TEMA 66: DISTRIBUCIONES DE PROBABILIDAD DE VARIABLE CONTINUA. CARACTERÍSTICAS Y TRATAMIENTO. LA DISTRIBUCIÓN NORMAL. APLICACIONES.

TIEMPO: 86 — 78

Esquema

1) Introducción
   1.1) Laplace
   1.2) Normal
   1.3) de Moivre
2 Definiciones
   2.1) Def: variable aleatoria, v.a. continua
   2.2) Def: función de distribución + propiedades
   2.3) Def: función de densidad + propiedades
   2.4) Teorema
3) Momentos de una v.a. continua
   3.1) Def: esperanza + propiedades
   3.2) Def: momentos centrales y no centrales
   3.3) Def: varianza + propiedades
   3.4) Def: mediana, moda, coeficiente de variación, f.g.m. + propiedades
   3.5) Desigualdad de Tchebycheff + Ley débil de los grandes números + Desigualdad de Markov
4) Algunas variables aleatorias continuas
   4.1) Distribución uniforme
   4.2) Distribución de Cauchy
5) Distribución Normal
   5.1) Introducción + aplicaciones
   5.2) Def: Normal + propiedades
   5.3) $E[\mathbb{X}]$, $\sigma^2[\mathbb{X}]$, f.g.m, función característica
   5.4) Normal tipificada
   5.5) Teorema Central del Límite
      5.1) Corolario (Binomial)
      5.2) Corolario (Poisson)
   5.6) Otras distribuciones

1) Introducción:

▷ Laplace: el científico francés Laplace (XVIII - XIX) fue la primera persona en definir el concepto de Probabilidad:"la teoría de probabilidad es, en el fondo, sólo sentido común expresado con números". En su "Teoría Analítica de las Probabilidades" (s.XIX) introdujo por primera vez el Análisis al estudiar fenómenos aleatorios. En esa misma obra aparecen también el principio de los números cuadrados, la (posteriormente) llamada Regla de Bayes así como el concepto de función generatriz.

▷ Normal: este tema, definiciones previas al margen, está centrado en la distribución Normal y su estudio básico (esperanza, varianza, Tª Central del Límite,...). En la literatura podemos encontrar otros nombres para esta distribución: distribución de Gauss, campana de Gauss o distribución de Gauss-Laplace $\left(\text{pues Laplace parece ser fue el primero en demostrar que } \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} = \sqrt{2\pi}\right)$.

▷ de Moivre: Por otra parte, otros autores consideran que el primero en trabajar sobre la Normal fue el francés Abraham de Moivre (XVII-XVIII), exiliado en Londres y amigo de Newton. A finales del s.XVIII publica una obra donde aparece, por primera vez, la distribución Normal (aunque obviamente él no la llamaba así).

2) Definiciones:

$\triangleright$ Sea un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ donde Ω es un conjunto, $\mathcal{A}$ una σ-álgebra sobre Ω y $P : A \mapsto \mathbb{R}$ una función de probabilidad sobre A.

$\triangleright$ **Definición**: una función $\mathbb{X} : \Omega \mapsto \mathbb{R}$ es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$ si se cumple que $\mathbb{X}^{-1}(B) \in A$, $\forall B \subset \mathbb{R}$ conjunto Borel.

$\triangleright$ **Definición**: aquellas variables aleatorias que toman un número finito o numerable de valores son llamadas discretas (Binomial, Poisson, Hipergeométrica,...). En caso contrario son llamadas continuas.

$\triangleright$ *Nota*: las variables aleatorias continuas sirven para medir la vida de un virus, peso de la población, alturas,... donde pueden tomar cualquier valor entre dos dados.

$\triangleright$ **Definición**: dada $\mathbb{X}$ v.a. definida sobre $(\Omega, \mathcal{A}, P)$. A la función $F : \mathbb{R} \mapsto \mathbb{R}$ dada por $x \mapsto f(x) = P[\mathbb{X} \leq x]$ la llamaremos función de distribución de $\mathbb{X}$.

$\triangleright$ *Nota*: así definida, $F(x)$ es la probabilidad del conjunto de sucesos de Ω tales que la imagen por $\mathbb{X}$ sea menor o igual a "x", es decir:

$$F(x) = P[\{\omega \in \Omega / \mathbb{X}(\omega) \leq x\}] = P[\{\omega \in \Omega / \omega \in \mathbb{X}^{-1}(-\infty, x]\}]$$

$\triangleright$ **Propiedades**:

1) $\forall x \in \mathbb{R}$, $0 \leq F(x) \leq 1$.

