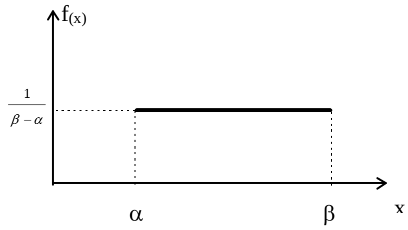
DISTRIBUCIONES CONTINUAS

1. MODELO UNIFORME

Se aplica a una variable que puede asumir con igual probabilidad cualquier valor real dentro de un intervalo acotado [α , β], con $\beta > \alpha$. Luego, la masa total de probabilidad se distribuye en forma uniforme o constante en todo el intervalo. También se la conoce con el nombre de distribución rectangular por la representación gráfica de su función de densidad.
La función de densidad de la variable uniforme tiene la siguiente expresión y representación gráfica:

$$f(x) = \begin{cases} \dfrac{1}{\beta-\alpha} & si\ \alpha \le x \le \beta \\ 0 & \forall x \end{cases}$$



Por lo tanto, la probabilidad de que la variable tome valores dentro de un intervalo incluido en el recorrido, o sea, (a , b) $\subset$ [α , β], es proporcional a la longitud de dicho intervalo. Si $\alpha < a < b < \beta$, con [α , β] = R_x entonces P(a < X < b)= $\dfrac{b-a}{\beta-\alpha}$ con lo cual se deduce que corresponden idénticas probabilidades a intervalos de igual amplitud, siempre que dichos intervalos estén incluidos en R_x =[α , β].
Se verifican las condiciones de no negatividad y de cierre, requisitos para ser considerada función de densidad::

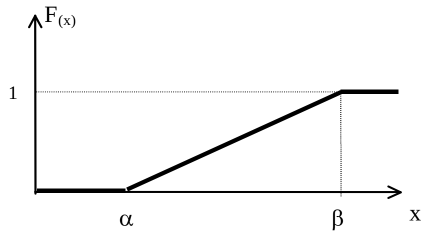
$$1.)\ f(x) = \frac{1}{\beta-\alpha} \ge 0\ \ \forall x \in R_x\ \ pues\ \beta > \alpha$$

$$2)\ \int_{-\infty}^{\infty} f(x)dx = \int_{\alpha}^{\beta} \frac{1}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}(\beta-\alpha) = 1$$

La función de probabilidad acumulada $F(x) = P(X \le x) = \int_{\alpha}^{x} \dfrac{1}{\beta-\alpha}ds = \dfrac{1}{\beta-\alpha} s\Big|_{\alpha}^{x} = \dfrac{x-\alpha}{\beta-\alpha}$ para $\alpha < x < \beta$

Luego:

$$F(x) = \begin{cases} 0 & si\ x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha} & si\ \alpha \le x \le \beta \\ 1 & si > \beta \end{cases}$$



Sus parámetros son α y β , y representan los límites del recorrido de la variable X.
La notación usada es X ~ U(α , β)

$$E(X) = \int_{\alpha}^{\beta} x \frac{1}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}\frac{x^2}{2}\Big|_{\alpha}^{\beta} = \frac{\beta^2-\alpha^2}{(\beta-\alpha).2} = \frac{(\beta-\alpha)(\beta+\alpha)}{(\beta-\alpha).2} = \frac{\beta+\alpha}{2}$$

$$V(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \int_{\alpha}^{\beta} x^2 \frac{1}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}\frac{x^3}{3}\Big|_{\alpha}^{\beta} = \frac{\beta^3-\alpha^3}{(\beta-\alpha).3} = \frac{(\beta-\alpha)(\beta^2+\beta\alpha+\alpha^2)}{(\beta-\alpha).3} = \frac{\beta^2+\beta\alpha+\alpha^2}{3}$$

$$\therefore V(X) == \frac{\beta^2+\beta\alpha+\alpha^2}{3} - \frac{(\beta+\alpha)^2}{4} = \frac{4\beta^2+4\beta\alpha+4\alpha^2-3\beta^2-6\beta\alpha-3\alpha^2}{12} = \frac{(\beta-\alpha)^2}{12}$$

Luego :

$$E(X) = \frac{\beta+\alpha}{2} \qquad V(X) = \frac{(\beta-\alpha)^2}{12} \qquad \Rightarrow \qquad \sigma_{(X)} = \frac{\beta-\alpha}{\sqrt{12}}$$

