DISTRIBUCIONES

1. DISTRIBUCIONES DISCRETAS

1.1DISTRIBUCIÓN UNIFORME

Una variable aleatoria discreta que toma todos sus valores $x_1, x_2, ..., x_k$ con igual probabilidad se dice que tiene una ***distribución uniforme discreta***.

La distribución de probabilidad de una ***variable aleatoria uniforme discreta*** está dada por

$$P(X = x) = \frac{1}{k} \text{ con } x = x_1, x_2, ..., x_k \quad \text{siendo } x_i \neq x_j \text{ para } i \neq j$$

Siendo k el número de elementos de la variable discreta.

La ***media*** y ***varianza*** de la distribución uniforme es

$$\mu = \frac{1}{k}\sum_{i=1}^{k} x_i \qquad \sigma^2 = \frac{1}{k}\sum_{i=1}^{k}(x_i - \mu)^2.$$

EJEMPLOS:

Para un dado de seis caras la probabilidad de obtener cualquier cara es 1/6.

Para una moneda todos los resultados {cara, cruz} tienen la misma probabilidad ½.

EJEMPLO 1:

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.

Solución:

Sea X la variable aleatoria que representa el número de tema seleccionado para el examen, como todos los temas tienen la misma probabilidad de ser seleccionado, X sigue una distribución uniforme discreta de 35 elementos.

- La probabilidad de salir el tema x es $P(X = x) = \frac{1}{35}$

El alumno aprueba el examen si le toca un tema del 1 al 25; así pues:

La probabilidad de salir un tema estudiado es:

$$P(X \leq 25) = \sum_{i=1}^{25} \frac{1}{35} = \frac{25}{35} \approx 0,7148$$

La media y varianza son los valores

$$\mu = \frac{1}{k}\sum_{i=1}^{k} x_i = \frac{1}{35}\sum_{i=1}^{35} i = \frac{35+1}{2} = 18$$

$$\sigma^2 = \frac{1}{k}\sum_{i=1}^{k} (x_i - \mu)^2 = \frac{1}{35}\sum_{i=1}^{35} (i-18)^2 = 102$$

1.2 DISTRIBUCIÓN BINOMIAL.

EJEMPLO 2:

Hallar la distribución de probabilidad del siguiente suceso: "número de caras que se obtienen en cinco lanzamientos sucesivos de una moneda".

Solución:

P(cara) = P(c) = ½ en un lanzamiento y P(cruz) = P(x) = ½ en un lanzamiento.

Los posibles valores de la variable son k = 0, 1, 2, 3, 4, 5.

La probabilidad de cada suceso:

xxxxx, que no salga cara $P(\xi = 0) = \binom{5}{0}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 = \frac{1}{32}$

xxxxc, que salga 1 vez $P(\xi = 1) = \binom{5}{1}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1 = \frac{5}{32}$

xxxcc que salga 2 veces $P(\xi = 2) = \binom{5}{2}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{10}{32}$

xxccc que salga 3 veces $P(\xi = 3) = \binom{5}{3}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 = \frac{10}{32}$

xcccc que salga 4 veces $P(\xi = 4) = \binom{5}{4}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 = \frac{5}{32}$

ccccc que salga 5 veces $P(\xi = 5) = \binom{5}{5}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5 = \frac{1}{32}$

x_i	0	1	2	3	4	5
$P(x_i)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Experimento de Bernoulli

Consideremos un experimento aleatorio que admite sólo dos resultados excluyentes:

Suceso A (éxito) con probabilidad P(A) = p.

Suceso $\overline{A}$ (fracaso) con probabilidad $P(\overline{A}) = 1-p = q$.

Supongamos que se realizan n pruebas de Bernoulli sucesivas e independientes, es decir, se realizan n experimentos independientes, en cada uno de los cuales un determinado suceso puede o no ocurrir.

A la variable aleatoria discreta ξ = "número de veces que ocurre el suceso A (éxito) en las n pruebas" se la denomina ***variable binomial***.

Para calcular la distribución de probabilidad nos fijamos en el suceso favorable a A en k veces, $A\overset{k)}{\ldots}A\overline{A}\overset{n-k)}{\ldots}\overline{A}$ cuya probabilidad por ser sucesos independientes es el producto de las probabilidades $p\overset{k)}{\ldots}p.q\overset{n-k)}{\ldots}q = p^k.q^{n-k}$. Resultado que se puede repetir $\binom{n}{k}$ veces, luego

$$P(\xi = k) = \binom{n}{k} p^k.(1-p)^{n-k}$$

Esta distribución discreta recibe el nombre de ***distribución binomial o de Bernouilli***.

La función de distribución correspondiente es: $F(x) = \sum_{k \le x} \binom{n}{k} p^k.(1-p)^{n-k}$

Si observamos el desarrollo de binomio de Newton:

$$(p+q)^n = \binom{n}{0} p^0.q^n + \binom{n}{1} p^1.q^{n-1} + \ldots + \binom{n}{n-1} p^{n-1}.q^1 + \binom{n}{n} p^n.q^0$$

entendemos el nombre de binomial y la simetría de la distribución de probabilidad. Los valores de $P(\xi = k)$ son laboriosos, por lo que se encuentran tabulados para valores de n de 0 a 10 y distintos valores de p. Una distribución binomial queda caracterizada cuando se conocen p y n y se escribe **B(n, p)**.

En el ejemplo anterior los parámetros son p = ½ y n = 5, por tanto se dice B(5,1/2).

La ***esperanza matemática***:

$$E[\xi] = \sum_{k=0}^{n} k\cdot P(\xi=k) = \sum_{k=0}^{n} k\cdot\binom{n}{k}p^k\cdot(1-p)^{n-k} = \sum_{k=1}^{n} k\cdot\binom{n}{k}p^k\cdot(1-p)^{n-k} =$$

$$= n\cdot p\cdot \sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}\cdot(1-p)^{n-k} = n\cdot p\cdot (p+1-p)^{n-1} = \mathbf{n\cdot p}$$

La ***varianza***:

$$E[\xi^2] = \sum_{k=0}^{n} k^2\cdot P(\xi=k) = \sum_{k=0}^{n} k^2\cdot\binom{n}{k}p^k\cdot(1-p)^{n-k} = \sum_{k=1}^{n} k^2\cdot\binom{n}{k}p^k\cdot(1-p)^{n-k} =$$

$$= n.p\cdot \sum_{k=1}^{n} k\cdot\binom{n-1}{k-1}p^{k-1}\cdot(1-p)^{n-k} =$$

$$= n\cdot p\cdot\left[\sum_{k=1}^{n}(k-1)\cdot\binom{n-1}{k-1}p^{k-1}\cdot(1-p)^{n-k} + \sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}\cdot(1-p)^{n-k}\right] =$$

$$= n\cdot p\cdot\left[(n-1)\cdot p\cdot(p+1-p)^{n-2} + (p+1-p)^{n-1}\right] = n.p.((n\text{-}1).p+1) = n^2\cdot p^2 - n\cdot p^2 + n\cdot p$$

$$V[\xi] = E[\xi^2] - \left(E[\xi]\right)^2 = n^2\cdot p^2 - n\cdot p^2 + n\cdot p - n^2\cdot p^2 = \mathbf{n\cdot p\cdot (1\text{-}p)}$$

Algunas distribuciones de probabilidad se ajustan a un modelo matemático. Conocido el modelo podemos escribir directamente las probabilidades de esas distribuciones. Si se observa que una variable estadística obtenida experimentalmente a partir de una muestra satisface las condiciones que conducen a una distribución Binomial, tendrá una distribución empírica que se aproximará a una distribución Binomial teórica.

El problema radica en determinar los parámetros de la Binomial que mejor se ajusta y el grado de dicho ajuste. Se puede demostrar que la que mejor se ajusta es la que tiene la misma media. Para ello se calcula la media muestral $\overline{x}$ y la hacemos coincidir con la de la población $\overline{\xi} = n.p = \overline{x}$, de donde $p = \frac{\overline{x}}{n}$ siendo n el número de elementos de la muestra. Por tanto, ajustaremos por una binomial $B\left(n, \frac{\overline{x}}{n}\right)$.