   Dem. Directamente de la definición. $\square$

2) $F(-\infty) = 0$, $F(+\infty) = 1$

   Dem. (1) $F(-\infty) = \lim_{x \to -\infty} P[\mathbb{X} \leq x] = P[\lim_{x \to -\infty} \{\omega \in \Omega / \mathbb{X}(\omega) \leq x\}] = P[\emptyset] = 0$
   (2) $F(+\infty) = \lim_{x \to +\infty} P[\mathbb{X} \leq x] = P[\lim_{x \to +\infty} \{\omega \in \Omega / \mathbb{X}(\omega) \leq x\}] = P[\Omega] = 1$ $\square$

3) $\forall x_1, x_2 \in \mathbb{R}$ con $x_1 < x_2 \longrightarrow F(x_1) \leq F(x_2)$

   Dem. $F(x_2) = P[\mathbb{X} \leq x_2] = P[\mathbb{X} \leq x_1] + P[x_1 \leq \mathbb{X} \leq x_2] \geq F(x_1)$ $\square$

4) $F(x)$ es continua por la derecha $\forall x \in \mathbb{R}$.

   Dem. Sea $h \to 0^+$, entonces: $\lim_{h \to 0^+} (F(x+h) - F(x)) = \lim_{h \to 0^+} P[x < \mathbb{X} \leq x + h] = P[\lim_{h \to 0^+} \{\omega \in \Omega / x < \mathbb{X}(\omega) \leq x + h\}] = P[\emptyset] = 0$ $\square$

▷ *Nota*: cuando consideramos variables aleatorias discretas la función de distribución cumplía las propiedades que acabamos de señalar y en el conjunto de puntos en los que la v.a. $\mathbb{X}$ tomaba valores, se producía un salto cuya amplitud era la probabilidad del punto donde se producía. Recordemos que una función es absolutamente continua si $\forall \epsilon > 0$, $\exists \delta > 0$ tal que $\forall$ $]a_1,\ b_1],...,]a_n,\ b_n]$ disjuntos verificando que $\displaystyle\sum_{j=1}^{n} |b_j - a_j| < \delta \Rightarrow \sum |F(b_j) - F(a_j)| < \epsilon$. Tenemos que si F es absolutamente continua $\longrightarrow F$ uniformemente continua $\longrightarrow F$ continua.

▷ **Definición**: diremos que la v.a. $\mathbb{X}$ definida sobre $(\Omega, \mathcal{A}, P)$ es continua si la función de distribución de $\mathbb{X}$ es absolutamente continua.

▷ *Nota*: para una v.a. continua se verifica que: $P[\mathbb{X} = x] = 0$; $P[x_1 \leq \mathbb{X} \leq x_2] = P[x_1 < \mathbb{X} \leq x_2] = P[x_1 \leq \mathbb{X} < x_2] = P[x_1 < \mathbb{X} < x_2]$. Dado que una función absolutamente continua es derivable salvo en un conjunto de medida cero, no debemos derivar sin cuidado. Sea $T = \{x \in \mathbb{R} \text{ tq } \exists F'(x)\}$.

▷ **Definición**: definimos la función de densidad de la v.a. continua $\mathbb{X}$ como la función:

$$f(x) = \begin{cases} F'(x) & \forall x \in T \\ 0 & \forall x \in \mathbb{R} - T \end{cases}$$

▷ **Propiedades**:

1) Como F es creciente y derivable en $T \longrightarrow f(x) = F'(x) \geq 0$, $\forall x \in T$. Esto nos dice que $f(x) \geq 0$, $\forall x \in \mathbb{R}$.

2) $F(x) = \displaystyle\int_{-\infty}^{x} f(y)dy$.

   Dem. $\displaystyle\lim_{\alpha \to -\infty} \int_{\alpha}^{y} f(x)dx = \lim_{\alpha \to -\infty} [F(x) - F(\alpha)] = F(x) - 0 = F(x)$ □

3) $P[x_1 \leq \mathbb{X} \leq x_2] = \displaystyle\int_{x_1}^{x_2} f(x)dx$.

   Dem. $F(x_2) - F(x_1) = \displaystyle\int_{-\infty}^{x_2} f(x)dx - \int_{-\infty}^{x_1} f(x)dx = \int_{x_1}^{x_2} f(x)dx$ □

4) $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$.

