DISTRIBUCIÓN NORMAL

Una situación es normal cuando se ajusta a unos parámetros determinados; pero normalidad no es lo mismo que uniformidad: respetemos lo diferente.

1. Distribuciones de probabilidad continua

Una variable estadística se llama continua cuando puede tomar todos los valores de un intervalo.
Así, por ejemplo, son variables estadísticas continuas, la estatura y el peso de los individuos, el tiempo de espera de un autobús, el tamaño de una determinada variedad de manzanas, etc.

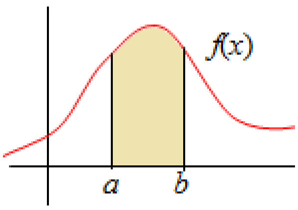
Para estas distribuciones, la probabilidad de un valor concreto es 0, pues el número de casos posibles es infinito. Por ejemplo, la probabilidad de que una persona mida exactamente 172,123456789… cm es 0; altura tan improbable como que mida exactamente 172,000… cm. En cambio, la probabilidad de que una persona mida entre 171,5 cm y 172,5 cm sí podrá calcularse.
Esto es, si X es la variable que mide la estatura de una persona, se tendrá:

$$P(X = 172{,}12345...) = 0;\quad P(X = 172{,}000...) = 0.$$

En cambio, $P(171{,}5 < X < 172{,}5) = ?$, valor que dependerá de la población de estudio.

Función de probabilidad

Esta función, que también se llama función de densidad, permite el cálculo de la probabilidad para sucesos asociados a distribuciones de variable continua.
Esa probabilidad se corresponde con el área del recinto comprendido entre la curva asociada a la función y el eje *OX*. En concreto, la probabilidad de que la variable tome valores en un intervalo, $P(a < X < b)$, será el área del recinto plano limitado por la función de densidad y el eje *OX* cuando $X \in (a,\ b)$.



La función $f(x)$ será de densidad si cumple:

1) $f(x) \geq 0$ para todo valor x de su dominio, pues la probabilidad nunca puede ser negativa.
→ Se corresponde con el axioma 1: Para cualquier suceso A se cumple que: $0 \leq P(A) \leq 1$.
2) El área limitada por la curva de $f(x)$ y el eje de abscisas vale 1.
→ Se corresponde con el axioma 2: La probabilidad del suceso seguro E es 1: $P(E) = 1$.

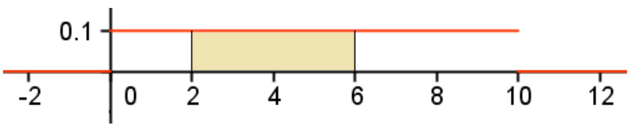
Observación: Dado que la probabilidad puntual vale 0, $P(X = a) = 0$, se verifica que:

$$P(X \leq a) = P(X < a);\quad P(a \leq X \leq b) = P(a < X < b) = P(X < b) - P(X < a).$$

Ejemplo:

La función $f(x) = \begin{cases} 0{,}1, & \text{si } 0 \leq x \leq 10 \\ 0, & \text{en otro caso} \end{cases}$, que puede describir el tiempo de espera, en minutos, hasta que llega un tren de cercanías, es una función de densidad.
En efecto, cumple:

1) $f(x) \geq 0$ para todo valor x de su dominio, que varía en 0 y 10 minutos.
2) El área limitada por la curva de $f(x)$ y el eje de abscisas, que es un rectángulo de base 10 y altura 0,1, vale 10 · 0,1 = 1.



→ La probabilidad de que haya que esperar entre 2 y 6 minutos es el área del rectángulo sombreado en la figura adjunta, que vale 4 · 0,1 = 0,4.

2. La distribución normal

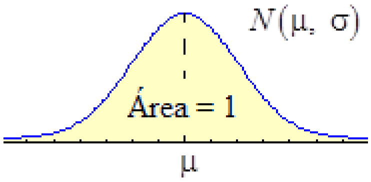
Es la distribución de probabilidad continua más frecuente. Está asociada a multitud de fenómenos naturales y cotidianos (cociente intelectual, talla o peso de las personas; tamaño de los frutos de cualquier tipo de árbol…), que se caracterizan porque la mayoría de los resultados tienden a agruparse en torno a su media, siendo muy improbable que la variable tome valores muy alejados de esa media.

- Una variable con distribución normal queda determinada cuando se conocen su media μ y su desviación típica σ. Se denota como $N(\mu, \sigma)$.

→ La expresión analítica de la <u>función de densidad</u> de la distribución normal es

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$, que efectivamente depende de μ y σ.

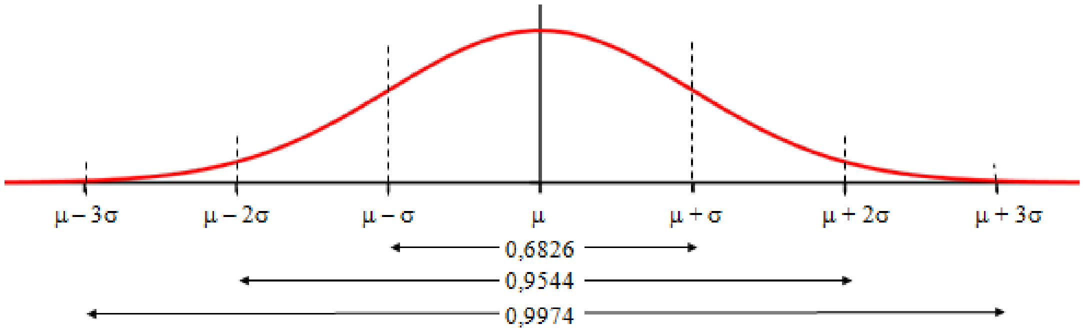
Su gráfica es la conocida "campana de Gauss".



Esta función cumple las siguientes propiedades:

- Está definida para todo número real, es decir, en el intervalo (–∞, +∞): la variable puede tomar cualquier valor; siendo $f(x)>0$ para todo x.
- El área del recinto plano limitado por la curva y el eje *OX* vale 1.
- Es simétrica respecto a su media μ.
- El eje de abscisas es una asíntota horizontal de la curva.
- Aunque la variable puede tomar cualquier valor entre –∞ y +∞, la probabilidad de que tome valores alejados de la media es prácticamente nula, pues se cumple que el área delimitada por la curva y el eje *OX* entre μ – σ y μ + σ es 0,6826; entre μ ± 2σ es de 0,9544; y entre μ ± 3σ es de 0,9974.

De hecho, a los valores que están a una distancia superior a 3,5σ de la media se les asigna una probabilidad 0.



Ejemplos:

a) Si la estatura de las jóvenes de 18 años de una determinada región es una variable estadística *X*, de media μ = 167 cm y deviación típica σ = 8 cm, entonces la distribución de estaturas de esa población puede estudiarse como una normal: $N(167, 8)$.

Esto significa que puede asegurase, con las probabilidades que se indican, que:

P(de que una de esas jóvenes mida menos de 167 cm) = $P(X < 167) = 0,5$ → (La mitad de las jóvenes tiene una estatura por debajo de la media; la otra mitad medirá más de 167 cm).