EJEMPLO 3:

a) ¿Cuál es la probabilidad de que una familia con seis hijos tenga dos hijos varones? b) ¿Cuál es la media de hijos varones de una familia de seis hijos? c) ¿Y la

desviación típica? d) ¿Cuál es la probabilidad de que una familia de seis hijos tenga como máximo cuatro hijos varones?

<u>Solución</u>:

Primeramente, observemos que tenemos una distribución binomial, cuya variable aleatoria es el número de hijos varones, B(6,1/2) siendo n=6=número de hijos, p=1/2= la probabilidad de que un hijo sea varón.

a) $P(\xi=2)=\binom{6}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4}=\frac{15}{64}$.

b) $E[\xi]=n.p=6.\frac{1}{2}=$ 3 hijos

c) $\sigma=\sqrt{V[\xi]}=\sqrt{n.p.q}=\sqrt{6.\frac{1}{2}.\frac{1}{2}}=\sqrt{\frac{3}{2}}$.

d) $F(4)=P(\xi\leq 4)=\sum_{k\leq 4}\binom{6}{k}p^{k}\cdot(1-p)^{6-k}=\sum_{i=0}^{4}\binom{6}{i}\left(\frac{1}{2}\right)^{i}\left(\frac{1}{2}\right)^{6-i}=\binom{6}{0}\left(\frac{1}{2}\right)^{6}+\binom{6}{1}\left(\frac{1}{2}\right)^{6}+\binom{6}{2}\left(\frac{1}{2}\right)^{6}+$

$+\binom{6}{3}\left(\frac{1}{2}\right)^{6}+\binom{6}{4}\left(\frac{1}{2}\right)^{6}=\left(\frac{1}{2}\right)^{6}\left[\binom{6}{0}+\binom{6}{1}+\binom{6}{2}+\binom{6}{3}+\binom{6}{4}\right]=\frac{57}{64}$

http://asignaturas.topografia.upm.es/matematicas/videos/Binomial.mp4

1.2. DISTRIBUCIÓN DE POISSON.

Es una distribución que se presenta cuando tenemos una población n grande y la probabilidad de que ocurra un suceso determinado, tiene una probabilidad muy pequeña (ley de casos raros).

Ejemplos:

- Llamadas a una central telefónica de emergencias.
- Desintegración de átomos radiactivos.
- Número de partos triples.
- Número de automóviles que llegan a un peaje.

Una variable aleatoria ξ tiene ***distribución de Poisson*** de parámetro λ, si toma los valores enteros 0, 1, ..., k, …. con probabilidad: $P(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ si $\lambda > 0$

Si en la distribución binomial, B(n,p), n es muy grande y p muy pequeña se puede aproximar mediante la distribución de Poisson, haciendo n.p = λ, suponiendo λ constante. Se observa que: $\lim_{n\to\infty} p = \lim_{n\to\infty} \frac{\lambda}{n} = 0$.

$$P(\xi = k) = \binom{n}{k} p^k . q^{n-k} = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k . \left(1 - \frac{\lambda}{n}\right)^{n-k} \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!}e^{-\lambda}$$

expresión de la probabilidad de que ocurra el suceso un número k de veces en un número grande de observaciones o experiencias.

Comprobemos que $\sum_{k=0}^{\infty} P(\xi = k) = 1$.

$$\sum_{k=0}^{\infty} P(\xi = k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda}.e^{\lambda} = 1$$

ya que el desarrollo en serie de e es: $e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + ... = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$

Por consiguiente, las probabilidades límites de las probabilidades binomiales cuando n tiende a infinito, siendo p variable y tal que n.p = λ = constante, constituyen la llamada ley de Poisson, **P(λ)**.

En la práctica la aproximación es aceptable para valores n>50 y p<0,1 ó n.p<5.

Existen tablas para calcular la probabilidad para distintos valores de k y λ.

Esperanza matemática:

$$E[\xi] = \sum_{k=0}^{\infty} k.P(\xi = k) = \sum_{k=0}^{\infty} k.\frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k.\frac{\lambda^k}{k!} e^{-\lambda} = \lambda.e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.e^{-\lambda}.e^{\lambda} = \lambda$$

Varianza:

$$E[\xi^2] = \sum_{k=0}^{\infty} k^2.P(\xi = k) = \sum_{k=0}^{\infty} k^2.\frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k^2.\frac{\lambda^k}{k!} e^{-\lambda} = \lambda.e^{-\lambda} \sum_{k=1}^{\infty} k.\frac{\lambda^{k-1}}{(k-1)!} =$$

$$= \lambda.e^{-\lambda}\left(\sum_{k=1}^{\infty}(k-1)\frac{\lambda^{k-1}}{(k-1)!}+\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) = \lambda.e^{-\lambda}\left(\lambda\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) = \lambda.e^{-\lambda}\lambda.e^{\lambda}+$$

$\lambda.e^{-\lambda}.e^{\lambda} = \lambda^2+\lambda$; luego $V[\xi]=E\left[\xi^2\right]-\left(E[\xi]\right)^2 = \lambda^2+\lambda-\lambda^2 = \lambda$

Para el ajuste de una distribución empírica por una distribución de Poisson, si la variable estadística experimental tiene la característica de una distribución de Poisson, se aproximará por una de este tipo y el parámetro λ que mejor se ajusta es $\lambda=\overline{x}$, ya que en este caso ambas distribuciones tienen la misma media.

EJEMPLO 4:

El número medio de llamadas telefónicas que llegan a una central de emergencias es de cuatro por minuto en las horas punta. Calcular: a) Probabilidad de que lleguen cuatro llamadas en un minuto. b) Número de llamadas que deberá ser capaz de soportar la central para que no se sature, con probabilidad 0,99.

Solución:

a) Sea la variable aleatoria ξ="número de llamadas que se reciben en un minuto (de las horas punta)", entonces ξ sigue una distribución de Poisson de parámetro $\lambda=4$. Nos piden: $P(\xi=4)=\frac{4^4}{4!}e^{-4} = 0,1954$.

b) Llamando x al número de llamadas que deberá soportar la centralita, nos dicen que: $P(\xi \le x)\ge 0,99$

Vamos sumando las probabilidades: $P(\xi=0)+P(\xi=1)+P(\xi=2)+\ldots$ hasta que la suma sea mayor o igual que 0,99.

Observando estas probabilidades en la tabla de la distribución P(4) tenemos: 0,0183+0,0733+0,1465+0,1954+0,1954+0,1563+0,1042+0,0595+0,0298 = 0,9798.

$P(\xi \le 9)=P(\xi \le 8)+P(\xi=9)$ = 0,9787 + 0,0132 = 0,9919 > 0,99.

Por tanto, x = 9.

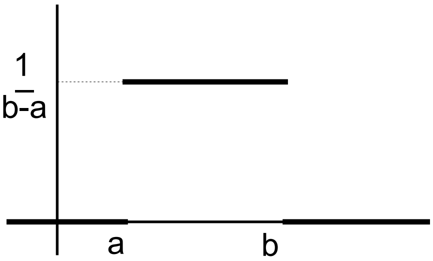
http://asignaturas.topografia.upm.es/matematicas/videos/Poisson.mp4

2. DISTRIBUCIONES CONTINUAS

2.1. DISTRIBUCIÓN UNIFORME O RECTANGULAR

Se dice que una variable continua sigue una ***distribución uniforme*** en el intervalo $[a,b]$ cuando su función de densidad es:

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



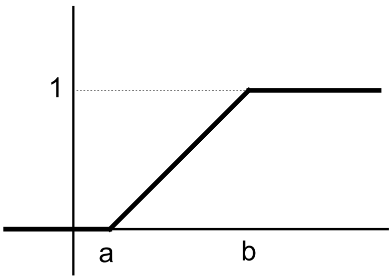
Gráficamente:

La función de distribución, será:

$$F(x) = \int_a^x f(x).dx = \int_a^x \frac{1}{b-a}dx = \left.\frac{x}{b-a}\right]_a^x = \frac{x-a}{b-a} \quad \text{si}$$

$a\leq x\leq b$, entonces:

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



cuya gráfica es:

obsérvese como la función de distribución es continua, por corresponder a una variable aleatoria continua, pero no lo es la función de densidad.