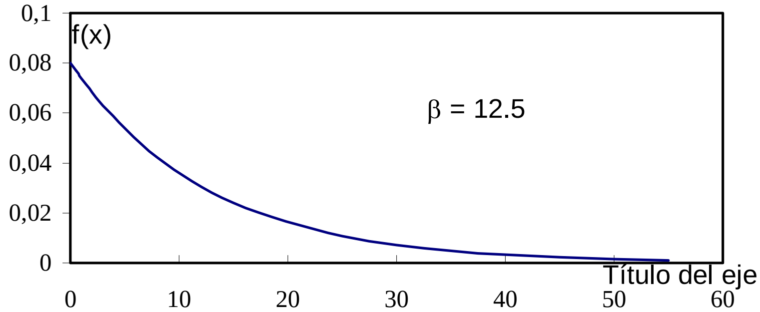
2. MODELO EXPONENCIAL

Se aplica a una variable aleatoria continua que puede asumir cualquier valor positivo, pero la densidad de probabilidad va disminuyendo conforme la variable asume valores mayores.

Su función de densidad y representación gráfica correspondiente es la siguiente:

$$f(x)=\begin{cases}\dfrac{1}{\beta}e^{-\frac{1}{\beta}x} & si\, x\geq 0\\ 0 & si\ x<0\end{cases}$$

:



Veamos que verifica con las condiciones de no negatividad y de cierre, requisitos para ser considerada función de densidad:

$$1.)\ f(x)=\frac{1}{\beta}e^{-\frac{1}{\beta}x}\geq 0 \quad \forall x\in R,\, pues\ \ \beta>0$$

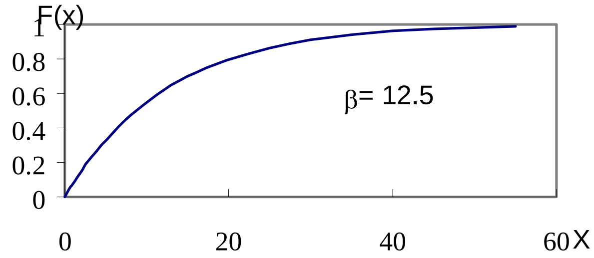
$$2)\ \int_{-\infty}^{\infty} f(x)dx=\int_{0}^{\infty}\frac{1}{\beta}e^{-\frac{1}{\beta}x}dx=-e^{-\frac{1}{\beta}x}\Big|_0^{\infty}=-(0-1)=1$$

La función de probabilidad acumulada F(x) = P(X ≤ x), para los valores positivos de x , está dada por:

$$F(x)=\int_0^x\frac{1}{\beta}e^{-\frac{1}{\beta}s}ds=-e^{-\frac{1}{\beta}x}\Big|_0^x=-\left(e^{-\frac{1}{\beta}x}-1\right)=1-e^{-\frac{1}{\beta}x}$$

Luego, la función de Distribución y su gráfico correspondiente es :

$$F(x)=\begin{cases}0 & si\ x<0\\ 1-e^{-\frac{1}{\beta}x} & si\ x\geq 0\end{cases}$$



Su único parámetro es β, y su notación es X ~ exp(β)

La esperanza y la variancia son :

$$E(X)=\int_{-\infty}^{\infty}x.f(x).dx=\int_0^{\infty}x.\frac{1}{\beta}e^{-\frac{1}{\beta}x}dx=-x.e^{-\frac{1}{\beta}x}\Big|_0^{\infty}+\int_0^{\infty}.e^{-\frac{1}{\beta}x}dx\ =0+\left(-\beta.e^{-\frac{1}{\beta}x}\right)\Big|_0^{\infty}=-\beta(0-1)=\beta$$

$$V(X)=E(X^2)-\left[E(X)\right]^2\quad donde\quad E(X^2)=\int_{-\infty}^{\infty}x^2f(x).dx=\int_0^{\infty}x^2.\frac{1}{\beta}e^{-\frac{1}{\beta}x}dx=\quad ....\quad =2\beta^2$$

$$\therefore\ \ V(X)=E(X^2)-\left[E(X)\right]^2=2\beta^2-\beta^2=\beta^2$$

Luego $\quad E(X)=\beta \qquad V(X)=\beta^2 \qquad$ y $\qquad \sigma(x)=\beta$

3. MODELO NORMAL

Es el más utilizado en las experiencias científicas, razón por la que se le da mayor importancia. La ecuación de esta curva fue originalmente publicada en 1733 por Moivre quien no supo aplicar sus resultados a las observaciones (como límite de la binomial). Al mismo resultado llegan dos astrónomos matemáticos Laplace y Gauss, independientemente uno del otro. Llegaron a ella estudiando la distribución de los errores de las observaciones, en particular los errores producidos en la medición de las órbitas de los astros. Existen importantes razones para afirmar que es la distribución más significativa del análisis estadístico.

