Distribuciones de probabilidad para variables aleatorias continuas

Héctor de la Torre Gutiérrez
hdelatorreg@up.edu.mx

Variables aleatorias continuas (v.a.c)

X es una v.a.c si su soporte (valores posibles) es infinito

Experimento	*v.a. X*	*Valores posibles*
Funcionamiento de un banco	Tiempo en minutos entre llegadas de clientes	$x \geq 0$
Llenar una lata de bebida (max. 12.1 onzas)	Cantidad de onzas	$0 \leq x \leq 12.1$
Proyecto para construir una nueva biblioteca	Porcentaje de termino de proyecto en seis meses	$0 \leq x \leq 100$
Ensayar un nuevo proceso quimico	Temperatura cuando se lleve a cabo la reaccion deseada (min 150 ^oF; max 212 ^oF)	$150 \leq x \leq 212$

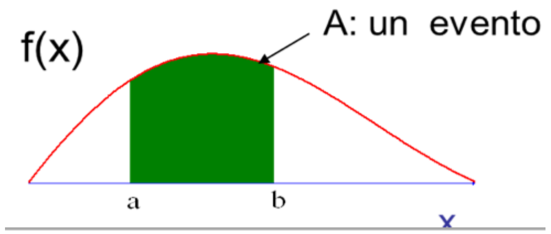
Función de densidad de probabilidades (fdp)

Una fdp de una v.a.c "X" satisface las siguientes propiedades:

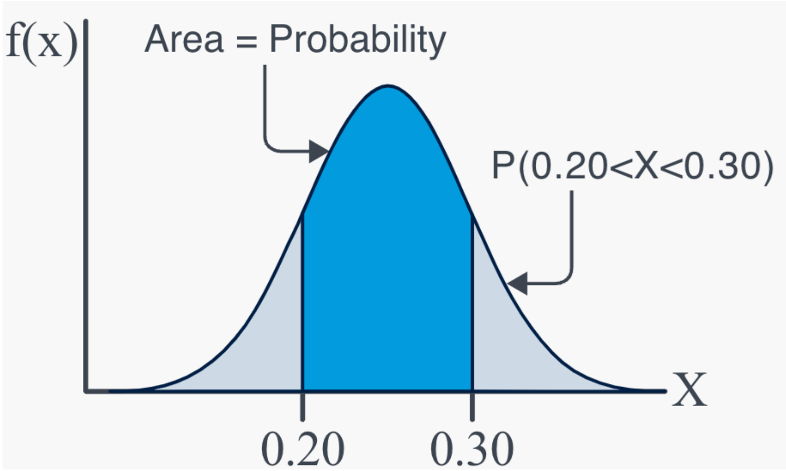
$$f_X(x) > 0 \ \forall X \in S_X$$

$$\int_{S_X} f_X(x)dx = 1$$

$$P(X \in A) = \int_A f_X(x)dx$$



Ejemplo: A=[0.20, 0.30]



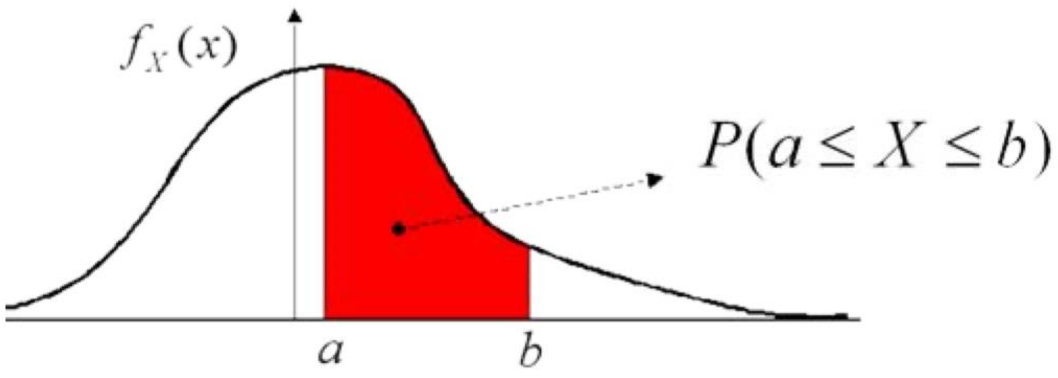
Función de Distribución - Acumulada

La **función de distribución acumulada** de la variable aleatoria X es la probabilidad de que X sea menor o igual a un valor específico x, esta dada por:

$$F(X) \equiv P(X \leq x) = \int_{-\infty}^{x} f(y)dy$$

por lo tanto

$$P(a \leq X \leq b) = \int_{a}^{b} f(y)dy = F(b) - F(a)$$



Modelos de Probabilidad Continuos (MPC)