Esperanza Matemática:

$$E[\xi]=\int_{-\infty}^{\infty} x.f(x).dx=\int_a^b x.\frac{1}{b-a}.dx=\left.\frac{1}{b-a}\frac{x^2}{2}\right]_a^b=\frac{1}{b-a}\frac{b^2-a^2}{2}=\frac{a+b}{2}$$

Varianza:

$$E\left[\xi^2\right]=\int_{-\infty}^{\infty} x^2\cdot f(x)\cdot dx=\int_a^b x^2\frac{1}{b-a}.dx=\left.\frac{1}{b-a}\frac{x^3}{3}\right]_a^b=\frac{1}{b-a}\frac{b^3-a^3}{3}=\frac{a^2+ab+b^2}{3}$$

y, por tanto,

$$V[\xi]=E[\xi^2]-(E[\xi])^2=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{a^2+ab+b^2}{3}-\frac{a^2+2a.b+b^2}{4}=$$

$$=\frac{a^2-2ab+b^2}{12}=\frac{(b-a)^2}{12}$$

Esta distribución depende de los parámetros a y b y se denota $U[a,b]$.

EJEMPLO 5:

De una estación parte un tren cada 10 minutos. Si designamos por X la variable aleatoria "tiempo de espera desde que un viajero entra en el andén hasta que parte", calcular:

a) La probabilidad de esperar menos de dos minutos.

b) El tiempo medio de espera.

c) La probabilidad de esperar dos minutos y medio.

Solución:

Si el viajero llega en cualquier momento y todos los momentos tienen la misma probabilidad, la variable aleatoria tiempo de espera sigue una distribución U(0,10).

Por tanto, la función de densidad de la variable X es

$$f(x)=\begin{cases} 0 & \text{si } x<0 \\ \frac{1}{10} & \text{si } 0\le x\le 10 \\ 0 & \text{si } 10<x \end{cases}$$

a) $P(X<2)=\int_0^2 \frac{1}{10}dx=\frac{1}{5}$

b) $\mu=E[X]=\int_{-\infty}^{\infty} xf(x)dx=\int_0^{10} x\frac{1}{10}dx=5$

c) Si el reloj aprecia hasta los segundos, el viajero espera dos minutos y medio si el tren llega entre 2,5-1/60 y 2,5+1/60 minutos, por tanto, $\int_{2,5-\frac{1}{60}}^{2,5+\frac{1}{60}} \frac{1}{10}dx=\frac{1}{300}$

2.2. DISTRIBUCIÓN NORMAL

Abraham de Moivre (1667-1754) presento en un artículo de año 1733 la función de densidad de la mal llamada "*campana de Gauss*". **Gauss** (1777-1855) y **Laplace** (1749-1827) llegaron a ella estudiando la distribución de los errores de observaciones. Es la distribución que aparece con más frecuencia en la práctica, hasta el punto que el nombre lo tiene porque se pensaba en un principio que todas las distribuciones estadísticas eran de este tipo, ya que era una ley de la naturaleza. Esta situación fue parafraseada por Lippmam "*todos creen en la ley normal de errores: los experimentadores, porque ellos piensan que puede ser probado por las matemáticas; los matemáticos, porque creen que se ha establecido por la observación*".

Hoy es evidente la exageración de tal creencia, aunque debe señalarse la extraordinaria frecuencia con que tal distribución se encuentra en las más variadas aplicaciones y también, como otras muchas distribuciones, pueden reducirse a ésta mediante un adecuado cambio de variable.

Ejemplos:

- Talla, peso y otras características físicas de las personas.
- Test psicológicos y sociológicos.
- Distribución de la renta en países desarrollados.
- Consumo de electricidad de un país.

Una variable aleatoria continua X se dice que tiene una ***distribución normal*** o de ***Laplace-Gauss*** de media μ y desviación típica σ, si: $\boxed{f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}}$ es su función de densidad.

Se verifica: $\int_{-\infty}^{\infty} f(x)\cdot dx = 1$

En efecto:

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

Cambio de variable

$$\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}=z \Rightarrow x-\mu=\sigma\sqrt{2}\sqrt{z} \Rightarrow dx=\sigma\sqrt{2}\frac{1}{2\sqrt{z}}dz \Rightarrow \begin{cases} x=\mu \Rightarrow z=0 \\ x\to\infty \Rightarrow z\to\infty \end{cases}$$

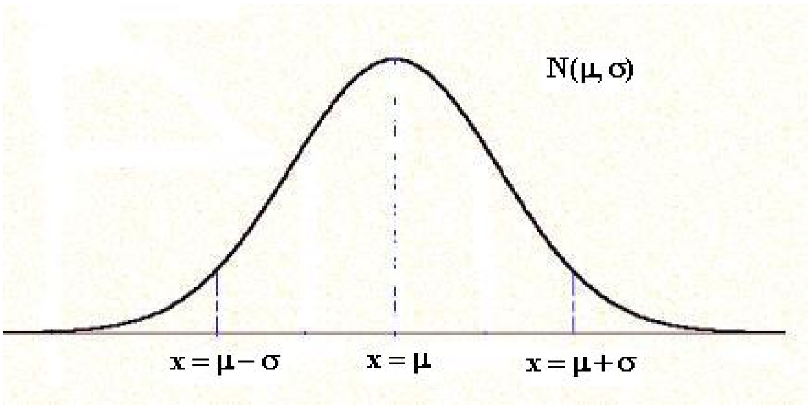
$$(*)=\frac{2}{\sigma\sqrt{2\pi}}\int_0^{\infty} e^{-z}\sigma\sqrt{2}\frac{1}{2\sqrt{z}}dz=\frac{1}{\sqrt{\pi}}\int_0^{\infty} e^{-z}z^{-1/2}dz=\frac{1}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right)=\frac{1}{\sqrt{\pi}}\sqrt{\pi}=1$$

Veamos el estudio de la gráfica de f(x):

1. $\lim\limits_{x\to\pm\infty} f(x)=0$
2. f(x)>0 para todo x finito.
3. f(x) es continua para todo x.
4. Es simétrica respecto de x= μ .
5. $f'(x)=-\frac{1}{2}2\frac{(x-\mu)}{\sigma^2}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} = -\frac{(x-\mu)}{\sigma^2}.f(x) = \begin{cases} \geq 0 \text{ si } x\leq\mu \\ \leq 0 \text{ si } x\geq\mu \end{cases}$ puesto que f(x)>0.

En x = μ pasa de creciente a decreciente, luego existe un máximo en (μ ,f(μ)) = $(\mu,\frac{1}{\sigma\sqrt{2\pi}})$ donde coinciden la moda, (por ser máximo), la mediana, (por simetría) y la media.

6. $f''(x)=-\frac{1}{\sigma^2}\left[1-\left(\frac{x-\mu}{\sigma}\right)^2\right].f(x)$ = 0 entonces $\left[1-\left(\frac{x-\mu}{\sigma}\right)^2\right]$= 0, resultando $x=\mu\pm\sigma$ puntos de inflexión.



es la llamada "***campana de Gauss***".

La función de distribución es: $\boxed{F(x)=P(\xi\leq x)=\int_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}.dx}$

La ***esperanza matemática*** es:

$$E[X]=\int_{-\infty}^{\infty} x\cdot f(x)dx=\int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}dx$$

y realizando un cambio de variable $\frac{x-\mu}{\sigma}=t\Rightarrow\frac{dx}{\sigma}=dt\Leftrightarrow x=\sigma t+\mu$; resulta,

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

la integral primera es inmediata y su resultado cero, por simetría; y en cuanto al la segunda es una integral impropia, que se identifica con la función gamma:

$\int_{-\infty}^{\infty}e^{-\frac{1}{2}t^2}dt=2\int_{0}^{\infty}e^{-\frac{1}{2}t^2}dt=2\frac{\sqrt{2}}{2}\Gamma\left(\frac{1}{2}\right)=\sqrt{2}\sqrt{\pi}=\sqrt{2\pi}$; sustituyendo, los dos resultados, se obtiene $E[X]=\frac{1}{\sqrt{2\pi}}(\sigma.0+\mu\sqrt{2\pi})=$ μ y la ***varianza*** es:

$V[X]=\int_{-\infty}^{\infty}(x-\mu)^2.f(x)dx=\int_{-\infty}^{\infty}(x-\mu)^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}.dx=\int_{-\infty}^{\infty}\sigma^2t^2\frac{1}{\sigma\sqrt{2\pi}}\sigma e^{-\frac{1}{2}t^2}dt$ que realizando por partes, se llega a

$$V[X]=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2e^{-\frac{1}{2}t^2}dt=\frac{\sigma^2}{\sqrt{2\pi}}\left(\left[-te^{-\frac{1}{2}t^2}\right]_{-\infty}^{\infty}+\int_{-\infty}^{\infty}e^{-\frac{1}{2}t^2}dt\right)=\frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi}=\sigma^2.$$

Por consiguiente, la ***media*** μ y la ***desviación típica*** σ son los dos parámetros de la distribución $N(\mu,\sigma)$.