- Una importante cantidad de poblaciones correspondientes a variables aleatorias continuas siguen la ley normal. Ej: estatura de individuos adultos, resistencia a la tracción de un gran número de muestras de acero, diámetro de arandelas, aumento de peso de animales sometidos a determinadas dietas.

- Esta distribución sirve como buena aproximación de muchas variables aleatorias discretas (binomial, poisson)

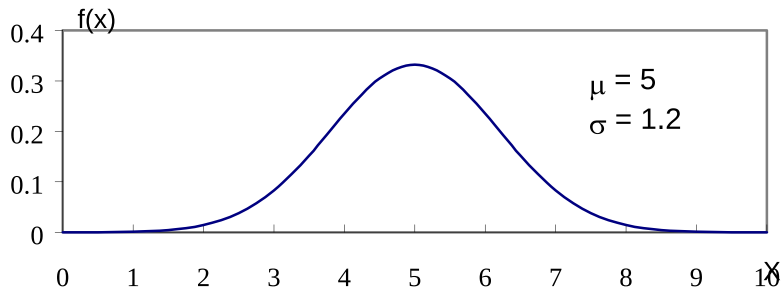
En teoría estadística, la hipótesis de normalidad de las variables aleatorias, permite solucionar problemas de tipo teórico, justificando o rechazando la aplicación del método.

<u>Definición</u>

Una variable aleatoria continua X tiene una distribución Normal si su función de densidad tiene la siguiente expresión:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} \qquad para \quad -\infty < x < \infty$$

siendo μ una constante cualquiera, $\mu \in \boldsymbol{R}$ y σ una constante positiva, $\sigma > 0$ .Es una curva simétrica en forma de campana, asintótica al eje de las abcisas, que alcanza su valor máximo en $x = \mu$, siendo $f(\mu)=\left(\sqrt{2\pi}\sigma\right)^{-1}$ y tiene dos puntos de inflexión: en $x = \mu - \sigma$ y en $x = \mu + \sigma$, y prácticamente $f(x)$ es nula para $x < \mu - 4.\sigma$ y para $x > \mu + 4.\sigma$.



Como $\sigma > 0$, entonces $\sqrt{2\pi}\sigma > 0$, luego $f(x) > 0$, $\forall\, x \in \boldsymbol{R}$.

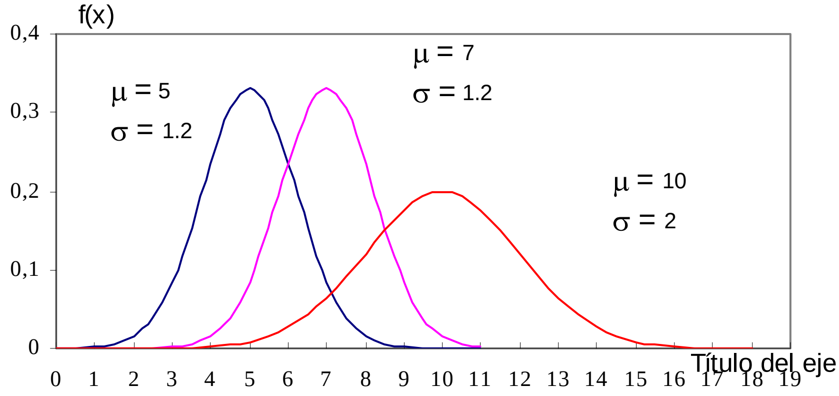
$$Como \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx = \sqrt{2\pi}\sigma \text{ ,entonces } \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx = 1 \; ; \int_{-\infty}^{\infty} f(x).dx = 1$$

Luego $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}$ es función de densidad.