$P(\mu-\sigma < X < \mu+\sigma) = P(167 - 8 < X < 167 + 8) = P(159 < X < 175) = 0,6826$ →

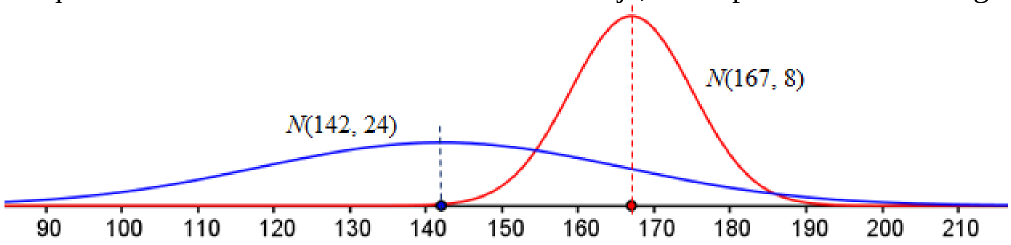
(El 68,26 % de las jóvenes de esa región tiene una estatura comprendida entre 159 y 175 cm).

$P(\mu-2\sigma < X < \mu+2\sigma) = P(167 - 16 < X < 167 + 16) = P(151 < X < 183) = 0,9544$ →

(El 95,44 % de las jóvenes de esa región tiene una estatura comprendida entre 151 y 183 cm).

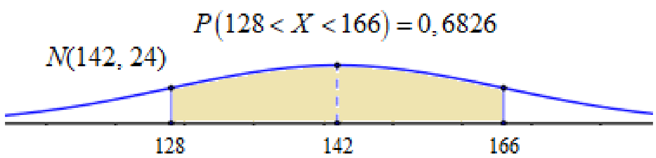
b) Para esa misma población de jóvenes, los niveles de colesterol se pueden estudiar como una normal $N(142, 24)$, pues según determinados estudios, la media de colesterol en mujeres menores de 19 años es de 142 mg/dL y su desviación típica es 24.

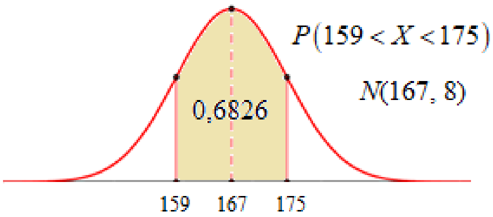
→ Los distintos valores de la media y de la desviación típica originan cambios en la curva, que se desplazará a izquierda o derecha o será más esbelta o más baja, como puede verse en la figura.



En la figura de arriba se representan las funciones de densidad asociadas a las distribuciones normales *N*(142, 24) y *N*(167, 8), correspondientes a los niveles de colesterol y a la estatura de chicas de 18 años. La curva *N*(142, 24) es más plana que la curva *N*(167, 8). Esto indica que la primera es una distribución más heterogénea: su desviación típica es mayor. En la normal *N*(167, 8) los datos están más agrupados, hay mayor homogeneidad.
Pero ambas son distribuciones normales, lo que significa que se mantienen las probabilidades teóricas. Esto es: $P(\mu-\sigma < X < \mu+\sigma) = 0,6826$ o $P(\mu-2\sigma < X < \mu+2\sigma) = 0,9544$; y así para cualquier número de desviaciones típicas que nos alejemos de la media. En concreto:
→ Para la *N*(142, 24), la probabilidad de que el nivel de colesterol de una chica, elegida al azar, esté entre 128 y 166 mg/dL (a una distancia máxima de σ = 24 puntos, de la media μ = 142), será:

$P(\mu-\sigma < X < \mu+\sigma)$ = $P(142 – 24 < X < 142 + 24)$ = $P(128 < X < 166) = 0,6826$.





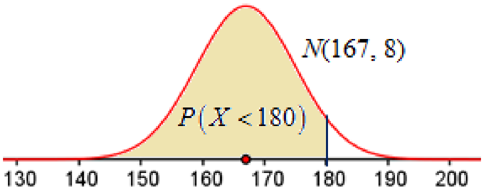
→ Para la *N*(167, 8), la probabilidad de que su estatura esté entre 159 y 175 cm (a una distancia máxima de σ = 8 cm de la media μ = 167), será:

$P(\mu-\sigma < X < \mu+\sigma)$ = $P(167 – 8 < X < 167 + 8)$ = $P(159 < X < 175) = 0,6826$.

(El área de las dos regiones coloreadas en ambas gráficas vale lo mismo: 0,6826).

Función de distribución

La función que asigna los valores de probabilidad "hasta un valor *x*" se llama <u>función de distribución</u>. Suele denotarse como $F(x)$ y es acumulativa, pues da el valor de probabilidad correspondiente al intervalo (–∞, *x*), que se mide por la superficie del recinto limitado por la función de densidad (la curva $y = f(x)$ definida más arriba) y el eje *OX*, desde –∞ hasta *x*.



Esto es: $F(x) = P(X < x)$ → probabilidad de que la variable *X* tome valores inferiores a *x*.

El área sombreada en la figura adjunta da el valor de $P(X < 180)$: la probabilidad de que una chica de 18 años mida menos de 180 cm de estatura.

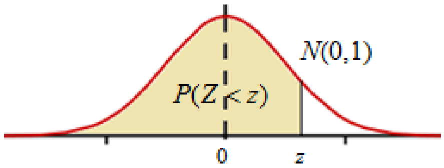
<u>Observaciones:</u>

1. Tradicionalmente la asignación de probabilidades se ha realizado con la **tabla normal *N*(0, 1**); así se hará aquí, siguiendo el proceso que se indicará en párrafo siguiente.
2. En la actualidad pueden encontrarse calculadoras que dan los valores de probabilidad requeridos introduciendo la media, la desviación típica y el valor de *x*. Se verá al final del tema.
3. En el punto **Uso de herramientas informáticas** se darán pautas para trabajar con GeoGebra y con Excel.

3. Distribución normal de media 0 y desviación típica 1: *N*(0, 1)

Si la media μ = 0 y la desviación típica σ = 1, la función de densidad es $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.

Para esta función se calcularon los valores de la función de distribución $F(x)$, que dan el área entre la gráfica de $f(x)$ y el eje *OX*, desde –∞ hasta *x*: la de la región sombreada en la figura adjunta. (Este proceso lo comenzó Moivre a partir de 1733; fue continuado por Laplace, Gauss, Pierce y Galton, entre otros.)



En la *N*(0, 1) se prefiere utilizar la letra *Z* en vez de *X*; luego, $F(z) = P(Z < z)$.

Esos valores se presentan en una tabla, secuenciados a intervalos de centésimas: 0,00; 0,01; 0,02…
La tabla normal *N*(0, 1) puede encontrarse fácilmente en internet; o en cualquier texto.

Tabla normal estándar: *N*(0, 1)

La tabla que se usará en este texto, de la que a continuación se muestra una parte de ella, da la probabilidad de que la variable *Z*, *N*(0, 1), tome valores entre –∞ y +z, $P(Z < z)$.