Si queremos calcular una probabilidad determinada:

$$P(a\le X\le b)=\frac{1}{\sigma\sqrt{2\pi}}\int_{a}^{b}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.dx=F(b)-F(a)$$

debemos hacer uso de métodos de integración aproximada o bien de tablas, pero sólo existen tablas de la N(0,1), es decir, de la distribución normal de media $\mu=0$ y desviación típica $\sigma=1$.

Distribución normal tipificada o estándar

Una transformación de la variable $X \equiv N(\mu, \sigma)$ a la variable $Z \equiv N(0,1)$ recibe el nombre de ***distribución normal tipificada***.

Hacemos $Z = \dfrac{X-\mu}{\sigma}$ que es el resultado de una traslación por un cambio de escala, cuya **media**:

$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{E[X]-\mu}{\sigma} = \frac{\mu-\mu}{\sigma} = 0$$

y su **varianza**:

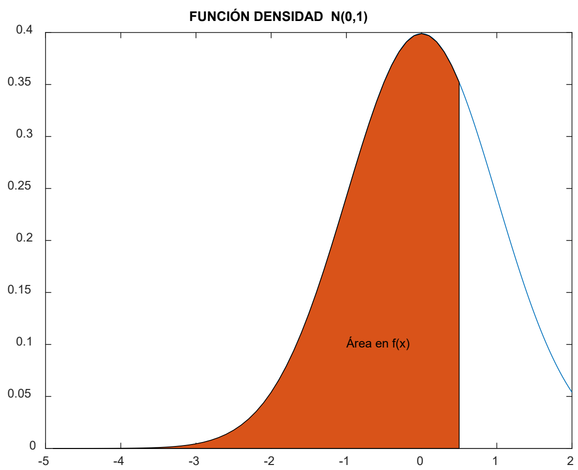
$$V[Z] = V\left[\frac{X-\mu}{\sigma}\right] = \frac{V[X]}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1$$

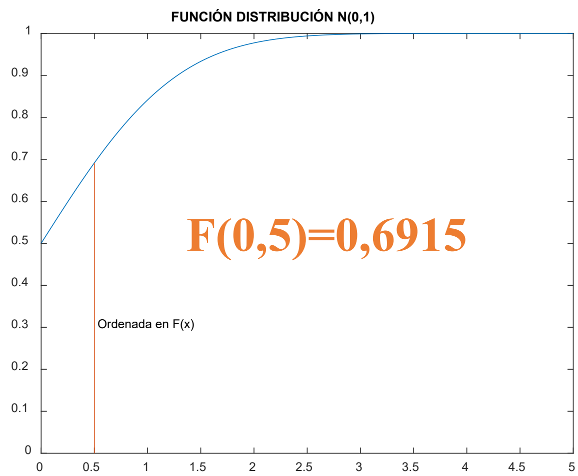
Efectivamente, la $N(0,1)$ con función de densidad: $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ y cuyos coeficientes de simetría y curtosis son nulos, dada la perfecta simetría de la distribución.

Cálculo de Probabilidades:

Podemos obtener el cálculo de la probabilidad a la izquierda x de un valor de la abscisa dado con la función de distribución:

$$P(X \le x) = F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$



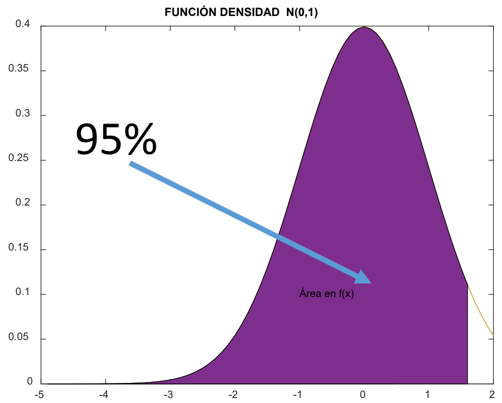


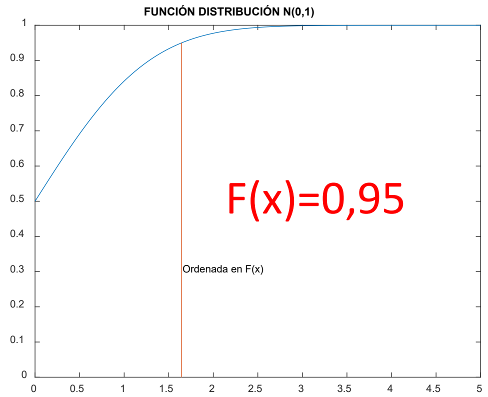
En particular: $P(Z \le \mathbf{0,5}) = F(\mathbf{0,5}) = \int_{-\infty}^{0,5} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \approx \mathbf{0,6914624612}$.

Cálculo de Percentiles:

Ahora, el problema inverso: dada la probabilidad α obtener el percentil P_α correspondiente:

$$F(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz=\alpha\Rightarrow ¿x?$$



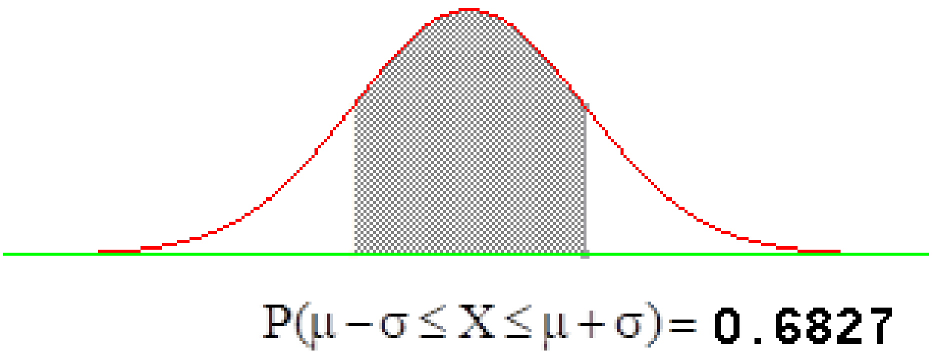


En particular: $0,95=F(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz\Rightarrow x\approx 1,6449$.

Es interesante observar la probabilidad del intervalo centrado en la media de valor cero:

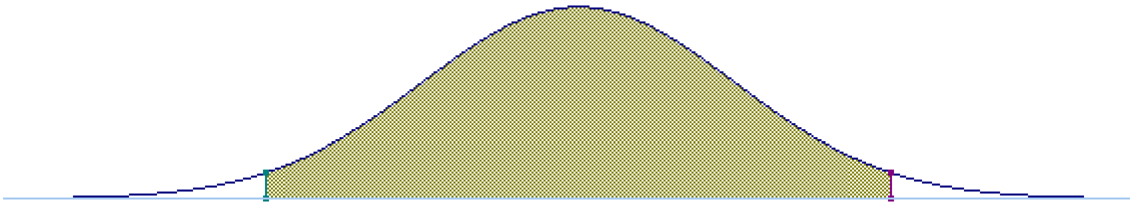
$$P(-a\le Z\le a)=F(a)-F(-a)=F(a)-(1-F(a))=2.F(a)-1$$

$$P(-1\le Z\le 1)=F(1)-F(-1)=\ \ 0,68$$



lo que equivale para una normal $N(\mu,\sigma)$: $P(\mu-\sigma\le X\le\mu+\sigma)\approx\ 0,68$

