EAN Escuela de Administración de Negocios

MEDIR

Probabilidad y Estadística

UNIVERSIDAD DE COSTA RICA
LUCEM ASPICIO

Términos estadísticos básicos

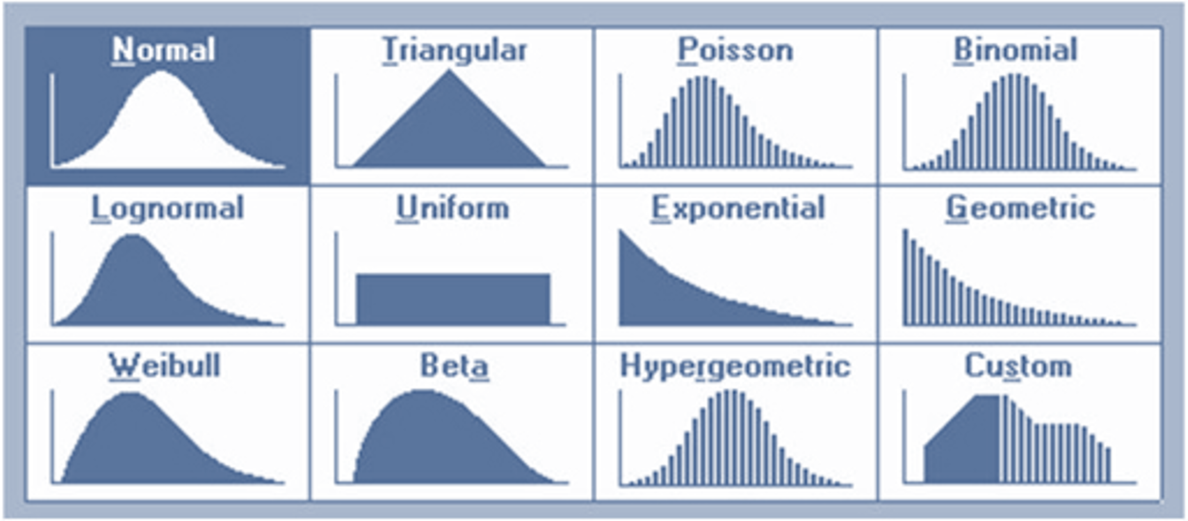
Distribuciones

CONTINUAS

Distribución que contiene una cantidad infinita de datos que pueden ser desplegados en una escala de medición.

DISCRETAS

Distribución que resulta de datos contados (atributos) que tienen una cantidad finita de posibles valores.



¿Cuáles son discretas y cuáles continuas?

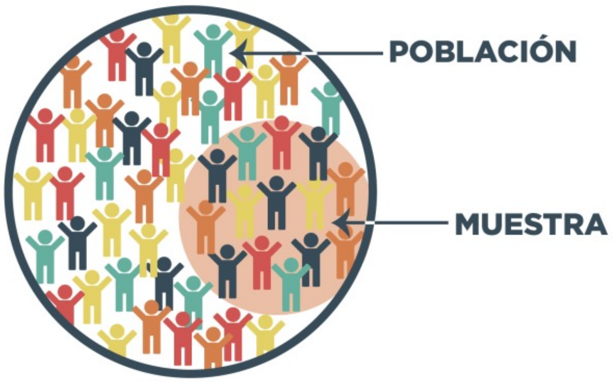
Población y Muestra

POBLACIÓN

Es el conjunto de individuos, objetos o fenómenos de los cuales se desea estudiar una o varias características.

MUESTRA

Un grupo seleccionado aleatoriamente de la población.



Parámetro y Estadístico

El valor numérico real de una población, estimado por un estadístico que es un valor numérico extraído de una muestra de la población.

Probabilidad

Es la medida numérica de la posibilidad de que un evento pueda ocurrir.

Por definición, la probabilidad se mide por un número entre cero y uno. Si un suceso no ocurre nunca, su probabilidad es cero, mientras que si ocurriese siempre su probabilidad es uno.

$$P(\text{evento}) = \frac{k}{N}$$

APROXIMACIÓN DE LA PROBABILIDAD POR FRECUENCIAS RELATIVAS

Método de asignar probabilidades con base en la experimentación o en datos históricos. La probabilidad del evento estará dada por la cantidad de veces que sucedió el evento durante el experimento.

$$P(A) = \frac{\text{Número de veces que ocurre A}}{\text{Número de veces que se repite el ensayo}}$$

Método Clásico de la Probabilidad

Parte de la consideración de que los resultados de un experimento son igualmente posibles, es decir, todos los componentes del espacio muestral tienen la misma probabilidad de ocurrir.

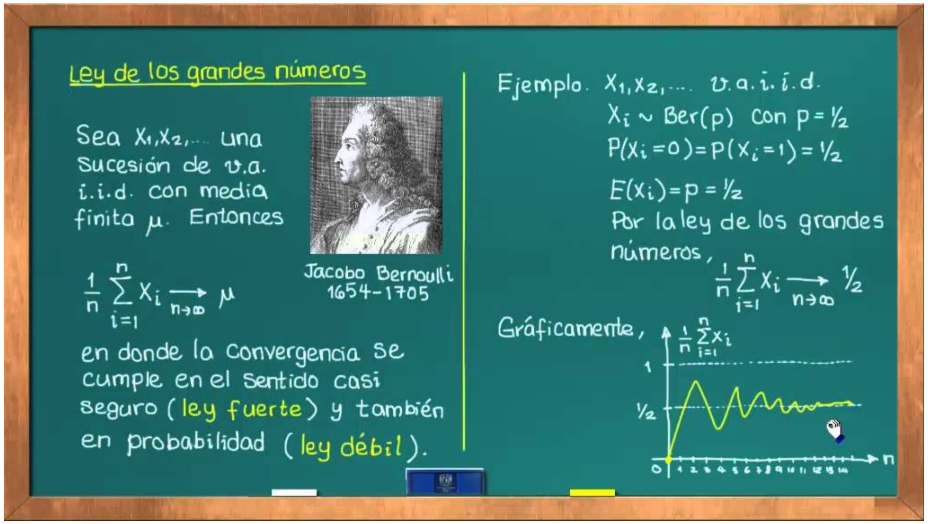
$$P(\text{evento}) = \frac{\text{Número de resultados favorables}}{\text{Número total de resultados posibles}}$$

Probabilidades subjetivas

Son las probabilidades que asigna una persona de que suceda un evento específico con base en cualquier información disponible. Es una probabilidad asignada bajo un criterio personal, con base en cualquier tipo de evidencia disponible, implica un grado de creencia personal.

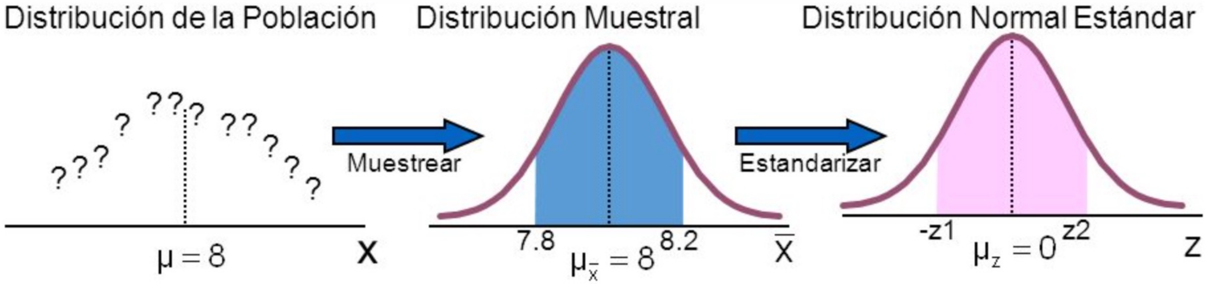
Ley de los Grandes Números

Conforme un procedimiento se repite una y otra vez, la probabilidad de frecuencias relativas de un suceso, tiende a aproximarse a la probabilidad real.



Teorema del Límite Central

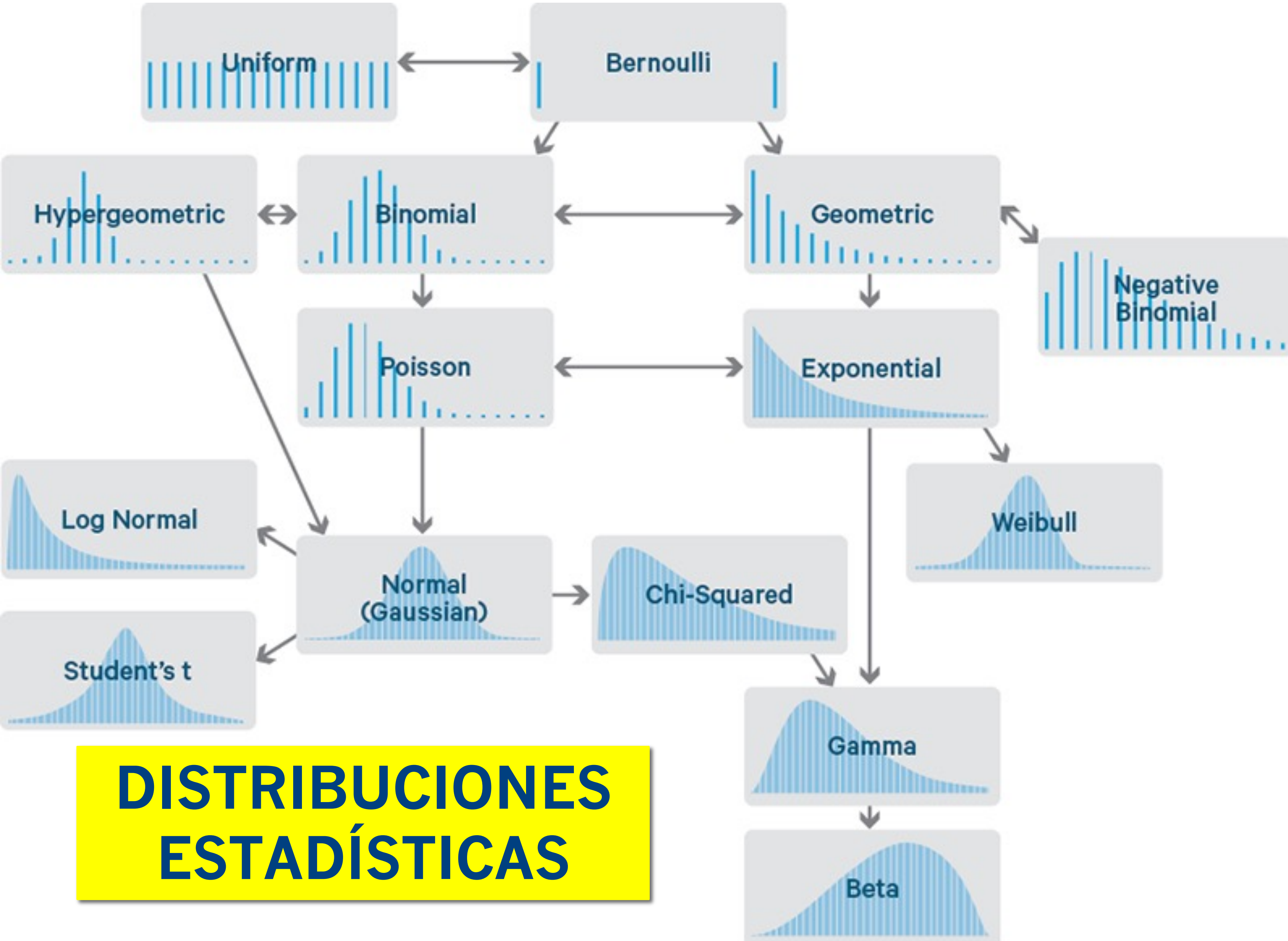
Sin tener en cuenta la forma de la población que se está estudiando, podemos seguir empleando la teoría normal para obtener inferencias sobre la media poblacional a condición de que obtengamos una muestra grande, porque la distribución muestral de x será aproximadamente normal cuando n sea grande.



