = \mu$$

Conclusión: En una distribución de Poisson, la varianza coincide con la media.

Para hallar el momento de tercer orden, se haya la derivada tercera:

$$\frac{d^3}{dt^3} G(t) = e^{-\mu} \cdot \mu \cdot e^t \cdot e^{\mu \cdot e^t} + e^{-\mu} \cdot \mu^2 \cdot (e^t)^2 \cdot e^{\mu \cdot e^t} = e^{-\mu} \cdot e^t \cdot e^{\mu \cdot e^t} \cdot \mu \cdot (1 + \mu \cdot e^t)$$

$$\frac{d^2}{dt^2} G(t) = e^{-\mu + t + \mu \cdot e^t} \cdot \mu \cdot (1 + \mu \cdot e^t)$$

$$\frac{d^3}{dt^3} G(t) = (1 + \mu \cdot e^t)^2 \cdot e^{-\mu + t + \mu \cdot e^t} \cdot \mu + e^{-\mu + t + \mu \cdot e^t} \cdot \mu^2 \cdot e^t$$

$$\frac{d^3}{dt^3} G(0) = (1 + \mu \cdot e^0)^2 \cdot e^{-\mu + 0 + \mu \cdot e^0} \cdot \mu + e^{-\mu + 0 + \mu \cdot e^0} \cdot \mu^2 \cdot e^0$$

$$\frac{d^3}{dt^3} G(0) = (1 + \mu \cdot 1)^2 \cdot e^{-\mu + \mu \cdot 1} \cdot \mu + e^{-\mu + \mu \cdot 1} \cdot \mu^2 \cdot 1 = \mu \cdot (1 + \mu \cdot 1)^2 + \mu^2$$

$$E(X^3) = \mu + 2 \cdot \mu^2 + \mu^3 + \mu^2 = \mu + 3 \cdot \mu^2 + \mu^3$$

Para hallar el sesgo, se debe encontrar el momento de tercer orden centrado:

$$E[(X - \mu)^3] = E(X^3) - 3 \cdot E(X^2) \cdot \mu + 3 \cdot E(X) \cdot \mu^2 - \mu^3$$

$$E[(X - \mu)^3] = (\mu + 3 \cdot \mu^2 + \mu^3) - 3 \cdot \mu \cdot (\mu + \mu^2) + 3 \cdot \mu^3 - \mu^3$$

$$E[(X - \mu)^3] = \mu + 3 \cdot \mu^2 + \mu^3 - 3 \cdot \mu^2 - 3 \cdot \mu^3 + 2 \cdot \mu^3 = \mu$$

Dado que:

$$\sigma^2 = \mu \quad \sigma^3 = \sigma^2 \cdot \sigma = \mu \cdot \sqrt{\mu}$$

Luego, el sesgo vale:

$$\gamma = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{\mu}{\mu \cdot \sqrt{\mu}} = \frac{1}{\sqrt{\mu}} \quad \text{sesgo}$$

Se ve que γ tiende a 0 cuando la media μ tiende a infinito, de lo cual se deduce que la distribución de Poisson es casi simétrica para μ grande.

Funciones densidad y distribución en Mathcad

Funciones para trabajar con densidades de probabilidad comunes que caen en las siguientes clases:

- **Densidades de probabilidad** – Estas dan la probabilidad de que una variable aleatoria tome un valor particular.
- **Distribuciones de probabilidad acumulativa** - Estas dan la probabilidad de que una variable aleatoria tome un valor **menor o igual** que uno especificado. Estas se obtienen por simple integración (o sumatoria cuando corresponda) de la densidad de probabilidad correspondiente sobre un rango apropiado.
- **Distribuciones de probabilidad acumulativa inversas** – Estas funciones toman una probabilidad como argumento y retornan un valor tal que la probabilidad que una variable aleatoria sea menor o igual que ese valor es la probabilidad que se suministró como argumento.

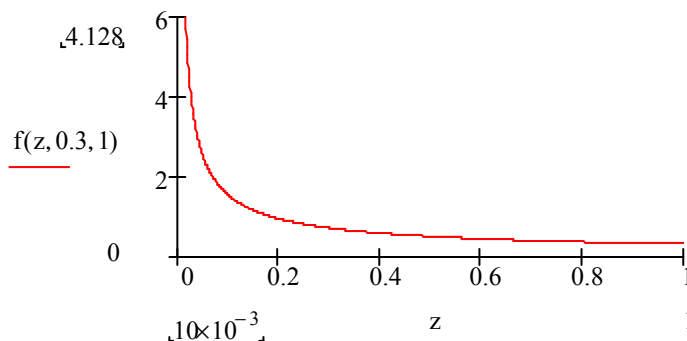
DISTRIBUCIONES CONTINUAS

Distribución Beta:

Responde a la siguiente función densidad:

$$f(z, \alpha, \beta) := \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \cdot (1 - z)^{\beta-1} \cdot z^{\alpha-1}$$

donde $\alpha > 0$ y $\beta > 0$ son parámetros de la distribución y z pertenece a los reales mayores que 0. Gráficamente:



Para encontrar el valor de probabilidad para un z dado, aplicando Mathcad, se recurre a la función built-in (**dbeta**, con $z=0.1$, $\alpha=0.3$ y $\beta=1$):

$$\text{dbeta}(0.1, 0.3, 1) = 1.504 \blacksquare$$

que coincide con la evaluación de la función en el mismo punto dado:

$$f(0.1, 0.3, 1) = 1.504 \blacksquare$$

Del mismo modo, para hallar la función distribución acumulativa de probabilidad,

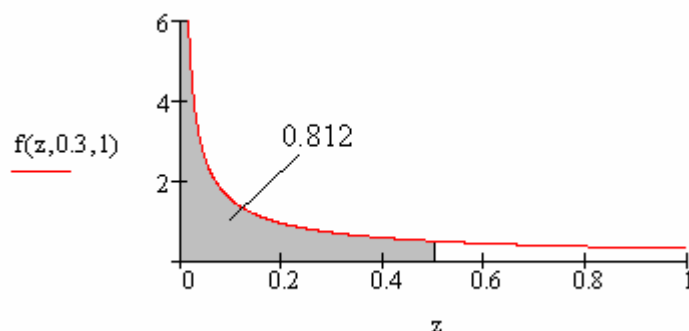
por definición:

$$F(x, \alpha, \beta) := \int_0^x \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \cdot (1 - z)^{\beta-1} \cdot z^{\alpha-1} dz$$

y evaluando para $x=0.5$, la función F y la built-in, **pbeta**:

$$F(0.5, 0.3, 1) = 0.812 \blacksquare \quad \text{pbeta}(0.5, 0.3, 1) = 0.812 \blacksquare$$

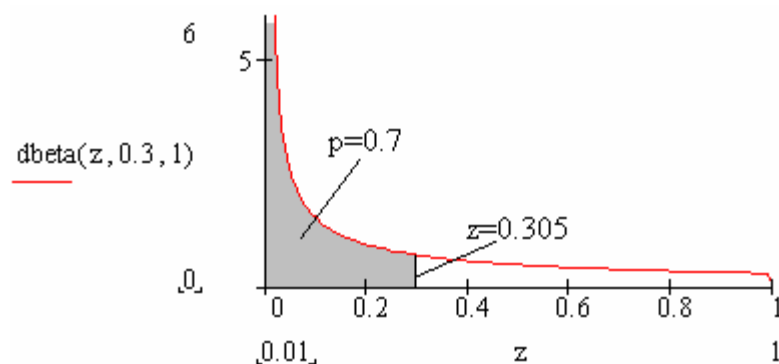
Esto es, la probabilidad que el valor de la variable aleatoria sea menor o igual que 0.5. En otras palabras, el área bajo la función densidad entre $x=0$ y $x=0.5$. Gráficamente:



También es posible encontrar una abscisa z , que deje un área bajo la curva entre 0 y z , de un valor de probabilidad especificado p :

$$qbeta(0.7, 0.3, 1) = 0.305 \blacksquare$$

Gráficamente:



O bien un vector aleatorio de m números aleatorios, que tiene la distribución beta:

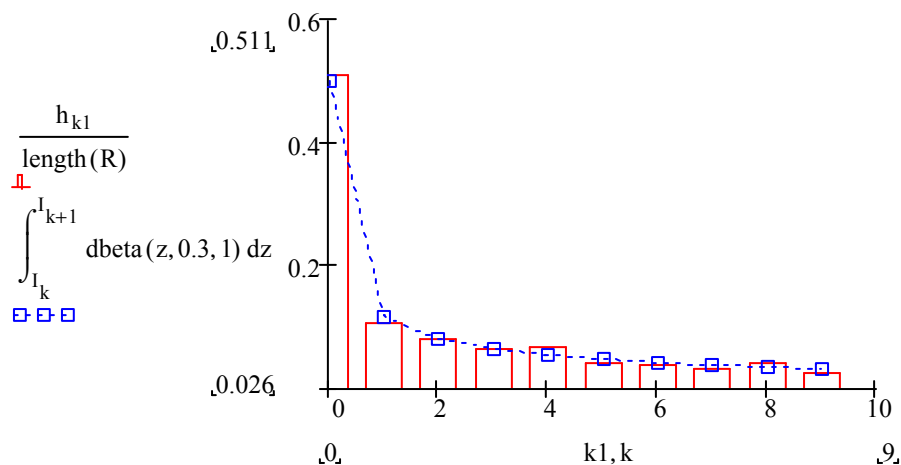
$$R := rbeta(1000, 0.3, 1)$$

Se puede demostrar que los elementos de este vector constituyen una muestra aleatoria con distribución beta, a través de un histograma:

$$i := 0..length(R) - 1 \quad k := 0..10$$

$$I_k := \frac{k}{10} \quad h := hist(I, R) \quad k1 := 0..9$$

Comparando el histograma con la función densidad teórica (línea punteada):