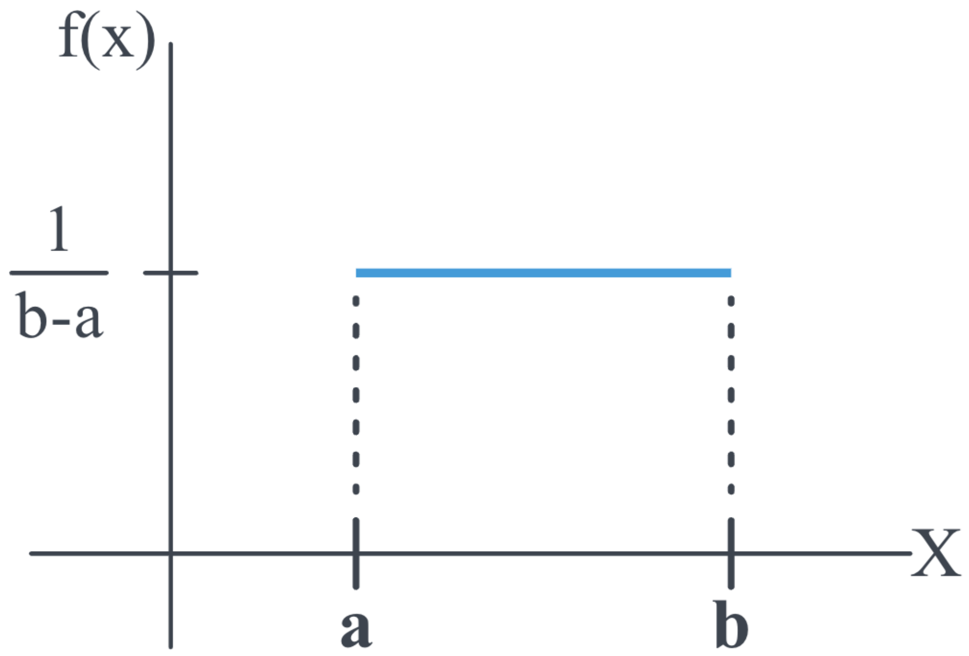
- Ahora revisaremos algunos MPC básicos:
  - Uniforme
  - Normal
  - t-student
  - Exponencial
  - Weibull

- Existen otras distribuciones continuas: gamma, beta, etc…… ver: https://seeing-theory.brown.edu/probability-distributions/es.html#section2

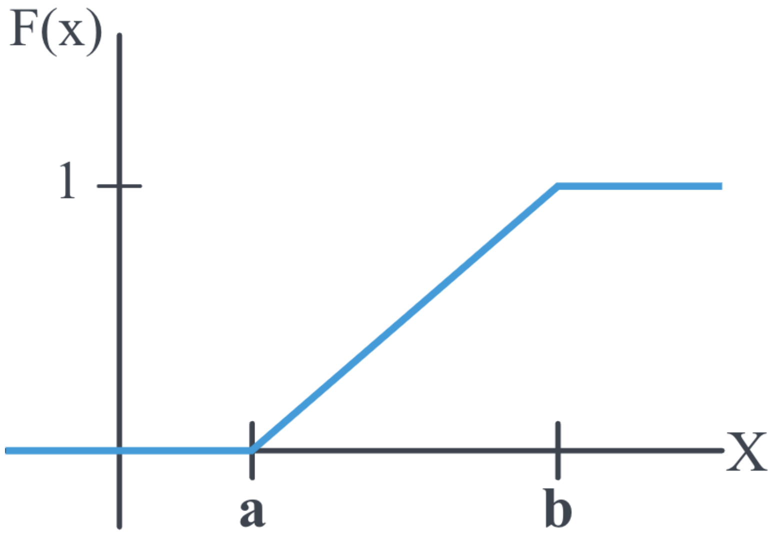
MPC- Uniforme

La función de densidad de una variable aleatoria uniformemente distribuida es:

$$f(x) = \frac{1}{b-a} \quad a \leq x \leq b \qquad (7)$$



$$F(x) = \frac{x-a}{b-a}$$



MPC- Normal

Su origen data del siglo XVIII. Gauss fue el primero quien la justificó.

