

FORMULARIO DE DISTRIBUCIONES DE PROBABILIDAD

Jorge M. Galbiati

	pág.
DISTRIBUCION BINOMIAL	2
DISTRIBUCION POISSON	4
DISTRIBUCION HIPERGEOMETRICA	5
DISTRIBUCION GEOMETRICA	7
DISTRIBUCION NORMAL	8
DISTRIBUCION JI-CUADRADO	11
DISTRIBUCION T DE STUDENT	13
DISTRIBUCION F DE SNEDECOR	15
DISTRIBUCION UNIFORME	17
DISTRIBUCION EXPONENCIAL	18
DISTRIBUCION GAMA	20
DISTRIBUCION BETA	23
TRANSFORMACION DE VARIABLES ALEATORIAS CONTINUAS	25

DISTRIBUCION BINOMIAL

Función de probabilidad:

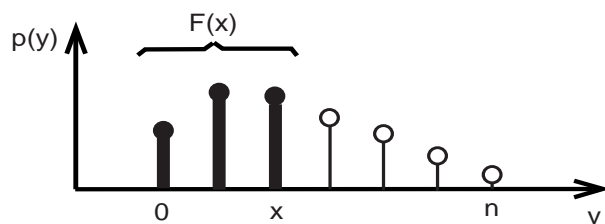
$$p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad \text{si } x = 0, 1, 2, \dots, n$$

Espacio paramétrico: $n \in \{1, 2, 3, \dots\}$ $p \in (0, 1)$

Valor esperado: np

Varianza: $np(1-p)$

Función generadora de momentos: $(1-p + p e^t)^n$



APROXIMACION NORMAL DE LA BINOMIAL

Si una variable aleatoria X tiene distribución **binomial** con parámetros n y p , entonces si n es grande y si p no es ni muy cercano a cero ni muy cercano a 1, la variable aleatoria $Z = \frac{X-np}{\sqrt{np(1-p)}}$ tiene distribución aproximada **normal estándar**. En la práctica, si n es grande y p no es ni muy pequeño ni muy grande, si se requiere la probabilidad acumulada $F(x)$ con F distribución **binomial**, se puede obtener su valor aproximado buscando en la tabla **normal**

$$F_N\left(\frac{x - 0,5 - np}{\sqrt{np(1-p)}}\right)$$

en que F_N es la distribución **normal estándar**. Se puede utilizar, como criterio, las condiciones simultáneas $n > 30$, $np > 5$ y $n(1-p) > 5$.

APROXIMACION POISSON DE LA BINOMIAL.

Si una variable aleatoria X tiene distribución **binomial** con parámetros n y p , entonces si n es grande, y p muy cercano a cero, la variable aleatoria X tiene distribución aproximada **poisson** con parámetro $\lambda = np$.

En la práctica, si n es grande y p cercano a cero, si se requiere la probabilidad acumulada $F(x)$ con F distribución **binomial**, se puede obtener su valor aproximado buscando en la tabla **poisson**

$$F_P(x) = \sum_{y=0}^x \frac{e^{-\lambda}(\lambda)^y}{y!}$$

en que F_P es la distribución **poisson** con parámetro $\lambda = np$. Se puede utilizar, como criterio, las condiciones simultáneas $n > 30$ y $np \leq 5$.

DISTRIBUCION POISSON

Función de probabilidad:

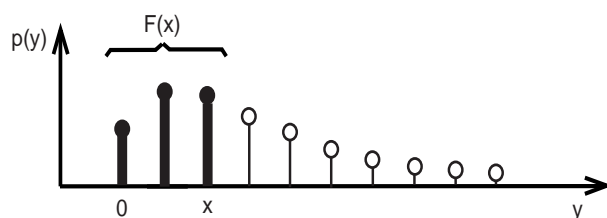
$$p(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad \text{si } x = 0, 1, 2, \dots$$

Espacio paramétrico: $\lambda \in (0, +\infty)$

Valor esperado: λ

Varianza: λ

Función generadora de momentos: $e^{\lambda(e^t - 1)}$



APROXIMACION NORMAL DE LA POISSON.

Si una variable aleatoria X tiene distribución **Poisson** con parámetro λ , entonces si λ es grande, la variable aleatoria $Z = \frac{X - \lambda}{\sqrt{\lambda}}$ tiene distribución aproximada **normal estándar**.