Los demás valores se obtienen teniendo en cuenta la simetría de la curva y que el área por debajo de la gráfica vale 1. Así, a partir del valor $P(Z < z)$, pueden obtenerse los valores de:

$P(Z > z)$, $P(Z < -z)$, $P(Z > -z)$, $P(0 < Z < z)$ y $P(-z < Z < z)$.

En esta tabla, la cifra de las unidades y de las décimas se muestran en la columna de la izquierda, la de las centésimas en la fila superior.

→ Los valores de *Z* están dados en desviaciones típicas. Así, *Z* = 1,26 indica 1,26 desviaciones típicas más que la media; *Z* = 0 indica que la variable se sitúa justamente en la media.

Z	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319

Ejemplos:

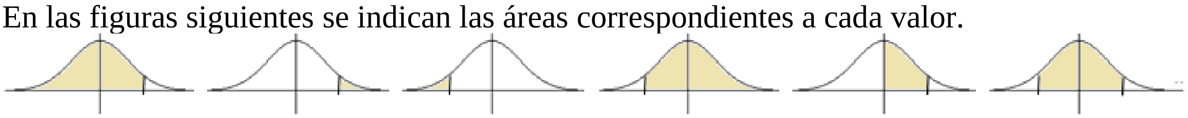
La probabilidad de que *Z* < 1,26 es: $P(Z < 1,26) = 0,8962$; valor dado en el cruce de líneas verdes.

A partir de ese valor se deducen:

$P(Z > 1,26) = 1 - 0,8962 = 0,1038$; $P(Z < -1,26) = 0,1038$; $P(Z > -1,26) = 0,8962$;

$P(0 < Z < 1,26) = 0,8962 - 0,5 = 0,3962$; $P(-1,26 < Z < 1,26) = 2 \cdot 0,3962 = 0,7924$.

En las figuras siguientes se indican las áreas correspondientes a cada valor.



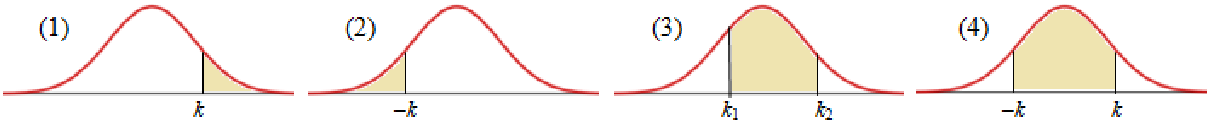
$P(Z < 1,26)$ $P(Z > 1,26)$ $P(Z < -1,26)$ $P(Z > -1,26)$ $P(0 < Z < 1,26)$ $P(-1,26 < Z < 1,26)$

→ En general se cumple:

(1) $P(Z>k)=1-P(Z<k)$; (2) $P(Z<-k)=P(Z>k)=1-P(Z<k)$, válido para todo k

(3) $P(k_1<Z<k_2)=P(Z<k_2)-P(Z<k_1)$;

(4) $P(-k<Z<k)=P(Z<k)-P(Z<-k)=P(Z<k)-\left[1-P(Z<k)\right]=2\cdot P(Z<k)-1$.



Ejemplos:
De la lectura de la tabla se obtiene:

a) $P(Z>0,8)=1-P(Z<0,8)=1-0,7881=0,2119$.

b) $P(Z<-075)=P(Z>0,75)=1-P(Z<0,75)=1-0,7734=0,2266$.

c) $P(-0,9<Z<1,25)=P(Z<1,25)-P(Z<-0,9)=0,8944-(1-P(Z<0,9))$ =

= 0,8944 – (1 – 0,8159) = 0,7103.

d) Como $P(Z<1)=0,8413$, puede deducirse que.

$P(-1<Z<1)=P(Z<1)-P(Z<-1)=0,8413-(1-0,8413)=0,6826$,

que es el valor de probabilidad que se apuntó anteriormente.

Cálculo del valor de *Z* a partir de su probabilidad asociada
La tabla normal se emplea también en sentido contrario, para hallar la abscisa (el valor de *Z*) correspondiente a una probabilidad determinada.
Esto es, igual que se sabe que $P(Z < 1) = 0,8413$, en sentido contrario la pregunta sería: ¿cuánto debe valer z para que $P(Z < z) = 0,8413$? La respuesta es evidente: el valor de z debe ser 1.

Ejemplos:
a) El valor de z_1 tal que $P(Z<z_1)$ = 0,9207 es z_1 = 1,41. Para determinarlo basta con buscar en la tabla normal el valor de *Z* correspondiente a una probabilidad de 0,9207: fila 1,4; columna 0,01.

b) Si la pregunta es: ¿cuánto debe valer Z para $P(-z_2<Z<z_2)=0,80$? Se procede así:

Como

$P(-z_2<Z<z_2)=P(Z<z_2)-P(Z<-z_2)=P(Z<z_2)-\left[1-P(Z<z_2)\right] \Rightarrow$

$\Rightarrow P(-z_2<Z<z_2)=2\cdot P(Z<z_2)-1$.

Así: $P(-z_2<Z<z_2)=0,80 \Rightarrow 2\cdot P(Z<z_2)-1=0,80 \Rightarrow P(Z<z_2)=0,90 \Rightarrow z_2=1,28$.

c) Un caso que se presenta con frecuencia es encontrar el intervalo $(-z, z)$ que contiene el 95 % de los datos de la variable estadística. Esto es, hallar el valor z tal que $P(-z<Z<z)=0,95$.

Por el ejemplo anterior:

$P(-z<Z<z)=0,95 \Rightarrow 2\cdot P(Z<z)-1=0,95 \Rightarrow P(Z<z)=0,9750 \Rightarrow z = 1,96$.

→ Si el valor de probabilidad no figura en la tabla se tomará el más cercano. Así se acaba de hacer en el ejemplo b): en la tabla el valor más cercano a 0,9000 es 0,8997. (Suponemos que una diferencia de 3 diezmilésimas no será problemática).
También puede optarse por la interpolación. Así, para $P(Z<z)=0,9950$ se toma z = 2,575, intermedio entre 2,57 y 2,58, cuyos valores de probabilidad respectivos son 0,9949 y 0,9951.

Tipificación

Las distribuciones normales con las que se trabaja en la práctica no son la estándar: la $N(0, 1)$. Son distribuciones con media μ (la que sea) y desviación típica σ: $N(\mu, \sigma)$.

Para calcular valores de probabilidad de una variable X, normal $N(\mu, \sigma)$, se hace el cambio de variable $Z = \dfrac{X-\mu}{\sigma}$, siendo Z la $N(0, 1)$.