Repasar

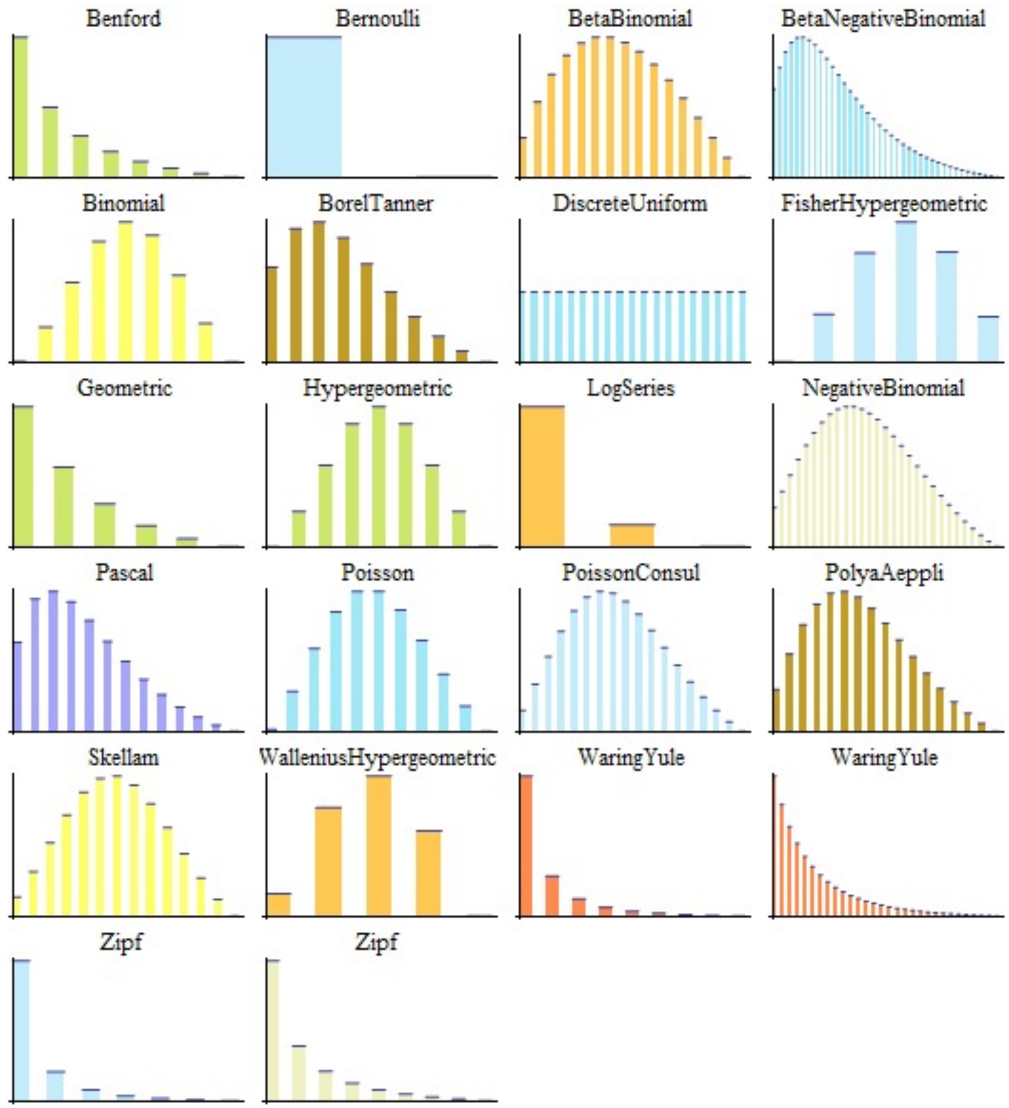
1. El complemento de una probabilidad.
2. Ley aditiva.
3. Ley aditiva completa.
4. Ley multiplicativa.
5. Probabilidad condicional.
6. Árbol de espacio muestral.

DISTRIBUCIONES ESTADÍSTICAS



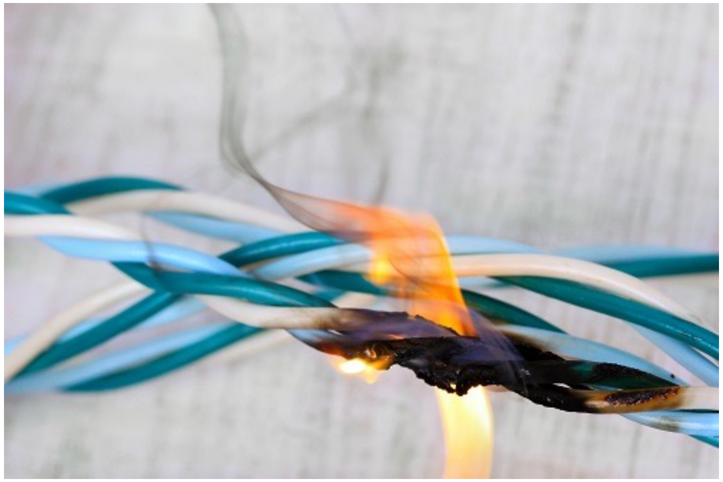
Distribuciones Discretas

Una distribución de probabilidades para una variable aleatoria discreta es un listado mutuamente excluyente de todos los resultados numéricos posibles para esa variable aleatoria tal que una probabilidad específica de ocurrencia se asocia con cada resultado.



Situaciones de SI-NO en la vida real.

Es frecuente que en control de calidad se den variables del tipo "pasa, no pasa". Por ejemplo, un artículo cumple con especificaciones o no, una pieza resiste cierta fuerza o no, una lámpara enciende o no.



Proceso de Bernoulli

Frecuentemente un experimento consiste en ensayos repetidos, cada uno con dos posibles resultados que pueden llamarse éxito o fracaso.

1. El experimento consiste en n intentos repetidos
2. Los resultados de cada uno de los intentos pueden clasificarse como un éxito o como un fracaso
3. La probabilidad de éxito representada por p, permanece constante en todos los intentos
4. Los intentos repetidos son independientes

Este procedimiento se conoce como *El proceso de Bernoulli*

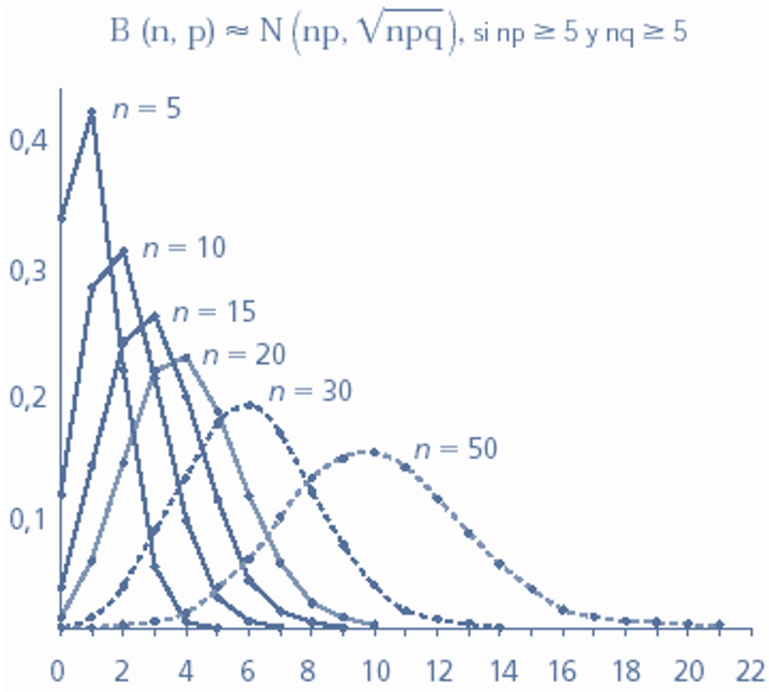
Distribución Binomial

El número X de éxitos en n experimentos de Bernoulli recibe el nombre de variable aleatoria binomial.

La distribución de probabilidad de esta variable aleatoria discreta se llama distribución binomial y sus valores se representan por: *b(x; n, p)*

$$b(x;\, n,\, p) = \binom{n}{x} p^x q^{n-x}$$

Se requieren dos parámetros, p y n.



Distribución Binomial

Probabilidad de x éxitos en n ensayos es:

$$\frac{n!}{x!\ (n-x)!}\ P^{x} Q^{n-x}$$

Se requieren dos parámetros, p y n.

Distribución Binomial

Media

$\mu = np$

Varianza

$\sigma^2 = npq$

Ejercicio

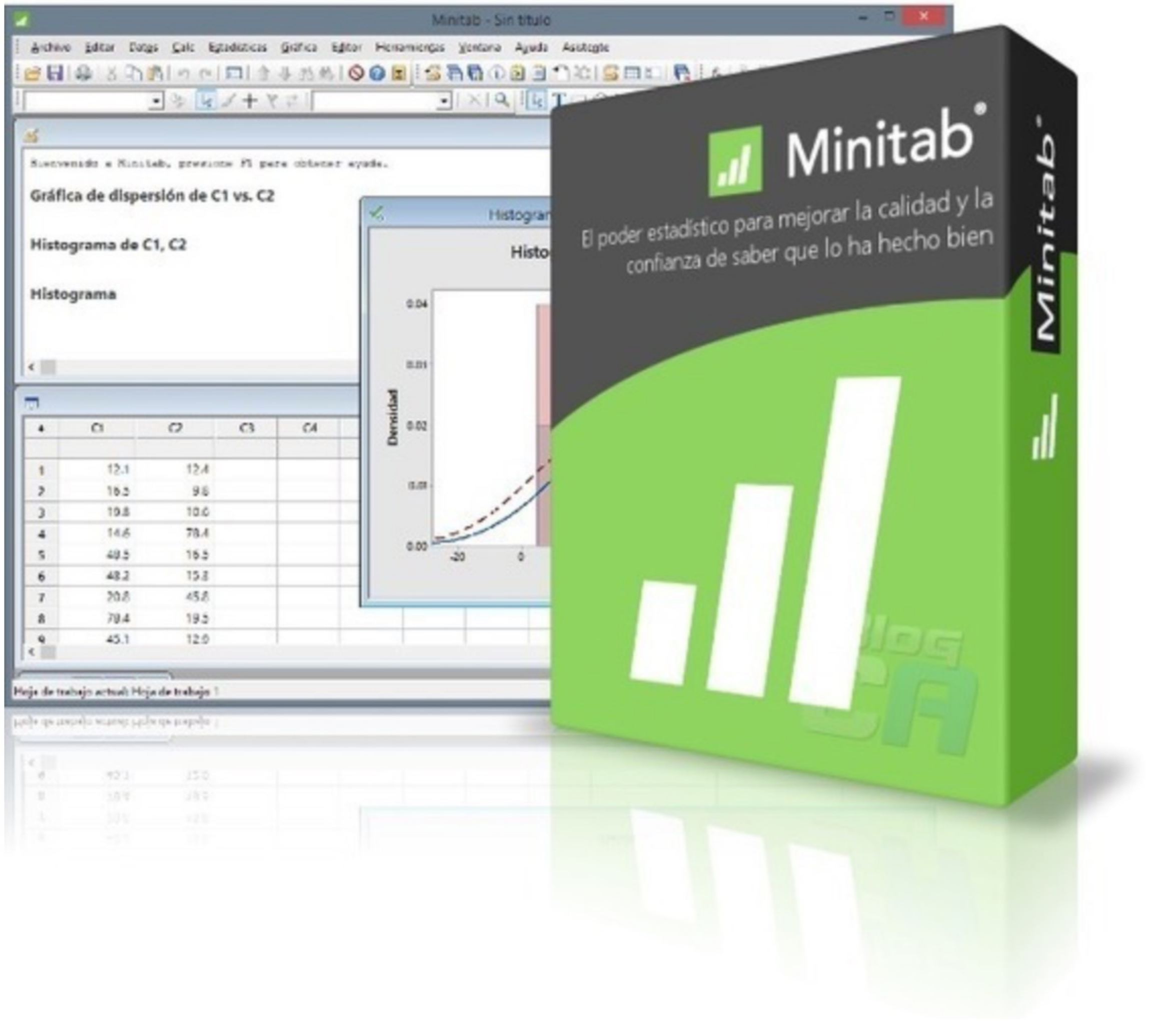
Una empresa que fabrica tornillos, del último lote de producción se escoge un grupo de ellos y se seleccionan 6 tornillos al azar. Si el 40% de los tornillos salen doblados (defectuosos) calcule:

1. La probabilidad de que exáctamente 4 salgan defectuosos.
2. Por lo menos 4 salgan defectuosos.
3. Entre 3 y 5 salgan defectuosos.