En la práctica, si λ es grande, si se requiere la probabilidad acumulada $F(x)$ con F distribución **Poisson**, se puede obtener su valor aproximado buscando en la tabla **normal**

$$F_N\left(\frac{x - \lambda}{\sqrt{\lambda}}\right)$$

en que F_N es la distribución **normal estándar**. Se puede utilizar, como criterio, la condición $\lambda > 36$.

DISTRIBUCION HIPERGEOMETRICA

Función de probabilidad:

$$p(x) = \frac{\frac{k!}{n!(n-k)!} \times \frac{(N-k)!}{(n-x)!(N-k-n+x)!}}{\frac{N!}{n!(N-n)!}} \quad \text{si } x = a, a + 1, a + 2, \dots, b$$

en que $a = \max(0; n + k - N)$ y $b = \min(k, n)$. x es el número de éxitos en la muestra.

Espacio paramétrico: N, k y n enteros positivos, tales que $k < N$, $n < N$ y $n < N - k$.

N es el tamaño de la población.

k es el número de éxitos en la población.

n es el tamaño de la muestra.

Valor esperado: $\frac{nk}{N}$

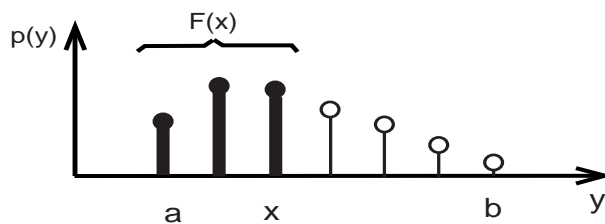
Varianza: $\frac{nk}{N} \left(1 - \frac{k}{N}\right) \left(\frac{N-n}{N-1}\right)$

Función generadora de momentos:

$$\frac{(N-n)!(N-k)!}{N!} H(-n; -k; N-k-n+1; e^t)$$

donde $H(p, q, r, z) = 1 + \frac{pq}{r} \frac{z}{1!} + \frac{p(p+1)q(q+1)}{r(r+1)} \frac{z^2}{2!} + \frac{p(p+1)(p+2)q(q+1)(q+2)}{r(r+1)(r+2)} \frac{z^3}{3!}$

(función hipergeométrica)



APROXIMACION BINOMIAL DE LA HIPERGEOMETRICA

Si una variable aleatoria X tiene distribución **hipergeométrica** con parámetros N , k y n , entonces si N es grande y si $\frac{k}{N}$ no es ni muy cercano a cero ni muy cercano a 1, X tiene distribución aproximada **binomial** con parámetros n y $p = \frac{k}{N}$.

DISTRIBUCION GEOMETRICA

Función de probabilidad:

$$p(x) = p(1 - p)^{x-1} \quad \text{si } x = 1, 2, 3, \dots, b$$

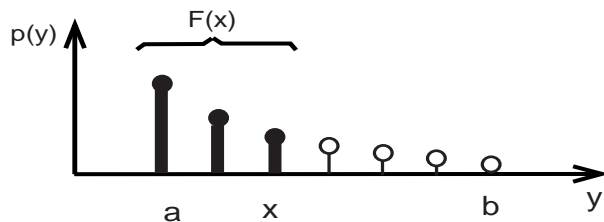
x es el número de intentos hasta lograr el primer éxito.

Espacio paramétrico: $p \in (0, 1)$, probabilidad de éxito en un intento.

Valor esperado: $\frac{1}{p}$

Varianza: $\frac{1-p}{p^2}$

Función generadora de momentos: $p \frac{e^t}{1 - (1-p)e^t}$ si $t < -\log(1 - p)$



DISTRIBUCION NORMAL

Función de probabilidad:

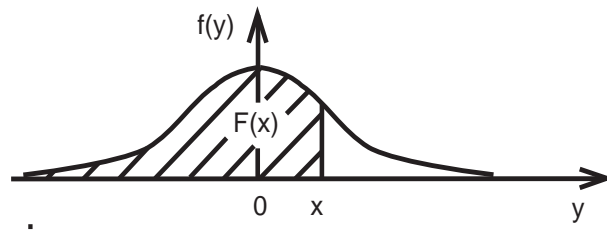
$$p(x) = \frac{1}{\sqrt{(2\pi)} \cdot \sigma} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right] \quad \text{para } x \in (-\infty, +\infty)$$

Espacio paramétrico: *media* $\mu \in (-\infty, +\infty)$ *varianza* $\sigma^2 \in (0, +\infty)$

Valor esperado: μ

Varianza: σ^2

Función generadora de momentos: $e^{(\mu t + \sigma^2 t^2 / 2)}$



DISTRIBUCION NORMAL ESTANDAR

Es un caso especial de la normal, en que $\mu = 0$ y $\sigma^2 = 1$.