   Dem. $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{x_1} f(x)dx + \int_{x_1}^{+\infty} f(x)dx = F(x_1) - F(-\infty) - F(x_1) + F(+\infty) = 1$ □

▷ *Nota*: otro camino para definir una v.a. continua a partir de una función "f" verificando que $f(x) \geq 0$, $\forall x \in \mathbb{R}$ y $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$ (es decir, propiedades **1** y **4**), es llamar $F(x) = \displaystyle\int_{-\infty}^{x} f(t)dt = P[x \leq x]$.

▷ **<u>Definición</u>**: sea $\mathbb{X}$ una v.a. continua. Diremos que es <u>simétrica respecto de x_0</u> si se cumple que:

$$P[\mathbb{X} \geq x_0 + x] = P[\mathbb{X} \leq x_0 - x]$$

Si $x_0 = 0$, diremos que es simétrica respecto del origen.

▷ **<u>Teorema</u>** sean $\mathbb{X}$ v.a. continua sobre $(\Omega, \mathcal{A}, P)$, $h : \mathbb{R} \mapsto \mathbb{R}$ medible Borel tal que $h'(x) > 0$, $\forall x \in \mathbb{R}$. Entonces $h \circ \mathbb{X}$ es una nueva v.a. continua con función de distribución $H(x) = F(h^{-1}(x))$ donde F es la función de distribución de $\mathbb{X}$.

3) Momentos de una variable aleatoria continua:

▷ **Definición**: sea $\mathbb{X}$ una v.a. continua con función de densidad $f(x)$.
Supongamos que $\int_{-\infty}^{+\infty} |x| \cdot f(x)dx < \infty$. Definimos la esperanza matemática, $E[\mathbb{X}]$, de la v.a. $\mathbb{X}$ como:

$$E[\mathbb{X}] = \int_{-\infty}^{+\infty} x \cdot f(x)dx,$$

que representa el promedio de la distribución.

▷ **Propiedades**:

1) Si $\mathbb{X}$ es simétrica respecto de $x_0 \in \mathbb{R}$, $E[\mathbb{X}] = x_0$ (en caso de que exista).

2) $E[a \cdot \mathbb{X} + b] = a \cdot E[\mathbb{X}] + b$, $\forall a,\, b \in \mathbb{R}$, es decir, es un operador lineal.

▷ **Definición**: sea $\mathbb{X}$ una v.a. continua con función de densidad $f(x)$.
Supongamos que $\int_{-\infty}^{+\infty} |x|^n \cdot f(x)dx < \infty$.
Llamaremos momento no centrado de orden "n"/momento de orden "n" respecto al origen al valor:

$$m_n = E[\mathbb{X}^n] = \int_{-\infty}^{+\infty} x^n \cdot f(x)dx$$

▷ **Definición**: bajo las mismas condiciones anteriores, si $\exists E[(\mathbb{X} - \mathbb{X})^n]$ lo llamaremos momento centrado de orden "n":

$$\mu_n = E[(\mathbb{X} - E[\mathbb{X}])^n] = \int_{-\infty}^{+\infty} (x - E[\mathbb{X}])^n \cdot f(x)dx$$

▷ **Definición**: μ_2, en caso de existir, se llama varianza de $\mathbb{X}$, lo notaremos por σ^2 y es la medida óptima de dispersión.

▷ **Definición**: llamaremos desviación típica a la raíz cuadrada positiva de la varianza:

$$\sigma = +\sqrt{\sigma^2}$$

▷ **Propiedades**:

1) Si existe el momento de orden "n"$\longrightarrow$ existen todos los anteriores.

2) Se verifica que $\sigma^2[a \cdot \mathbb{X} + b] = a^2 \cdot \sigma^2[\mathbb{X}]$.

3) Se cumple la siguiente relación entre los momentos: $\mu_n = E[(\mathbb{X} - \mathbb{X})^n] = \sum_{k=0}^{n} (-1)^k \binom{n}{k} m_{n-k} m_1^k$

4) Teorema de Köning: $\sigma^2 = \mu_2 = E[\mathbb{X}^2] - (E[\mathbb{X}])^2$

▷ **Definición**: si $F(x)$ es la función de distribución de la v.a. continua $\mathbb{X}$, llamamos mediana de $\mathbb{X}$ al calor de la variable "x" tal que $F(x) = \frac{1}{2}$, es decir, aquel que divide a la población en dos partes iguales.