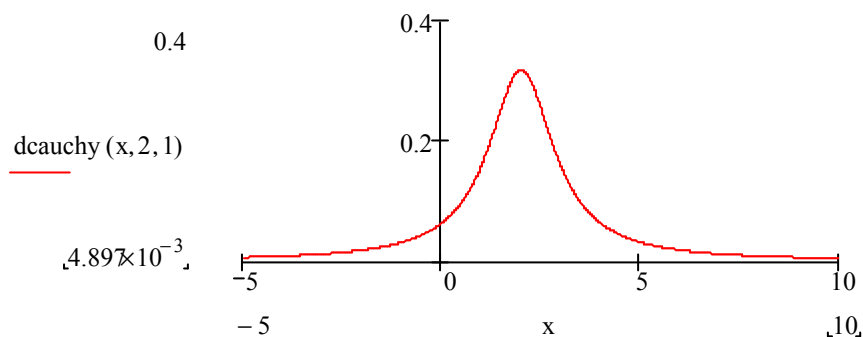


Distribución Cauchy:

Responde a la siguiente función densidad:

$$f_c(x, \alpha, \beta) := \frac{\beta}{\pi \cdot [\beta^2 + (x - \alpha)^2]}$$

donde α perteneciente a los reales y $\beta > 0$ son parámetros de la distribución. Además, z pertenece a los reales. Gráficamente:



Del mismo modo que en el caso anterior, se pueden encontrar:

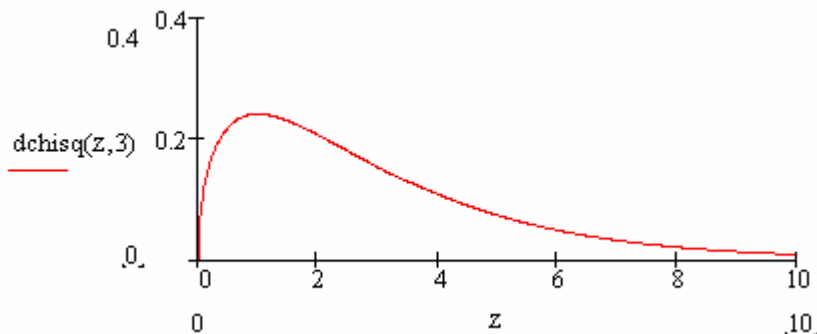
- **dcauchy(z,α,β)** – valor de la densidad de probabilidad para la distribución de Cauchy de parámetros α,β.
- **pcauchy(z,α,β)** – valor del área (probabilidad) comprendida entre $-\infty$ y z para la distribución de Cauchy de parámetros α,β.
- **qcauchy(p,α,β)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre $-\infty$ y dicha abscisa, para la distribución de Cauchy de parámetros α,β.
- **rcauchy(m,α,β)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución de Cauchy de parámetros α,β.

Distribución Chi-cuadrado:

Responde a la siguiente función densidad:

$$f_{\text{chi}}(z, v) := \frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)} \cdot z^{\frac{v}{2}-1} \cdot \exp\left(-\frac{z}{2}\right)$$

donde v es un parámetro llamado grados de libertad y z pertenece a los reales mayores o iguales a 0.



Del mismo modo que en los casos anteriores, se pueden encontrar:

- **dchisq(z,v)** – valor de la densidad de probabilidad para la distribución Chi-cuadrado de parámetro v.
- **pchisq(z,v)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Chi-cuadrado de parámetro v.
- **qchisq(p, v)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Chi-cuadrado de parámetro v.

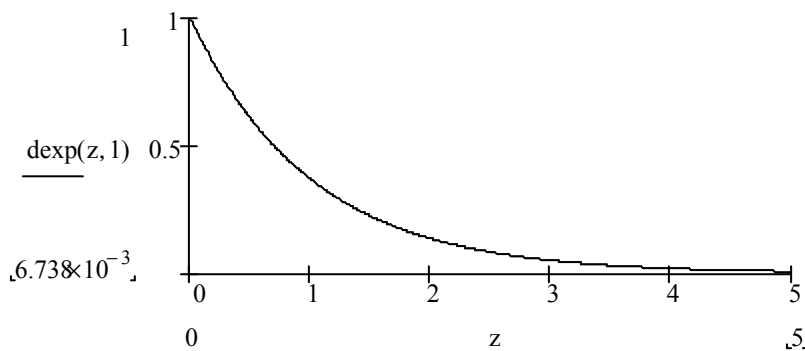
- **rchisq (m, v)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Chi-cuadrado de parámetro v.

Distribución Exponencial:

Responde a la siguiente función densidad:

$$f_{\text{exp}}(z, \lambda) := \lambda \cdot e^{-\lambda z}$$

donde λ es un parámetro y z pertenece a los reales mayores o iguales que 0.



Del mismo modo que en los casos anteriores, se pueden encontrar:

- **dexp(z, λ)** – valor de la densidad de probabilidad para la distribución Exponencial de parámetro λ .
- **pexp(z, λ)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Exponencial de parámetro λ .
- **qexp(p, λ)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Exponencial de parámetro λ .
- **rexp(m, λ)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Exponencial de parámetro λ .

