LEYES BINOMIAL Y NORMAL
Resumen

1) Distribución binomial
Realizamos n pruebas de Bernoulli, esto es, n experimentos aleatorios independientes en los que cabe, en cada uno de ellos, dos resultados mutuamente excluyentes (E = "éxito" ó F = "fracaso"), con probabilidades respectivas $p = P(E)$ $q = 1 - p = P(F)$. Si consideramos la variable aleatoria:

$$X = \text{nº de éxitos en las } n \text{ realizaciones}$$

entonces, la probabilidad de obtener en ellas k éxitos en total ($0 \le k \le n$) es:

$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

donde el número combinatorio $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$, siendo $a! = a(a-1)\cdot\ldots\cdot 2\cdot 1$; $1! = 1$; $0! = 1$. ($a!$ = "a factorial" ó "factorial de a", y el número combinatorio se lee "n sobre k").

Se dice que X sigue una Distribución Binomial de parámetros n y p: $X \in B(n, p)$. Sus media y desviación típica poblacionales valen:

$$\mu = np; \quad \sigma = \sqrt{npq}$$

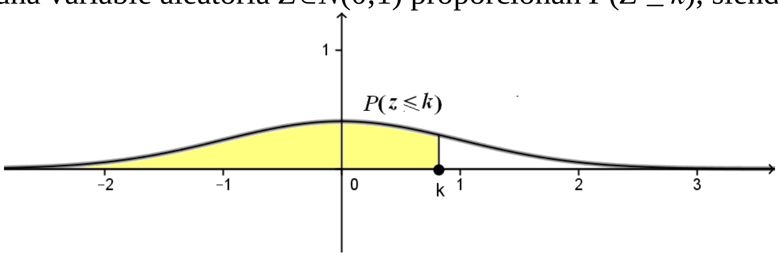
Una Binomial es una variable aleatoria *discreta* (toma valores aislados: entre dos valores que puede tomar la variable, hay valores que no puede tomar).

2) Distribución normal
Es una distribución de probabilidad *continua* (puede tomar cualquier valor, dentro del rango en que puede moverse). Es el modelo probabilístico más importante.

La notación estándar es $X \in N(\mu; \sigma)$, siendo μ = media poblacional y σ = desviación típica poblacional.

En los problemas nos van a decir que el experimento aleatorio se ajusta a una distribución Normal. Esto se comprueba con determinados contrastes que no vamos a estudiar.

Normal $N(0;1)$: Uso de las tablas
Las tablas de una variable aleatoria $Z \in N(0;1)$ proporcionan $P(Z \le k)$, siendo $k \ge 0$.



Uso:

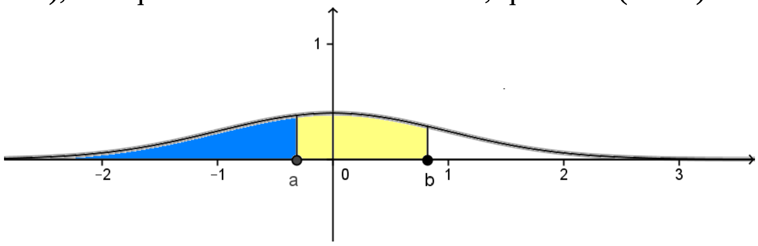
1) Si conocemos k, para hallar $P(Z \le k)$ buscamos la fila correspondiente a la unidad y primer decimal y la columna correspondiente al segundo decimal. El número del interior de la tabla, donde se cruzan, es la probabilidad $P(Z \le k)$.