▷ **Definición**: llamaremos moda de $\mathbb{X}$ al valor de la variable que maximiza $f(x)$.

▷ **Definición**:si $E[\mathbb{X}] \neq 0$, se llama coeficiente de variación de $\mathbb{X}$ al cociente: $\dfrac{\sigma}{E[\mathbb{X}]}$.

▷ **Definición**: sea la v.a. continua $\mathbb{X}$, llamamos función generatriz de momentos, f.g.m. a la función: $M_{\mathbb{X}}(t) = E[e^{t\mathbb{X}}]$ en caso de que exista para $t \in]-\epsilon, \epsilon[$ entorno abierto de cero.

▷ **Propiedades**:

1) Si dos v.a. continuas tienen la misma f.g.m. en un entorno de cero $\longrightarrow$ tienen la misma función de densidad.

2) Si $\exists M_{\mathbb{X}}(t)$ en $]-\epsilon, \epsilon[\longrightarrow$ es infinitamente derivable en $t=0$ y se cumple que: $M_{\mathbb{X}}^{n/}(0) = E[\mathbb{X}^n] = m_n$, $\forall n \in \mathbb{N}$.

▷ **Desigualdad de Tchebycheff**: sea $\mathbb{X}$ una v.a. con $E[\mathbb{X}]$ y σ^2 finitas. Entonces, $\forall \epsilon > 0$:

$$P\Big[|\mathbb{X} - E[\mathbb{X}]| > \epsilon\Big] \leq \frac{\sigma^2}{\epsilon^2}$$

▷ **Teorema** (ley débil de los grandes números): sea $\mathbb{X}_1, \ldots, \mathbb{X}_n$ una sucesión de v.a. independientes con igual esperanza y varianza. Sea $S_n = \dfrac{\mathbb{X}_1 + \ldots + \mathbb{X}_n}{n}$. Entonces, $\forall \epsilon > 0$ se cumple que:

$$\lim_{n\to\infty} P\Big[|S_n - E[\mathbb{X}]| > \epsilon\Big] = 0 \Longleftrightarrow \lim_{n\to\infty} P\Big[|S_n - E[\mathbb{X}]| \leq \epsilon\Big] = 1$$

Dem. Como $E[S_n] = \dfrac{E[\mathbb{X}_1] + \ldots + E[\mathbb{X}_n]}{n} = \dfrac{n \cdot E[\mathbb{X}]}{n} = E[\mathbb{X}]$ y $\sigma^2[\mathbb{X}_1 + \ldots + \mathbb{X}_n] = n \cdot \sigma^2$, tenemos que $\sigma^2[S_n] = \sigma^2[\mathbb{X}_1 + \ldots + \mathbb{X}_n] \cdot \dfrac{1}{n^2} = \dfrac{\sigma^2}{n}$.
Aplicando Tchebytcheff: $P\Big[|S_n - E[S_n](= E[\mathbb{X}])| > \epsilon\Big] \leq \dfrac{\frac{\sigma^2}{n}}{\epsilon^2} = \dfrac{\sigma^2}{\epsilon} \cdot \dfrac{1}{n}$. Tomamos límites para obtener: $\lim\limits_{n\to\infty} P\Big[|S_n - E[\mathbb{X}]| > \epsilon\Big] = 0$

$\square$

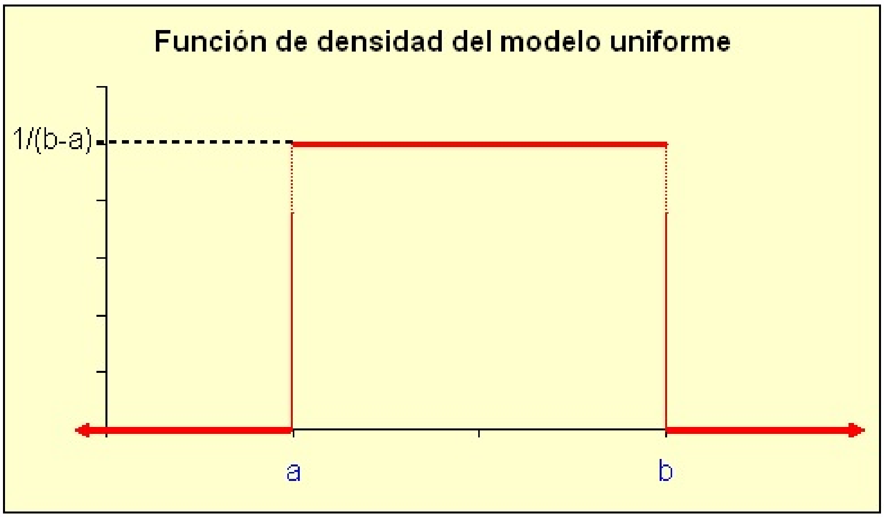
▷ **Desigualdad de Markov**: sea $\mathbb{X}$ v.a. tal que existe $E[|\mathbb{X}|^k]$ tq $k \in \mathbb{R}^+$. Entonces, $\forall \epsilon > 0$:

$$P\Big[|\mathbb{X}| > \epsilon\Big] \leq \frac{E\Big[|\mathbb{X}|^k\Big]}{\epsilon^k}$$

4) Algunas variables aleatorias continuas

▷ **Distribución uniforme o rectangular**: su función de densidad es:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & \forall x \notin [a,b] \end{cases}$$



$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{a}^{x} \frac{1}{b-a}dx = \frac{x-a}{b-a}$$

$$E[\mathbb{X}] = \frac{a+b}{2}$$

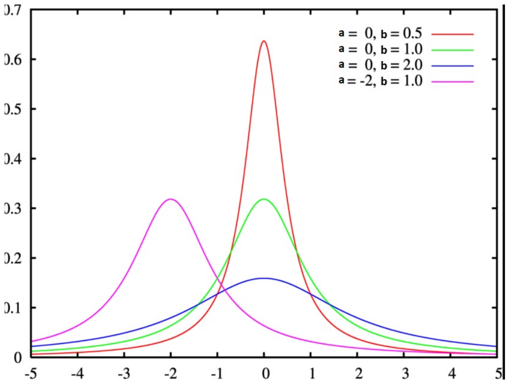
$$M_{\mathbb{X}}(t) = \frac{\mathrm{e}^{t\cdot b} - \mathrm{e}^{t\cdot a}}{(b-a)t}$$

▷ Por ejemplo: una aguja gira libremente alrededor de un eje centrado en un círculo marcado del 0 al 12. Cuando para se anota su valor. Claramente $x \in [0,12]$ y $P[\mathbb{X} = k] = 0$. Sin embargo, $P[x \in [a,b]] = C \cdot (b-a)$ con $C \equiv$ cte (cuanto más grande sea el intervalo, más probable es que caiga la aguja en él).

$P[x \in [0,12] = C \cdot (12-0) = 1 \longrightarrow C = \dfrac{1}{12}$, luego:

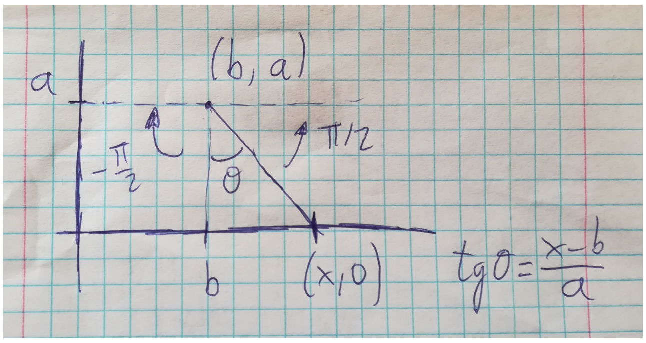
$$f(x) = \begin{cases} \dfrac{1}{12} & 0 \leq x \leq 12 \\ 0 & \forall x \notin [0,12] \end{cases}$$

▷ **Distribución de Cauchy**: decimos que $\mathbb{X}$ sigue un distribución de Cauchy, $\mathcal{C}(a,b)$ de parámetros "a" (>0) y "b" ($\in \mathbb{R}$) si su función de densidad es: $f(x) = \dfrac{1}{\pi} \cdot \dfrac{a}{a^2 + (x-b)^2}$



Al parámetro "a" lo llamaremos parámetro de aplastamiento.
El parámetros "b" nos indica que $x = b$ es eje de simetría de la función de densidad.
$\displaystyle\int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\pi} \cdot \left[\arctan(\frac{x-b}{a}) \right]_{-\infty}^{+\infty} = 1$

▷ Podemos definir la distribución de Cauchy de la siguiente manera: fijemos un punto del plano (b,a) con $a > 0$. Tomemos un ángulo al azar $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ y entonces obtenemos un punto de corte $(x,0)$ con el eje X. Si $\theta \to \frac{\pi}{2} \rightsquigarrow x \to +\infty$ y si $\theta \to -\frac{\pi}{2} \rightsquigarrow x \to -\infty$.