Tabla 10. Sumas de probabilidades binomial $\sum_{X=0} b(x;n,p)$

		p									
n	r	.10	.20	.25	.30	.40	.50	.60	.70	.80	.90
1	0	.9000	.8000	.7500	.7000	.6000	.5000	.4000	.3000	.2000	.1000
	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	0	.8100	.6400	.5625	.4900	.3600	.2500	.1600	.0900	.0400	.0100
	1	.9900	.9600	.9375	.9100	.8400	.7500	.6400	.5100	.3600	.1900
	2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0	.7290	.5120	.4219	.3430	.2160	.1250	.0640	.0270	.0080	.0010
	1	.9720	.8960	.8438	.7840	.6480	.5000	.3520	.2160	.1040	.0280
	2	.9990	.9920	.9844	.9730	.9360	.8750	.7840	.6570	.4880	.2710
	3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0	.6561	.4096	.3164	.2401	.1296	.0625	.0256	.0081	.0016	.0001
	1	.9477	.8192	.7383	.6517	.4752	.3125	.1792	.0837	.0272	.0037
	2	.9963	.9728	.9492	.9163	.8208	.6875	.5248	.3483	.1808	.0523
	3	.9999	.9984	.9961	.9919	.9744	.9375	.8704	.7599	.5904	.3439
	4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0	.5905	.3277	.2373	.1681	.0778	.0312	.0102	.0024	.0003	.0000
	1	.9185	.7373	.6328	.5282	.3370	.1875	.0870	.0308	.0067	.0005
	2	.9914	.9421	.8965	.8369	.6826	.5000	.3174	.1631	.0579	.0086
	3	.9995	.9933	.9844	.9692	.9130	.8125	.6630	.4718	.2627	.0815
	4	1.000	.9997	.9990	.9976	.9898	.9688	.9222	.8319	.6723	.4095
	5		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
6	0	.5314	.2621	.1780	.1176	.0467	.0156	.0041	.0007	.0001	.0000
	1	.8857	.6554	.5339	.4202	.2333	.1094	.0410	.0109	.0016	.0001
	2	.9841	.9011	.8306	.7443	.5443	.3438	.1792	.0705	.0170	.0013
	3	.9987	.9830	.9624	.9295	.8208	.6563	.4557	.2557	.0989	.0158
	4	.9999	.9984	.9954	.9891	.9590	.8906	.7667	.5798	.3447	.1143
	5	1.000	.9999	.9998	.9993	.9959	.9844	.9533	.8824	.7379	.4686
	6		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Con Minitab



Minitab
El poder estadístico para mejorar la calidad y la confianza de saber que lo ha hecho bien
Minitab
Minitab - Sin título
Gráfica de dispersión de C1 vs. C2
Histograma de C1, C2
Histograma

Distribución Geométrica

La distribución geométrica es un modelo adecuado para aquellos procesos en los que se repiten pruebas hasta la consecución del éxito, es decir, hasta que se alcance el resultado deseado, por ejemplo, supongamos que interesa encontrar un producto defectuoso en un lote de muchos artículos.

Las probabilidades p y q son constantes en todas las pruebas ,por lo tanto, las pruebas ,son independientes (si se trata de un proceso de "extracción" éste se llevará a cabo con la devolución del individuo extraído) .

Distribución Geométrica

Esta distribución se puede hacer derivar de un proceso de Bernouilli ***independiente*** en el que tengamos las siguientes características:

1. El proceso consta de un número no definido de pruebas o experimentos separados o separables. El proceso concluirá cuando se obtenga por primera vez el resultado deseado (éxito).
2. Cada prueba puede dar dos resultados mutuamente excluyentes : A y no A
3. La probabilidad de obtener un resultado A en cada prueba es p y la de obtener un resultado que no sea A es q siendo (p + q = 1).

Distribución Geométrica

La Función de probabilidad:

$$f_X(x) = P(X=k) = p(1-p)^{k-1}$$

Donde k es el número de intentos o fallas antes del primer éxito.

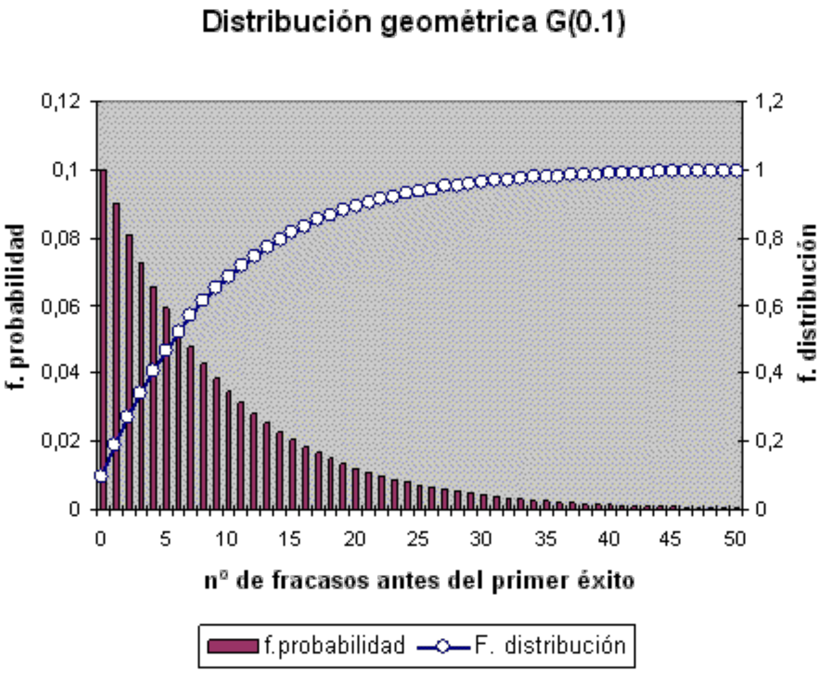
k es una variable aleatoria discreta y p es la probabilidad de éxito.

Distribución Geométrica

- **Media & Varianza**

$$\mu = E(X) = \sum_{n=1}^{\infty} n(1-p)^{n-1} p = \frac{1}{p}$$

$$\sigma^2 = E(X^2) - [E(X)]^2 = \frac{1-p}{p^2}$$



Ejercicio

Sí la probabilidad de que cierto dispositivo de medición muestre una desviación excesiva es de 0.05, ¿cuál es la probabilidad de que;

a) El sexto de estos dispositivos de medición sometidos a prueba sea el primero en mostrar una desviación excesiva.

b) El séptimo de estos dispositivos de medición sometidos a prueba, sea el primero que no muestre una desviación excesiva.

c) Entre el segundo y el cuarto de estos dispositivos de medición sometidos a prueba aparezca el primero en mostrar una desviación excesiva.

Distribución Hipergeométrica

Hasta ahora se han analizado distribuciones que modelan situaciones en las que se realizan pruebas que implican una dicotomía (proceso de Bernouilli) de manera que en cada experimento la probabilidad de obtener cada uno de los dos posibles resultados se mantenía constante. Si el proceso consiste en una serie de extracciones ello implica la reposición de cada extracción, o la consideración de una población muy grande.

Pero si la población es pequeña y las extracciones no se remplazan las probabildades no se mantendrán constantes, en ese caso las distribuciones anteriores no sirven para modelar esta situación, es aquí cuando se aplica la distribución hipergeométrica.

Distribución Hipergeométrica

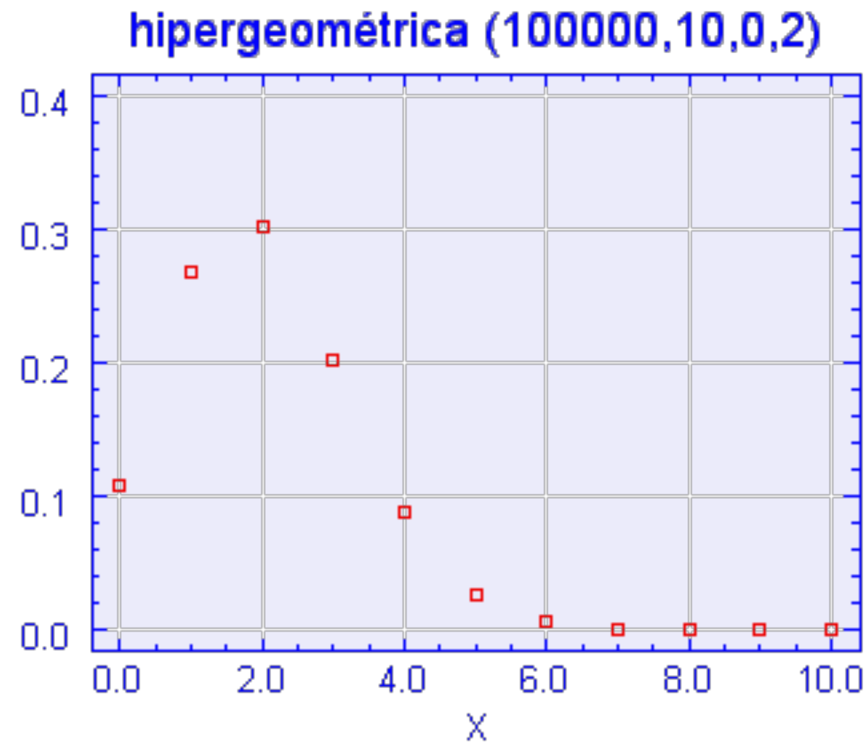
Da la probabilidad de obtener X éxitos en n experimentos Bernoulli, donde la probabilidad de éxito cambia de un experimento al siguiente.

$$p_X(k) = \frac{\binom{r}{k}\binom{N-r}{m-k}}{\binom{N}{m}}$$

En donde:
N=Tamaño de la población
m=Tamaño de la muestra
r=cantidad de elementos que cumplen lo deseado
k=cantidad de éxitos

Distribución Hipergeométrica

$$
\begin{aligned}
E(X) &= m\frac{r}{N} \\
var(X) &= m\frac{r}{N}\frac{(N-r)}{N}\frac{(N-m)}{(N-1)}
\end{aligned}
$$



Es fundamental en el estudio de muestras pequeñas de poblaciones pequeñas.
Tiene grandes aplicaciones en Control de Calidad y Diseño de Experimentos

Ejercicio

De un lote de 10 proyectiles, 4 se seleccionan al azar y se disparan. Si el lote contiene 3 proyectiles defectuosos que no explotarán, ¿cuál es la probabilidad de que:

a) Los 4 exploten.

b) Al menos 2 no exploten.

Distribución Multinomial

La distribución multinomial es un caso particular de la distribución binomial. Estas dos distribuciones se diferencian porque en la binomial cada uno de los experimentos puede ser éxito o no éxito, mientras que en la distribución multinomial hay muchas más opciones, precisamente k opciones. A cada una de estas opciones se las representa con la letra A y el subíndice hasta el k. Cabe remarcar que obligatoriamente al hacer uno de los experimentos, tiene que suceder alguno de los eventos A.