Función de densidad:

$$f(x) = \frac{1}{\sqrt{(2\pi)}} \exp\left[-\frac{x^2}{2}\right] \quad \text{para } x \in (-\infty, +\infty)$$

Valor esperado: 0

Varianza: 1

Función generadora de momentos: $e^{t^2/2}$

RELACION CON LA NORMAL ESTANDAR

Los valores de la función de distribución de la normal con parámetros μ y σ^2 se obtienen de la tabla de distribución **normal estándar** (en que $\mu = 0$ y $\sigma^2 = 1$) como se muestra a continuación. Por esa razón sólo se entrega la tabla de la **normal estándar**.

Si se requiere la probabilidad acumulada hasta la cuantila x , se efectúa la transformación $z = \frac{x-\mu}{\sigma}$ y se busca la probabilidad asociada a la cuantila z en la tabla de distribución **normal estándar**.

Al revés, si se quiere saber a qué cuantila corresponde una probabilidad acumulada dada, $F(z)$, se busca la cuantila z asociada a $F(z)$ en la tabla de distribución **normal estándar**. Entonces la correspondiente cuantila de la normal con parámetros μ y σ^2 es $x = \sigma z + \mu$.

FUNCIONES LINEALES DE NORMALES

1.- Si X es una variable aleatoria **normal** con valor esperado μ y varianza σ^2 , si a y b son constantes, entonces la variable aleatoria $a + bX$ tiene distribución **normal**, con valor esperado $a + b\mu$ y varianza $b^2\sigma^2$.

Como caso particular, la variable aleatoria **estandarizada** $Z = \frac{X-\mu}{\sigma}$ tiene distribución **normal estándar**.

2.- Si X_1 y X_2 son variables aleatorias **normales** (pág. 50), estadísticamente **independientes**, con valores esperados respectivos μ_1 y μ_2 , con varianzas respectivas σ_1^2 y σ_2^2 , y si a y b son dos números reales, entonces la variable aleatoria $aX_1 + bX_2$ tiene distribución **normal** con valor esperado $a\mu_1 + b\mu_2$ y varianza $a^2\sigma_1^2 + b^2\sigma_2^2$.

3.- Si X_1, X_2, \dots, X_n son n variables aleatorias **normales** con valor esperado μ , y varianza σ^2 entonces el promedio $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ tiene distribución **normal** con valor esperado μ y varianza σ^2/n .

TEOREMA DEL LIMITE CENTRAL

Si X_1, X_2, \dots, X_n son n variables aleatorias estadísticamente **independientes**, con valor esperado μ y varianza σ^2 y cualquier distribución probabilística, continua o discreta, entonces si n es grande, la variable aleatoria $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ tiene distribución aproximada **normal estándar**

DISTRIBUCION JI CUADRADO

Función de densidad:

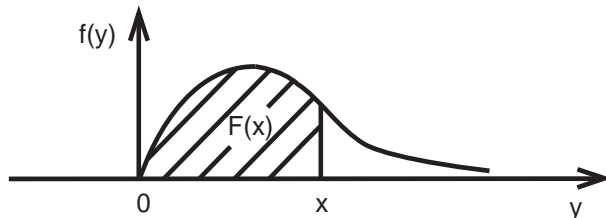
$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2} \quad \text{si } x > 0$$

Espacio paramétrico: *Grados de libertad* $k \in \{1, 2, 3, \dots\}$

Valor esperado: k

Varianza: $2k$

Función generadora de momentos: $\left(\frac{1}{1-2t}\right)^{k/2}$ para $t < 1/2$



APROXIMACION NORMAL DE LA JI-CUADRADO.