A cada valor arbitrario de θ le corresponde un punto de la recta horizontal $Y = 0$. Podemos asignarle, pues, una distribución uniforme:

$$f(\theta) = \begin{cases} \dfrac{1}{\pi} & -\dfrac{\pi}{2} \leq x \leq \dfrac{\pi}{2} \\ 0 & \forall \theta \notin [-\dfrac{\pi}{2}, \dfrac{\pi}{2}] \end{cases}$$

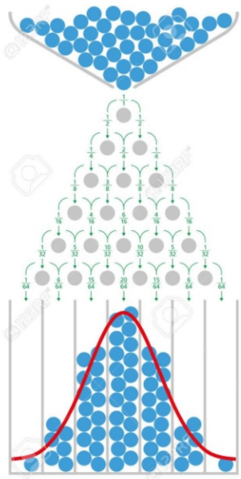
Tomamos $\mathbb{X} = \tan(\theta)$. Por un teorema anterior, la función tangente $\tan :]-\frac{\pi}{2}, \frac{\pi}{2}[\mapsto \mathbb{R}$ es Borel medible y como $\tan'(x) > 0$, $\forall x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, haciendo $h(x) = f(\theta)\dfrac{d\theta}{dx} = \dfrac{1}{\pi} \cdot \dfrac{a}{a^2+(x-b)^2}$ que es $\mathcal{C}(a,b)$.

▷ La distribución de Cauchy no tiene función generatriz de momentos y, por lo tanto, no tiene momentos de ningún orden.

5) Distribución Normal

▷ Francis Galton (XIX - XX)hizo una primera aproximación a la distribución Normal mediante un dispositivo mecánico: un tablero en posición vertical con clavos distribuidos regularmente por donde pasan unas bolas que caen a un depósito colocado en la parte superior. Las bolas se alejan con una cierta amplitud mayor o menor de la línea central marcada por la posición del depósito ya mencionada. Las bolas se recogen en compartimentos estancos situados en la parte inferior del tablero: la altura alcanzada por las bolas nos da una idea de la Distribución Normal.

▷ Hace unos años se construyó "el demostrador mecánico de la probabilidad", esencialmente como el de Galton pero sustituyendo los clavos por obstáculos hexagonales de forma que la bola vaya a izquierda o derecha, tras chocar con éstos, con probabilidad igual a 1/2.



▷ Se llegó a pensar que la Distribución Normal regía todos los fenómenos a largo plazo pues existen numerosos ejemplos de la vida real en que ello ocurre. Esto, dicho así, es falso. Sí es cierto que existen la "Ley Fuerte de los Grandes Números" y el "Teorema Central del Límite" que aseguran que, a la larga, muchos fenómenos pueden aproximarse por una Normal sin que ello afecte a la fiabilidad del resultado. Este teorema, clave en Estadística y Probabilidad, fue demostrado en 1.901.
Como las hipótesis se cumplen en multitud de casos prácticas, el uso (y estudio) de la Normal es fundamental para la Psicología, la Estadística, la Biología,...

▷ Hay que decir que el primero en obtener la Normal como límite de una Binomial fue de Moivre. Posteriormente Gauss y Laplace la relacionarían con sus estudios astronómicos y el análisis de los errores.
El primero que la usó desde el punto de vista de los fenómenos sociales fue el estadístico belga Quetelet (s.XIX).

▷ **Definición**: sea una v.a. continua $\mathbb{X}$. Decimos que $\mathbb{X}$ sigue una distribución normal de parámetros μ y σ si su función de densidad es:

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

y la notaremos por $\mathbb{X} \sim N(\mu, \sigma)$

▷ *Nota*: es una función de densidad pues $f(x) \geq 0$ trivialmente ($f(x) > 0$ de hecho) $\forall x \in \mathbb{R}$ y $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = \{x - \mu = \sigma \cdot z \to dx = \sigma dz\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{z^2}{2}} dz = 1$, donde la última integral se calcula usando el Teorema de los Residuos.