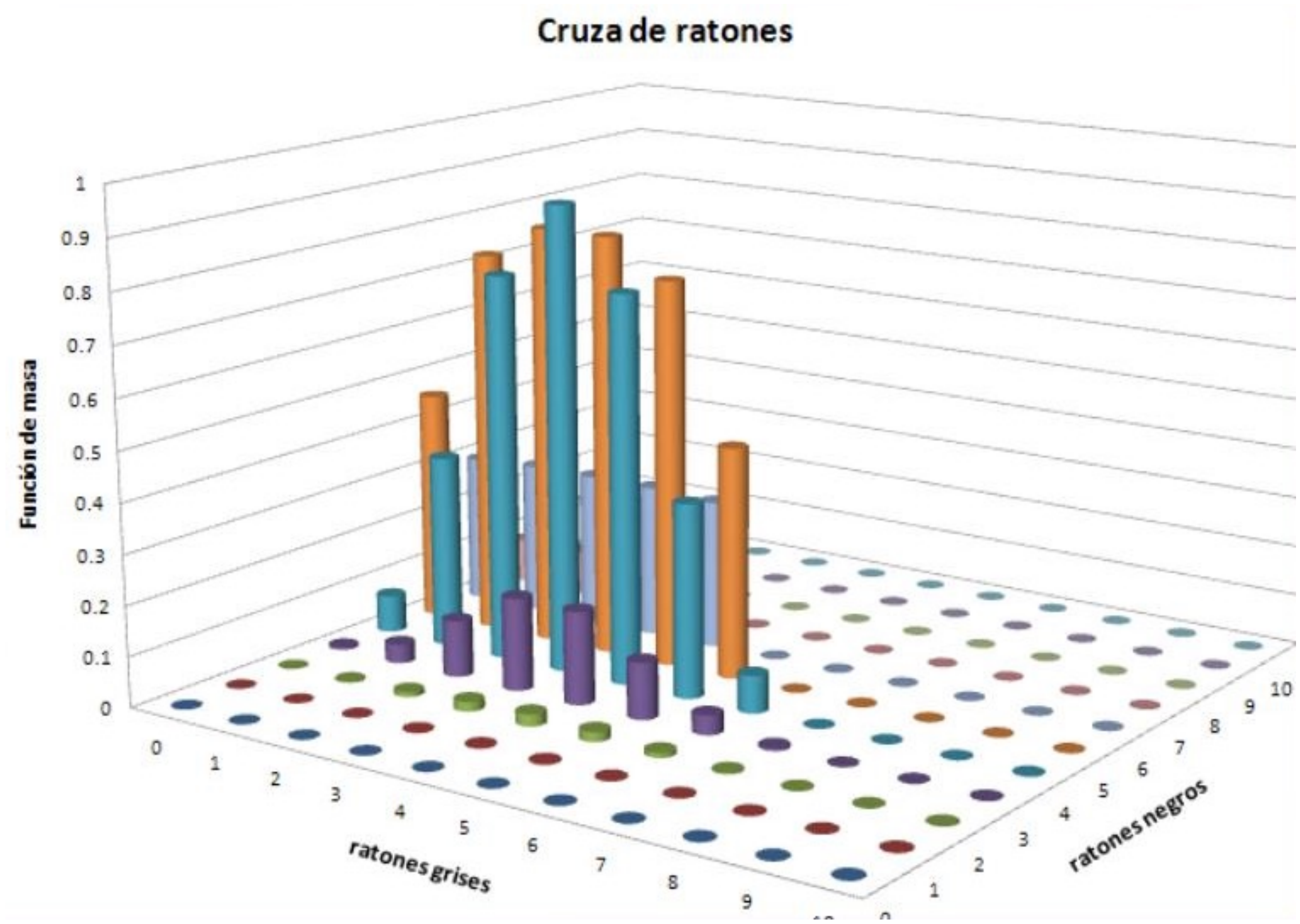
Distribución Multinomial

- Supongamos que el resultado de una determinada experiencia puede ser r valores distintos: A1, A2,..., Ar cada uno de ellos con probabilidad p1, p2,..., pr, respectivamente.

- Si repetimos la experiencia n veces en condiciones independientes, podemos preguntarnos la probabilidad de que el suceso A1 aparezca k1 veces, el suceso A2, k2 veces y así sucesivamente.

$$f(x_1,x_2,x_3,x_k)= \frac{n!}{x_1!x_2!x_3!\ldots x_k!} p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3} \ldots\, p_k^{x_k}$$

Distribución Multinomial



Ejercicio

Las probabilidades de que un delegado llegue por aire, en autobús, en automóvil o en tren a una cierta convención son de 0.40, 0.20, 0.30 y 0.10, respectivamente. ¿Cuál es la probabilidad de que entre 9 delegados seleccionados aleatoriamente en esta convención:

a) 3 hayan llegado por aire, 3 en autobús, 1 en auto y 2 en tren?

b) 4 hayan llegado por aire, 1 en autobús, 2 en auto y 2 en tren?

c) 5 hayan llegado en auto?

NO TIENE SOLUCIÓN EN MINITAB NI EXCEL

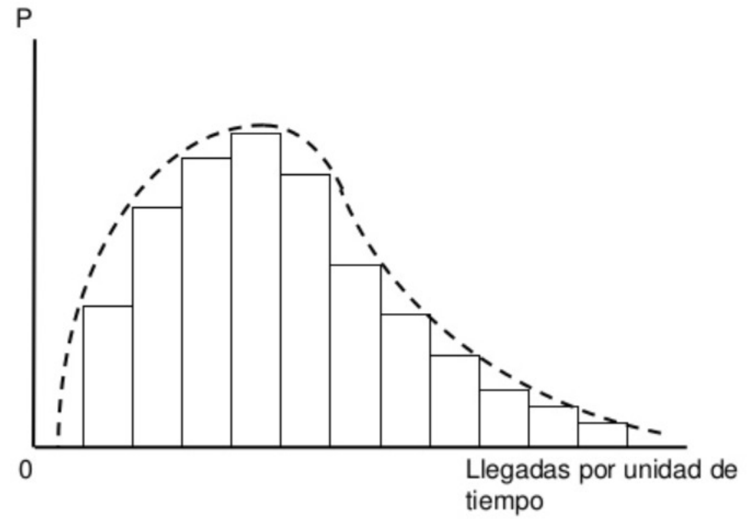
Distribución de Poisson

Es una distribución que se basa en el conteo de las veces que se presenta un evento dentro de un área de oportunidad dada. El área de oportunidad es una unidad continua o intervalo de tiempo, volumen o área en donde se puede presentar más de un evento.

(Berenson, Mark L., Levine D. Página 166).

Distribución de Poisson

La distribución fue descubierta por Siméon-Denis Poisson, proponiendo que el trabajo estaba enfocado en ciertas variables aleatorias N que cuentan, entre otras cosas, un número de ocurrencias discretas (muchas veces llamadas "arribos") que tienen lugar durante un intervalo de tiempo de duración determinada.



Si el número esperado de ocurrencias en este intervalo es λ, entonces la probabilidad de que haya exactamente k ocurrencias (siendo k un entero no negativo, k = 0, 1, 2, ...) es igual a:

$$p(k;\lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!}$$

Donde:

e es el base del logaritmo natural (*e* = 2.71828...),

k! es el factorial de *k*,

k es el número de ocurrencias de un evento,

λ es un número real positivo, equivalente al número esperado de ocurrencias durante un intervalo dado. Por ejemplo, si los eventos ocurren de media cada 4 minutos, y se está interesado en el número de eventos ocurriendo en un intervalo de 10 minutos, se usaría como modelo una distribución de Poisson con λ = 2.5. Esto quiere decir que landa es la media y la variancia en Poisson

Ejercicio

En la inspección de hojalata producida por un proceso electrolítico continuo, se identifican 0.2 imperfecciones en promedio por minuto. Determine las probabilidades de identificar:

a) una imperfección en 3 minutos

b) al menos dos imperfecciones en 5 minutos

c) a los sumo una imperfección en 15 minutos.

TABLA 7. Distribución de Poisson acumulada

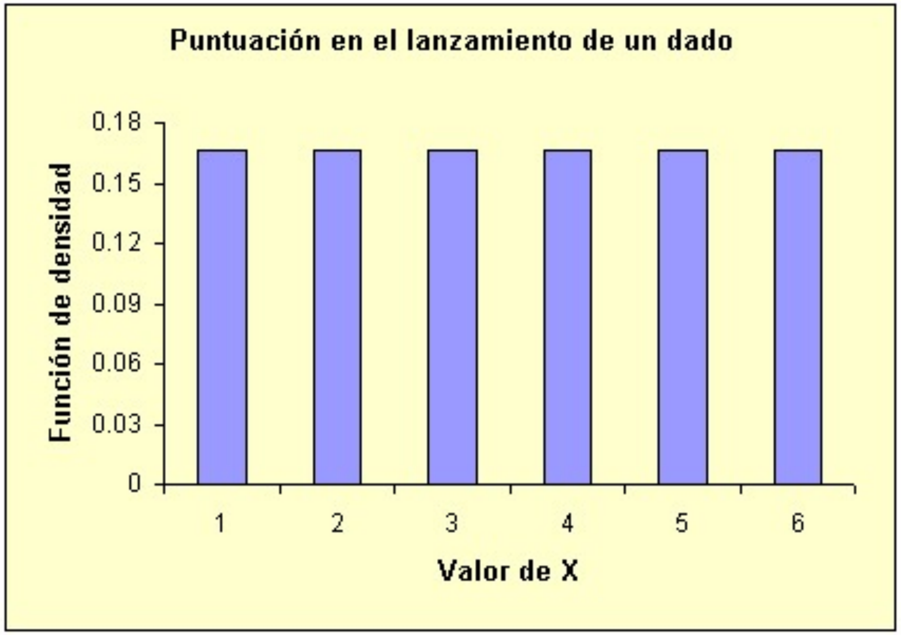
VALORES DE X

LANDA	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0.05	0.950	0.999																		
0.10	0.905	0.995	0.999																	
0.15	0.861	0.990	0.999																	
0.20	0.819	0.982	0.999																	
0.25	0.779	0.973	0.998	0.999																
0.30	0.741	0.963	0.996	0.999																
0.35	0.705	0.951	0.994	0.999																
0.40	0.670	0.938	0.992	0.999																
0.45	0.638	0.924	0.989	0.998	0.999															
0.50	0.607	0.910	0.986	0.998	0.999															
0.55	0.577	0.894	0.981	0.997	0.999															
0.60	0.549	0.878	0.977	0.997	0.999															
0.65	0.522	0.861	0.972	0.996	0.999															
0.70	0.497	0.844	0.966	0.994	0.999															
0.75	0.472	0.827	0.959	0.993	0.999															
0.80	0.449	0.809	0.953	0.991	0.998	0.999														
0.85	0.427	0.791	0.945	0.989	0.998	0.999														
0.90	0.407	0.772	0.937	0.987	0.998	0.999														
0.95	0.387	0.754	0.929	0.984	0.997	0.999														
1.00	0.368	0.736	0.920	0.981	0.996	0.999														
1.05	0.350	0.717	0.910	0.978	0.995	0.999														
1.10	0.333	0.699	0.900	0.974	0.995	0.999														
1.15	0.317	0.681	0.890	0.970	0.993	0.999														
1.20	0.301	0.663	0.879	0.966	0.992	0.998	0.999													
1.25	0.286	0.645	0.868	0.962	0.991	0.998	0.999													
1.30	0.273	0.627	0.857	0.957	0.989	0.998	0.999													
1.35	0.259	0.609	0.845	0.952	0.988	0.997	0.999													
1.40	0.247	0.592	0.833	0.946	0.986	0.997	0.999													
1.45	0.235	0.575	0.821	0.940	0.984	0.996	0.999													