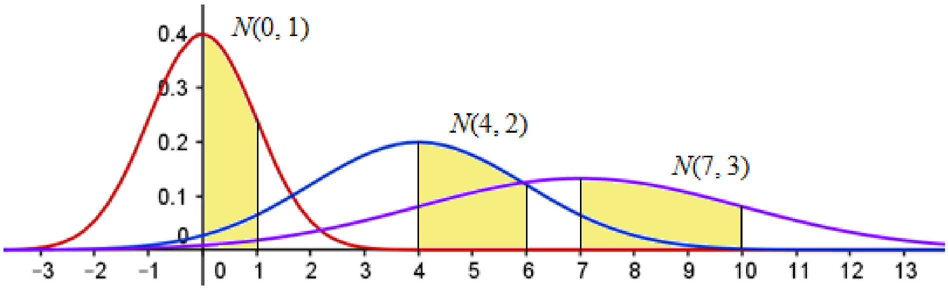
Con esto, las diferencias de X respecto de su media μ se "tipifican": se calculan en desviaciones típicas, pues lo significativo no es el valor que tome X, sino cuántas desviaciones típicas es mayor o menor que su media. (En cada caso se utiliza una unidad relativa que es la desviación típica σ).

Ejemplo:

En la normal $N(0, 1)$, el valor de $Z = 2$ (dos unidades típicas a la derecha de la media $\mu = 0$) está, relativamente, igual de alejado de la media 0, que en la normal $N(45, 7)$ el valor 59, que también es 2 desviaciones típicas mayor que $\mu = 45$, pues $59 = 45 + 2 \cdot 7$. Al tipificar, el valor $X = 59$ se transforma en $Z = 2$.

En efecto, para una $N(45, 7)$, si $X = 59$, el valor de Z tipificado es $Z = \dfrac{59-45}{7} = \dfrac{14}{7} = 2$.

- El proceso de tipificación se puede explicar gráficamente con ayuda de la siguiente figura. En ella, los recintos coloreados tienen la misma área.



→ En el caso de la $N(0, 1)$, el área coloreada mide la probabilidad de que la variable Z tome valores entre 0 y 1: $P(0 < Z < 1)$.

Esa probabilidad vale:

$$P(0 < Z < 1) = P(Z < 1) - P(Z < 0) = 0{,}8413 - 0{,}5000 = 0{,}3413.$$

→ Para la variable X, $N(4, 2)$, el área coloreada entre la curva y los valores $X = 4$ y $X = 6$, mide la probabilidad $P(4 < X < 6)$. Haciendo el cambio de variable $Z = \dfrac{X-4}{2}$ se tiene:

$$P(4 < X < 6) = P\left(\frac{4-4}{2} < Z < \frac{6-4}{2}\right) = P(0 < Z < 1) = P(Z < 1) - P(Z - 0) = 0{,}3413.$$

→ Y lo mismo para la variable X, $N(7, 3)$: el área coloreada entre la curva y los valores $X = 7$ y $X = 10$, mide la probabilidad $P(7 < X < 10)$, que también vale 0,3413. El cambio es $Z = \dfrac{X-7}{3}$.

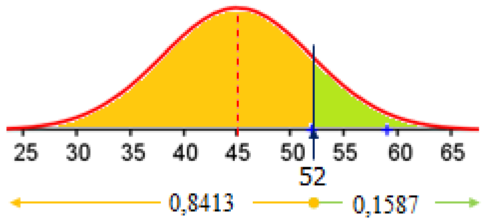
- En general, para la variable $X \approx N(\mu, \sigma)$, la probabilidad de que $X < k$, se calcula así:

$$P(X < k) = P\left(\frac{X-\mu}{\sigma} < \frac{k-\mu}{\sigma}\right) = P\left(Z < \frac{k-\mu}{\sigma}\right)$$, donde Z es $N(0, 1)$.

Ejemplos:

a) Si X es una variable normal $N(45, 7)$, la probabilidad de que X tome valores menores de 52, mayores de 52; o entre 52 y 59 es:

$$P(X<52)=P\left(Z<\frac{52-45}{7}\right)=P(Z<1)=0,8413;$$

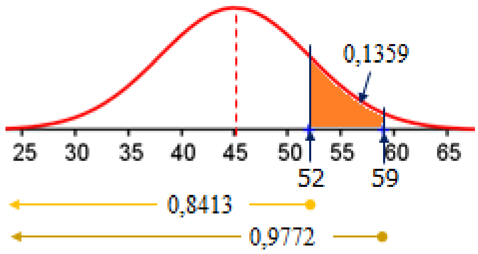


$$P(X>52)=1-P(Z<52)\ =1-0,8413=0,1587.$$

$$P(52<X<59)=P\left(\frac{52-45}{7}<Z<\frac{59-45}{7}\right)=$$

$$=P(1<Z<2)=P(Z<2)-P(Z-1)=$$

$$= 0,9772 – 0,8413 = 0,1359.$$



b) En sentido inverso, también se puede encontrar el valor de X a partir de la probabilidad asociada. Esto es, para la misma $N(45, 7)$, si se desea conocer el valor de k tal que, por ejemplo, $P(X<k)=0,9772$, se resuelve

$$P(X<k)=P\left(Z<\frac{k-45}{7}\right)=0,9772\Rightarrow\frac{k-45}{7}=2\Rightarrow k=45+2\cdot 7=59.$$

→ De manera análoga se procederá si se plantea determinar algún valor de k de especial relevancia; como, por ejemplo, el correspondiente al percentil 90: el valor de k tal que $P(X<k)=0,90$.

Se resuelve

$$P(X<k)=P\left(Z<\frac{k-45}{7}\right)=0,90\Rightarrow\frac{k-45}{7}=1,28\Rightarrow k=45+1,28\cdot 7=53,96.$$

(El valor más próximo a 0,9000 en la tabla normal es 0,8997, que está asociado a Z = 1,28).

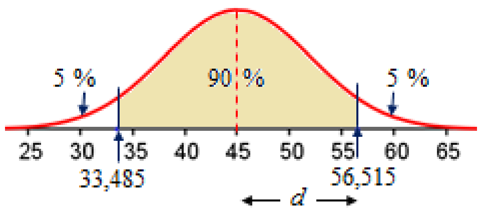
En consecuencia, todos los valores de X > 53,96 están por encima de percentil 90 de la distribución.

→ Otras veces se desea conocer un intervalo centrado en la media y que contenga un porcentaje determinado de los valores de la distribución. Por ejemplo, el intervalo que contiene el 90 % de los valores centrales de la distribución. Esto es, hallar el valor de d tal que $P(\mu-d<X<\mu+d)=0,90$.

Si el 90 % de los datos quedan dentro de ese intervalo, entonces, por cada uno de los lados (cola izquierda y derecha) quedará un 5 % de ellos:

$$P(X<\mu-d)=0,05 \text{ y } P(X>\mu+d)=0,05.$$

Lo que significa que $P(X<\mu+d)=0,95$.



Para la $N(45, 7)$ hay que encontrar el valor d tal que $P(X<45+d)=0,95$:

$$P(X<45+d)=P\left(Z<\frac{45+d-45}{7}\right)=0,95\ \Rightarrow\ \frac{d}{7}=1,645\Rightarrow d=11,515.$$

(El valor de probabilidad 0,95 no aparece en la tabla $N(0, 1)$, pero como es el intermedio de 0,9495 y 0,9505, asociados a Z = 1,64 y Z = 1,65, respectivamente, entonces, a 0,95 se le asocia Z = 1,645, la media aritmética de 1,64 y 1,65).