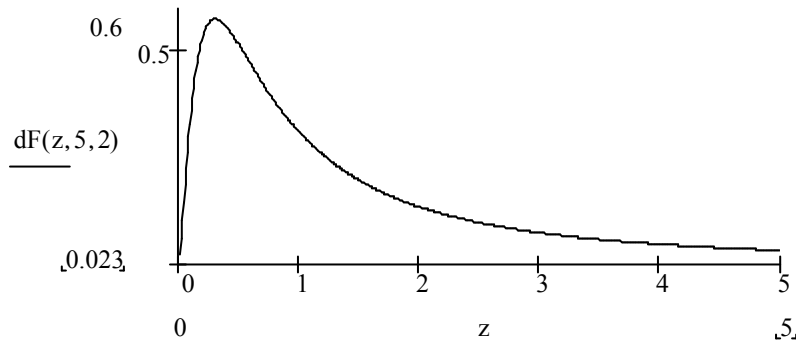
Distribución F:

Responde a la siguiente función densidad:

$$f_F(z, v_1, v_2) := \frac{\Gamma\left(\frac{v_1 + v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right) \cdot \Gamma\left(\frac{v_2}{2}\right)} \cdot \left(v_1 \frac{v_1}{2} \cdot v_2 \frac{v_2}{2}\right) \cdot \frac{z^{\frac{v_1}{2}-1}}{(v_1 \cdot z + v_2)^{\frac{v_1+v_2}{2}}}$$

donde v1 es un parámetro llamado grados de libertad del numerador, v2 son los grados de

libertad del denominador y z pertenece a los reales mayores o iguales a 0.



Del mismo modo que en los casos anteriores, se pueden encontrar:

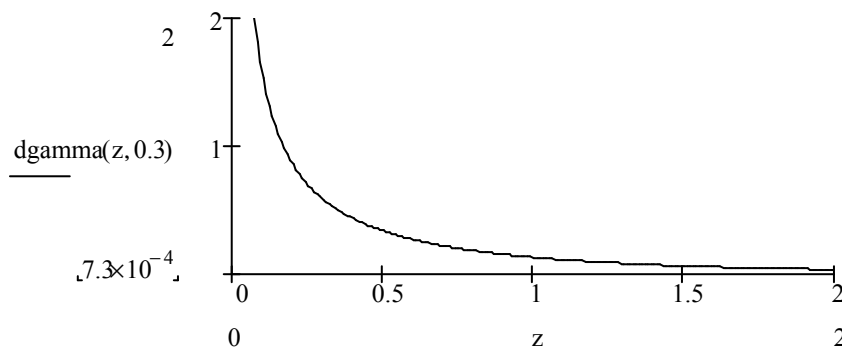
- **dF(z,v1,v2)** – valor de la densidad de probabilidad para la distribución F de parámetros $v1$ y $v2$.
- **pF (z,v1,v2)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución F de parámetros $v1$ y $v2$.
- **qF (p,v1,v2)** – valor de la abscisa que deja un área (probabilidad) p , bajo la curva comprendida entre 0 y dicha abscisa, para la distribución F de parámetros $v1$ y $v2$.
- **rF (m,v1,v2)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución F de parámetros $v1$ y $v2$.

Distribución Gamma:

Responde a la siguiente función densidad:

$$f_{\Gamma}(z, \alpha) := \frac{1}{\Gamma(\alpha)} \cdot \exp(-z) \cdot z^{\alpha-1}$$

donde α es un parámetro y z pertenece a los reales mayores o iguales que 0.



Del mismo modo que en los casos anteriores, se pueden encontrar:

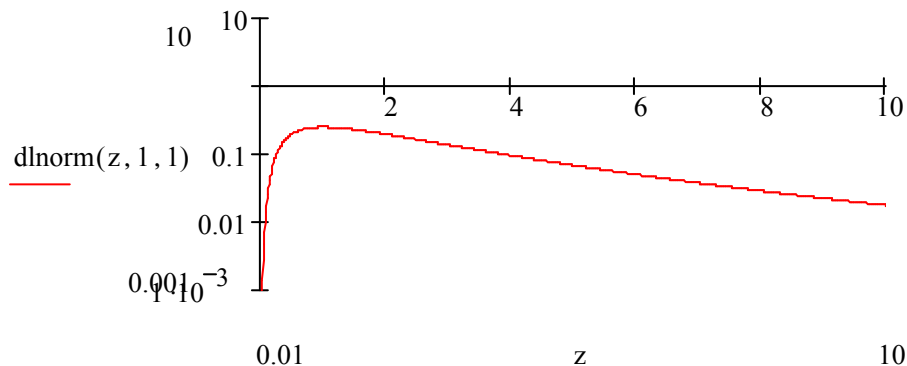
- **dgamma(z,α)** – valor de la densidad de probabilidad para la distribución Gamma de parámetro α.
- **pgamma(z,α)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Gamma de parámetro α.
- **qgamma(z,α)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Gamma de parámetro α.
- **rgamma(z,α)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Gamma de parámetro α.

Distribución Log Normal:

Responde a la siguiente función densidad:

$$f_{\ln}(z, \mu, \sigma) := \frac{1}{\sigma \cdot z \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\left[\frac{(\ln(z) - \mu)^2}{2 \cdot \sigma^2} \right]}$$

donde μ es un parámetro real y σ positivo. Además, z pertenece a los reales mayores a 0.