LANDA	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2.85	0.058	0.223	0.458	0.681	0.840	0.930	0.973	0.991	0.997	0.999										
2.90	0.055	0.215	0.446	0.670	0.832	0.826	0.971	0.999	0.997	0.999										
2.95	0.052	0.207	0.434	0.658	0.824	0.921	0.969	0.989	0.997	0.999										
3.00	0.050	0.199	0.423	0.647	0.815	0.916	0.966	0.988	0.996	0.999										
3.05	0.047	0.192	0.415	0.636	0.807	0.911	0.964	0.987	0.996	0.999										
3.10	0.045	0.185	0.401	0.625	0.798	0.906	0.961	0.988	0.995	0.999										
3.15	0.043	0.178	0.390	0614	0.789	0.900	0.958	0.984	0.995	0.998	0.999									
3.20	0.041	0.171	0.380	0.603	0.781	0.895	0.955	0.983	0.994	0.998	0.999									
3.25	0.039	0.165	0.370	0.591	0.772	0.889	0.952	0.982	0.994	0.998	0.999									
3.30	0.037	0.159	0.359	0.580	0.763	0.883	0.949	0.980	0.993	0.998	0.999									
3.35	0.035	0.153	0.349	0.569	0.753	0.877	0.946	0.979	0.992	0.998	0.999									
3.40	0.033	0.147	0.340	0.559	0.744	0.871	0.942	0.977	0.992	0.997	0.999									
3.45	0.032	0.141	0.330	0.547	0.735	0.864	0.934	0.975	0.991	0.997	0.999									
3.50	0.030	0.136	0.321	0.537	0.725	0.858	0.935	0.973	0.990	0.997	0.999									
3.55	0.029	0.131	0.312	0.526	0.716	0.851	0.931	0.971	0.989	0.996	0.999									
3.60	0.027	0.126	0.303	0.515	0.706	0.844	0.927	0.969	0.988	0.996	0.999									
3.65	0.026	0.121	0.294	0.505	0.697	0.837	0.922	0.967	0.987	0.996	0.999									
3.70	0.025	0.116	0.285	0.494	0.687	0.830	0.918	0.965	0.988	0.995	0.988	0.999								
3.75	0.024	0.112	0.277	0.484	0.677	0.823	0.914	0.962	0.985	0.995	0.998	0.999								
3.80	0.022	0.107	0.267	0.473	0.668	0.816	0.909	0.960	0.983	0.994	0.998	0.999								
3.85	0.021	0.103	0.261	0.463	0.658	0.808	0.904	0.957	0.983	0.994	0.998	0.999								
3.90	0.020	0.099	0.253	0.453	0.648	0.800	0.899	0.955	0.981	0.993	0.998	0.999								
3.95	0.019	0.095	0.246	0.443	0.639	0.793	0.894	0.952	0.980	0.993	0.997	0.999								
4.00	0.018	0.092	0.238	0.433	0.629	0.785	0.889	0.949	0.979	0.992	0.997	0.999								
4.05	0.017	0.088	0.231	0.424	0.619	0.777	0.884	0.946	0.977	0.991	0.997	0.999								
4.10	0.017	0.084	0.224	0.414	0.609	0.769	0.879	0.943	0.975	0.990	0.997	0.999								
4.15	0.015	0.081	0.217	0.405	0.600	0.761	0.873	0.939	0.974	0.990	0.996	0.999								
4.20	0.015	0.080	0.210	0.395	0.590	0.753	0.867	0.936	0.972	0.989	0.996	0.999								
4.25	0.014	0.075	0.204	0.386	0.580	0.745	0.862	0.933	0.970	0.988	0.996	0.998	0.999							
4.30	0.014	0.072	0.197	0.377	0.570	0.737	0.856	0.929	0.968	0.987	0.995	0.998	0.999							

Distribución Uniforme Discreta

En teoría de la probabilidad, la distribución uniforme discreta es una distribución de probabilidad que asume un número finito de valores con la misma probabilidad.



Distribución Uniforme

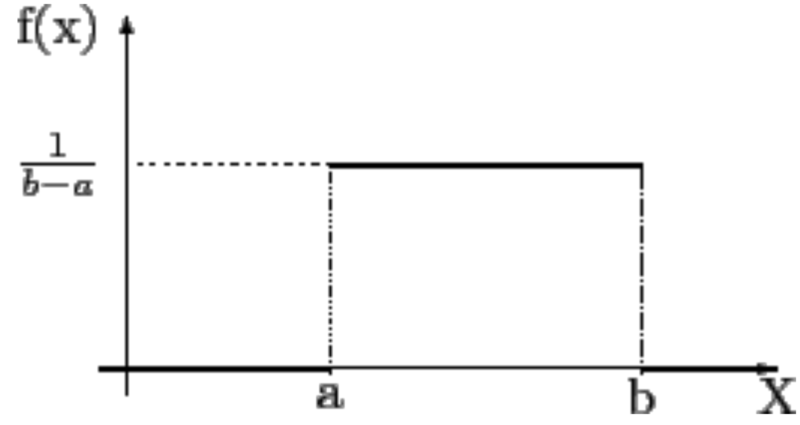
Sea X una variable uniforme que puede tomar los *n* siguientes números reales: x_1, x_2, x_3, ….., x_n entonces:

$$P(X=x_i) = \frac{1}{n}$$

Distribución Uniforme

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \mu^2$$

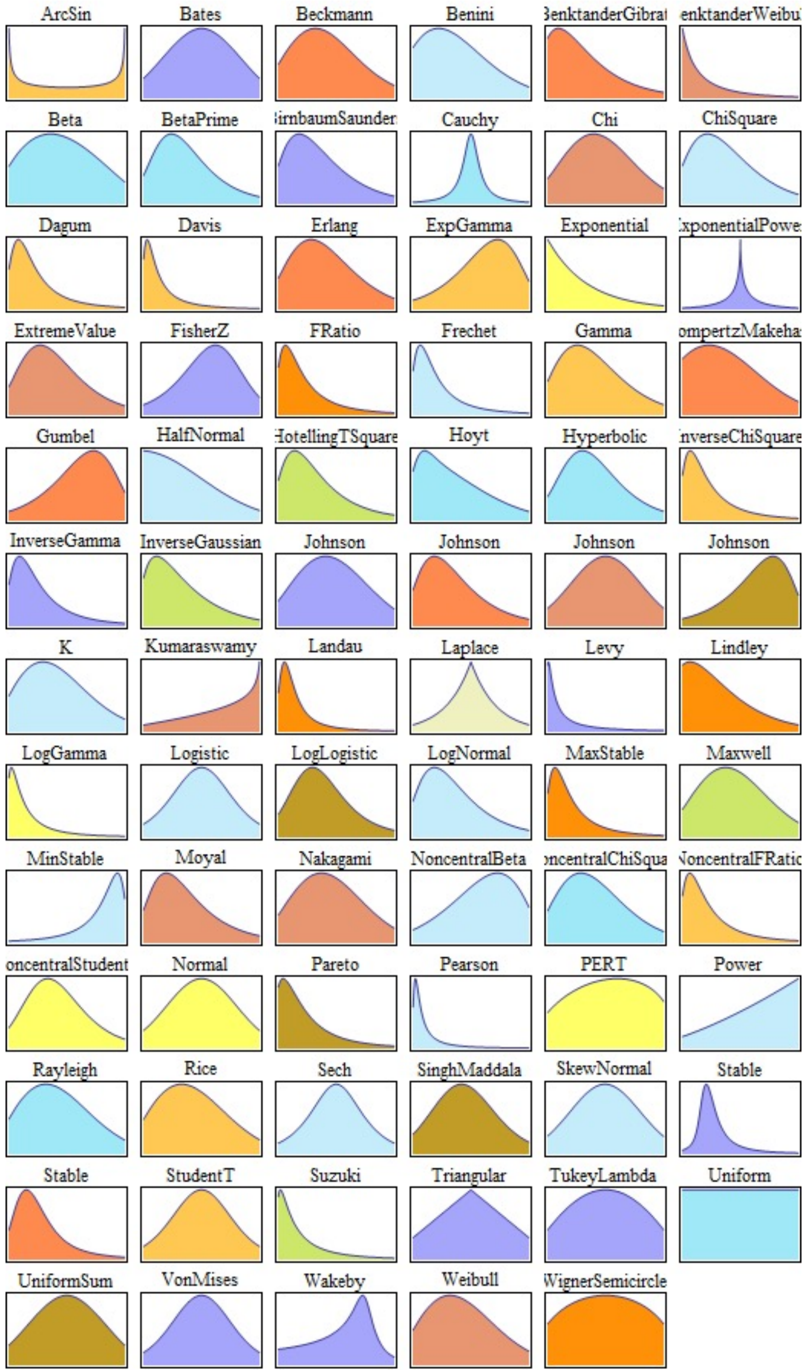


Ejercicio

El temario para un examen consta de 35 temas, de los cuales se elegirá uno al azar. Si un alumno no ha estudiado los 10 últimos temas ¿Cuál es la probabilidad de que el alumno sepa el tema elegido para el examen? Hallar la media y varianza.

Distribuciones Continuas

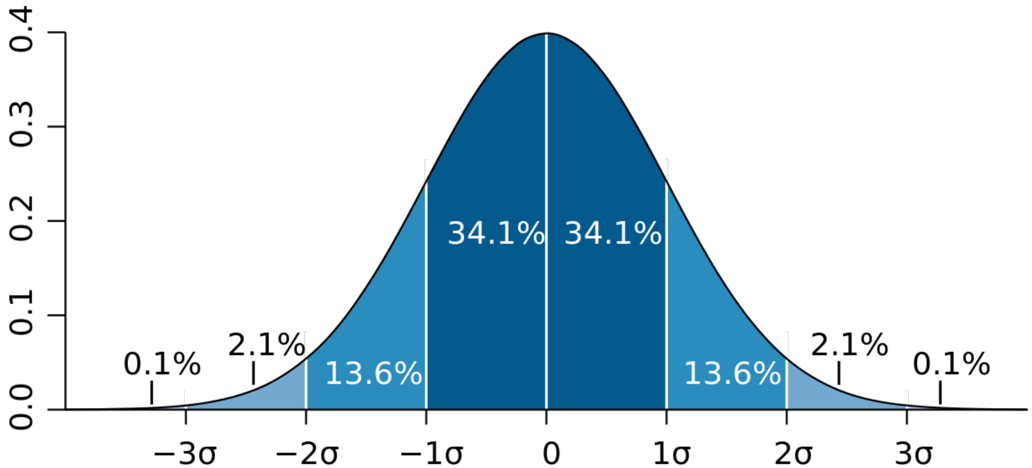
Para una variable continua hay infinitos valores posibles de la variable. En estas condiciones no es posible deducir la probabilidad de un valor puntual de la variable; como se puede hacer en el caso de variables discretas, pero es posible calcular la probabilidad acumulada hasta un cierto valor (función de distribución de probabilidad), y se puede analizar como cambia la probabilidad acumulada en cada punto (estos cambios no son probabilidades sino otro concepto, la función de densidad).



Distribución Normal

En estadística y probabilidad se llama distribución normal, distribución de Gauss, distribución gaussiana o distribución de Laplace-Gauss, a una de las distribuciones de probabilidad de variable continua que con más frecuencia aparece en estadística y en la teoría de probabilidades.

La gráfica de su función de densidad tiene una forma acampanada y es simétrica respecto de un determinado parámetro estadístico. Esta curva se conoce como campana de Gauss y es el gráfico de una función gaussiana.



Distribución Normal

Toma valores continuos que vienen desde -∞ hasta +∞ con un valor central que es igual para la Media, la Mediana y la Moda.

Los demás valores se alejan de igual forma a derecha e izquierda de la media (μ).

Conforme nos separamos de la μ la probabilidad va decreciendo dependiendo de las desviación estándar (σ).

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos.

