INSTITUTO de ENSEÑANZAS a DISTANCIA de ANDALUCÍA

PAU

Mayores de 25 años

Contenidos

Matemáticas Aplicadas a las Ciencias Sociales
Estadística y Probabilidad: Distribuciones de probabilidad. Variable continua.

1. Variable aleatoria continua

Lanzamos una moneda 400 veces, y nos preguntamos por la probabilidad de que salgan justamente 169 caras y 231 cruces.

Recordemos lo que estudiamos en el tema de las variables aleatorias discretas, y en concreto en la distribución binomial. Sabemos que estamos ante una **B(400, 0.5)**, y nos preguntamos conocer la probabilidad p(X=169).

Si aplicamos la fórmula, tenemos que dicha probabilidad es igual a $\binom{400}{169} \cdot 0.5^{169} \cdot 0.5^{231}$. Haciendo uso de una calculadora, incluso una hoja de cálculo, nos dará error, ya que el número combinatorio es inmensamente grande, y las potencias de 0.5 son números decimales con muchos ceros.

Gracias a algunas aplicaciones como GeoGebra podemos saber que esa probabilidad es 0.0003. Un número pequeño, pero si nos alejamos de un número cercano a 200 caras o cruces, dicha cantidad seguirá disminuyendo.

Y es que en un número tan grande de lanzamientos no tiene sentido preguntarnos por una cantidad concreta de caras o cruces, sino por un intervalo de apariciones. Por ejemplo, **cuál es la probabilidad de que obtengamos más de 200 caras**.

Imagen de geralt en Pixabay. Pixabay License

En este tema estudiaremos distribuciones con funciones de probabilidad que no tiene sentido considerarlas puntualmente, sino en intervalos.

1.1. Definición

En octubre de 2012 el Instituto Nacional de Estadística (INE) publicaba una encuesta sobre la estructura salarial entre los trabajadores del estado español. De los datos recogidos se extraía el siguiente gráfico. En el eje horizontal se representa el **sueldo anual** y en el vertical el **número de asalariados**.

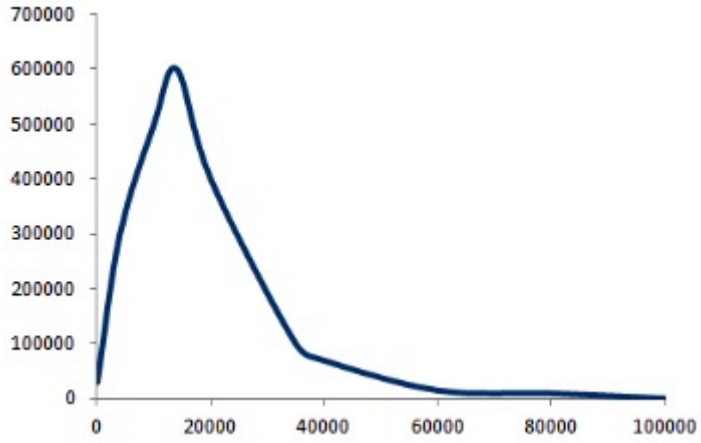


Imagen de elaboración propia

Podemos considerar la **variable aleatoria continua** X que a cada sueldo mensual le asocia el número de asalariados que tienen ese sueldo. La gráfica anterior sería la **función de probabilidad** de X.

En las distribuciones continuas tiene sentido preguntarse por la probabilidad de un intervalo de la variable, pero no un valor puntual de ella.

Por ejemplo, por el estudio del INE sabemos que el **salario mediano** es de 19.000 euros, por tanto podemos afirmar que $p(X < 19000) = 0.5$, ya que el valor mediano divide a una distribución justamente por la mitad.