Ejemplo: $P(Z < 1{,}96) = 0{,}975$
Nota: Por tratarse de una distribución continua, $P(Z = \text{número}) = 0$. Por tanto, $P(Z < 1{,}96) = P(Z \le 1{,}96)$

	CENTÉSIMAS			
k	…	…	6	…
…	…	…	…	…
1,9	…	…	0,9750	…

Casos que pueden presentarse usando las tablas de la $N(0;1)$ (no se distingue entre $<$ y $\le$, ni entre $>$ y $\ge$, como hemos dicho en el ejemplo; $k \ge 0$):

1. $P(Z < k)$: Lo buscamos en las tablas de la forma descrita.
2. $P(Z > k) = 1 – P(Z < k)$ (el área bajo la totalidad de la curva es 1).
3. $P(Z < –k) = P(Z > k) = 1 – P(Z < k)$ (por simetría de la curva).
4. $P(Z > –k) = P(Z < k)$ (por simetría de la curva).
5. $P(a < Z < b) = P(Z < b) – P(Z < a)$ $a < b$, cualesquiera. En el gráfico siguiente, $P(Z < b)$ es el área bajo la curva desde $-\infty$ hasta b (zona azul + amarilla), a la que restaríamos el área azul, que es $P(Z < a)$:



2) Si conocemos la probabilidad $P(Z \le k)$ y queremos hallar k, buscamos la probabilidad conocida en el interior de la tabla; k es la fila y columna que al cruzarse dan dicho número (la fila da la unidad y primer decimal; la columna, el segundo decimal). Si $P(Z \le k) < 0.5$, las tablas no nos proporcionan el valor de k. Por simetría (si k es negativo $\Rightarrow -k$ es positivo): $P(Z \le k) = P(Z \ge -k) = 1 - P(Z \le -k)$. Calculamos $-k$ en las tablas.
<u>Ejemplo</u>: Hallar k tal que $P(Z \le k) = 0.3$. Se tiene: $P(Z \le k) = P(Z \ge -k) = 1 - P(Z \le -k) = 0.3 \Rightarrow P(Z \le -k) = 1 - 0.3 = 0.7$, resultando, aproximadamente, $-k = 0{,}525 \Rightarrow k = -0{,}525$.

<u>Normal N(μ;σ)</u>

Si $X \in N(\mu;\sigma) \Rightarrow z = \dfrac{X-\mu}{\sigma} \in N(0;1)$ ("tipificación de la normal")

<u>Ejemplo</u>: Si $X \in N(18;4) \Rightarrow P(11 \le X < 25) = P\left(\dfrac{11-18}{4} \le z < \dfrac{25-18}{4}\right) =$

$= P(-1.75 \le z < 1.75) = P(z<1.75) - P(z<-1.75) = P(z<1.75) - P(z>1.75) =$
$= P(z<1.75) - [1 - P(z<1.75)] = 2P(z<1.75) - 1 = 2\cdot 0.9599 - 1 = 0.9198$

<u>Aproximación de la Binomial por una Normal.</u>

Cuando aumentan los datos observados, muchas distribuciones de probabilidad se pueden aproximar mediante una normal bajo ciertas condiciones. Concretamente, para la <u>binomial</u> usamos:

- **Teorema de Moivre-Laplace (caso particular del *Teorema Central del Límite*)**: "Si $X \in B(n, p)$, entonces, X se puede aproximar mediante una normal $N(np, \sqrt{npq})$". La aproximación es, en general, buena, cuando $n \ge 30$. (A veces, se pide también que $n \cdot p \ge 5$ y $n \cdot q \ge 5$).
- **Corrección por continuidad de Yates (o de Fisher)**: La aproximación de una variable discreta X por una continua, a la que llamaremos X', genera un cierto error. Si la variable discreta X toma valores enteros y consecutivos, cuando aproximamos mediante la continua X', cada uno de ellos lo sustituimos por el intervalo que va desde 0.5 unidades menos a 0.5 unidades más. Así:

$P(X = a) = P(a - 0{,}5 \le X' \le a + 0{,}5)$. Y, por tanto:

- $P(X \le a) = P(X' \le a + 0{,}5)$
- $P(X < a) = P(X' \le a - 0{,}5)$
- $P(X > a) = P(X' \ge a + 0{,}5)$
- $P(X \ge a) = P(X' \ge a - 0{,}5)$
- $P(a < X < b) = P(a + 0.5 \le X' \le b - 0.5)$
- $P(a \le X \le b) = P(a - 0.5 \le X' \le b + 0.5)$