▷ **Propiedades**:

1) Su dominio es todo $\mathbb{R}$.

2) Es simétrica respecto a μ pues: $f(\mu + x) = f(\mu - x)$.

3) $f'(x) = -f(x) \cdot \dfrac{x-\mu}{\sigma^2} = 0 \longleftrightarrow x = \mu \longrightarrow f(\mu)$ es la moda (máximo).

4) $x = \mu + \sigma$ y $x = \mu - \sigma$ son los puntos de inflexión.

5) $\lim_{x\to-\infty} f(x) = \lim_{x\to+\infty} f(x) = 0$

6) $E[\mathbb{X}] = \mu$

Dem. $E[\mathbb{X}] = \displaystyle\int_{-\infty}^{+\infty} x \cdot f(x)dx = \left\{\begin{matrix} x-\mu & = \sigma\cdot z \\ dx & = \sigma \cdot dz \end{matrix}\right\} = \frac{1}{\sqrt{2\pi}}\Big[\int_{-\infty}^{+\infty} \sigma z \mathrm{e}^{-\frac{z^2}{2}}dz + \int_{-\infty}^{+\infty} \mu \mathrm{e}^{-\frac{z^2}{2}}dz\Big] =$

$= 0$ (integral de función impar sobre $\mathbb{R}$) $+ \mu = \mu$ $\square$

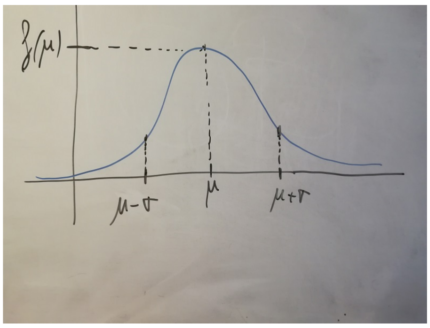
7) Varianza $= E[(\mathbb{X} - E[\mathbb{X}])^2] = \sigma^2$

Dem. $E[(\mathbb{X}-E[\mathbb{X}])^2] = \left\{\begin{matrix} x-\mu & = \sigma\cdot z \\ dx & = \sigma \cdot dz \end{matrix}\right\} = \dfrac{1}{\sigma\sqrt{2\pi}}\Big[\displaystyle\int_{-\infty}^{+\infty} \sigma^3 \cdot z^2 \cdot \mathrm{e}^{-\frac{z^2}{2}}dz\Big] = \sigma^2$ $\square$

8) La f.g.m: $M_{\mathbb{X}}(t) = \mathrm{e}^{t\cdot\mu + \frac{\sigma^2 \cdot t^2}{2}}$

Dem. Vamos a ver cuál es para la $\mathbb{Y} \sim N(0,1)$ y luego generalizaremos.
$M_{\mathbb{Y}}(t) = E[\mathrm{e}^{t\cdot\mathbb{Y}}] = \frac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{y^2}{2}+t\cdot y}dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{1}{2}(-y^2+2ty+t^2-t^2)}dy =$
$= \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{1}{2}(t-y)^2+\frac{t^2}{2}}dy = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{t^2}{2}}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{1}{2}(t-y)^2}dy = \mathrm{e}^{\frac{t^2}{2}} = M_{N(0,1)}(t)$
$M_{a\cdot\mathbb{X}+b}(t) = E[\mathrm{e}^{(a\cdot\mathbb{X}+b)\cdot t}] = \mathrm{e}^{t\cdot b}E[\mathrm{e}^{(a\cdot t)\mathbb{X}}] = \mathrm{e}^{t\cdot b}M_{\mathbb{X}}(at)$
Luego, si $\mathbb{Y} = \dfrac{\mathbb{X}-\mu_{\mathbb{X}}}{\sigma_{\mathbb{X}}} \longrightarrow \mathbb{Y} \sim N(0,1) \Rightarrow M_Y(t) = \mathrm{e}^{\frac{t^2}{2}}$. Despejando $\mathbb{X}$ obtenemos:
$M_{\mathbb{X}}(t) = M_{\sigma_{\mathbb{X}}\mathbb{Y}+\mu_{\mathbb{X}}}(t) = \mathrm{e}^{t\cdot\mu}\cdot \mathrm{e}^{\frac{\sigma^2 t^2}{2}} = \mathrm{e}^{t\cdot\mu+\frac{\sigma^2 t^2}{2}}$, $\forall t \in \mathbb{R}$ $\square$