Distribución Normal

Si X es una variable aleatoria normal, entonces su función de densidad de probabilidades está dada por:

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

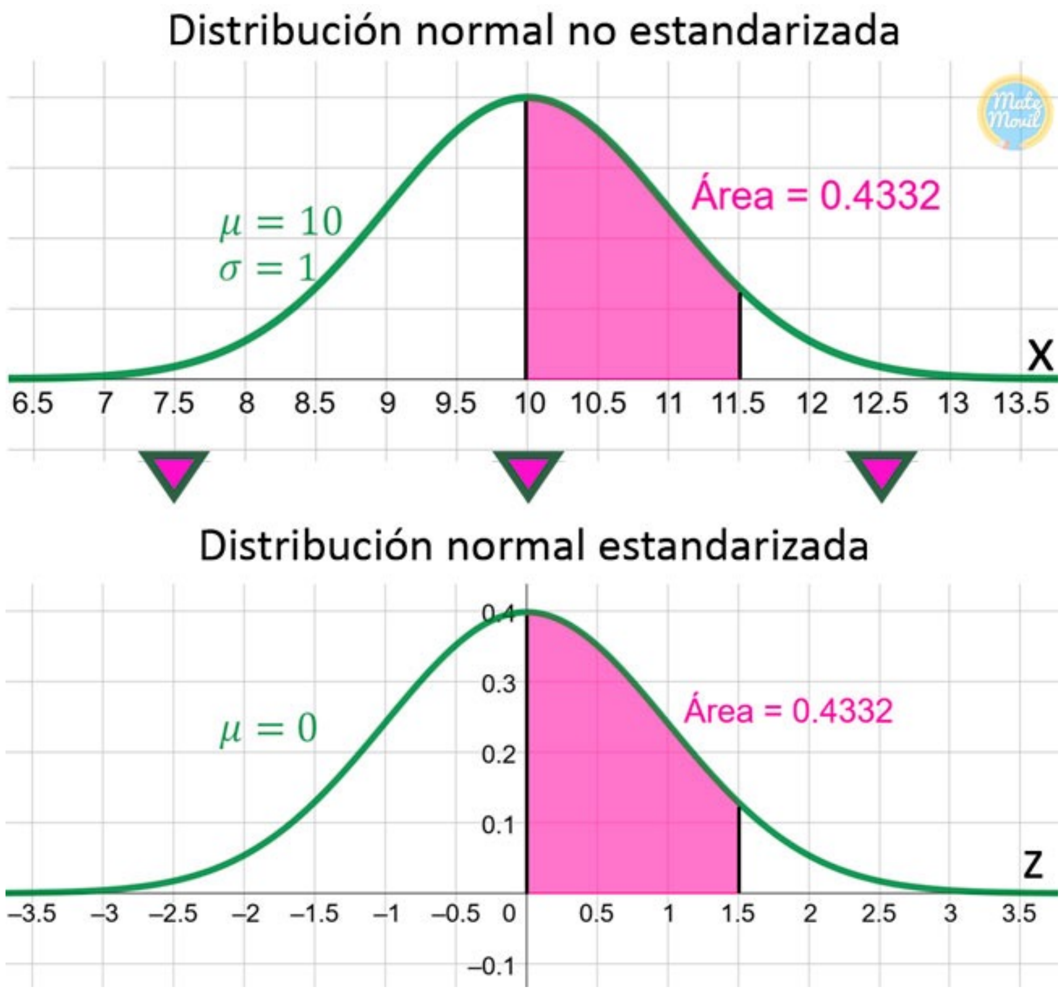
La forma de la campana depende de los parámetros μ y σ

Distribución Normal Estándar

A Z se le denomina función tipificada de x, y a su curva de función de densidad se le conoce como curva normal estándar.

Es una distribución normal con μ de 0 y σ de 1.

Todas las variables que se distribuyen normalmente se pueden transformar a la distribución normal estándar utilizando la fórmula de conversión para calcular el valor de Z.



Distribución Normal Estándar

Fórmula de cálculo de la Z para convertir una distribución normal con μ = x y una σ = x a una distribución normal estándar:

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

Ejercicio

El fabricante de una empresa laser informa que la cantidad media de páginas que imprime un cartucho antes de reemplazarlo es de 12200. La distribución de páginas impresas por cartucho se aproxima a la distribución de probabilidad normal y la desviación estándar es de 820 páginas. Calcule:

a) La probabilidad de que el cartucho logre imprimir al menos 14000 páginas.

b) La probabilidad de que el cartucho logre imprimir entre 12000 y 13000 páginas.

c)La probabilidad de que el cartucho logre imprimir 13500 páginas.

Tabla 2. (Continuación) Áreas bajo la curva normal

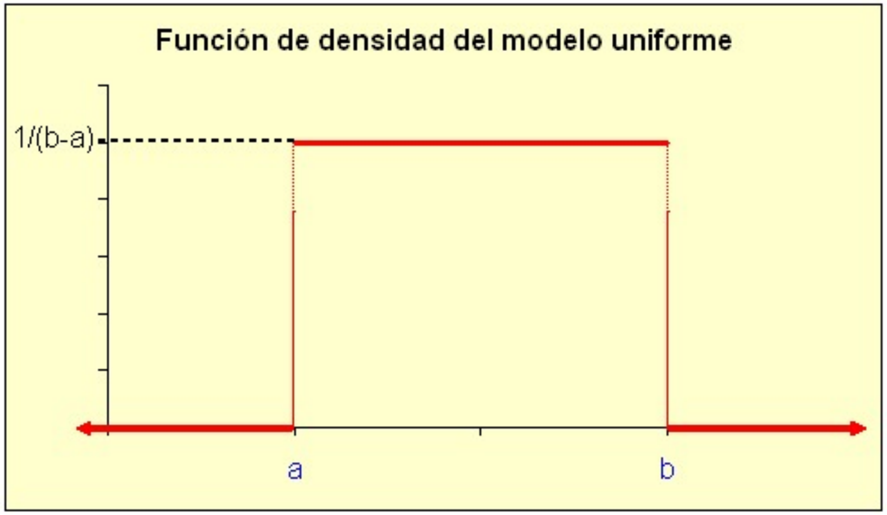
z	.00	.01	.02	.03	.04	.05	.06	.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6065	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	..9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9788	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9834	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9871	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9901	.9901	.9904	,9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9925	.9925	.9927	.9929	.9931	.9932	.9934	.9936

Tabla 2. Áreas bajo la curva normal

z	.00	.01	.02	.03	.04	.05	.06	.07	0.08	0.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	0.125	0.122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	0.375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Distribución Uniforme Continua

En teoría de probabilidad y estadística, la distribución uniforme continua es una familia de distribuciones de probabilidad para variables aleatorias continuas, tales que para cada miembro de la familia, todos los intervalos de igual longitud en la distribución en su rango son igualmente probables.



Distribución Uniforme Continua

La función de densidad debe tomar el mismo valor para todos los puntos dentro del intervalo (a,b).

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in (a,b) \\[2ex] 0 & \text{si } x \notin (a,b) \end{cases}$$

Distribución Uniforme Continua

La función de distribución se obtiene integrando la fución de densidad.

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x \leq a \\ \dfrac{x-a}{b-a} & \text{si } x \in (a,b) \\ 1 & \text{si } x \geq b \end{cases}$$

Ejercicio

En una cierta fábrica se producen cada día una media de 40 mil metros de cable. Unos días más, otros días menos. Aunque el mínimo seguro que siempre se fabrica es de 30 mil metros. La variable aleatoria que recoge el número de metros de cable fabricados en un día sigue una distribución uniforme continua.

a) ¿Cuál es el máximo de metros fabricados en un día?

b) ¿Qué porcentaje de días se fabrican más de 34 mil metros de cable?

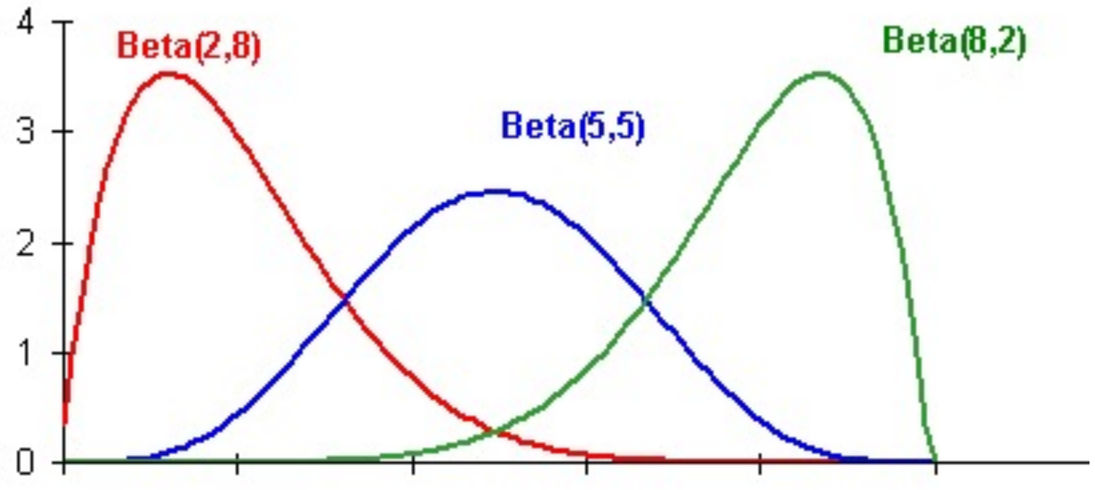
c) ¿Qué porcentaje de días se fabricarán menos de 33 mil metros de cable?

Distribución Beta

Es una extensión de la distribución uniforme. La distribución beta es una familia de distribuciones de probabilidad continuas definidas en el intervalo de 0 a 1, con dos parámetros positivos que determinan la forma típica notados como α y β.

Distribución Beta

Generalmente utilizada cuando no existen datos históricos sólidos en los cuales basar una estimación de actividades, se emplea para variables aleatorias continuas que no son negativas por lo que su gráfica está sesgada a la derecha.



Distribución Beta

Función de densidad, media y varianza:

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\mathbb{I}_{(0,1)}(x) \qquad \text{con } a, b > 0$$

$$\begin{aligned} E(X) &= \frac{a}{a+b} \\ var(X) &= \frac{ab}{(a+b)^2(a+b+1)} \end{aligned}$$

Ejercicio

La proporción de ítems defectuosos en una línea de producción sigue una Distribución Beta con parámetros $\boldsymbol{\alpha} = 1$ y $\boldsymbol{\beta} = 5$.

a) Determine la Media y Varianza.

b) ¿Cuál es la probabilidad de que 10% de los productos fabricados sean defectuosos?

Problemas de Teoría de Colas o de Confiabilidad

Se toma como variable aleatoria el tiempo que tarda en producirse un hecho, (tiempo que transcurre mientras llega un cliente, tiempo que transcurre mientras ocurre un fallo)

Distribución Gama

Si se esta interesado en la ocurrencia de un evento generado por un proceso de Poisson de media λ, la variable que mide el tiempo transcurrido hasta obtener n ocurrencias del evento sigue una distribución gamma.

La función de densidad es:

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{I}_{(0,+\infty)}(x) \qquad \text{con } \lambda > 0,\ \alpha > 0$$

Recordemos que el s ímbolo $\Gamma(\alpha)$ representa a la *función gama* que se define por

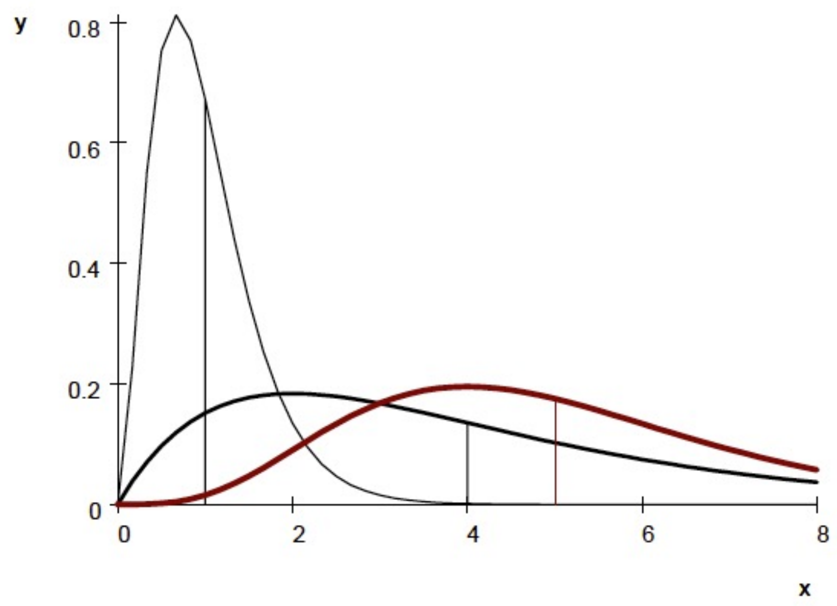
$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \qquad \text{si } y > 0$$

Distribución Gama

$$E(X) = \frac{\alpha}{\lambda}$$

$$var(X) = \frac{\alpha}{\lambda^2}$$

$$\varphi_X(t) = \frac{1}{\left(1 - \frac{it}{\lambda}\right)^{\alpha}}$$



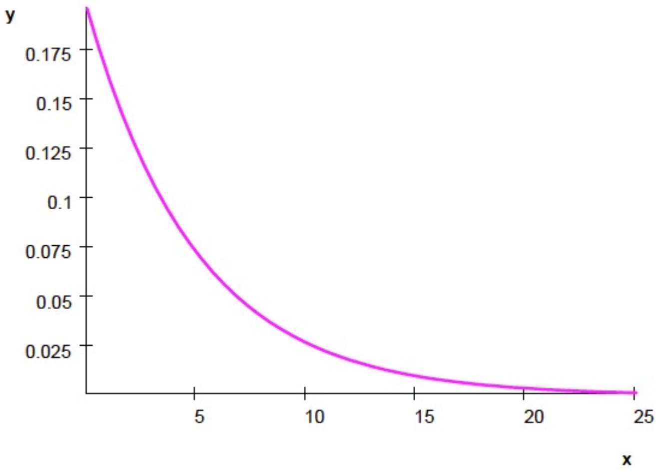
Ejercicio

El tiempo en horas que semanalmente requiere una máquina para mantenimiento es una variable aleatoria con distribución Gamma con parámetros $\boldsymbol{\alpha}=3$, $\boldsymbol{\beta}=2$.