Los parámetros son μ y σ, que representan la media y el desvío standard de la variable X.
La notación utilizada es $X \sim N(\mu , \sigma)$
La esperanza y la variancia son:

$$E(X)= \mu \qquad\qquad V(X)= \sigma^2$$

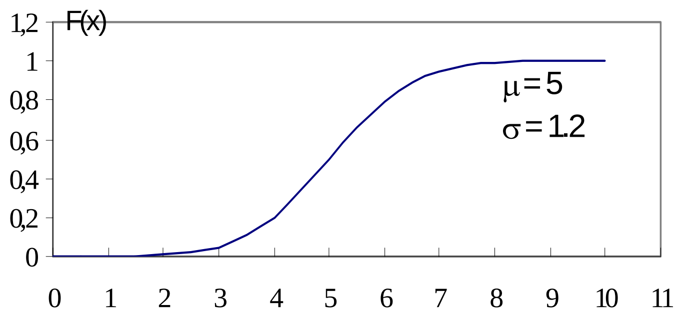
Representación gráfica de distribuciones normales para distintos valores de los parámetros



<u>Función de Distribución (o de probabilidad acumulada) :</u>

$$F(x) = \int_{-\infty}^{x} f(s)ds = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left[\frac{s-\mu}{\sigma}\right]^2} ds$$ para la cual no existe una expresión matemática.

Su representación gráfica es la siguiente:



No contar con una expresión matemática de F(x), hace impracticable el cálculo inmediato de las probabilidades. Por esta razón, se han tabulado los resultados de la función de distribución de una variable normal con media nula (E(X) = 0) y variancia igual a la unidad (V(X) =1). Esta variable recibe el nombre de variable normal standart; y cualquier variable normal puede ser transformada en una normal standart mediante operaciones matemáticas. Ese proceso de transformación se denomina estandarización.

<u>Variable Normal Standart</u>

Sea X ~ N (μ , σ) por lo tanto E(x)= μ y V(x)= σ^2

La variable normal standard la simbolizamos con Z y se define:

$$Z = \frac{X-\mu}{\sigma} \quad y \quad tiene \quad E(Z) = 0 \qquad y \quad V(Z) = 1$$

La función de densidad de esta variable Z está dada por :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad para \; -\infty < z < \infty$$

y su función de distribución o de probabilidad acumulada se simboliza universalmente por $\Phi(z) = P(Z \leq z)$.

Existen distintos tipos de tablas que nos brindan estas probabilidades. La más completa contiene todos los valores de $\Phi(z)$ para $-4 \leq z \leq 4$ (para valores de z con dos decimales), ya que se considera que $\Phi(z) = 0$ para $z < -4$ y $\Phi(z) = 1$ para $z > 4$.

Otra de las tablas contiene los valores de $\Phi(z)$ sólo para los valores positivos de z $(0 \leq z \leq 4)$, ya que por causa de la simetría de la función f(z) no es necesario disponer también para valores negativos.

Y un tercer tipo de tabla contiene, para los valores positivos de z, $(0 \leq z \leq 4)$, las probablidades:

$P(0 \leq Z \leq z) = \Phi(z) - 0.5$.

4. DISTRIBUCIÓN CHI CUADRADO (χ^2)

Surge de la suma de los cuadrados de variables independientes normales estandarizadas.

Sean Z_1 , Z_2 , Z_3 , ... , Z_n iid (independiente e idénticamente distribuidas) N(0,1) y sea la variable

$$X = \sum_{i=1}^{n} Z_i^2 \qquad \therefore \quad X = Z_1^2 + Z_2^2 + \ldots\ldots + Z_n^2 \qquad y \qquad X \sim \chi_n^2$$

La función de densidad de esta variable aleatoria X está dada por la siguiente expresión :

$$f(x) = \frac{x^{\frac{n}{2}-1}\; e^{-x/2}}{2^{n/2}\;\Gamma(n/2)} \qquad para \quad x \geq 0$$

El único parámetro de esta distribución es n que representa el número de grados de libertad , y equivale a la cantidad de variables independientes que intervienen en la suma .