Dicho de otro modo, el área que hay bajo la gráfica de la función desde el salario 0 y hasta el salario 19.000 euros es igual que el área que hay bajo la curva desde 19.000 y hasta 100.000 euros, por tanto también se tiene que $p(X > 19000) = 0.5$

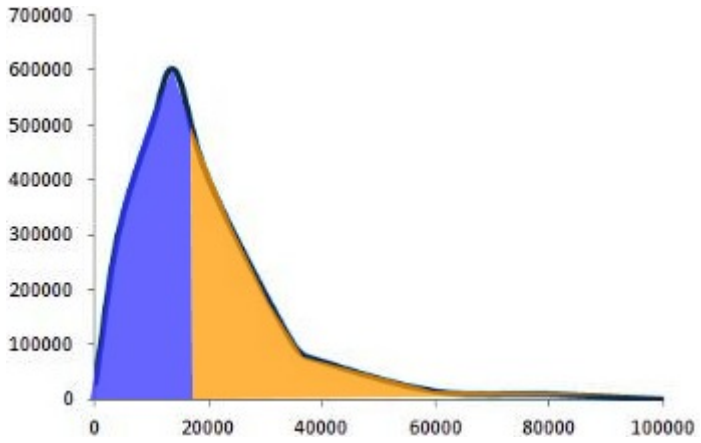


Imagen de elaboración propia

Importante

Una variable aleatoria continua X queda determinada por una función f que llamaremos **función de probabilidad** o **función de densidad**.

f debe ser siempre mayor o igual que cero y el área debajo de su gráfica tiene que ser 1.

La probabilidad estará definida como el área delimitada por la gráfica de f y el eje horizontal.

En el caso de $X=a$, dicha probabilidad es cero, $p(X=a)=0$, ya que no se encierra ningún área.

Para un intervalo, la $p(a\leq X\leq b)$ será igual al área que hay bajo la gráfica de f en el intervalo $[a,b]$.

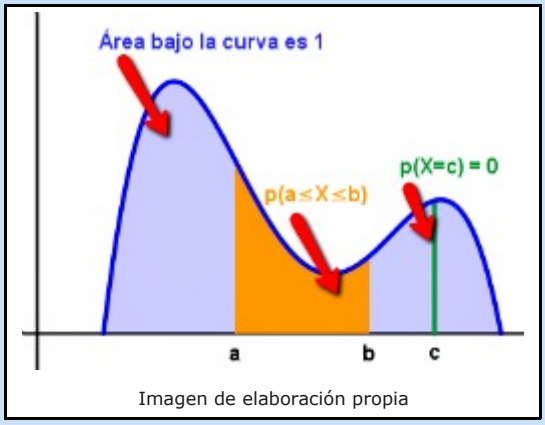


Imagen de elaboración propia

Como acabamos de ver la $p(X=a)=0$, si X es una variable aleatoria continua. Por tanto, el hecho de que los extremos de un intervalo estén o no incluidos en la probabilidad que hay que calcular, no afecta su valor.

Es decir $p(a<X<b)=p(a\leq X\leq b)=p(a\leq X<b)=p(a<X\leq b)$, ya que la probabilidad de que X sea igual a a o b es cero.

Ejercicio resuelto

Una línea de autobuses pasa cada 20 minutos. Consideramos la variable aleatoria X que determina el tiempo que hay que esperar, en función del tiempo que hace que pasó el autobús la última vez.

La gráfica de su función de densidad sería:

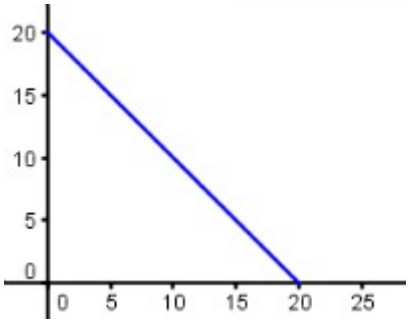


Imagen de elaboración propia

Determina las siguientes probabilidades:

$$p(5<X<10) \text{ y } p(X\geq 10)$$

Lo que se nos pregunta es la probabilidad de tener que esperar entre 5 y 10 minutos, y la de tener que esperar al menos 10 minutos.