Del mismo modo que en los casos anteriores, se pueden encontrar:

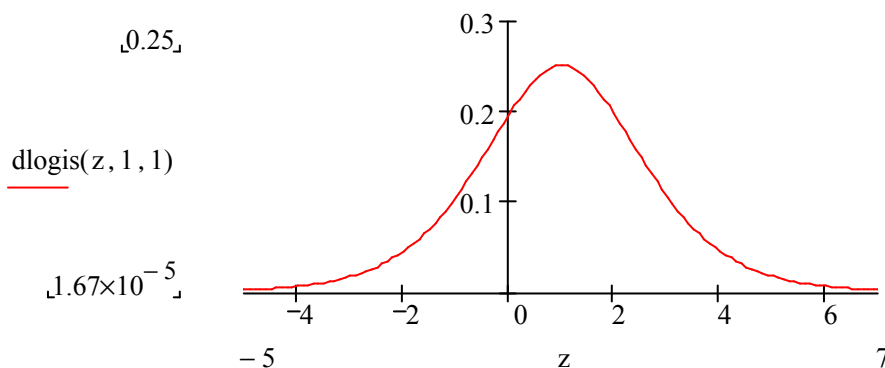
- **dlnorm(z,μ,σ)** – valor de la densidad de probabilidad para la distribución Log Normal de parámetros μ y σ .
- **plnorm(z, μ,σ)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Log Normal de parámetros μ y σ .
- **qlnorm(p, μ,σ)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Log Normal de parámetros μ y σ .
- **rlnorm(m, μ,σ)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Log Normal de parámetros μ y σ .

Distribución Logística:

Responde a la siguiente función densidad:

$$f_{lg}(z, l, s) := \frac{\frac{(z-l)}{s}}{|s| \cdot \left[1 + e^{\frac{(z-l)}{s}} \right]^2}$$

donde l , el llamado parámetro de localización, es un número real; y s , el parámetro de escala, real positivo. Además, z pertenece a los reales.



Del mismo modo que en los casos anteriores, se pueden encontrar:

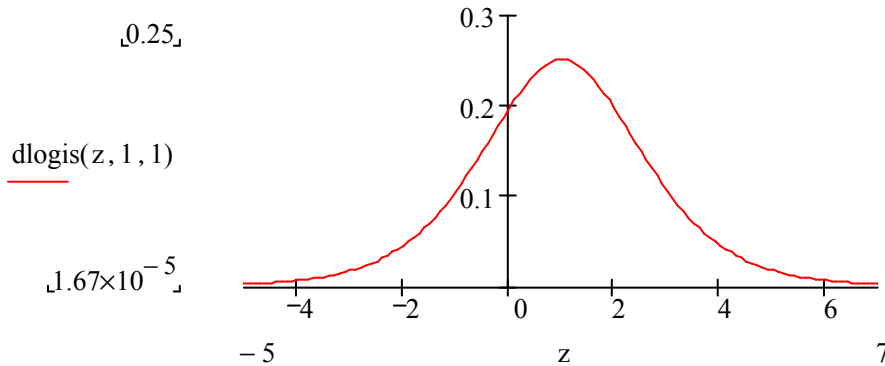
- **dlogis(z,l,s)** – valor de la densidad de probabilidad para la distribución Logística de parámetros l y s .
- **plogis(z,l,s)** – valor del área (probabilidad) comprendida entre $-\infty$ y z para la distribución Logística de parámetros l y s .
- **qlogis(p,l,s)** – valor de la abscisa que deja un área (probabilidad) p , bajo la curva comprendida entre $-\infty$ y dicha abscisa, para la distribución Logística de parámetros l y s .
- **rlogis(m,l,s)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Logística de parámetros l y s .

Distribución Normal:

Responde a la siguiente función densidad:

$$f_N(z, \mu, \sigma) := \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\left[\frac{(z-\mu)^2}{2 \cdot \sigma^2} \right]}$$

donde μ , media, es un número real; y σ , desviación estándar, real positivo. Además, z pertenece a los reales.



Del mismo modo que en los casos anteriores, se pueden encontrar:

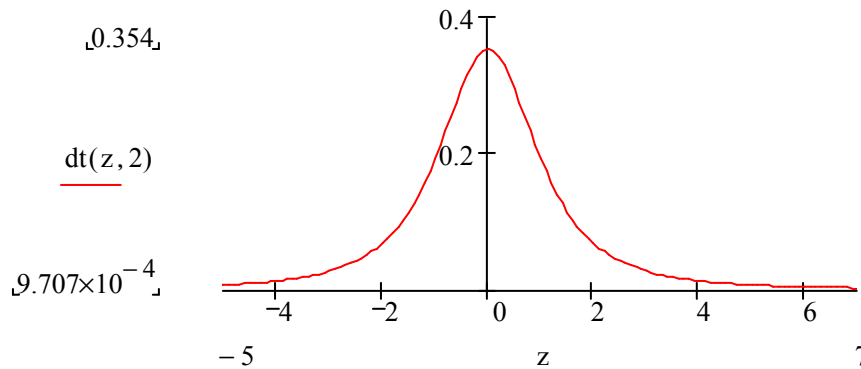
- **dnorm(z,μ,σ)** – valor de la densidad de probabilidad para la distribución Normal de parámetros μ y σ .
- **pnorm(z,μ,σ)** – valor del área (probabilidad) comprendida entre $-\infty$ y z para la distribución Normal de parámetros μ y σ .
- **qnorm(p,μ,σ)** – valor de la abscisa que deja un área (probabilidad) p , bajo la curva comprendida entre $-\infty$ y dicha abscisa, para la distribución Normal de parámetros μ y σ .
- **rnorm(m,μ,σ)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Normal de parámetros μ y σ .