El intervalo buscado será: (45 – 11,515, 45 + 11,515) = (33,485, 56,515).
Luego: el 90 % de los valores centrales de la distribución está entre 33,485 y 56,515; el 5 % de los valores está por debajo de 33,485; el otro 5 % supera el valor 56,515.

4. Aproximación de la distribución binomial mediante una normal

Al estudiar la distribución binomial, $B(n, p)$, donde n indica el número de ensayos, p es la probabilidad de éxito en cada ensayo y $q = 1 - p$ la de fracaso, entonces, si X es la variable aleatoria que cuenta el número de éxitos en esos n ensayos, se cumple:

1. La media y varianza de la distribución $B(n, p)$ se obtienen a partir de sus parámetros, siendo:

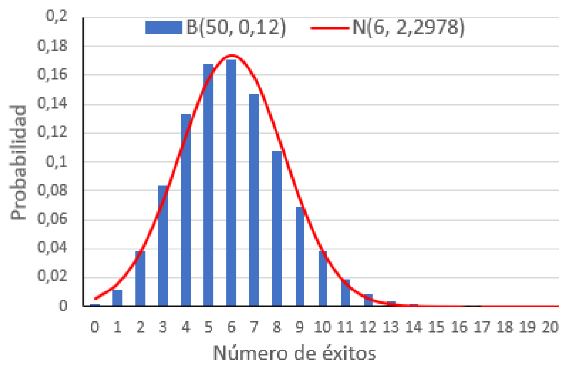
- Media: $\mu = n{\cdot}p$.
- Varianza: $\sigma^2 = n{\cdot}p{\cdot}q \Rightarrow$ la desviación típica vale $\sigma = \sqrt{npq}$.

2. La función de probabilidad que mide el número r de éxitos viene dada por:

$$P(X = r) = \binom{n}{r} p^r \cdot q^{n-r}, \quad r = 0, 1, 2, ..., n; \; q = 1 - p.$$

Condiciones para la aproximación

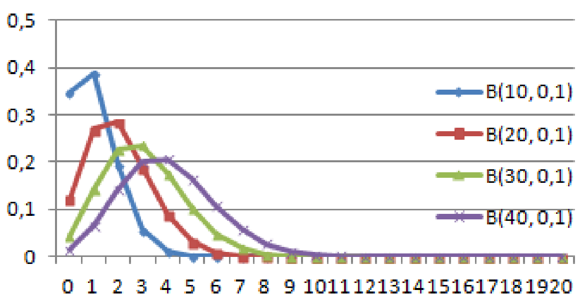
Para valores grandes de n el cálculo de probabilidades, aplicando la fórmula anterior, se hace engorroso; además, los valores de probabilidad pueden resultar muy pequeños. Así, por ejemplo, para la binomial $B(50, 0{,}12)$, las probabilidades para $r = 6$ y $r = 12$ son $P(X = 6) = 0{,}17118594$ y $P(X = 12) = 0{,}0084088$, respectivamente.

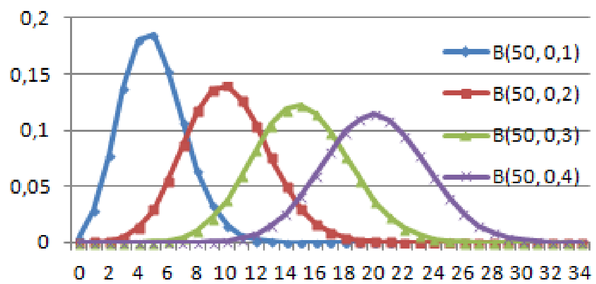


Más complicado resulta, por ejemplo, calcular la probabilidad de que $X > 10$ o de que $4 < X < 12$. Tales inconvenientes pueden obviarse, pues, para valores grandes de n, la distribución binomial se aproxima bastante bien mediante una distribución normal

En el gráfico adjunto, la altura de cada barra da la probabilidad de cada uno de los sucesos asociados a la $B(50, 0{,}12)$: $P(X = r)$. La línea continua representa la función de densidad de la normal de media $\mu = 50 \cdot 0{,}12 = 6$ y desviación típica $\sigma = \sqrt{50{\cdot}0{,}12{\cdot}0{,}88} \approx 2{,}2978$, $N(6, 2{,}2978)$.

En las siguientes figuras puede observarse la evolución de las poligonales de frecuencias de la distribución binomial $B(n, 0{,}1)$ cuando n va creciendo (figura izquierda), y de la distribución binomial $B(50, p)$ cuando p se acerca a 0,5 (figura derecha).





Como ves, cuando n aumenta la poligonal se parece más a una campana de Gauss; y es mucho más evidente cuando n es grande y p se acerca más a 0,5.

→ <u>En general, se admite que la aproximación es buena cuando $n \geq 25$ y el producto $np \geq 5$</u>.

Así, para las poligonales de la figura de la izquierda, la única que puede admitirse como aceptable es la $B(40, 0{,}1)$, aunque $np = 4$. En cambio, las binomiales representadas en la figura de la derecha se aproximan bastante bien a campanas de Gauss.

→ Cuando el ajuste sea posible, la distribución normal que mejor se aproxima a la $B(n, p)$ es la que tiene por media y desviación típica la de la distribución binomial. Como la media y desviación típica de la variable $X \approx B(n, p)$ son $\mu = np$ y $\sigma = \sqrt{npq}$, el ajuste se hace por la variable $X' \approx N\left(np,\ \sqrt{npq}\right)$, que se tipifica haciendo $Z = \dfrac{X'-\mu}{\sigma} = \dfrac{X'-np}{\sqrt{npq}}$.

Con esto, cuando sea preciso calcular probabilidades de una variable X binomial $B(n, p)$ puede hacerse recurriendo a la variable normal X' asociada. Por tanto:

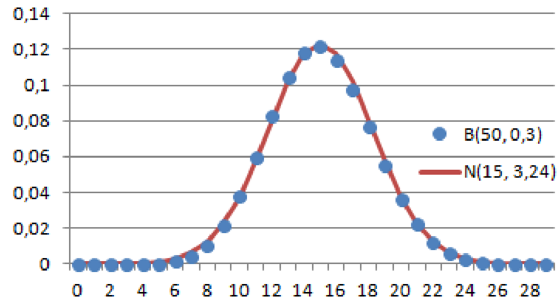
$$P(X < k) = P(X' < k) = P\left(Z < \frac{k-np}{\sqrt{npq}}\right).$$

Ejemplo:

En la figura adjunta se han dibujado la $B(50, 0{,}3)$ y la $N\left(50 \cdot 0{,}3,\ \sqrt{50 \cdot 0{,}3 \cdot 0{,}7}\right) = N(15,\ 3{,}24)$.

Las probabilidades binomiales se representan por puntos; la normal es la curva continua. Es evidente la gran coincidencia.