9) La función característica: $E[\mathrm{e}^{it\mathbb{X}}] = M_{\mathbb{X}}(i\cdot t) = \mathrm{e}^{i\cdot t\cdot\mu - \frac{\sigma^2 \cdot t^2}{2}}$



▷ La importancia de la $N(0,1)$, la llamada Normal tipificada, es que con ella nos vale para hallar los valores $P[\mathbb{X} \leq x]$ cuando $\mathbb{X} \sim N(\mu,\sigma)$. Sea $\mathbb{X} \sim N(\mu,\sigma)$, entonces:

$P[\mathbb{X} \leq a] = P\Big[\dfrac{\mathbb{X}-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\Big] = \Big\{\mathbb{X} \sim N(\mu,\sigma) \longrightarrow \mathbb{Y} = \dfrac{\mathbb{X}-\mu}{\sigma} \sim N(0,1)\Big\} = P[\mathbb{Y} \leq \alpha]$ y este valor podremos buscarlo en la tabla.

Luego, si $\mathbb{X} \sim N(\mu,\sigma)$ y queremos hallar $P[a \leq \mathbb{X} \leq b]$ nos basta con:
$P\Big[\dfrac{a-\mu}{\sigma} \leq \mathbb{Y} \leq \dfrac{b-\mu}{\sigma}\Big] = P[\alpha \leq \mathbb{Y} \leq \beta]$, con $\mathbb{Y} \sim N(0,1)$.

$\triangleright$ Usando la desigualdad de Tchebycheff: $P[|\mathbb{X} - \mu| > \epsilon \cdot \sigma] \leq \dfrac{1}{\epsilon^2} \longleftrightarrow P\left[\left|\dfrac{\mathbb{X} - \mu}{\sigma}\right| > \epsilon\right] \leq \dfrac{1}{\epsilon^2}$

obtenemos los siguiente (para $\mathbb{Y} \sim N(0,1)$).

a) Si $\epsilon = 2/3$, el 50% de las observaciones están entre $\mu - \frac{2}{3}\sigma$ y $\mu + \frac{2}{3}\sigma$

b) Si $\epsilon = 3$, el 99% de las observaciones están entre $\mu - 3\sigma$ y $\mu + 3\sigma$

$\triangleright$ <u>**Teorema Central del Límite** (o ley fuerte de los grandes números)</u>: sea $\{\mathbb{X}_n\}$ una sucesión de v.a. independientes e idénticamente distribuidas de media μ y desviación típica σ (ambas finitas). Construimos una nueva v.a. $\rho = \dfrac{\mathbb{X}_1 + \mathbb{X}_2 + \ldots + \mathbb{X}_n - n \cdot \mu}{\sigma\sqrt{n}}$, que resultará ser asintóticamente equivalente a la $N(0,1)$, es decir:

$$\lim_{n\to\infty} \rho_n \to N(0,1), \text{ es decir: } \lim_{n\to\infty} F_n(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{z^2}{2}} dz$$

o que $\mathbb{X}_1 + \mathbb{X}_2 + \ldots + \mathbb{X}_n \rightsquigarrow N(n \cdot \mu, \sigma\sqrt{n})$

$\triangleright$ <u>**Corolario**</u>: sea $\mathbb{X} \sim B(n,p)$. Para "n" suficientemente grande ($n \geq 25$) se verifica que:

$$\mathbb{X} \approx N(n \cdot p, \sqrt{n \cdot p \cdot (1-p)})$$

$\triangleright$ <u>**Corolario**</u>: sea $\mathbb{X} \sim P(\lambda)$. Para "$n$" suficientemente grande ($n \geq 10$) se verifica que:

$$\mathbb{X} \approx N(\lambda, \sqrt{\lambda})$$

$\triangleright$ <u>*Nota*</u>: como en una v.a. continua tenemos que $P[\mathbb{X} = x] = 0$, cuando aproximamos discretas por la Normal hacemos: $P[\mathbb{X} = a] = P[a - \epsilon \leq \mathbb{X} \leq a + \epsilon]$ y hacemos $\epsilon \to 0$. Generalmente se toma $\epsilon = 0,5$ para mirar en las tablas. Hay más distribuciones continuas: la Normal multidimensional, la distribución Gamma, Beta, la $\mathcal{X}^2$ de Pearson, la t de Student, la F de Snedecor,...