Si una variable aleatoria X tiene distribución **ji-cuadrado** con k grados de libertad, entonces si k es grande la variable aleatoria $Z = \frac{X-k}{\sqrt{(2k)}}$ tiene distribución aproximada **normal standard**.

En la práctica, si k es grande, si se requiere la probabilidad acumulada $F(x)$ con F distribución **ji-cuadrado**, se puede obtener su valor aproximado buscando en la tabla **normal**

$$F_N\left(\frac{x-k}{\sqrt{(2k)}}\right)$$

en que F_N es la distribución **normal estándar**. Se puede utilizar, como criterio, la condición $k > 200$.

CONSTRUCCION DE UNA JI-CUADRADO A PARTIR DE NORMALES

1.- Si Z_1, Z_2, \dots, Z_n son n variables aleatorias **normales estándar** estadísticamente **independientes**, entonces la variable aleatoria $\sum_{i=1}^n Z_i^2$ tiene distribución **ji-cuadrado** con n grados de libertad.

2.- Si X_1, X_2, \dots, X_n son n variables aleatorias normales con valor esperado μ y varianza σ^2 , **independientes**, entonces la variable aleatoria $\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}$ tiene distribución **ji-cuadrado** con $n - 1$ grados de libertad.

Además esta expresión es estadísticamente **independiente** del promedio \bar{X} .

DISTRIBUCION T DE STUDENT

Función de densidad:

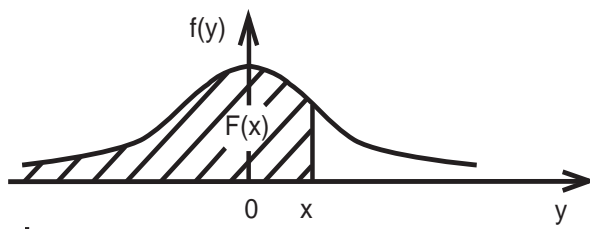
$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma(k/2)} \cdot \frac{1}{\sqrt{(k\pi)}} \cdot \frac{1}{\left(1 + \frac{x^2}{k}\right)^{\frac{k+1}{2}}} \quad \text{para } x \in (-\infty, +\infty)$$

Espacio paramétrico: *Grados de libertad* $k \in \{1, 2, 3, \dots\}$

Valor esperado: 0 para $k > 1$

Varianza: $\frac{k}{k-2}$ para $k > 2$

Función generadora de momentos: *no existe*



VALORES DE PROBABILIDAD MENORES QUE 0.5

Por la simetría de la distribución **t de student**, rige la igualdad $F(-x) = 1 - F(x)$. Por esa razón, la tabla sólo tiene probabilidades mayores que 0.5, asociadas a cuantiles positivos.

Si se requiere el cuantil asociado a una probabilidad acumulada P menor que 0.5, se ingresa a la tabla el valor de probabilidad acumulada $1 - P$; al correspondiente cuantil x obtenido de la tabla se le pone signo menos, quedando $-x$ como el cuantil requerido.

APROXIMACION NORMAL DE LA T DE STUDENT

Si una variable aleatoria X tiene distribución **t de student** con k grados de libertad, entonces si k es grande la variable aleatoria X tiene distribución aproximada **normal standard**.

En consecuencia, si k es grande, si se requiere la probabilidad acumulada $F(x)$ con F distribución **t de student**, se puede obtener su valor aproximado buscando en la tabla **normal** el valor $F_N(x)$, en que F_N es la distribución **normal standard**. Se puede utilizar, como criterio, la condición $k > 200$.

CONSTRUCCION DE UNA T DE STUDENT A PARTIR DE UNA NORMAL Y UNA JI-CUADRADO

1.- Si Z es una variable aleatoria **normal estándar** y V es una variable aleatoria **ji-cuadrado** con n grados de libertad, ambas estadísticamente **independientes**, entonces la variable aleatoria $\frac{Z}{\sqrt{V/n}}$ tiene distribución **t de student** con n grados de libertad.

2.- Si X_1, X_2, \dots, X_n son n variables aleatorias **normales** con valor esperado μ y varianza σ^2 , estadísticamente **independientes**, entonces la variable aleatoria $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ tiene distribución **t de student** con $n - 1$ grados de libertad, en que \bar{X} es el promedio y $s^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}$ es la varianza muestral.