Por tanto, las probabilidades asociadas a la variable binomial $X = B(50, 0{,}3)$ pueden aproximarse mediante la variable normal $X' = N(15,\ 3{,}24)$. Así, por ejemplo:



$$P(X < 18) = P(X' < 18) = P\left(Z < \frac{18-15}{3{,}24}\right) = P(Z < 0{,}93) = 0{,}8238.$$

$$P(15 < X < 18) = P(X' < 18) - P(X' < 15) = P(Z < 0{,}93) - P(Z < 0) = 0{,}8238 - 0{,}5 = 0{,}3238.$$

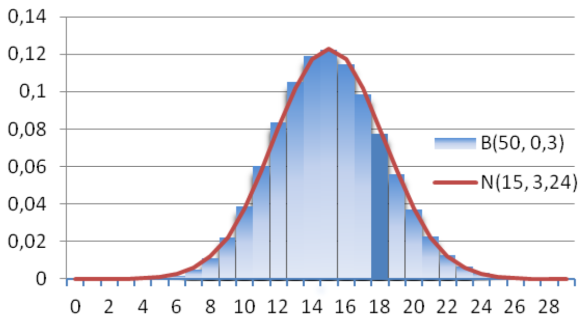
Corrección de continuidad

En la distribución normal la probabilidad de que la variable tome un valor concreto es 0, pero en una binomial no es así. Por seguir con el ejemplo anterior:

- Para la normal $X' = N(15,\ 3{,}24)$, $P(X' < 18) = P(X' \le 18)$, pues $P(X' = 18) = 0$.
- Para la binomial $X = B(50, 0{,}3)$, $P(X < 18) < P(X \le 18)$, pues $P(X = 18) = 0{,}077247062$, que

es el resultado de $P(X = 18) = \binom{50}{18} \cdot 0{,}3^{18} \cdot 0{,}7^{32}$ (puede hallarse con la calculadora).

Este valor coincide con el área de la barra sobre $X = 18$ (sombreada con mayor intensidad), que se muestra en el gráfico adjunto.



Esta anomalía puede paliarse mediante la llamada corrección de continuidad, por la que probabilidades puntuales (de valor 0 en las distribuciones continuas) son sustituidas por probabilidades de intervalo de la forma siguiente:

→ $P(X = k) = P(k - 0{,}5 < X' < k + 0{,}5)$

→ $P(X \le k) = P(X' < k + 0{,}5)$ → $P(X < k) = P(X' < k - 0{,}5)$

→ $P(X \ge k) = P(X' > k - 0{,}5)$ → $P(X > k) = P(X' > k + 0{,}5)$

Ejemplos:

Para la binomial $X = B(50,\ 0{,}3) \approx X' = N(15,\ 3{,}24)$, se tiene:

a) $P(X=15) = P(14{,}5 < X' < 15{,}5) = P\left(\dfrac{14{,}5-15}{3{,}24} < Z < \dfrac{15{,}5-15}{3{,}24}\right) = P(-015 < Z < 0{,}15) =$

$= P(Z < 0{,}15) - P(Z < -0{,}15) = 0{,}5596 - (1-0{,}5596) = 0{,}1192$.

b) $P(X=18) = P(17{,}5 < X' < 18{,}5) = P\left(\dfrac{17{,}5-15}{3{,}24} < Z < \dfrac{18{,}5-15}{3{,}24}\right) =$

$= P(Z < 1{,}08) - P(Z < 0{,}77) = 0{,}8599 - 0{,}7794 = 0{,}0805$.

c) $P(15 < X < 18) = P(X' < 17{,}5) - P(X' < 15{,}5) = P(Z < 1{,}08) - P(Z < 0{,}15) =$

$= 0{,}8599 – 0{,}5596 = 0{,}3003$.

Observación: En todos los casos, los valores de probabilidad obtenidos mediante la normal son aproximados. El resultado binomial exacto, obtenidos con Excel, es:

$P(X=15) = 0{,}12234686$; $P(X=18) = 0{,}07724706$. (La diferencia es de milésimas).

Ejercicio 2. Admitamos que el peso, en kg, de los habitantes adultos de una gran ciudad sigue una distribución normal de media 60 kg y desviación típica 5 kg. Si se elige una de las personas al azar, ¿qué probabilidad hay de que pese?:

a) Menos de 50 kg. b) Entre 52 y 65 kg.

c) Elegidas 100 personas, ¿cuántas cabe esperar que pesarán más de 65 kg?

Solución:

La variable X, que mide el peso de esas personas, se distribuye según la $N(60,\ 5)$. Se tipifica haciendo el cambio $Z = \dfrac{X-60}{5}$. Con esto:

a) $P(X < 50) = P\left(Z < \dfrac{50-60}{5}\right) = P(Z < -2) = 1 - P(Z < 2) = 1 - 0{,}9772 = 0{,}0228$.

b) $P(52 < X < 65) = P\left(\dfrac{52-60}{5} < Z < \dfrac{65-60}{5}\right) = P(-1{,}6 < Z < 1) = P(Z < 1) - P(Z < -1{,}6) =$

$= 0{,}8413 – (1 – 0{,}9452) = 0{,}7865$.

c) La probabilidad de que una persona pese más de 65 kg es:

$P(X > 65) = P\left(Z > \dfrac{65-60}{5}\right) = P(Z > 1) = 1 - P(Z < 1) = 1 - 0{,}8413 = 0{,}1587$.

Entonces, para 100 personas: $100 \cdot 0{,}1587 = 15{,}87 \approx 16$ pesarán más de 65 kilos.

Ejercicio 3. En un determinado país, se ha estimado en un 22 % el porcentaje de hombres con problemas de alopecia (calvicie). Si se toma una muestra aleatoria de 40 hombres:

a) ¿Cuál es la probabilidad de que exactamente 10 de ellos tengan problemas de alopecia?

b) ¿Y de que haya más de 10?

Solución:

La variable X que computa el número de personas con alopecia es una variable $B(40,\ 0{,}22)$ que puede aproximarse por la normal X': $N(40 \cdot 0{,}22, \sqrt{40 \cdot 0{,}22 \cdot 0{,}78}\,) = N(8{,}8,\ 2{,}62)$. Con esto, haciendo la corrección de continuidad y tipificando:

a) $P(X=10) = P(9{,}5 < X' < 10{,}5) = P\left(\dfrac{9{,}5-8{,}8}{2{,}62} < Z < \dfrac{10{,}5-8{,}8}{2{,}62}\right) = P(0{,}27 < Z < 0{,}65) =$

$= 0{,}7422 – 0{,}6064 = 0{,}1358$.

b) $P(X > 10) = P(X' > 10{,}5) = P(Z > 0{,}65) = 1 - P(Z < 0{,}65) = 1 - 0{,}7422 = 0{,}2578$.