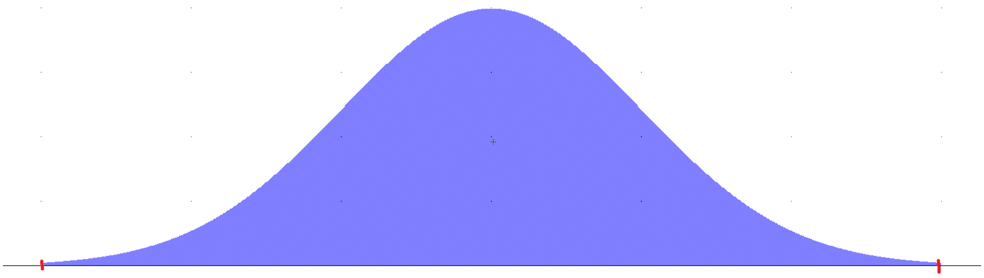
$$P(-2 \le Z \le 2) = F(2) - F(-2) = 0{,}95$$



$$P(\mu - 2\sigma \le X \le \mu + 2\sigma) = 0.9545$$

lo que equivale para una normal $N(\mu,\sigma)$: $P(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 0{,}95$

$$P(-3 \le Z \le 3) = F(3) - F(-3) = 0{,}99$$



$$P(\mu - 3\sigma \le X \le \mu + 3\sigma) = 0{,}9968$$

lo que equivale para una normal $N(\mu,\sigma)$: $P(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 0{,}99$

que nos indica la concentración de valores, en torno a la media, de tal forma que el 99% de la distribución está situado entre -3σ y $+3\sigma$ entorno a la media μ.

Paso al límite de la distribución binomial a la normal

Al estudiar la distribución límite de la binomial cuando n tiende a ∞ y p es fijo, se observa que $\binom{n}{k} p^k q^{n-k} \to 0$. Además, la media n.p y la varianza n.p.(1-p) tienden a ∞. Se trata, pues, de hacer un cambio de variable sencillo que nos haga tender la nueva variable, cuando n tiende a ∞, hacia una variable que tenga momentos finitos.

$$Z = \frac{\xi - n.p}{\sqrt{n.p.(1-p)}} \equiv N(0,1) \text{ siendo } \xi \equiv B(n,p)$$

En la práctica, es aceptable la aproximación si $0{,}1 < p < 0{,}9$ y $n > 50$. También si $p < 0{,}1$ y $n.p > 5$.

Ajuste de una distribución empírica mediante una distribución normal

El procedimiento de determinación de los parámetros consiste en igualar la media $\overline{x}$ y la varianza S^2 de la distribución estadística con la media μ y la varianza σ^2 de la normal. Es decir: $\mu = \overline{x},\ \sigma^2 = S^2$.

EJEMPLO 6:

Sabiendo que el número de alumnos que pasan cada año al tercer curso de una escuela sigue una distribución normal N(120,10), calcular: a) Probabilidad de que pasen más de 140. b) Probabilidad de que pasen entre 100 y 130. c) ¿Qué número de plazas debe haber disponibles en tercero para que se puedan matricular todos los que lo deseen con probabilidad 0,98?

Solución:

Sea $X =$ "número de alumnos que pasan a tercero" $\equiv N(120,10)$

a) $P(X > 140) = 1 - P(X \le 140) = 1 - F(140) \approx 0,0228$

b) $P(100 \le X \le 130) = P(X \le 130) - P(X < 100) = F(130) - F(100) \approx 0,8185$

c) $x =$ número de plazas

En este caso, conviene tipificar para poder calcular el percentil 98:

$$P(X \le x) = 0,98 \Leftrightarrow P\left(\frac{X-120}{10} \le \frac{x-120}{10}\right) = 0,98$$

$\Leftrightarrow P\left(Z \le \frac{x-120}{10}\right) = F_{N(0,1)}\left(\frac{x-120}{10}\right) = 0,98$, le corresponde al valor 2,05, luego $\frac{x-120}{10} = 2,05$ de donde x=140,5, luego el número de plazas disponibles debe ser de 141.

EJEMPLO 7:

El 10% de los jóvenes tallados para realizar el servicio militar, miden más de 1,80 m; el 30% mide menos de 1,65m. Suponiendo que la altura de las personas sigue una distribución normal, calcular: a) Altura media y desviación típica de la población. b) Porcentaje de jóvenes que mide 1,70 m.

Solución:

$$X = \text{“altura de los jóvenes”} \equiv N(\mu, \sigma)$$

a) P(X >1,80) = 0,1, puesto que son el 10% y P(X<1,65) = 0,3, ya que representan el 30%.

Tipificando,

$$P\left(\frac{X-\mu}{\sigma} > \frac{1,80-\mu}{\sigma}\right) = 0,1 \Leftrightarrow F\left(\frac{1,80-\mu}{\sigma}\right) = P\left(Z \le \frac{1,80-\mu}{\sigma}\right) = 1 - P\left(Z > \frac{1,80-\mu}{\sigma}\right) = 1 - 0,1 = 0,9$$

y observando el percentil 90, se tiene que: $\frac{1,80-\mu}{\sigma} = 1,282$ de donde

$$\boxed{\mu = 1,80 - 1,282.\sigma} \qquad (1)$$

Por otra parte, P(X <1,65) = 0,3

$$P\left(\frac{X-\mu}{\sigma} < \frac{1,65-\mu}{\sigma}\right) = 0,3 \Leftrightarrow P\left(Z < \frac{1,65-\mu}{\sigma}\right) = F\left(\frac{1,65-\mu}{\sigma}\right) = 0,3$$

Y observando el percentil 30: $\frac{1,65-\mu}{\sigma} = -0,525$ de donde $\boxed{\mu = 1,65 + 0,525.\sigma}$ (2)

Igualando (1) y (2) $\mu = 1,80 - 1,282.\sigma = 1,65 + 0,525.\sigma$ resulta $\sigma = 0,083$ y sustituyendo en (2) $\mu = 1,693$.

En resumen: la altura media es 1,693 m y la desviación típica 0,083.

b) Nos piden $P(X = 1,70)$, por ser la distribución continua, la probabilidad en un punto es cero. Ahora bien, es obvio, que no podemos negar la existencia de jóvenes que midan 1,70 m, entonces, ¿cómo podemos saber cuántos son?

Pues, dando un intervalo que contenga a 1.70 y sea lo más pequeño posible, por ejemplo, (1.695,1.705). Evidentemente, éstos no miden 1.70, pero con los instrumentos de medida habituales que suelen tener de precisión 1 cm, decimos que una persona mide 1.70, cuando su altura está en el intervalo (1.695,1.705).

La probabilidad de este intervalo es:

$$P\left(1,695 < X < 1,705\right) = P(X \le 1,705) - P(X \le 1,695) =$$
$$= 0.55747808 - 0.50961213 = 0.04786595$$

lo que quiere decir que el **4,786%** aproximadamente de los jóvenes miden entre 1,695 y 1,705.

http://asignaturas.topografia.upm.es/matematicas/videos/Normal.mp4

2.3. ADICIÓN DE DISTRIBUCIONES NORMALES

Sean $X_1 \equiv N(\mu_1, \sigma_1)$ y $X_2 \equiv N(\mu_2, \sigma_2)$ independientes. Entonces la variable $X = X_1 + X_2$ corresponde a una normal de media $\mu = \mu_1 + \mu_2$ y desviación típica $\sigma = \sqrt{\sigma_1{}^2 + \sigma_2{}^2}$, es decir, $X = X_1 + X_2 \equiv N(\mu_1 + \mu_2, \sqrt{\sigma_1{}^2 + \sigma_2{}^2})$.

Se puede generalizar a $X_i \equiv N(\mu_i, \sigma_i)$, n variables independientes, resultando:

$$\sum X_i \equiv N\left(\sum \mu_i, \sqrt{\sum \sigma_i^2}\right)$$

EJEMPLO 8:

El peso de un limón está distribuido según una distribución N(125, 10) en gramos. ¿Cuál es la probabilidad de que una caja con 50 limones pese menos de 6 kg? Si llenamos bolsas de ocho limones, ¿cuál es la probabilidad de que las bolsas pesen entre 975 y 1025 gramos?