Distribución t-Student:

Responde a la siguiente función densidad:

$$f_t(z, v) := \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v \cdot \pi} \cdot \Gamma\left(\frac{v}{2}\right)} \cdot \frac{1}{\left(1 + \frac{z^2}{v}\right)^{\frac{v+1}{2}}}$$

donde v , grados de libertad, es un número real positivo; y z pertenece a los reales.



Del mismo modo que en los casos anteriores, se pueden encontrar:

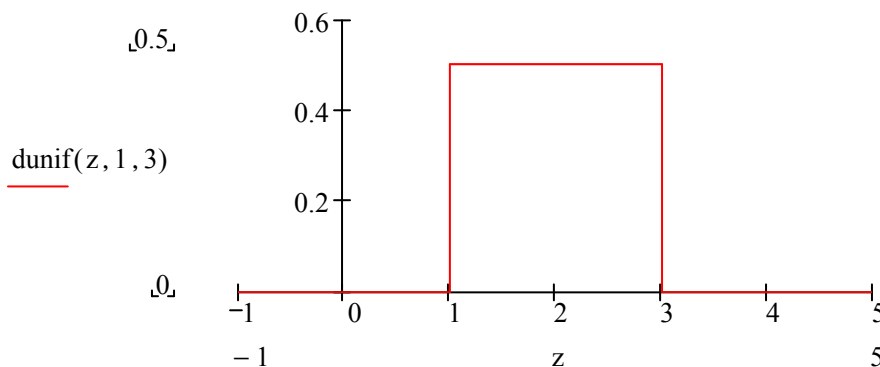
- **dt(z,v)** – valor de la densidad de probabilidad para la distribución t-Student de parámetro v.
- **pt(z,v)** – valor del área (probabilidad) comprendida entre $-\infty$ y z para la distribución t-Student de parámetro v.
- **qt(p,v)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre $-\infty$ y dicha abscisa, para la distribución t-Student de parámetro v.
- **rt(m,v)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución t-Student de parámetro v.

Distribución Uniforme:

Responde a la siguiente función densidad:

$$f_U(z, a, b) := \begin{cases} \frac{1}{b-a}, & a < z \leq b \\ 0, & \text{otherwise} \end{cases}$$

donde a y b son dos parámetros reales y z pertenece a los reales.



Del mismo modo que en los casos anteriores, se pueden encontrar:

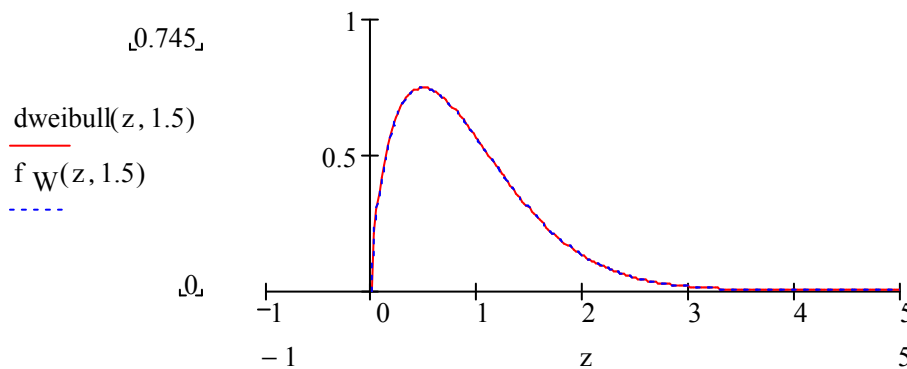
- **dunif(z,a,b)** – valor de la densidad de probabilidad para la distribución Uniforme de parámetros a y b.
- **punif(z,a,b)** – valor del área (probabilidad) comprendida entre $-\infty$ y z para la distribución Uniforme de parámetros a y b.
- **qunif(p,a,b)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre $-\infty$ y dicha abscisa, para la distribución Uniforme de parámetros a y b.
- **runif(m,a,b)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Uniforme de parámetros a y b.
- **rnd(z)**: Retorna un número aleatoriamente distribuido entre 0 y z.

Distribución de Weibull:

Responde a la siguiente función densidad:

$$f_W(z, s) := s \cdot z^{s-1} \cdot e^{-z^s}$$

donde s, parámetro, es un número real positivo; y z pertenece a los reales positivos.



Del mismo modo que en los casos anteriores, se pueden encontrar:

- **dweibull(z,s)** – valor de la densidad de probabilidad para la distribución Weibull de parámetro s.
- **pweibull(z,s)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Weibull de parámetro s.
- **qweibull(p,s)** – valor de la abscisa que deja un área (probabilidad) p, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Weibull de parámetro s.

- **rweibull (m,s)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Weibull de parámetro s.