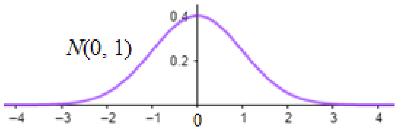
5. Uso de herramientas computacionales

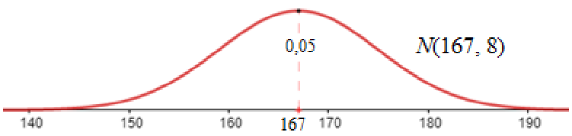
El cálculo de los valores de probabilidad a partir de la Tabla $N(0, 1)$, tal y como se ha hecho aquí, resulta engorroso, pero, de momento, hay que seguir estudiándolo así. Aunque, como sucedió con las tablas trigonométricas y logarítmicas, la tabla $N(0, 1)$ quedará en desuso, pues ya son muy accesible las calculadoras y ordenadores que facilitan los cálculos. No obstante, vuelvo a advertir que lo importante son los conceptos; en este caso, el comportamiento de las distribuciones normal y binomial, conocer lo que estás haciendo y saber interpretar los resultados.

1. Con **GeoGebra** puede obtenerse:

- La representación gráfica de la $N(\mu, \sigma)$ y el valor de probabilidad para cualquier valor de X.

Para la curva $N(0, 1)$, teclea: Normal(0, 1, x, false). (La escala de ejes EjeX : EjeY puede ser 5 a 1).
Para la curva $N(167, 8)$, teclea: Normal(167, 8, x, false). (EjeX : EjeY → 300 a 1).
Así, se obtienen las gráficas que siguen:





- Los valores de probabilidad. Así, por ejemplo:

Para la curva $N(0, 1)$, $P(Z < 1,7)$: teclear Normal(0, 1, 1.7, true); se obtiene 0,9554.

Para la curva $N(167, 8)$, $P(X < 180)$: teclear Normal(167, 8, 180, true); se obtiene 0,9479.

2. Algunas calculadoras dan los valores de probabilidad requeridos, introduciendo la media, la desviación típica y el valor de x. Por ejemplo, en este enlace puedes obtener esos valores.

Para la $N(0, 1)$, $P(-0,5 < Z < 1,5) = 0,624731553$.

Para la $N(167, 8)$, $P(X < 175) = 0,158447781$.

Para la $N(0, 1)$, $P(Z < z_a) = 0,82 \Rightarrow z_a = 0,9153$.

Para la $N(167, 8)$, $P(X < k) = 0,90 \Rightarrow k = 177,2512$, que da la altura mínima del 10 % de las chicas más altas. El 90 % de esas chicas mide menos de 177,25 cm.

3. Con **Excel** se han elaborado los gráficos de carácter binomial de este tema. Para ello es necesario introducir los datos relacionados con la $B(n, p)$ en cuestión. Esos datos se generan indicando los comandos oportunos. Así, por ejemplo, para la $B(50, 0,3)$, vista anteriormente, hay que confeccionar la tabla adjunta. Para ello:

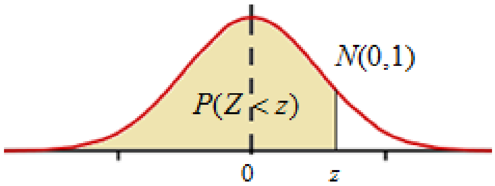
	A	B
1	B(50, 0,3)	
2	éxitos	probab.
3	0	1,7985E-08
4	1	3,8539E-07
5	2	4,0465E-06
6	3	2,7748E-05
7	4	0,00013973

→ En A3 escribir 0; en A4 anotar =A3+1. Arrastrar A4 hacia abajo, hasta llegar a 50 (en A53). Se genera la columna A del número de éxitos.
→ En B3 escribir =DISTR.BINOM(A3;50;0,3;FALSO), aparece el valor 1,7985E-08. Arrastrar B4 hacia abajo, hasta llegar a B53, 5,6329E-27. Se genera la columna B.
→ Abarcando las dos columnas de datos se elige uno de los gráficos recomendados.
(El valor de probabilidad correspondiente a 15 éxitos es 0,12234686; puede obtenerse de forma independiente escribiendo =DISTR.BINOM(15;50;0,3;FALSO)).

Nota. Los valores de probabilidad pueden diferir (en la cifra de las diezmilésimas) si se obtienen a partir de la tabla normal o con ordenador. Así, por ejemplo:
Con la tabla $N(0, 1)$: $P(-1 < Z < 1) = 0,6826$. Con ordenador: $P(-1 < Z < 1) = 0,68277703$.

ÁREAS BAJO LA DISTRIBUCIÓN DE PROBABILIDAD NORMAL ESTÁNDAR

Los valores en la tabla representan el área bajo la curva normal hasta un valor positivo *z*.



Z	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

Ejercicio 4. Utilizando la tabla normal *N*(0, 1) calcula el valor de *k* tal que:

a) $P(Z<1{,}5)=k$ b) $P(Z<-0{,}5)=k$ c) $P(Z<k)=0{,}67$

Solución:

a) $P(Z<1{,}5)=0{,}9332 \Rightarrow k=0{,}9332$.

b) $P(Z<-0{,}5)=1-P(Z<0{,}5)=1-0{,}6915=0{,}3085 \Rightarrow k=0{,}3085$.

c) $P(Z<k)=0{,}67 \Rightarrow k=0{,}44 \rightarrow$ cruce de la fila **0,4** con la columna **0,04**.

PROBLEMAS PROPUESTOS

1. Sea Z una variable normal estándar, $N(0, 1)$; halla las probabilidades:
a) $P(Z \le 2{,}22)$. b) $P(Z \le -2{,}22)$. c) $P(-1{,}5 < Z < 3)$.

2. Utilizando la Tabla Normal estándar, $N(0, 1)$, halla las probabilidades:
a) $P(-1 < Z < 1)$. b) $P(-2 < Z < 2)$. c) $P(-3 < Z < 3)$.
Halla los mismos valores de probabilidad utilizando recursos informáticos.

3. Siendo X una variable que se distribuye $N(4,\ 1{,}5)$.
a) Tipificando la variable calcula las probabilidades: (1) $P(X < 7)$; (2) $P(X < 5{,}5)$; (3) $P(X < 1{,}5)$.
b) (Optativo) Halla esas mismas probabilidades directamente con ordenador.

4. Si X es una variable continua $N(28, 5)$, halla:
a) $P(X > 31)$. b) $P(28 < X < 35{,}5)$. c) $P(20 < X < 38)$.

5. Utilizando la tabla normal $N(0, 1)$, determina el valor de k que cumple:
a) $P(Z < k) = 0{,}9115$. b) $P(Z < k) = 0{,}9452$.
c) $P(Z < k) = 0{,}1587$. d) $P(Z < k) = 0{,}95$.
(Optativo) Halla esos mismos valores de k utilizando recursos informáticos.

6. Para una distribución normal $N(60, 5)$, determina el valor de k que cumple:
a) $P(X < k) = 0{,}90$. b) $P(X > k) = 0{,}95$. c) $P(60 - k < X < 60 + k) = 0{,}9544$.
Halla esos mismos valores de k utilizando recursos informáticos.

7. Las calificaciones de un grupo numeroso de estudiantes que han realizado un examen siguen una distribución normal de media 20 y desviación típica 10. Calcula:
a) La probabilidad de que un estudiante elegido al azar obtenga una calificación entre 15 y 25.
b) La calificación que solo superan o igualan el 20 % de los estudiantes.