Encuentre la probabilidad que en alguna semana el tiempo de mantenimiento sea mayor a 8 horas.

Distribución Exponencial

Es un caso partiular de la distribución Gama, y al igual que esta, tiene una gran utilidad práctica ya que podemos considerarla como un modelo adecuado para la distribución de probabilidad del tiempo de espera entre dos hechos que sigan un proceso de Poisson.



Densidad Exponencial con $\lambda = 1/5$

Distribución Exponencial

$$f_X(x) = \lambda e^{-\lambda x}\mathbb{I}_{(0,+\infty)}(x) \qquad \text{con } \lambda > 0$$

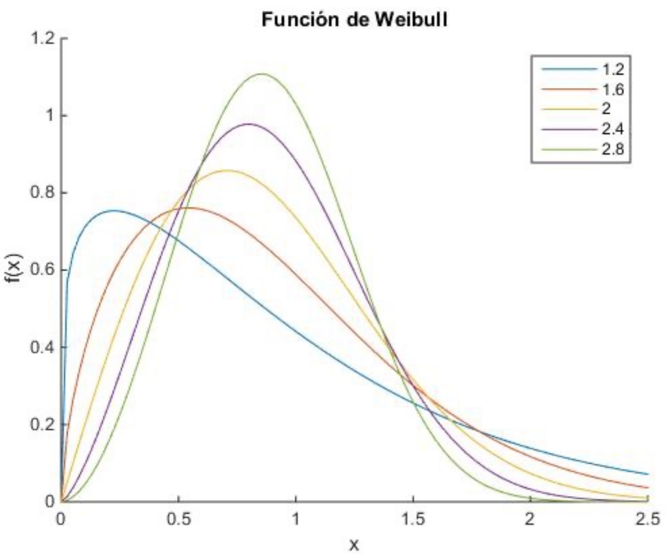
$$\begin{aligned} E(X) &= \frac{1}{\lambda} \\ var(X) &= \frac{1}{\lambda^2} \\ \varphi_X(t) &= \frac{\lambda}{\lambda - it} \end{aligned}$$

Ejercicio

Suponga que un ensamble contiene cierto componente cuyo tiempo de falla en años está dado por la variable aleatoria X, dsitribuida exponencialmente con tiempo promedio de falla $\beta = 5$. Si 5 de estos componentes se instalan en un producto, ¿cuál es la probabilidad de que uno continue funcionando después de 8 años?

Distribución Weibull

Es una distribución aplicable al estudio de la confiabilidad en problemas realtivos a la fatiga y vida de componentes y materiales. Esta distribución es utilizada con gran eficacia en los modelos de fallas, cuenta con dos parámetros, β llamado el parámetro de forma y α corresponde al parámetro de escala.



Distribución Weibull

El uso de esta función se debe principalmente a la gran diversidad de formas que este modelo puede tomar, esto nos permite usar un mismo modelo, independientemente de en qué forma cambie la tasa de fallos del componente estudiado simplificando la tarea del análisis de resultados.

Al aplicar Weibull se obtiene la distribución de fallos de conjunto de donde proviene la muestra, únicamente ajustando parámetros del modelo al conjunto de componentes ensayados.

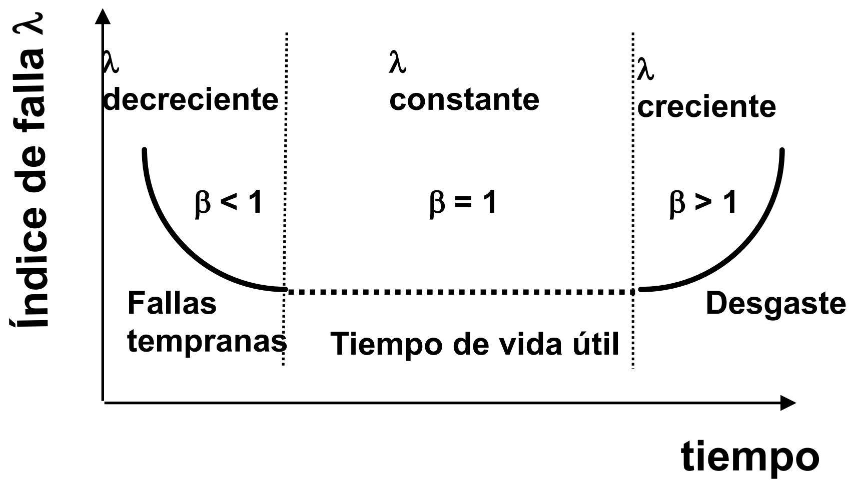
Distribución de Weibull

La distribución de Weibull es una de las más utilizadas en confiabilidad y estadística.

La versión de dos parámetros forma y escala (que representa la vida característica) no incluye el parámetro de localización, es cero.

La versión de tres parámetros tiene una parámetro de localización cuando hay un tiempo de falla diferente de cero para la primera falla

Distribución Weibull



Las tres porciones de la curva de tina de la bañera tienen diferentes índices de falla. Las fallas tempranas se caracterizan por un índice de falla decreciente, la vida útil por un índice de falla constante y el desgaste se caracteriza por un índice de falla creciente. La distribución de Weibull puede modelar matemáticamente estas tres situaciones.

β < 1 disminuye la tasa de riesgo, implica mortalidad infantil
β = 1 tasa de riesgo constante, fallas aleatorias
1< **β** < 4 aumenta la tasa de riesgo, fallas por corrosión, erosión
β > 4 aumenta rápidamente la tasa de riesgo, implica fallas por desgaste y envejecimiento

La Distribución Weibull - Interpretación

$\beta < 1$ (Tasa de riesgo decreciente)

- Implica mortalidad infantil
- Si esto ocurre, puede existir:
  - -Carga, inspección o prueba inadecuada
  - -Problemas de Manufactura
  - -Problemas de reparación
- Si un componente sobrevive la mortalidad infantil , la resistencia a fallar mejora con la edad.

$1 < \beta < 4$ (Tasa de Riesgo creciente)

- Si esto ocurre
  - -La mayoría de los baleros y engranes fallan
  - -Corrosión o Erosión
  - -El reemplazo programado puede ser efectivo en costo
  - $-\ \beta = 3.44 \Rightarrow$ aprox. Normal, $\beta = 2 \Rightarrow$ Rayleigh

$\beta = 1$ (Tasa de riesgo constante)

- Implica fallas aleatorias(Distribución Exponencial)
- Una parte vieja es tan buena como una nueva
- Si esto ocurre:
  - -Mezcla de modos de falla
  - -Las fallas pueden deberse a eventos externos, como:luminosidad o errores humanos
  - -Fundido y removido antes de su desgaste

$\beta > 4$ (La tasa de riesgo crece rápidamente)

- Implica edad avanzada y rápido desgaste
- Si esto ocurre, sospeche de:
  - -Propiedades del material
  - -Materiales frágiles como la cerámica
  - -Variabilidad pequeña en manufactura o material

Distribución Weibull

Funcion de densidad:

$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Distribución Weibull

Su media y varianza son

$$\mathrm{E}(X) = \lambda\Gamma\left(1 + \frac{1}{k}\right)$$

y

$$\mathrm{var}(X) = \lambda^2\left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right].$$

$\Gamma(k) = (k - 1)!$ (el factorial de $k - 1$).

Ejercicio

El tiempo T en segundos que tarda en conectarse a un servidor durante un día laborable sigue una distribución de Weibull de parámetros $\alpha = 0.6$ y $\beta = 1/4$.

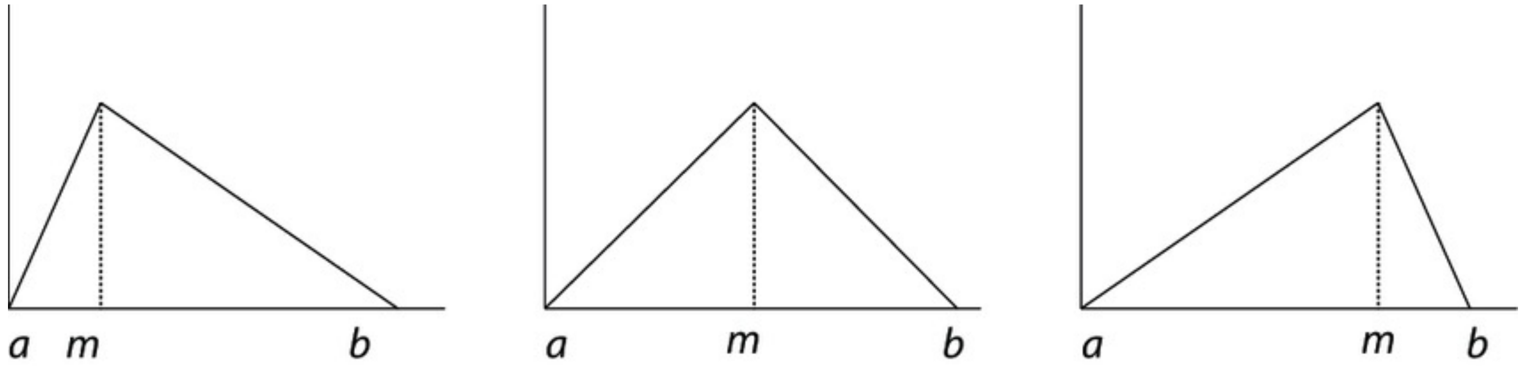
Se quiere saber:

a) La probabilidad de tardar más de 10 segundos en realizar la conexión.

b) Si se tienen ya 5 segundos esperando la conexión. ¿cuál es la probabilidad de que se demore 10 segundos más?

Distribución Triangular

En probabilidad y estadística, la distribución triangular es la distribución de probabilidad continua que tiene un valor mínimo a, un valor máximo b y una moda c, de modo que la función de densidad de probabilidad es cero para los extremos (a y b), y afín entre cada extremo y la moda, por lo que su gráfico es un triángulo.



Distribución Triangular

Se utiliza para describir una población sobre los cuales existen datos de muestra limitados, o para modelar procesos estocásticos o de riesgo.

Se utiliza sólo cuando el comportamiento de un proceso no se ajusta a otro tipo de distribución.

Distribución Triangular

La función densidad de probabilidad es

$$f(x|a,b,c) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{para } a \le x < c, \\ \frac{2}{b-a} & \text{para } x = c, \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{para } c < x \le b, \\ 0 & \text{para otros casos} \end{cases}$$

- Función de probabilidad acumulada: =
$$\begin{cases} 0 & \text{para } x \le a, \\ \frac{(x-a)^2}{(b-a)(c-a)} & \text{para } a < x \le c, \\ 1 - \frac{(b-x)^2}{(b-a)(b-c)} & \text{para } c < x < b, \\ 1 & \text{para } b \le x. \end{cases}$$
- Media: $\dfrac{a+b+c}{3}$
- Mediana: $\begin{cases} a + \frac{\sqrt{(b-a)(c-a)}}{\sqrt{2}} & \text{para } c \ge \frac{a+b}{2} \\ b - \frac{\sqrt{(b-a)(b-c)}}{\sqrt{2}} & \text{para } c \le \frac{a+b}{2} \end{cases}$
- Moda: c
- Varianza: $\dfrac{a^2+b^2+c^2-ab-ac-bc}{18}$

Ejercicio

Las horas laboradas sin paros por desperfectos han arrojado que en el pasado la máquina trabajó como mínimo 3000 horas, como máximo 7000 horas pero la mayoría del tiempo trabajó 5000 horas continuas antes del fallo. Calcule la probabilidad de:

a) Falle a las 6000 horas de uso.

b) Falle entre las 4500 y las 5500 horas de uso.

c)Trabaje más de 6500 horas sin parar.