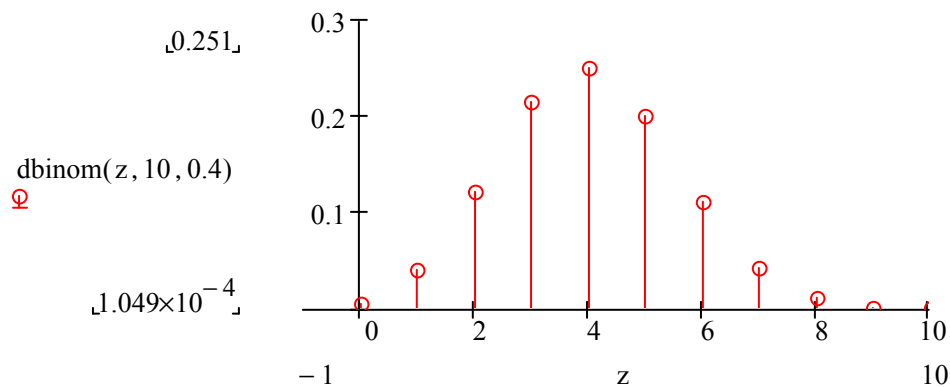
DISTRIBUCIONES DISCRETAS

Distribución Binomial:

Responde a la siguiente función densidad:

$$f_B(z, n, p) := \frac{n!}{z! \cdot (n - z)!} \cdot p^z \cdot (1 - p)^{n-z}$$

donde n es un parámetro entero positivo, p es un real positivo entre 0 y 1 y z pertenece a los enteros positivos (incluyendo el cero).



Del mismo modo que en los casos anteriores, se pueden encontrar:

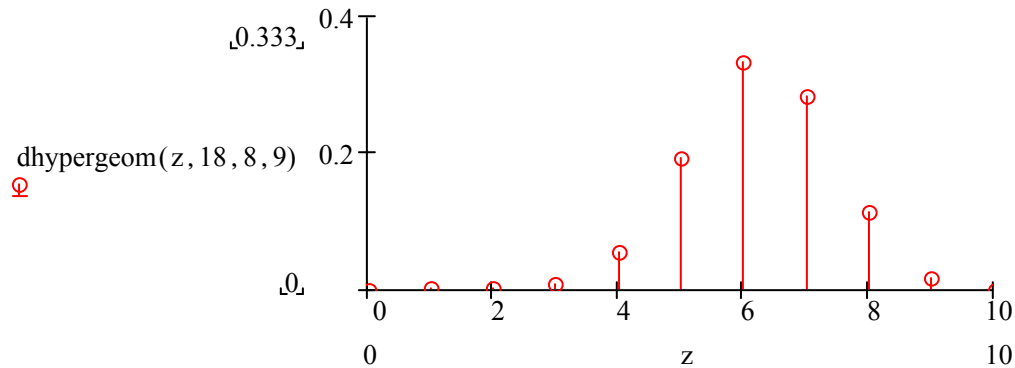
- **dbinom(z,n,p)** – valor de la densidad de probabilidad para la distribución Binomial de parámetros n y p.
- **pbinom(z,n,p)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Binomial de parámetros n y p.
- **qbinom(p1,n,p)** – valor de la abscisa que deja un área (probabilidad) p1, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Binomial de parámetros n y p.
- **rbinom(m,n,p)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Binomial de parámetros n y p.

Distribución Hipergeométrica:

Responde a la siguiente función densidad:

$$f_H(z, a, b, n) := \frac{\frac{a!}{z! \cdot (a-z)!} \cdot \frac{b!}{(n-z)! \cdot (b-n+z)!}}{\frac{(a+b)!}{n! \cdot (a+b-n)!}}$$

donde a , b y n son parámetros enteros positivos y z pertenece a los enteros positivos (incluyendo el cero).



Del mismo modo que en los casos anteriores, se pueden encontrar:

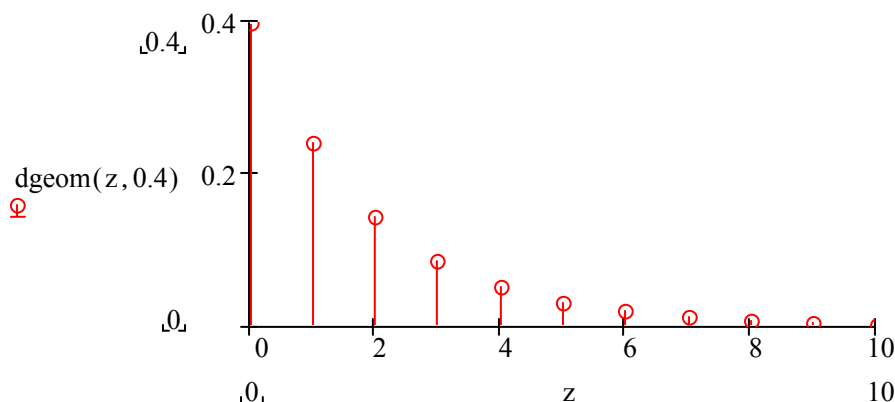
- **dhypergeom (z,a,b,n)** – valor de la densidad de probabilidad para la distribución Hipergeométrica de parámetros a , b y n .
- **phypergeom(z,a,b,n)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Hipergeométrica de parámetros a , b y n .
- **qhypergeom(p,a,b,n)** – valor de la abscisa que deja un área (probabilidad) p , bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Hipergeométrica de parámetros a , b y n .
- **rhypergeom m,a,b,n** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Hipergeométrica de parámetros a , b y n .