Es un modelo que se ajusta “bien” a la mayoría de las variables de interés práctico, o al menos es un modelo que las aproxima regularmente “bien”.

Esta razón es fundamental en el desarrollo de métodos estadísticos que suponen de normalidad de las variables sobre las que se desean hacer inferencias sobre:

Medias de poblaciones.

Varianzas de poblaciones.

Pruebas aproximadas para proporciones.

La justificación sobre la suposición de normalidad se fortalece con el resultado conocido como el Teorema del Límite Central, el cual demuestra la convergencia del promedio de las distribuciones más usuales a la distribución normal.

Se dice que una v.a ***X*** sigue una distribución normal con media "$\boldsymbol{\mu}$" y varianza "$\boldsymbol{\sigma}^2$" si su fdp está dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$-\infty < x < +\infty$$

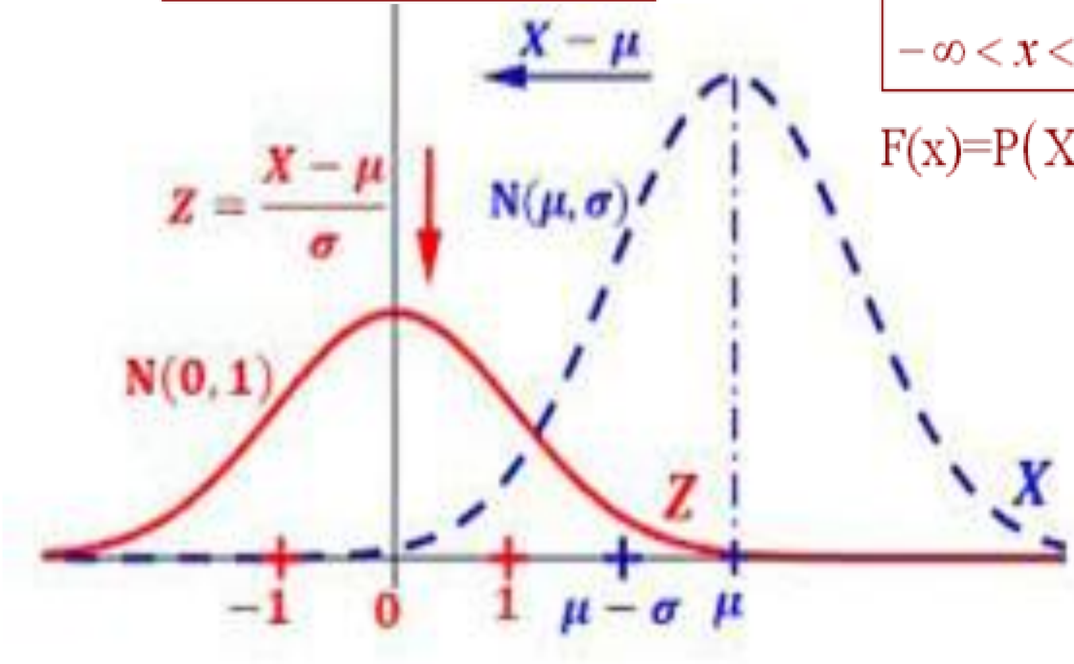
$$F(x) = P(X \leq x)$$

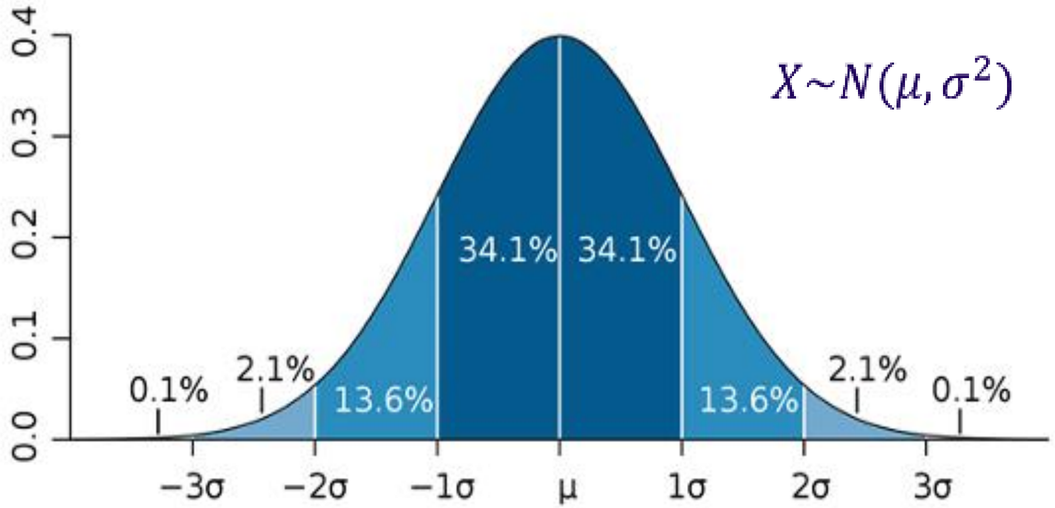
Donde μ y σ^2 son los parámetros

$\mu \in R$ y $\sigma^2 > 0$; $x \in R$

Notación para una variable aleatoria Normal

$X \sim N(\mu, \sigma^2)$





Ayuda: https://en.wikipedia.org/wiki/Normal_distribution

Características de la normal

Dominio: $Dom\, f = \mathbb{R}$

Máximo: $\left(\mu, \frac{1}{\sigma\sqrt{2\pi}}\right)$

P. inflexión: *en* $x = \mu + \sigma$ *y* $x = \mu - \sigma$

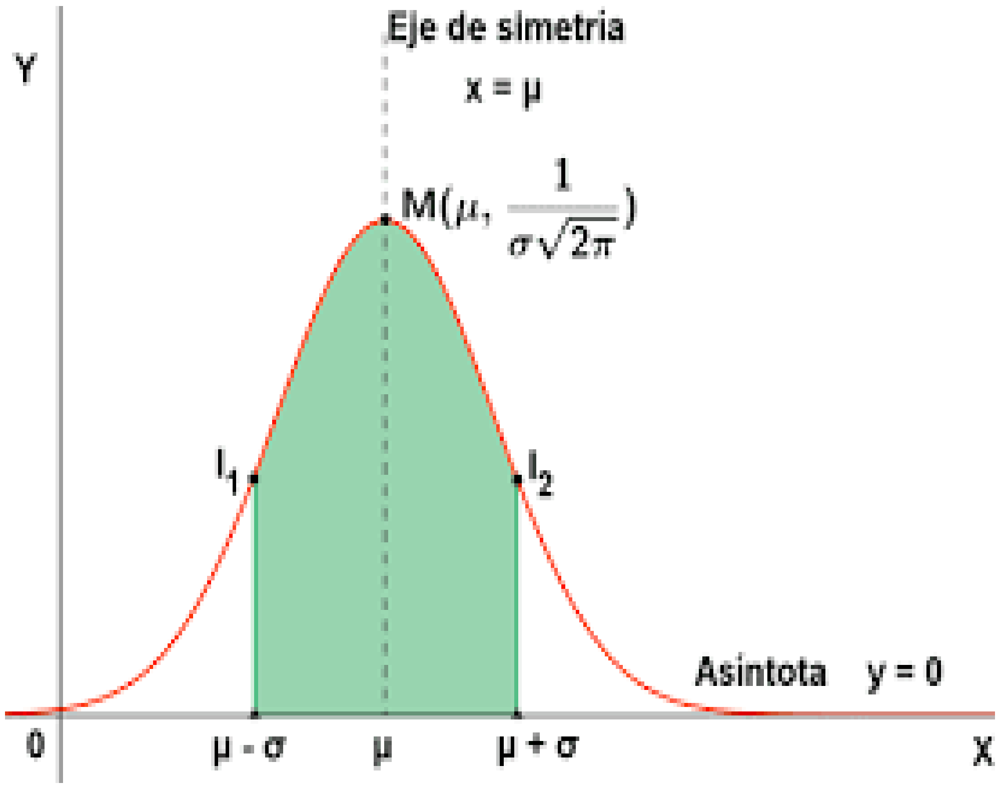
Asíntotas: *el eje OX es una asítota horizontal*