Solución:

X="peso de 50 limones"$\equiv N(\mu, \sigma)$

El peso de cada limón sigue una distribución Normal, luego 50 limones se corresponde con la suma de las 50 distribuciones normales, resultando:

$$N\left(\sum \mu_i, \sqrt{\sum \sigma_i^2}\right) = N\left(50 \cdot 125, \sqrt{50 \cdot 10^2}\right) = N\left(6250, 50\sqrt{2}\right)$$
$$P\left(X < 6000\right) = F(6000) \approx \boxed{0.000203476}$$

Ahora será: X="peso de 8 limones" $\equiv N(\mu, \sigma)$

Para bolsas de 8 limones $N\left(\sum \mu_i, \sqrt{\sum \sigma_i^2}\right) = N\left(8 \cdot 125, \sqrt{8 \cdot 10^2}\right) = N\left(1000, 20\sqrt{2}\right)$

$$P\left(975 < X < 1025\right) = F(1025) - F(975) \approx \boxed{0.6232408821}$$

Nota: $X_1 - X_2 \equiv N(\mu_1 - \mu_2, \sqrt{\sigma_1{}^2 + \sigma_2{}^2})$

Ejemplo: $\mathbf{N(0,1) - N(0,1) = N(0, \sqrt{1^2 + 1^2}) = N(0, \sqrt{2})}$

2.4. TEOREMA CENTRAL DEL LÍMITE.

Si una variable es suma de n variables independientes, con medias μ_i y varianzas σ_i^2, su distribución tenderá a ser normal, con media igual a la suma de medias y varianza igual a la suma de las varianzas, cuando n, número de sumandos, tienda a infinito. Es decir, la variable se aproximará a $N\left(\sum_i \mu_i, \sqrt{\sum_i \sigma_i^2}\right)$.

El teorema central del límite da una justificación teórica del hecho de que en la naturaleza se encuentren con frecuencia variables aleatorias con distribución aproximadamente normal, ya que dichas variables pueden ser consideradas frecuentemente como suma de gran número de variables, todas ellas de poca influencia sobre la conducta de la variable considerada, pero que contribuyen, aunque individualmente sea con poca intensidad, a la cuantificación de sus valores y merced al teorema central del límite, estas variables podrán también ser tratadas por aproximación, como normales.

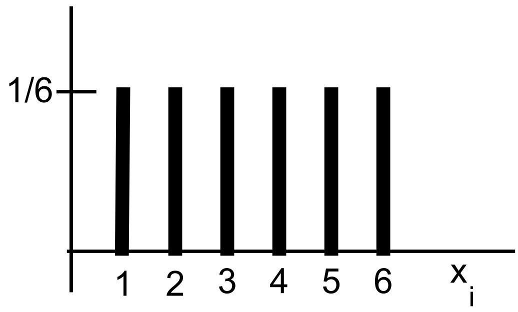
Demostrado en 1922 por Lindeberg, se puede comprobar empíricamente simulando distribuciones de probabilidad con ordenador. Se puede usar un generador de números aleatorios para obtener n valores, $x_1, x_2, ..., x_n$ de una U(0,1). Según sabemos las x_i tienen de media ½ y varianza 1/12, luego $S_n = \sum_{i=1}^{n} x_i$ tiene de media n/2 y varianza n/12. La variable tipifica $S_n^* = \dfrac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$ tiene por el teorema central de límite, una distribución aproximadamente normal tipificada para n grande. Si bien la distribución límite de S_n^* no depende de las distribuciones de las variables x_i, la "rapidez" con la que S_n^* se aproxima a la normal depende en gran medida de la forma de dichas distribuciones. Sin embargo, si las distribuciones de las x_i son muy asimétricas, n tiene que ser a veces muy grande para obtener una aproximación satisfactoria.

El teorema central de límite, garantiza la aproximación normal de la distribución binomial.

EJEMPLO

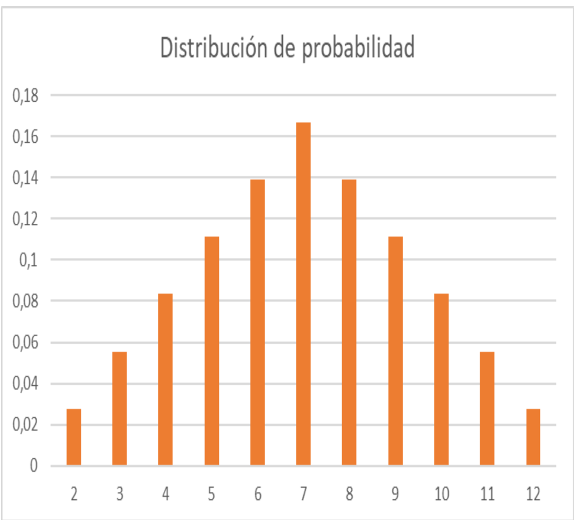
Si consideramos el lanzamiento de un dado y la variable aleatoria X="puntos obtenidos". **E** = {1, 2, 3, 4, 5, 6}. Se trata de la distribución uniforme U(1,6).

x_i	$\xi(x_i) = x_i$	$P(\xi = x_i)$
1	1	1/6
2	2	1/6
3	3	1/6
4	4	1/6
5	5	1/6
6	6	1/6

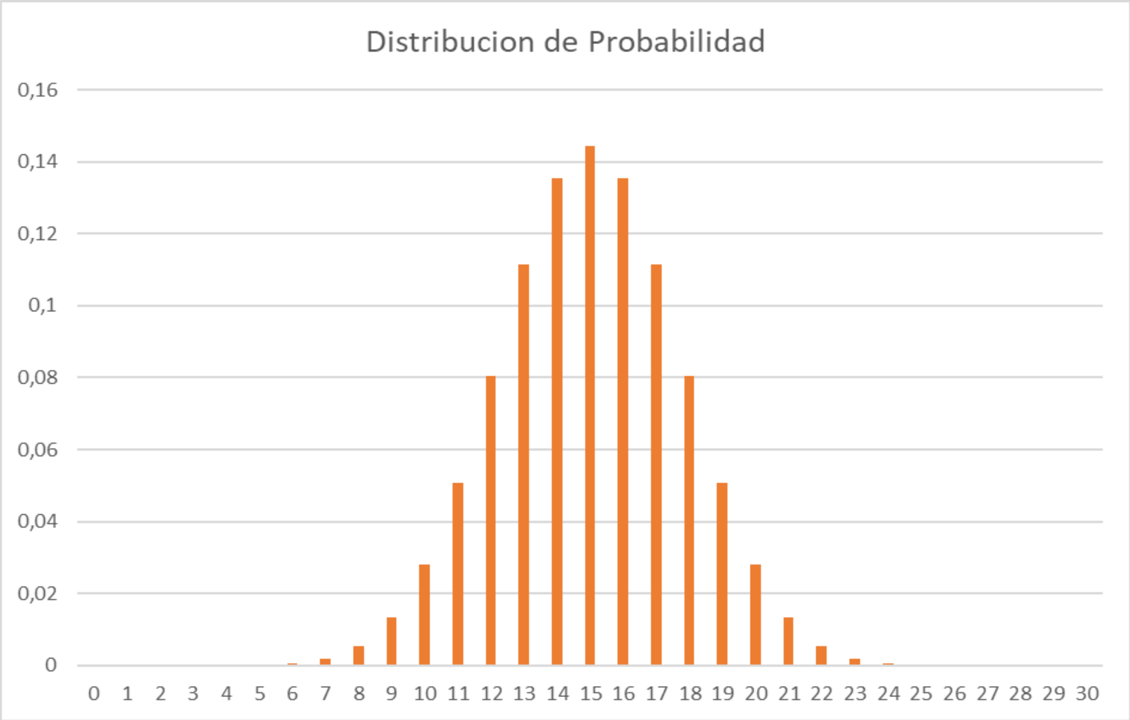


Si consideramos el lanzamiento de dos dados y la variable aleatoria $X=X_1+X_2$="puntos obtenidos". **E** = {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

x_i	$\xi(x_i) = x_i$	$P(\xi = x_i)$
2	2	1/36
3	3	2/36
4	4	3/36
5	5	4/36
6	6	5/36
7	7	6/36
8	8	5/36
9	9	4/36
10	10	3/36
11	11	2/36
12	12	1/36