8. En una distribución normal, halla el porcentaje de valores que distan de la media:
a) Menos de 1,2 desviaciones típicas.
b) Entre 0,5 y 1 desviación típica.

9. Supongamos que la estatura media de las alumnas de bachillerato se distribuye normalmente con media μ = 166 cm y desviación típica 9 cm. Si se elige una alumna al azar halla la probabilidad de que su estatura sea:
a) Superior a 175 cm. b) Inferior a 155 cm. c) Esté entre 155 cm y 175 cm.

10. Las alturas de 500 estudiantes varones están distribuidas normalmente con media 172 cm y desviación típica 12 cm. Aproximadamente, ¿cuántos estudiantes tienen una altura?
a) Igual a 170 cm. b) Menor que 170 cm. c) Entre 175 y 1 90 cm.

11. La longitud de cierto tipo de peces sigue una distribución normal de media 100 mm y desviación típica 9 mm. ¿Cuál es la probabilidad de que uno de esos peces mida entre 82 mm y 91 mm?

12. Una compañía de telefonía fabrica teléfonos móviles cuya vida útil (obsolescencia programada) se ajusta a una distribución normal de media 30 meses y desviación típica 3 meses. Elegido un teléfono móvil al azar:
a) Calcula la probabilidad de que funcione bien durante más de 32 meses.
b) ¿Qué porcentaje de teléfonos móviles funcionan bien entre 24 y 36 meses?
c) La compañía garantiza el correcto funcionamiento del teléfono durante un tiempo *t* (en meses); si se estropea antes, la compañía ofrece un teléfono nuevo. Si se desea afrontar un máximo de un 1 % de reposiciones, ¿cuál será el tiempo de garantía?

13. El diámetro de las ciruelas de una determina variedad se distribuye normalmente con media 4,5 cm y desviación típica 0,3 cm. Si se desea seleccionar, para su exportación, el 10 % de las más grandes, ¿a partir de qué tamaño hay que cogerlas?

14. El peso de los varones adultos de cierta comunidad sigue una normal de media $\mu = 75$ kg y desviación típica $\sigma = 5$ kg.
a) Si se elige al azar un hombre adulto de esa población, halla la probabilidad de que tenga un peso entre 70 y 80 kg.
b) Si se admite que un adulto de esa población tiene sobrepeso si pesa más de 82 kg, ¿qué porcentaje de esa población tiene sobrepeso?
c) ¿Cuál es el peso máximo para estar entre el 25 % de los más delgados en esa población?

15. La temperatura de una ciudad en verano se ajusta a una distribución normal de media 30 grados centígrados (ºC) y desviación típica 6 ºC. Si se elige un día de verano al azar, calcula la probabilidad de que:
a) Su temperatura supere los 39 ºC.
b) Su temperatura esté entre 25º y 30º.
c) ¿A partir de qué temperatura uno de esos días estará entre el 10 % de los días más calurosos?

16. Los resultados de un examen realizado a 500 opositores se ajustan a una distribución normal de media 40 puntos y desviación típica 10 puntos.
a) ¿Qué porcentaje de opositores obtuvo una puntuación entre 30 y 60 puntos?
b) Si solo hay 100 plazas, que se asignan a los 100 mejores resultados de ese examen, ¿cuántos puntos hay que obtener para conseguir una de esas plazas?

17. La edad de los habitantes de cierta ciudad se distribuye normalmente, con una media de 40 años. Se sabe además que el 2,28 % de los habitantes tiene más de 60 años.
a) ¿Cuál es la desviación típica?
b) ¿Cuál es el porcentaje de habitantes con menos de 35 años?

18. Según datos de la Compañía Metropolitana de una ciudad, el tiempo de espera de los usuarios del metro de esa ciudad sigue una distribución de media 7,5 minutos con 2 minutos de desviación típica. De acuerdo con esa distribución, qué porcentaje de usuarios esperarán:
a) Más de 9 minutos. b) Entre 7 y 10 minutos.

19. Se sabe que el cociente intelectual de la población de un país sigue una distribución normal de media 100 y desviación típica 20, $N(100, 20)$.
a) ¿Qué porcentaje de esa población tendrá un cociente intelectual entre 95 y 105?
b) Si se considera que una persona es superdotada cuando su cociente intelectual es mayor que 160, ¿qué porcentaje de superdotados hay en esa población?

20. Si X es variable $N(\mu, \sigma)$ y se tiene que $P(X < 4) = 0,2549$ y $P(X < 7) = 0,9082$, halla los valores de μ y σ.

21. Se supone que la estatura de una población sigue una distribución normal, $N(\mu, \sigma)$. Sabiendo que el 1 % de los individuos de esa población supera los 185 cm de estatura, mientras que el 3 % no llega a los 160 cm, calcula la media y la desviación típica de esa distribución.

22. Mediante la aproximación normal de la binomial $B(50, 0{,}12)$ calcula:
a) $P(X=6)$. b) $P(X=12)$. c) $P(6<X\le 12)$.
(Optativo) Halla esas mismas probabilidades utilizando Excel.

23. En una determinada ciudad, el 20 % de sus habitantes tiene los ojos azules. Si se eligen 40 personas al azar, calcula la probabilidad de que:
a) Exactamente 6 de ellos tengan los ojos azules.
b) Más de 10 tengan los ojos azules.

24. Supongamos que el 80 % de los estudiantes de bachillerato posee ordenador personal.
a) Si se eligen al azar 100 estudiantes de bachillerato, ¿cuál es la probabilidad de que más de 75 tengan ordenador?
b) Si se eligen al azar 225 estudiantes de bachillerato, ¿cuál es la probabilidad de que tengan ordenador entre 170 y 190 estudiantes, ambos valores incluidos?

25. En una ciudad, el 45 % de los mayores de 60 años dice estar "bastante bien de salud". Si en esa ciudad se eligen al azar 100 personas mayores de 60 años:
a) Indica cómo podría calcularse la probabilidad de que, entre esas 100 personas, "exactamente *r* estén bastante bien de salud".
b) Con ayuda de calculadora o de ordenador, halla la probabilidad de entre esas 100 personas, "exactamente 40 estén bastante bien de salud".
c) Calcula dicha probabilidad aproximándola mediante una distribución normal.

26. El 42 % de los habitantes de un pueblo pasa cada día por la calle mayor. Elegidos 60 habitantes al azar, ¿qué probabilidad hay de que más de 30 de ellos pasen ese día por la calle mayor?

27. Un examen de respuesta múltiple consta de 80 preguntas, cada una con 4 opciones, una de ellas correcta y erróneas las otras tres. Si un estudiante contesta al azar, ¿cuál es la probabilidad de que acierte 25 o más preguntas? ¿Y menos de 10?

28. A lo largo del mes de noviembre, la probabilidad de que un día llueva en una ciudad del norte de España es 0,4.
a) ¿Cómo puede describirse y estudiarse el modelo de probabilidad que proporciona los días de lluvia durante ese mes?
b) Halla la probabilidad de que durante ese mes llueva exactamente 8 días.
c)) Halla la probabilidad de que durante ese mes llueva entre 10 y 20 días.