Simetrías: *respecto a la recta* $x = \mu$

Monotonía: *creciente* $(-\infty, \mu)$, *decreciente* $(\mu, +\infty)$

Signo: *es siempre positiva*

P. Corte: $OY \rightarrow \left(0, \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}}\right)$



$E(X) = \mu$ y $Var(X) = \sigma^2$

Normal estándar para calcular probabilidades

$$P(Z > a) = 1 - P(Z \le a)$$

$$P(Z \le -a) = P(Z > a) = 1 - P(Z \le a)$$

$$P(Z > -a) = P(Z \le a)$$

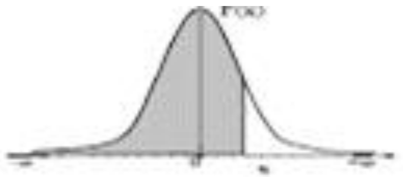
$$P(a \le Z \le b) = P(Z \le b) - P(Z \le a)$$

$$P(-a \le Z \le -b) = P(b \le Z \le a)$$

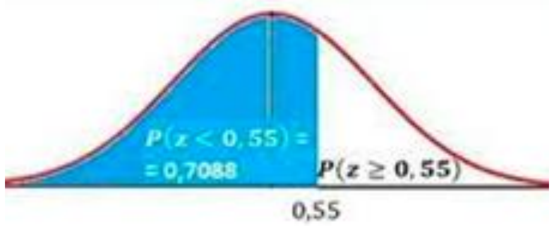
Distribución

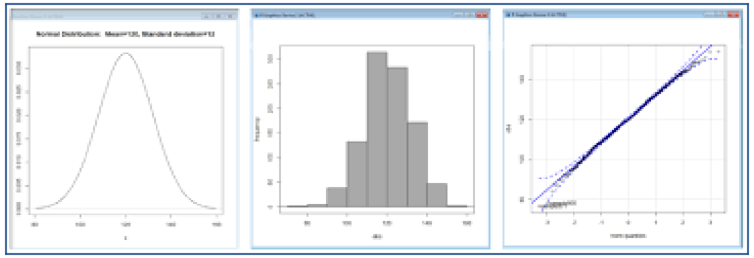
Normal $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$