DISTRIBUCION F DE SNEDECOR

Función de densidad:

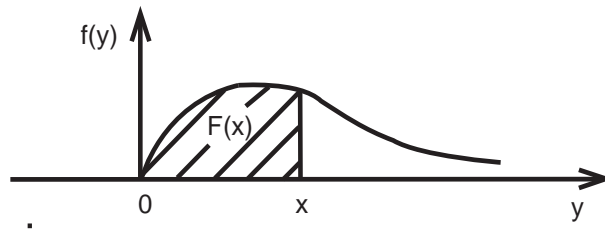
$$f(x) = \frac{\Gamma\left(\frac{n+d}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \cdot \Gamma\left(\frac{d}{2}\right)} \cdot \left(\frac{n}{d}\right)^{n/2} \cdot \frac{x^{n/2-1}}{\left(1 + \frac{n}{d}x\right)^{\frac{n+d}{2}}} \quad \text{si } x > 0$$

Espacio paramétrico: grados de libertad del numerador n y grados de libertad del denominador d ambos enteros positivos.

Valor esperado: $\frac{d}{d-2}$ para $d > 2$

Varianza: $\frac{2d^2(n+d-2)}{n(d-2)^2(d-4)}$ para $d > 4$

Función generadora de momentos: *no existe*



INVERSION DE LA F DE SNEDECOR

Se puede usar la siguiente relación para calcular valores que no aparecen en la tabla:

Si la variable aleatoria X tiene distribución **F** con n grados de libertad del numerador y d grados de libertad del denominador, entonces $1/X$ tiene distribución **F**, con d grados de libertad del numerador y n grados de libertad del denominador.

Por lo tanto se pueden obtener más valores de los que aparecen en la tabla, mediante en la relación $F_{n,d}(x) = 1 - F_{d,n}\left(\frac{1}{x}\right)$ en que F es el valor de probabilidad acumulada de la tabla, el primer subíndice corresponde a los grados de libertad del numerador, el segundo a los grados de libertad del denominador.

CONSTRUCCION DE UNA F DE SNEDECOR A PARTIR DE DOS
JI-CUADRADO

1.- Si X es una variable aleatoria **ji-cuadrado** con n grados de libertad e Y es una variable aleatoria **ji-cuadrado** con d grados de libertad, estadísticamente **independientes**, entonces el cociente $\frac{X/n}{Y/d}$ tiene distribución **F de Snedecor** con n grados de libertad en el numerador y d grados de libertad en el denominador.

2.- También $\frac{Y/d}{X/n}$ tiene distribución **F de Snedecor** con d grados de libertad en el numerador y n grados de libertad en el denominador.

DISTRIBUCION UNIFORME

Función de densidad:

$$f(x) = \frac{1}{b-a} \quad \text{si } a < x \leq b$$

Espacio paramétrico:

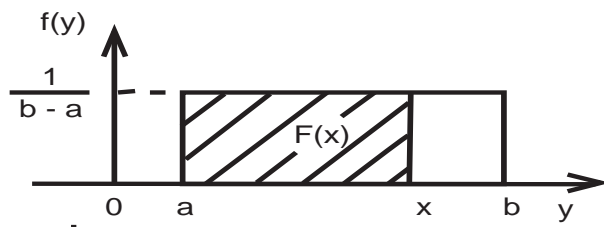
$$-\infty < a, b < \infty \quad a < b$$

Valor esperado: $\frac{a+b}{2}$

Varianza: $\frac{(b-a)^2}{12}$

Función generadora de momentos:

$$\frac{e^{bt} - e^{at}}{(b-a)t}$$



VALORES DE LA DISTRIBUCION UNIFORME

La función de distribución de la **uniforme** se puede calcular analíticamente mediante la fórmula

$$F(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a < x \leq b \\ 1 & \text{si } x > b \end{cases}$$

DISTRIBUCION EXPONENCIAL

Función de densidad:

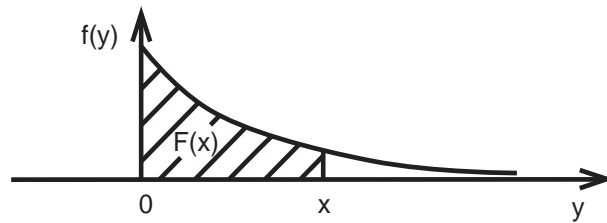
$$f(x) = \lambda \cdot e^{-\lambda x} \quad \text{si } x > 0$$