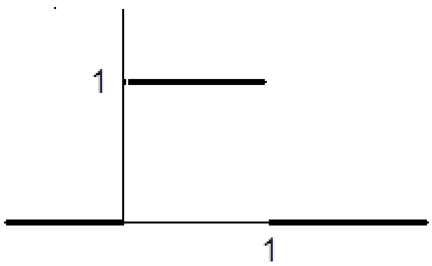


Ahora para el caso del lanzamiento de 30 monedas y observar el número de caras Binomial de parámetros n=30, p=1/2:

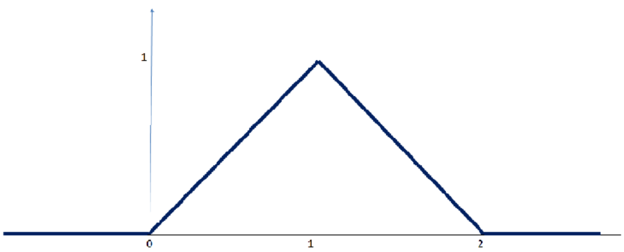


Para el caso continuo, consideramos dos distribuciones uniformes U(0,1) y a continuación, la suma:

$$U(0,1) \Leftrightarrow f(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 1 & \text{si} \quad 0 \le x \le 1 \\ 0 & \text{si} \quad 1 < x \end{cases}$$



$$U(0,1) + U(0,1) \Rightarrow f(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ x & \text{si} \quad 0 \le x < 1 \\ 2 - x & \text{si} \quad 1 \le x < 2 \\ 0 & \text{si} \quad 2 \le x \end{cases}$$



2.5 DISTRIBUCIONES DERIVADAS DE LA NORMAL.

2.5.1. Distribución χ^2 de Pearson

Sean $Z_1, Z_2, ..., Z_n$, n variables aleatorias N(0,1) e independientes, entonces la expresión: $\chi_n^2 = Z_1^2 + Z_2^2 + ... + Z_n^2$ es una variable aleatoria que recibe el nombre de ***ji-dos o chi-cuadrado de Pearson*** (1857,1936). El número de variables normales sumadas recibe el nombre de grados de libertad, y la representamos por χ_n^2.

Función de densidad:

La función de densidad de la variable aleatoria χ_n^2 es:

$$f(x) = \begin{cases} \dfrac{x^{(\frac{n}{2}-1)}}{2^{n/2}\Gamma(n/2)} e^{-x/2} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Algunas propiedades de la función de densidad de χ_n^2:

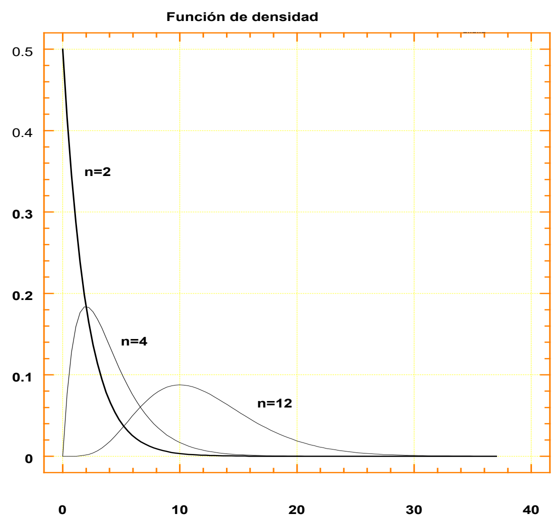
1.- El dominio de χ_n^2 es $R^+ \cup \{0\}$ $\forall n \in N$.

2.- f(x) presenta un único máximo si $n \geq 2$.

3.- f(x) no es simétrica.

4.- f(x) depende del número de grados de libertad n.

Algunas gráficas:



Función de distribución:

La función de distribución de una variable aleatoria χ_n^2 viene dada por:

$$F(x) = P(\chi_n^2 \leq x) = \int_0^x f(x)dx \ .$$

Características de la variable aleatoria χ_n^2.

La ***media, varianza*** y ***desviación típica*** de una variable aleatoria ji-cuadrado con n grados de libertad es $\mu = n$, $\sigma^2 = 2n$, $\sigma = \sqrt{2n}$ respectivamente.

Propiedades de la variable aleatoria χ_n^2 .

1.- Si χ_1^2, χ_2^2 ,..........,χ_m^2 son m variables aleatorias ji-cuadrado con n_1, n_2,...n_m grados de libertad respectivamente, entonces, la variable aleatoria $\chi_1^2 + \chi_2^2 + \ldots\ldots + \chi_m^2$ es una variable aleatoria χ_n^2 con $n=n_1+n_2+..+n_m$ grados de libertad.

2.- Si el número de grados de libertad n es mayor que 30 entonces, la variable aleatoria $(2\chi_n^2)^{1/2}$ sigue una distribución $\mathbf{N(\sqrt{(2n-1)},1)}$. Esta propiedad es útil en caso de usar tablas de χ_n^2.

La distribución de Pearson es de gran aplicación en la Teoría de Muestras o Inferencia Estadística para conocer si una determinada variable real observada empíricamente a través de una muestra se ajusta en mayor o menor grado a una distribución teórica.

Es habitual hallar el área encerrada por f(x) en $\left[\chi_{1-\alpha}, \infty\right)$ esto es, $P(\chi^2 \leq \chi_{1-\alpha}) = 1-\alpha$. Al valor α se suele denominar ***nivel de significación***. Y el reciproco obtener el percentil α, es decir la abscisa χ_α.

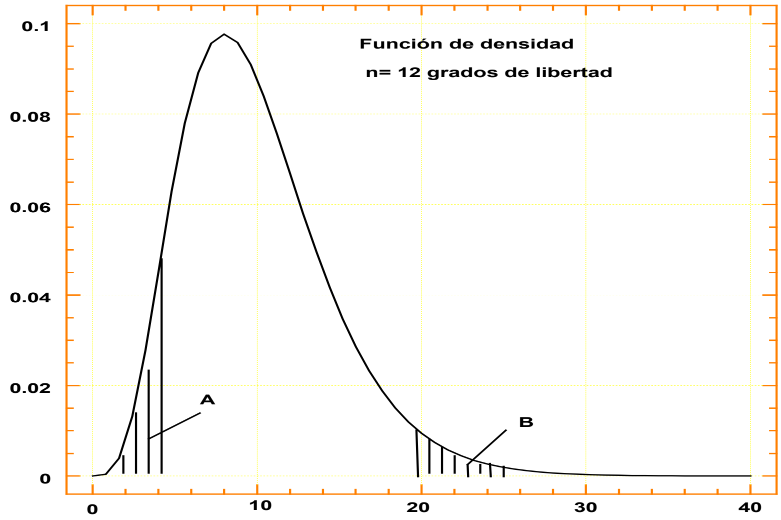
http://asignaturas.topografia.upm.es/matematicas/videos/Pearsom.mp4

EJEMPLO 9:

Dada una distribución χ^2_{12} y siendo A y B las áreas rayadas del gráfico, calcular los valores χ_α tales que:

a) El área A sea igual a 0.05.

b) El área B sea igual a 0.025.



Solución:

a) $P(\chi^2_{n=12} \le \chi_\alpha) = 0,05 \Rightarrow$ se obtiene $\chi_\alpha = 5.23$.

b) $P(\chi^2_{n=12} \le \chi_\alpha) = 1 - 0,025 = 0,975 \Rightarrow \chi_\alpha = 23,30$.