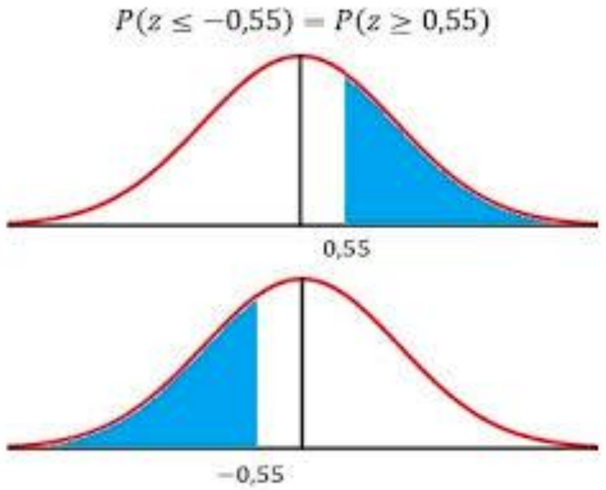
profesor10demates



	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141







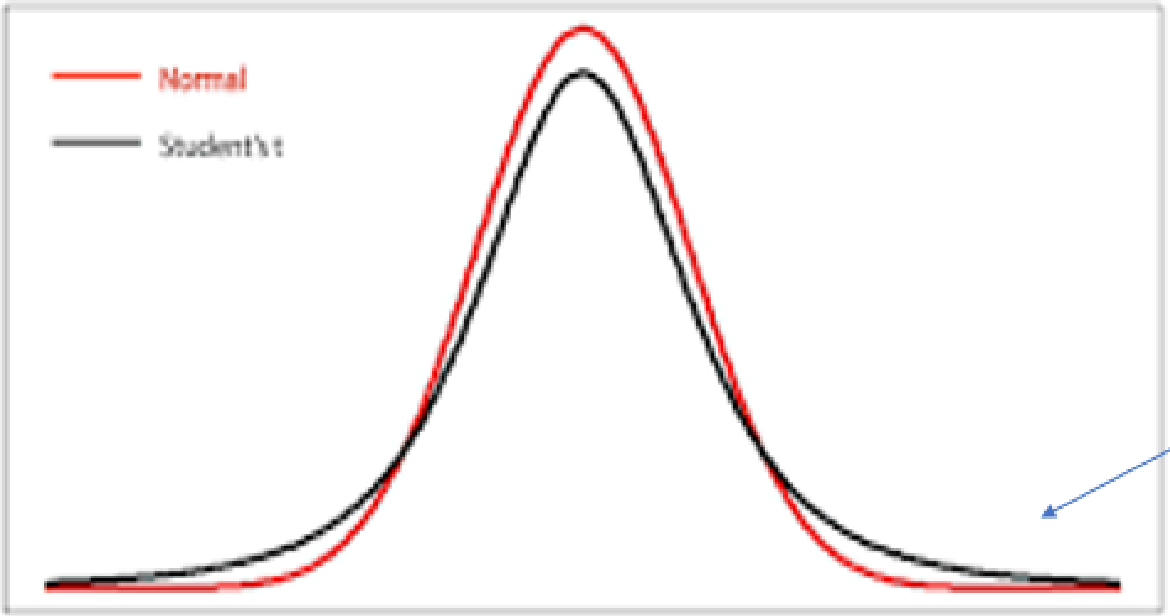
MPC- t-Student

- La Distribución t _ Student es básica en la prueba de hipótesis de promedios, una y dos medias.
- Se genera a partir de la división entre dos variables aleatorias independientes; una con distribución Z Normal Estándar y la segunda es la raíz de una variable aleatoria S^2 *Ji*-Cuadrada con n grados de libertad.
- La transformación es:

$$t = \frac{Z}{\sqrt[2]{\frac{S^2}{n}}} \sim t_{(n)}$$

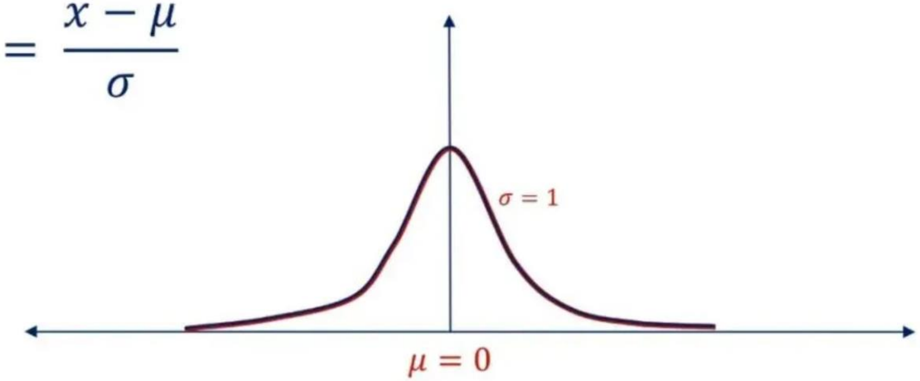
- El parámetro de la distribución, similar al de la distribución *Ji* –Cuadrada, son los grados de libertad, lo que se hereda precisamente de la variable aleatoria *Ji* –Cuadrada.