Espacio paramétrico: *Tasa media de ocurrencia* $\lambda > 0$

Valor esperado: $\frac{1}{\lambda}$

Varianza: $\frac{1}{\lambda^2}$

Función generadora de momentos: $\frac{\lambda}{\lambda-t}$ *para* $t < \lambda$



VALORES DE LA DISTRIBUCION EXPONENCIAL

La función de distribución de la **exponencial** se puede calcular analíticamente mediante la fórmula $F(x) = 1 - e^{-\lambda x}$ para $x > 0$.

RELACION ENTRE UNA POISSON Y UNA EXPONENCIAL

1.- Si X es una variable aleatoria **Poisson** con parámetro λ , que describe el número de ocurrencias de un fenómeno por unidad de tiempo, entonces la variable aleatoria que describe el tiempo entre ocurrencias tiene distribución **exponencial** con parámetro λ . En tal caso el parámetro λ es la "tasa media de ocurrencias" por unidad de tiempo, y $\theta = 1/\lambda$ es el "tiempo medio entre ocurrencias".

2.- En forma recíproca, si Y es una variable aleatoria **exponencial** con parámetro λ , que describe el tiempo entre ocurrencias de un fenómeno, entonces el número de veces que ocurre el fenómeno en una unidad de tiempo, es una variable aleatoria con distribución **Poisson**, con el mismo parámetro, que representa la "tasa media de ocurrencias" por unidad de tiempo.

DISTRIBUCION GAMA

Función de densidad: Hay dos formas usuales de parametrizar esta distribución.

Primera parametrización (Par. 1):

$$f(x) = \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} \quad \text{si } x > 0$$

Segunda parametrización (Par. 2):

$$f(x) = \frac{1}{\theta^p \Gamma(p)} x^{p-1} e^{-\frac{x}{\theta}} \quad \text{si } x > 0$$

Espacio paramétrico:

Par. 1: *Parametro de escala* $\lambda > 0$
Parametro de forma $p > 0$

Par. 2: *Parametro de escala* $\theta > 0$
Parametro de forma $p > 0$

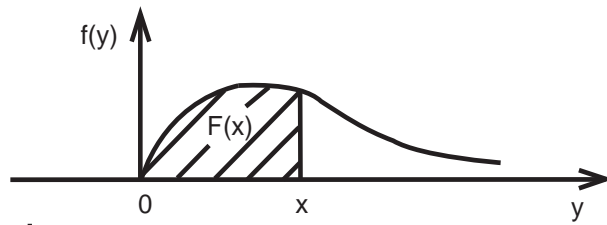
Valor esperado: Par. 1: $\frac{p}{\lambda}$ Par. 2: $p \theta$

Varianza: Par. 1: $\frac{p}{\lambda^2}$ Par. 2: $p \theta^2$

Función generadora de momentos:

Par. 1: $\left(\frac{\lambda}{\lambda-t}\right)^p$ para $t < \lambda$

Par. 2: $\frac{1}{(1-\theta t)^p}$ para $t < \frac{1}{\theta}$



Casos particulares:

1) Si $p=1$ (parámetro de forma) entonces la **gama** se convierte en una **exponencial** cuyo parámetro es igual al parámetro de escala de la **gama**, λ (Par. 1) o equivalentemente θ (Par. 2).

2) Si $p=\frac{1}{2} k$, en que k es cualquier número entero positivo, y si $\lambda=\frac{1}{2}$ (Par. 1) o equivalentemente $\theta=2$ (Par. 2), entonces la **gama** se convierte en una **ji-cuadrado** cuyo parámetro *grados de libertad* es igual a k .

La función de distribución **gama** no se puede calcular analíticamente, salvo en casos especiales.

RELACIONES ENTRE GAMAS

Lo siguiente se expresa en términos de la primera parametrización, con el parámetro λ . Es equivalente para la segunda parametrización, con el parámetro θ . Sólo se debe sustituir λ por θ .