Distribución Geométrica:

Responde a la siguiente función densidad:

$$f_G(z, p) := p \cdot (1 - p)^z$$

donde p es un real positivo entre 0 y 1 y z pertenece a los enteros positivos (incluyendo el cero).



Del mismo modo que en los casos anteriores, se pueden encontrar:

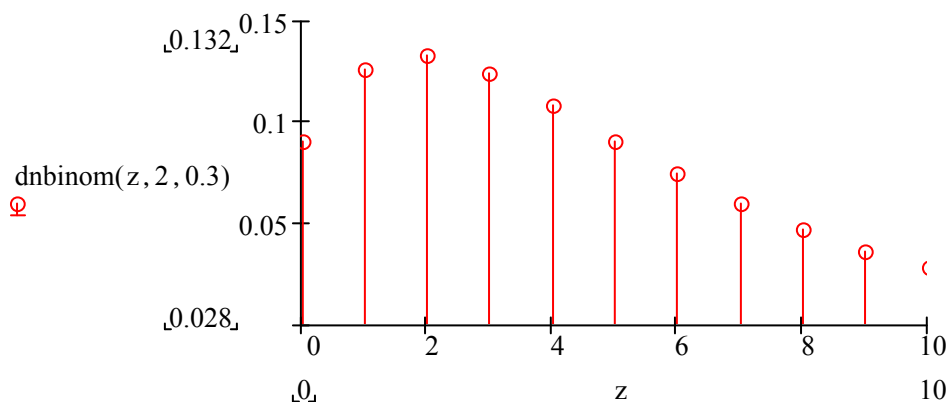
- **dgeom(z,p)** – valor de la densidad de probabilidad para la distribución Geométrica de parámetro p.
- **pgeom(z,p)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Geométrica de parámetro p.
- **qbinom(p1,n,p)** – valor de la abscisa que deja un área (probabilidad) p1, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Geométrica de parámetro p.
- **rgeom(m,p)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Geométrica de parámetro p.

Distribución Binomial Negativa (o de Pascal):

Responde a la siguiente función densidad:

$$f_{BN}(z,r,p) := \frac{(z+r-1)!}{(r-1)! \cdot z!} \cdot p^r \cdot (1-p)^z$$

donde r es un parámetro entero positivo, p es un real positivo entre 0 y 1 y z pertenece a los enteros positivos (incluyendo el cero).



Del mismo modo que en los casos anteriores, se pueden encontrar:

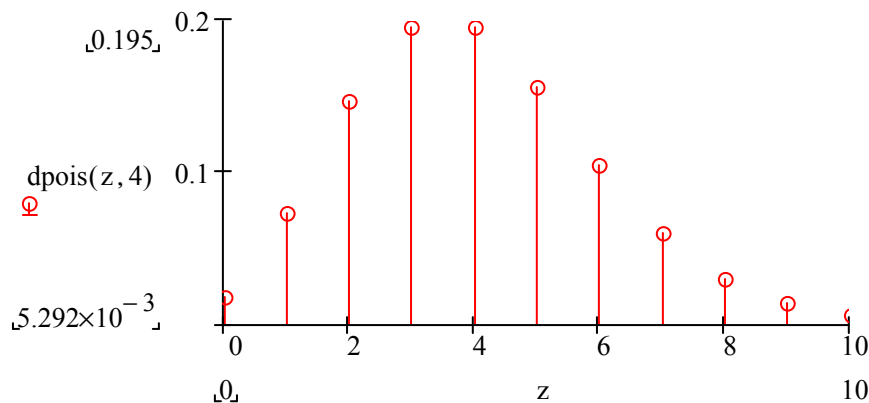
- **dnbinom(z,r,p)** – valor de la densidad de probabilidad para la distribución Binomial Negativa de parámetros r y p.
- **pnbinom(z,r,p)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Binomial Negativa de parámetros r y p.
- **qnbinom(p1,r,p)** – valor de la abscisa que deja un área (probabilidad) p1, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Binomial Negativa de parámetros r y p.
- **rnbinom(m,r,p)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Binomial Negativa de parámetros r y p.

Distribución de Poisson:

Responde a la siguiente función densidad:

$$f_p(z, \lambda) := \frac{\lambda^z}{z!} \cdot e^{-\lambda}$$

donde λ es un real positivo y z pertenece a los enteros positivos (incluyendo el cero).



Del mismo modo que en los casos anteriores, se pueden encontrar:

- **dpois(z,λ)** – valor de la densidad de probabilidad para la distribución Poisson de parámetro λ .
- **ppois(z,λ)** – valor del área (probabilidad) comprendida entre 0 y z para la distribución Poisson de parámetro λ .
- **qpois(p1,λ)** – valor de la abscisa que deja un área (probabilidad) p1, bajo la curva comprendida entre 0 y dicha abscisa, para la distribución Poisson de parámetro λ .
- **rgeom(m,λ)** – retorna un vector de m valores aleatorios cuyos elementos tienen una distribución Poisson de parámetro λ .