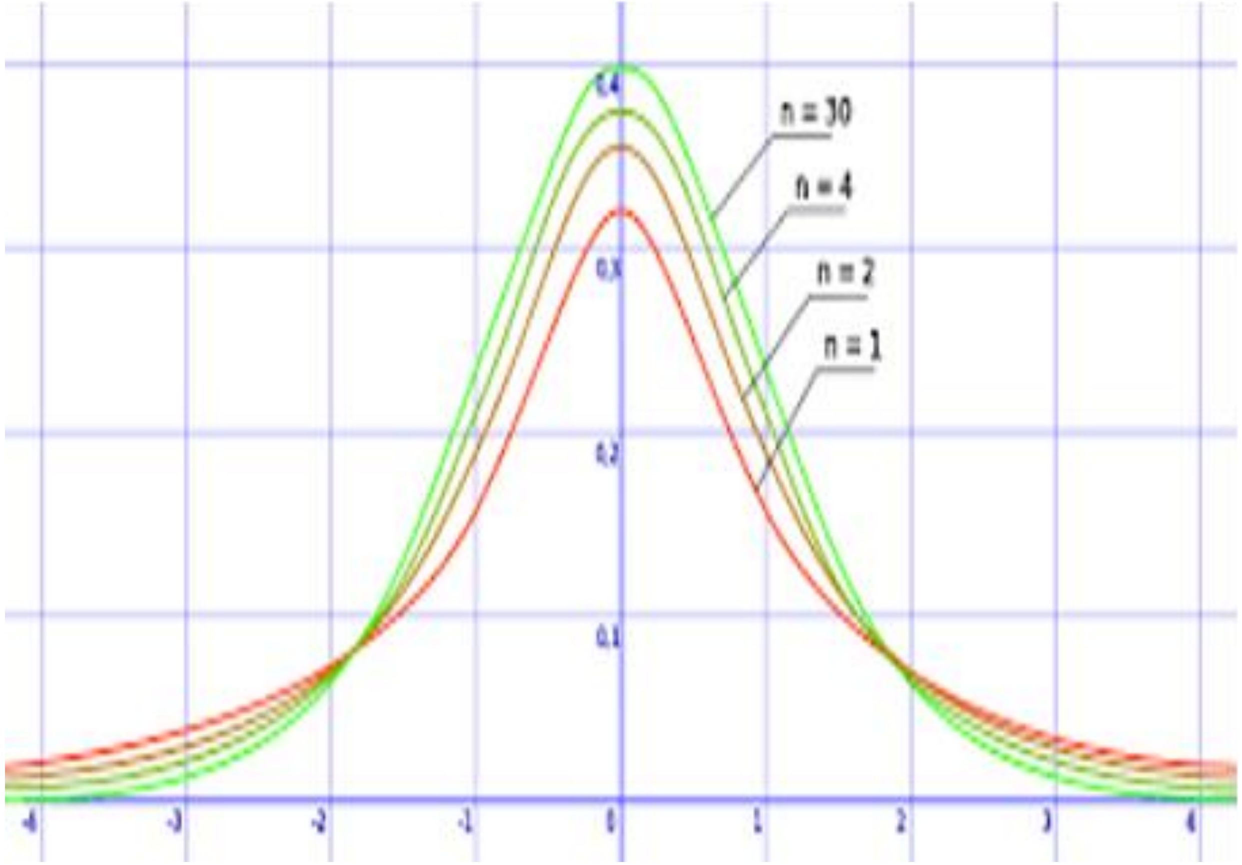
Comparación: z Vs t



La distribución *t* – Student
Tiene probabilidades más
"pesadas" en las colas de la
distribución

$$z = \frac{x - \mu}{\sigma}$$





Conforme aumentan los grados de libertad, la distribución se parece más a la distribución Normal Estándar

Es simétrica alrededor de 0

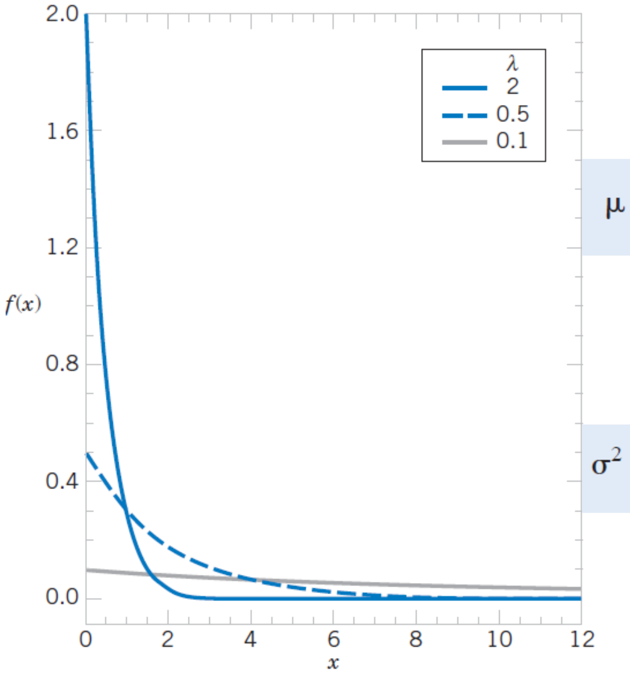
Distribución Exponencial (1)

Sea X la variable aleatoria igual a la distancia entre eventos sucesivos de un proceso de Poisson con número medio de eventos λ>0 por unidad de intervalo una variable aleatoria Exponencial con parámetro λ.