Propiedades

1.- Si $X \sim N(0,1)$ entonces $X^2 \sim \chi_1^2$

2.- Si X_1 y X_2 son dos variables independientes con:

$$X_1 \sim \chi_n^2 \quad y \quad X_2 \sim \chi_m^2 \quad entonces \;\; X_1 + X_2 \sim \chi_{n+m}^2 \qquad \text{(propiedad aditiva)}$$

3.- De las dos primeras propiedades se deduce que si $X \sim N(0,1)$ y X_1 , X_2 , ... , X_n son ***n*** observaciones de la variable que constituyen una muestra al azar, entonces $\sum_{i=1}^{n} X_i^2 \sim \chi_n^2$.

4.- De esta ultima propiedad (3), se deduce que si $X \sim N(\mu, \sigma)$ y X_1 , X_2 , ... , X_n son ***n*** observaciones de la variable que constituyen una muestra al azar, entonces:

$$\sum_{i=1}^{n}\left[\frac{X_i - \mu}{\sigma}\right]^2 \sim \chi_n^2 \qquad ya \;\; que \qquad \frac{X_i - \mu}{\sigma} \sim N(0,1) \quad \forall \; i = 1..n$$

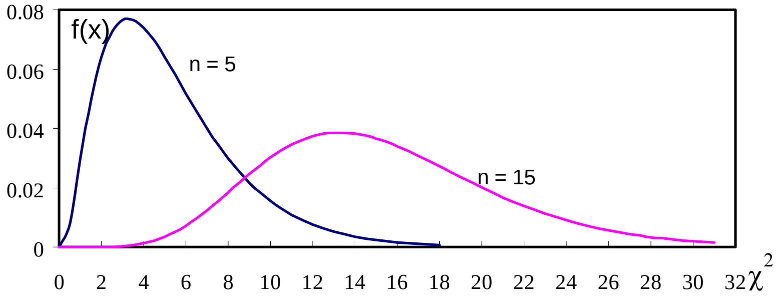
5.- Si $X \sim \chi_n^2$, entonces $E(X) = n$ y $V(X) = 2.n$ $\therefore$ la distribución queda completamente definida por sus grados de libertad, que constituyen su único parámetro.

6.- Esta variable, por ser una suma de cuadrados, siempre toma valores positivos, $\therefore$ varía de *cero* a *infinito*.

7.- La distribución de la variable es positivamente asimétrica. Sin embargo, cuando ***n*** aumenta, se aproxima a la distribución Normal. Es decir:

$$\chi_n^2 \xrightarrow{D} N(n, \sqrt{2n}) \quad cuando \quad n \to \infty$$

$$\therefore \; si \quad X \sim \chi_n^2 , \quad entonces \quad \frac{X - n}{\sqrt{2n}} \xrightarrow{D} N(0,1) \quad cuando \quad n \to \infty$$



5.- DISTRIBUCIÓN T - STUDENT (*T*)

Surge del cociente de una variable normal standart y la raíz cuadrada de una variable chi cuadrado dividido sus grados de libertad , siendo ambas variables independientes.

Sean Z e Y dos variables aleatorias *independientes* con Z ~ N(0,1) e Y ~ χ_n^2 y sea la variable

$$X = \frac{Z}{\sqrt{Y/n}} \quad \therefore \; X \sim t_n$$

La función de densidad de esta variable aleatoria X está dada por la siguiente expresión:

$$f(x) = \frac{1}{\sqrt{n\pi}} \cdot \frac{\Gamma\left[\frac{n+1}{2}\right]}{\Gamma\left[\frac{n}{2}\right]} \cdot \left[1 + \frac{x^2}{n}\right]^{-\frac{n+1}{2}} \quad , \; -\infty < x < \infty$$

El único parámetro de esta distribución es *n* que representa el número de grados de libertad, y coincide con el número de *grados de libertad* de la variable chi cuadrado.

Propiedades

1.- La variable *t*, (como la Normal), toma valores de -∞ a +∞ .

2.- Si $X_0, X_1, X_2, ...X_n$ $iid \; N(0,1)$, *entonces*: $\dfrac{X_0}{\sqrt{\frac{1}{n}(X_1^2 + X_2^2 + X_3^2 + ... + X_n^2}} \sim t_n$