1.- Si X_1, X_2, \dots, X_n son n variables aleatorias **gama** estadísticamente **independientes**, con parámetros de forma respectivos p_1, p_2, \dots, p_n , y con parámetro de escala común λ , entonces la variable aleatoria $Y = \sum_{i=1}^n X_i$ tiene distribución **gama** con parámetro de forma $p = \sum_{i=1}^n p_i$ y parámetro de escala λ .

2.- Si X_1 y X_2 son variables aleatorias **gama** estadísticamente **independientes**, con parámetros de forma respectivos p_1 y p_2 y parámetro de escala común λ entonces las variables aleatorias $U = X_1 + X_2$ y $V = \frac{X_1}{X_1 + X_2}$ son **independientes**, U tiene distribución **gama** con parámetro de forma $p_1 + p_2$ y de escala λ , y V tiene distribución **beta** (pág. 104) con parámetros $r = p_1$ y $s = p_2$.

3.- Caso especial de 1. Si X_1, X_2, \dots, X_n son n variables aleatorias **exponenciales** estadísticamente **independientes**, con parámetro común λ , entonces la variable aleatoria $Y = \sum_{i=1}^n X_i$ tiene distribución **gama** con parámetro de forma $p = n$ y parámetro de escala λ .

A esta forma especial de **gama**, con parámetro de forma entero, se le suele dar el nombre de distribución **erlang**.

4.- Caso especial de 1. Si X_1, X_2, \dots, X_n son n variables aleatorias **ji-cuadrado** estadísticamente **independientes**, con parámetros respectivos (grados de libertad) k_1, k_2, \dots, k_n , entonces la variable aleatoria $Y = \sum_{i=1}^n X_i$ tiene distribución **ji-cuadrado** con $k = \sum_{i=1}^n k_i$ grados de libertad.

DISTRIBUCION BETA

Función de densidad:

$$f(x) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} x^{r-1} (1-x)^{s-1} \quad \text{si } 0 < x < 1$$

Espacio paramétrico:

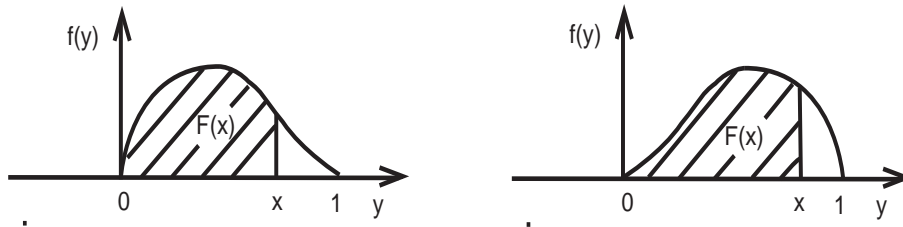
$$r > 0, s > 0$$

Valor esperado: $\frac{r}{r+s}$

Varianza: $\frac{rs}{(r+s)^2 (r+s+1)}$

Momentos: La función generadora de momentos no tiene una forma analítica. Sin embargo, el momento m -ésimo puede obtenerse directamente, mediante la fórmula

$$\mu_m = \frac{(r+s+1)! (r+m)!}{(r+s+m+1)! r!} \quad \text{para } m=1, 2, \dots$$



En la figura de la izquierda, $r < s$, mientras que en la figura de la derecha, $r > s$. Si r y s son iguales, la densidad es simétrica.

La función de distribución **beta** no se puede calcular analíticamente, salvo en casos especiales.

Caso particular:

Si $r=1$ y $s=1$ entonces la **beta** se convierte en una **uniforme** con parámetros $a=0$ y $b=1$.

TRANSFORMACION DE VARIABLES ALEATORIAS CONTINUAS

1.- Si X es una variable aleatoria continua con función de densidad $f_X(x)$ y $g(\cdot)$ es una función creciente, entonces la nueva variable aleatoria $Y = g(X)$ tiene función de densidad dada por la fórmula

$$f_Y(y) = \frac{1}{|g'[g^{-1}(y)]|} f_X[g^{-1}(y)]$$

en que $||$ denota el valor absoluto, g' es la derivada y g^{-1} es la inversa de la función g .

2.- Caso especial de 1. Si la función $g(x)$ del párrafo 1 es una función lineal $g(x) = a + bx$, en que a y b son constantes, $b \neq 0$, entonces la variable aleatoria $Y = g(X)$ tiene densidad

$$f_Y(y) = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right)$$