La función de distribución de probabilidad está dada por:

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \leq x < \infty$$

Distribución exponencial (2)



$$\mu = E(X) = \frac{1}{\lambda}$$

$$\sigma^2 = V(X) = \frac{1}{\lambda^2}$$

- Tiempos de espera para la ocurrencia de un cierto evento
- cantidad de tiempo hasta que ocurre algún evento específico.
- La cantidad de tiempo que transcurre hasta que se produce algún evento específico

Distribución exponencial (3). Ejemplo

- En una gran red informática corporativa, los inicios de sesión de los usuarios en el sistema pueden modelarse como un proceso de Poisson con una media de 25 inicios de sesión por hora. ¿Cuál es la probabilidad de que no se produzca ninguna conexión en un intervalo de 6 minutos?
- Sea X el tiempo en horas transcurrido desde el inicio del intervalo hasta la primera conexión. Entonces, X tiene una distribución exponencial con $\lambda = 25$ conexiones por hora.
- Nos interesa la probabilidad de que X supere los 6 minutos. Dado que λ se da en log-ons por hora, expresamos todas las unidades de tiempo en horas. Es decir, 6 minutos = 0.1 hora.

Distribución exponencial (4)

- ¿Cuál es la probabilidad de que el tiempo del siguiente inicio de sesión se encuentre entre 2 y tres minutos?
- Determine el interval de tiempo tal que la probabilidad de que no haya inicios de session sea 0.90. $P(X > x) = 0.90$

Distribución Weibull (1)

La distribución de Weibull se utiliza a menudo para modelar el tiempo hasta el fallo de muchos sistemas físicos diferentes.

Los parámetros de la distribución ofrecen una gran flexibilidad para modelar sistemas en los que:

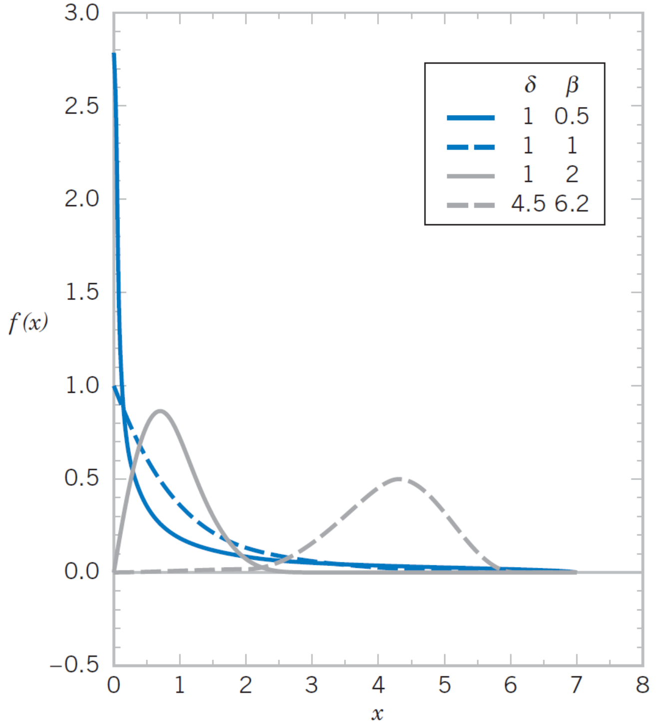
- El número de fallos aumenta con el tiempo (desgaste de rodamientos).
- Disminuye con el tiempo (algunos semiconductores) o
- Se mantiene constante con el tiempo (fallos causados por choques externos al sistema).

Distribución Weibull (2)

La función de densidad está dada por:

$$f(x) = \frac{\beta}{\delta}\left(\frac{x}{\delta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\delta}\right)^{\beta}\right]$$

Donde δ>0 es el parámetro de escala y β>0 el parámetro de forma.



Con función de distribución acumulada:

$$F(x) = 1 - e^{-\left(\frac{x}{\delta}\right)^{\beta}}$$

$$\mu = E(X) = \delta\Gamma\left(1 + \frac{1}{\beta}\right)$$

$$\sigma^2 = V(X) = \delta^2\Gamma\left(1 + \frac{2}{\beta}\right) - \delta^2\left[\Gamma\left(1 + \frac{1}{\beta}\right)\right]^2$$

Distribución Weibull (3)

El tiempo hasta el fallo (en horas) de un balero en un eje mecánico se modela satisfactoriamente como una variable aleatoria de Weibull con $\beta=1/2$ y $\delta = 5000$ horas.

1. Determina el tiempo medio hasta la falla.
2. Determina la probabilidad de que un balero dure por lo menos 6000 horas sin fallar.