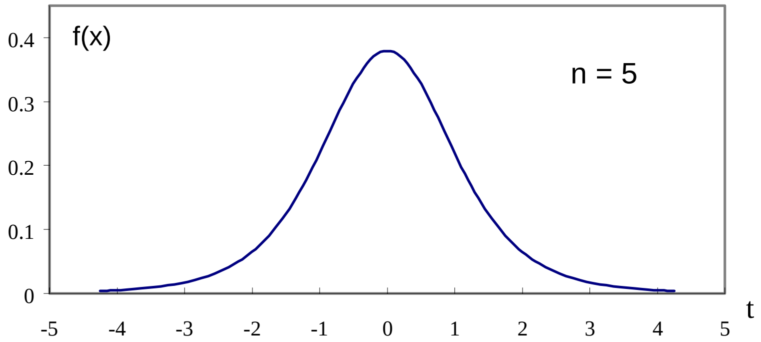
3.- La distribución *t*-Student es simétrica con media y variancia

$$E(X) = 0 \quad , \quad n > 1$$

$$V(X) = \frac{n}{n-2} \quad , \; n > 2$$

4.- La variable *t*-Student tiene mayor dispersión que la Normal standart $\left[V(x)=\frac{n}{n-2}>1\right]$; pero su variancia tiende a 1 a medida que aumenta *n*. ∴ la distribución *t*-Student tiende a la distribución Normal Standard a medida que aumentan los *grados de libertad*. Es decir:

$$t_n \xrightarrow{D} N(0,1) \;\; para \;\; n \to \infty$$



6.- DISTRIBUCIÓN F-SNEDECOR (*F*)

Surge del cociente de dos variables chi-cuadrado independientes, cada una dividida por sus respectivos grados de libertad.

Sean Z e Y dos variables aleatorias independientes con:

$$Z \sim \chi^2_m \quad y \quad Y \sim \chi^2_n \quad y \;\; sea \;\; la \;\; variable \;\; X = \frac{Z/m}{Y/n} \;\; \therefore \;\; X \sim F_{m,n}$$

La función de densidad de esta variable aleatoria X está dada por la siguiente expresión:

$$f(x) = \frac{\Gamma\left[\frac{n+m}{2}\right]}{\Gamma\left[\frac{n}{2}\right]\Gamma\left[\frac{m}{2}\right]}\left[\frac{m}{n}\right]^{\frac{m}{2}} .x^{\frac{n}{2}-1}\left[1+\frac{mx}{n}\right]^{-\frac{n+m}{2}} \;\; , \quad x \geq 0$$

Esta distribución tiene dos parámetros: *m* y *n*, que representan los *grados de libertad*, y estos números coinciden con los grados de libertad de las variables chi cuadrado del numerador y denominador, respectivamente.

Propiedades

1.- La variable *F* toma valores de 0 a +∞, por ser cociente de dos variables que asumen valores positivos.

2.- Si Y_1 , Y_2 , ... , Y_m iid N(0,1) y Z_1 , Z_2 , ... , Z_n iid N(0,1), entonces:

$$\frac{\frac{1}{m}(Y_1^2+Y_2^2+Y_3^2+........+Y_m^2)}{\frac{1}{n}(Z_1^2+Z_2^2+Z_3^2+........+Z_m^2)} \sim F_{m,n}$$

3.- Si X ~ $F_{m,n}$, entonces su media y variancia son :

$$E(X)=\frac{n}{n-2},\, n>2 \qquad\qquad V(X)=\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)},\, n>4$$

4.- De la segunda propiedad resulta que: si Y_1 , Y_2 , ... , Y_m iid $N(\mu_y,\sigma_y)$ y

Z_1 , Z_2 , ... , Z_n iid $N(\mu_z,\sigma_z)$, entonces con Y_i y Z_i independientes $\forall$ i,j entonces:

$$\frac{\frac{1}{m}\sum_{i=1}^{m}\left[\frac{Y_i-\mu_y}{\sigma_y}\right]^2}{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{Z_i-\mu_z}{\sigma_z}\right]^2}\sim F_{m,n} \qquad ya\ \ que \qquad \sum_{i=1}^{m}\left[\frac{Y_i-\mu_y}{\sigma_y}\right]^2\sim\chi^2_m \quad y \quad \sum_{i=1}^{n}\left[\frac{Z_i-\mu_z}{\sigma_z}\right]^2\sim\chi^2_n$$

5.- La distribución *F*-Snedecor es positivamente asimétrica, pero su asimetría se reduce a medida que aumentan los grados de libertad *n* y *m*.

6.- Si $X\sim F_{m,n},\ \ entonces\ \ \frac{1}{X}\sim F_{n,m}$

Representación gráfica