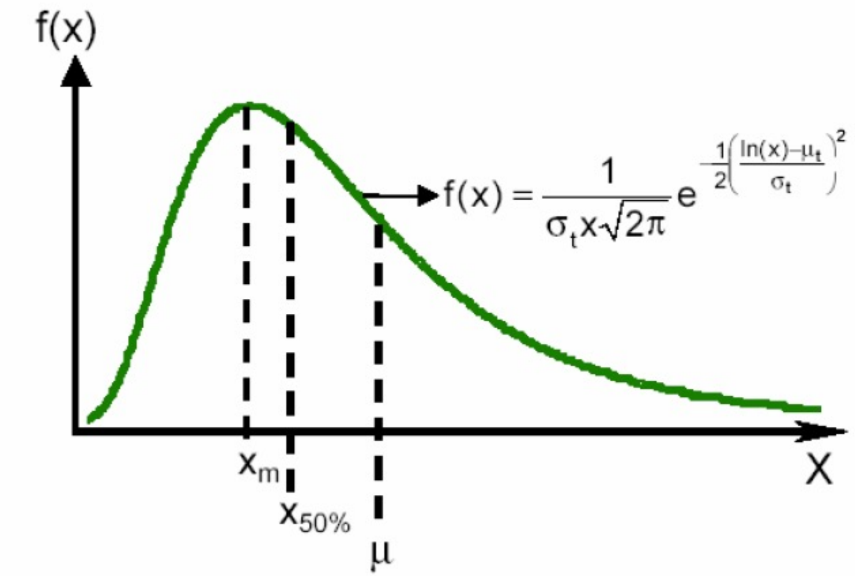
Distribución Log Normal

En probabilidades y estadísticas, la distribución normal logarítmica es una distribución de probabilidad de una variable aleatoria cuyo logaritmo está normalmente distribuido. Es decir, si X es una variable aleatoria con una distribución normal, entonces exp(X) tiene una distribución log-normal.

La base de una función logarítmica no es importante, ya que $\log_a X$ está distribuida normalmente si y solo si $\log_b X$ está distribuida normalmente, solo se diferencian en un factor constante.

Distribución Log Normal

Ampliamente utilizada para variables que ocurren cerca del valor mínimo, por ejemplo muchas variables físicas y procesos de deterioro pueden ser representados con esta distribución.



Distribución Log Normal

Distribución de Densidad de Probabilidad : f(x)

$$f(x) = \frac{1}{\sigma_t x \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu_t}{\sigma_t}\right)^2}$$

Distribución de Probabilidad Acumulada : F(x)

$$F(x) = \frac{1}{\sigma_t \sqrt{2\pi}} \int_0^{+\infty} \frac{1}{x} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu_t}{\sigma_t}\right)^2} dx$$

Distribución Log Normal

Parámetros:

Media Logarítmica: $$\mu_t = \frac{1}{n}\sum_{i=1}^{n} \ln(x_i)$$

Desviación Estándar Logarítmica: $$\sigma_t = \sqrt{\left(\left(\frac{1}{n}\right)\left(\sum_{i=1}^{n} (\ln(X_i) - \mu_t)^2\right)\right)}$$

Ejercicio

Unos cilindros de concreto tienen una resistencia a la compresión con distribución log-normal con β de 1 y α de 5.

Calcule la probabilidad de que la resistencia sea mayor a 300 Kg/cm^3

Distribuciones derivadas del Muestreo

A continuación se describen la distribución Chi-cuadrada, la T de Student y la F, las cuales son distribuciones que se derivan del muestreo porque algunos estadísticos muestrales siguen estas distribuciones, y esto permite hacer inferencias acerca de la población con base en muestras

Aplicaciones

Chi-cuadrado, T-Student y F se forman por combinaciones de variables aleatorias, por lo que no se emplean para modelar fenómenos físicos, pero sí para tomar decisiones y establecer intervalos de confianza.

Por lo que son muy utilizadas en diferentes pruebas de hipótesis.

Distribución de Chi Cuadrada

Sean Z1, Z2, …, Zn variables aleatorias independeintes distribuidas normalmente, cada una media de 0 y varianza de 1. La variable aleatoria es:

$$\chi^2 = Z_1^2 + Z_2^2 + \cdots + Z_k^2$$

Se dice que es una variable aleatoria chi cuadrado con k grados de libertad y su función de densidad está dada por:

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

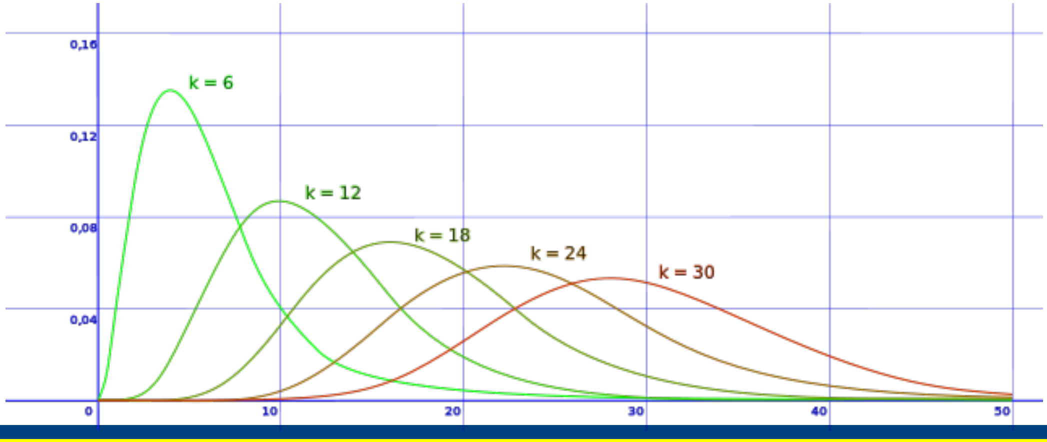
donde $\Gamma(\cdot)$ es la función gamma que está definida por

$$\Gamma\left(\alpha\right) = \int_0^\infty y^{\alpha-1} e^{-y} dy$$

Distribución de Chi Cuadrada

Tiene múltiples aplicaciones entre las cuales podemos mencionar tres de ellas:

1. Como prueba de bondad de ajuste
2. Para determinar si las poblaciones son o no homogéneas
3. Para determinar la dependencia e independencia de las variables a analizar.



Distribución t de student

En probabilidad y estadística, la distribución t (de Student) es una distribución de probabilidad que surge del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño.

Aparece de manera natural al realizar la prueba t de Student para la determinación de las diferencias entre dos varianzas muestrales y para la construcción del intervalo de confianza para la diferencia entre las partes de dos poblaciones cuando se desconoce la desviación típica de una población y esta debe ser estimada a partir de los datos de una muestra.

Distribución t de student

Sea Z una variable aleatoria con distribución normal estándar y sea V una variable aleatoria con distribución chi cuadrado con k grados de libertad. Si Z y V son independientes entonces la variable aleatoria:

$$T = \frac{Z}{\sqrt{V/k}}$$

tiene una distribución T con k grados de libertad, cuya función de densidad de probabilidad está dada por:

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

Distribución t de student

Una de las principales aplicaciones es fundamentar las inferencias sobre la media de una población. Si se obtiene una muestra aleatoria de tamaño n de una población cuya distribución es normal, el estadístico es:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

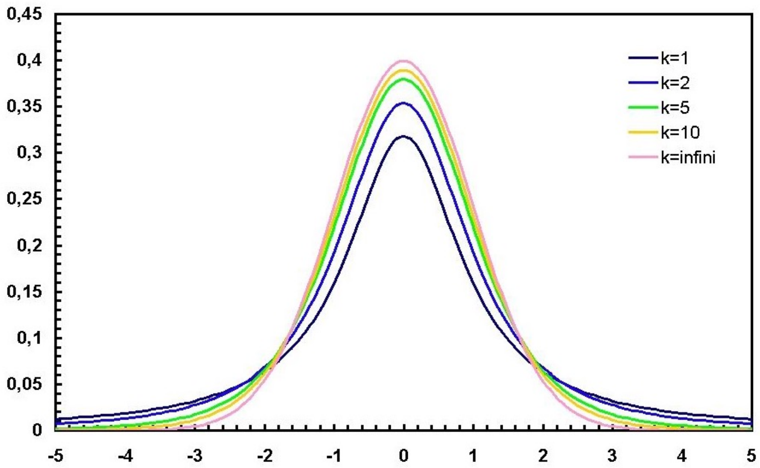
Sigue una distribución T de Student con n – 1 grados de libertad.

Distribución t de student

Es similar a la distribución normal estándar, excepto que tiene colas más pesadas.

Tiende a la normal estándar cuando el tamaño de la muestre es mayor o igual a 30.

Por eso su mayor aplicación es cuando las muestras son pequeñas, es decir menores a 30.



Distribución F

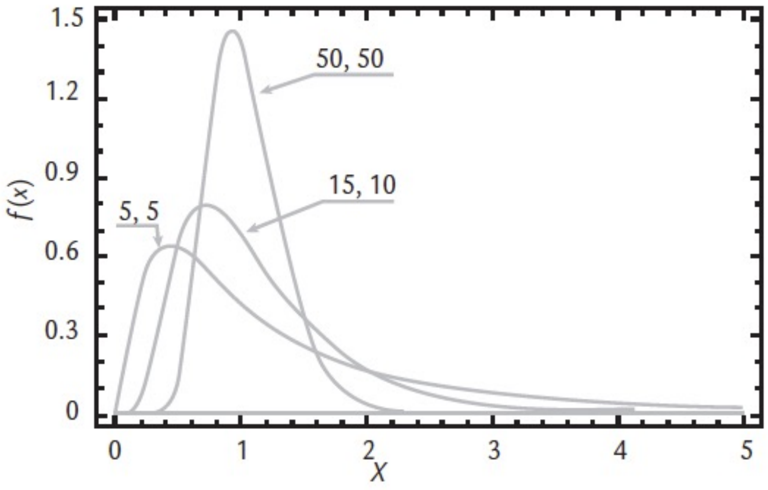
Una variable F se define como el cociente entre dos variables chi-cuadrado divididas por sus correspondientes grados de libertad.

Esta distribución de probabilidad se usa como estadística prueba en varias situaciones.

$$F = \frac{s^2{}_1}{s^2{}_2}$$

Distribución F

Una variable con distribución F es siempre positiva. La distribución de la variable es asimétrica, pero su asimetría disminuye cuando aumentan los grados de libertad del numerador y denominador. Hay una distribución F por cada par de grados de libertad.



Distribución F

Esta es la distribución de probabilidad de la razón de dos varianzas provenientes de dos poblaciones diferentes. Por medio de esta distribución es posible determinar la probabilidad de ocurrencia de una razón específica con $v_1=n_1-1$ y $v_2=n_2-1$ grados de libertad en muestras de tamaño n_1 y n_2.

Es la distribución más importante en experimentación pues permite hacer cálculos sobre varianzas diseminadas determinando si las diferencias mostradas son significativas y por lo tanto atribuibles a cambios importantes en el comportamiento de las poblaciones en estudio.

Distribución F

Función de densidad:

$$f(x) = \frac{\Gamma\left(\frac{u+v}{2}\right)\left(\frac{u}{v}\right)^{u/2} x^{(u/2)-1}}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{v}{2}\right)\left[\left(\frac{u}{v}\right)x+1\right]^{(u+v)/2}}, \quad 0 < x < \infty$$

Los parámetros son *u* grados de libertad en el numerador y *v* grados de libertad en el denominador.

Distribución F

La media y la varianza:

$$E(X) = v/(v-2) \text{ para } v > 2$$

$$\sigma^2 = \frac{2v^2(u+v-2)}{u(v-2)^2(v-4)}, \quad v > 4$$

Los parámetros son *u* grados de libertad en el numerador y *v* grados de libertad en el denominador.