EJEMPLO 10:

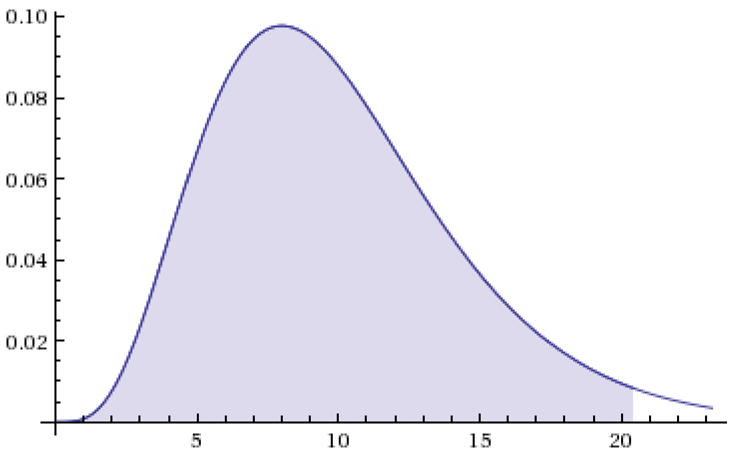
Calcular

a) $P(\chi^2_{n=10} \le 20.4)$

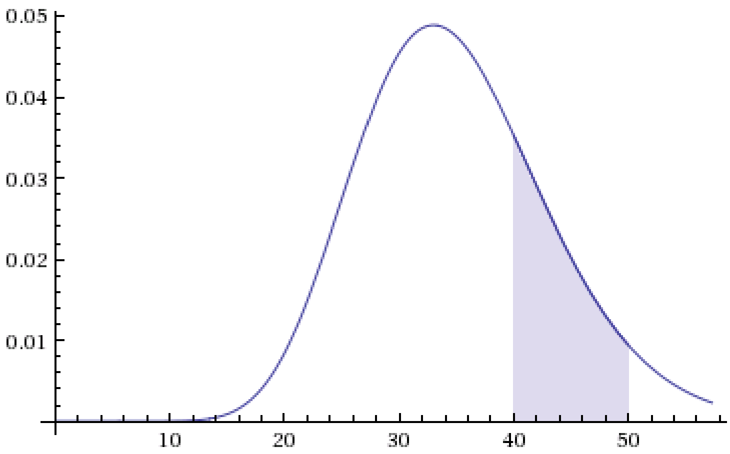
b) $P(35,56 < \chi^2_{n=26} < 37)$

Solución:

a) $P(\chi^2_{n=10} \le 20,4) = F(20,4) \approx 0,974312$



b) $P(40 < \chi^2_{n=35} < 50) = P(\chi^2_{n=35} < 50) - P(\chi^2_{n=35} \le 40) = F(50) - F(40) \approx 0,209684$



2.5.2. Distribución t de Student

Sean $X_1, X_2, ..., X_n$ y X, n+1 variables aleatorias independientes entre sí con distribución $N(0, \sigma)$ cada una de ellas. Entonces la variable $t_n = \dfrac{X}{\sqrt{\dfrac{X_1^2 + ... + X_n^2}{n}}}$ se denomina ***t de Student*** con n grados de libertad.

Student fue el seudónimo de William Sealy GOSSET (1876,1937), el estadístico y químico que descubrió la forma de la distribución t mediante una combinación de trabajo matemático y trabajo empírico con números aleatorios, una aplicación temprana de lo que ahora se llama el método de MonteCarlo.

Si tipificamos $X_1, X_2, ..., X_n$ y X, dividiendo por σ se obtienen las variables $Z_1, Z_2, ..., Z_n$ y Z que son N(0,1), obteniendo: $t_n = \dfrac{Z}{\sqrt{\dfrac{Z_1^2 + ... + Z_n^2}{n}}} = \dfrac{Z}{\sqrt{\dfrac{\chi_n^2}{n}}}$, puesto que $Z_1^2 + ... + Z_n^2 = \chi_n^2$.

Esto significa que:

1.- Se podía haber definido la variable aleatoria $\mathbf{t_n}$, suponiendo las $X_1, X_2, ..., X_n$ y X, variables N(0,1).

2.- Que la $\mathbf{t_n}$ de Student se caracteriza por eliminar la influencia de la varianza lo que tiene una gran importancia en la Inferencia Estadística.

Función de densidad:

$$f(x)=\frac{1}{\sqrt{n\pi}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

Algunas propiedades de la función de densidad de t_n:

1.- El dominio es R.

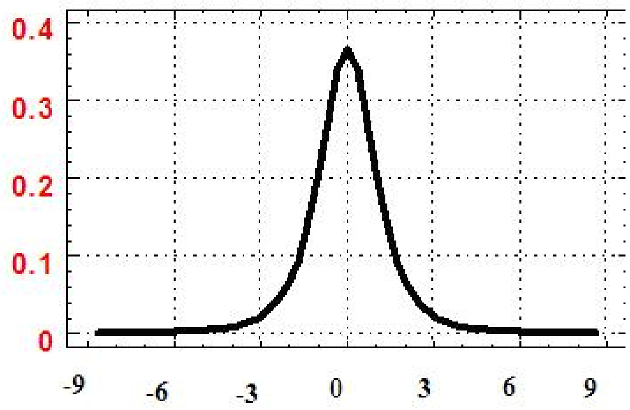
2.- f(x) presenta un único máximo si $\mathbf{n \geq 1}$.

3.- f(x) es simétrica.

4.- f(x) depende del número de grados de libertad (n).

5.- Al aumentar n se va haciendo cada vez más apuntada la función de densidad siendo el límite $(n \to \infty)$ la curva normal tipificada.

Veamos a continuación la gráfica de la función de densidad para n=3 grados de libertad: $\mathbf{f(x)=\frac{1}{\sqrt{n\pi}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}=\frac{1}{\sqrt{3\pi}}\frac{\Gamma(2)}{\Gamma\left(\frac{3}{2}\right)}\left(1+\frac{x^2}{3}\right)^{-2}}$



Características de la variable aleatoria t_n.

Su ***media*** $\mu=0$ y su ***varianza*** $\sigma^2=\frac{n}{n-2}$.

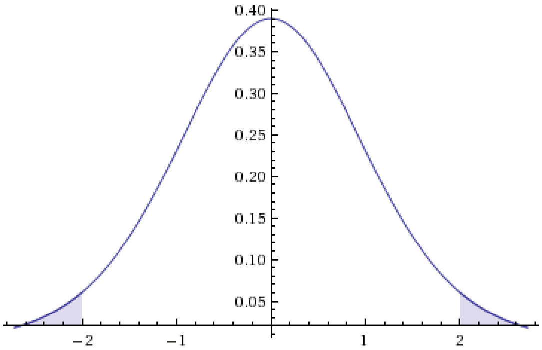
Tiene su aplicación en la estimación de las características de una población con distribución normal mediante los llamados intervalos de confianza.

EJEMPLO 11:

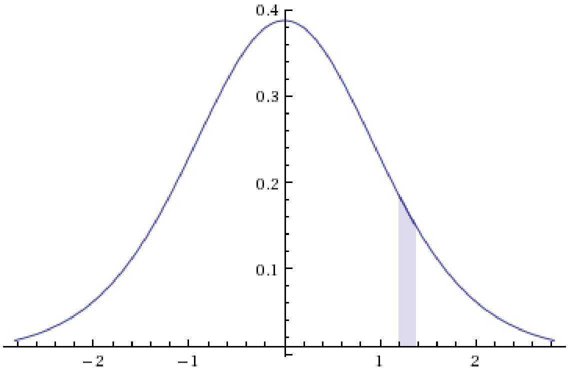
Calcular: a) $P(|t_{n=11}|>2)$ **. b)** $P(1,2<t_{n=9}<1,37)$ **. c) Valor de x tal que** $P(t_{n=10}<x)=0,9$

<u>Solución</u>:

a) $P(|t_{n=11}|>2)=2\cdot(1-F(2))\approx 0{,}070803955$.



b) $P(1,2<t_{n=9}<1,37)=F(1{,}37)-F(1{,}2)\approx 0{,}02844171293$



c) Valor de x tal que: $P(t_{n=10}<x)=0,9\Rightarrow x=1{,}372183635$

http://asignaturas.topografia.upm.es/matematicas/videos/Student.mp4

2.5.3. Distribución F de Snedecor

Si χ^2_n y χ^2_m son dos χ^2 de Pearson, respectivamente con n y m grados de libertad, independientes entre sí, entonces la variable: $\mathbf{F_{n,m} = \dfrac{\chi^2_n / n}{\chi^2_m / m}}$ se la denomina ***F de Snedecor*** con n y m grados de libertad.

La distribución F toma valores positivos por ser cociente de dos distribuciones positivas.

Características de la variable aleatoria $F_{n,m}$.

Su ***media*** $\mu = \dfrac{m}{m-2}$ **si m>2** y su ***varianza*** $\sigma^2 = \dfrac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$ si m>4.

Una aplicación importante de esta distribución es la comparación de dos estimadores independientes de varianzas (ANOVA).