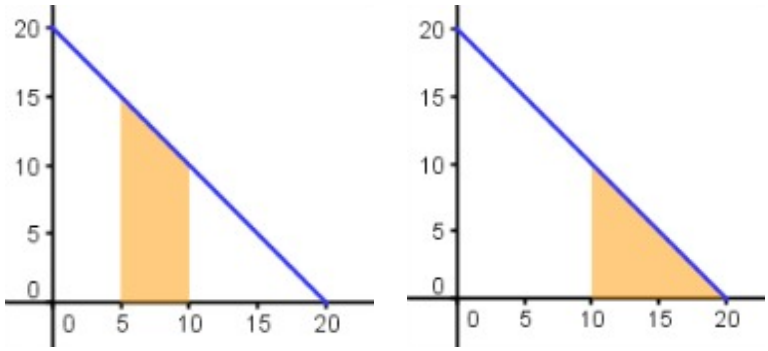
la de tener que esperar al menos 10 minutos.

Mostrar retroalimentación

Como la gráfica del tiempo de espera corresponde a una función de densidad, el área que hay debajo de ella será igual a la unidad. Por tanto, consideramos que el área del triángulo es 1.

$p(5<X<10)$ es igual al área encerrada debajo de la gráfica en el intervalo $[5,10]$. Si el área total es 1, el de la franja debajo de la gráfica correspondiente al intervalo $[5,10]$ es $\frac{5}{16}=0.3125$. La mejor forma de hallarlo es dividir el triángulo rectángulo en 16 triángulos rectángulos pequeños, y ver cuántos corresponden a las franjas sombreadas.

Para hallar $p(X\geq 10)$, basta observar que dicha franja está compuesta 4 triángulos pequeños, por tanto dicha probabilidad será igual a $\frac{4}{16}=0.25$.



Imágenes de elaboración propia

La **media,** μ, y la **desviación típica,** σ, indican los mismos parámetros que en las distribuciones estadísticas. La media indica el centro de gravedad de la distribución y la desviación típica la medida de dispersión.

Comprueba lo aprendido

Responde verdadero o falso a las siguientes cuestiones que se te plantean.

Si X es una variable aleatoria continua, entonces:

1. $p(X=5)=0.2$

○ Verdadero ○ Falso

Falso

En una variable aleatoria continua, la probabilidad de un valor concreto siempre vale 0, sea cual sea ese valor.

2. $p(X>5)<0$

○ Verdadero ○ Falso

Falso

La función de probabilidad nunca puede ser negativa.

3. El área máxima de la región encerrada entre la gráfica de su función de densidad y el eje OX vale 1.

○ Verdadero ○ Falso

Verdadero

Es una de las dos características principales que cumple la función de densidad. Es así, porque la probabilidad máxima o total de cualquier variable vale 1.

4. $p(X>5)=p(X\geq 5)$

○ Verdadero ○ Falso

Verdadero

No hay más que recordar que $p(X=5)=0$.

5. Su función de densidad puede tomar valores por debajo del eje OX , es decir, valores negativos.

○ Verdadero ○ Falso

Falso

La segunda característica fundamental de la función de densidad es que es una función positiva, es decir, está siempre sobre o por encima del eje OX .

2. Distribución Normal

Las distribuciones de probabilidad de una variable aleatoria continua no son más que idealizaciones de distribuciones estadísticas asociadas a experimentos aleatorios.

Hemos visto que experimentos que en un principio están asociados a variables discretas se van transformando en continuos cuando, por ejemplo, el número de monedas lanzadas aumenta o el de niveles de una máquina de Galton es grande.

Y en ambos casos, los rectángulos de las funciones de probabilidad se transforman en gráficas con una silueta muy similar y parecida a una campana. De ahí el nombre que en algunos ocasiones se le suele dar, campana de Gauss.

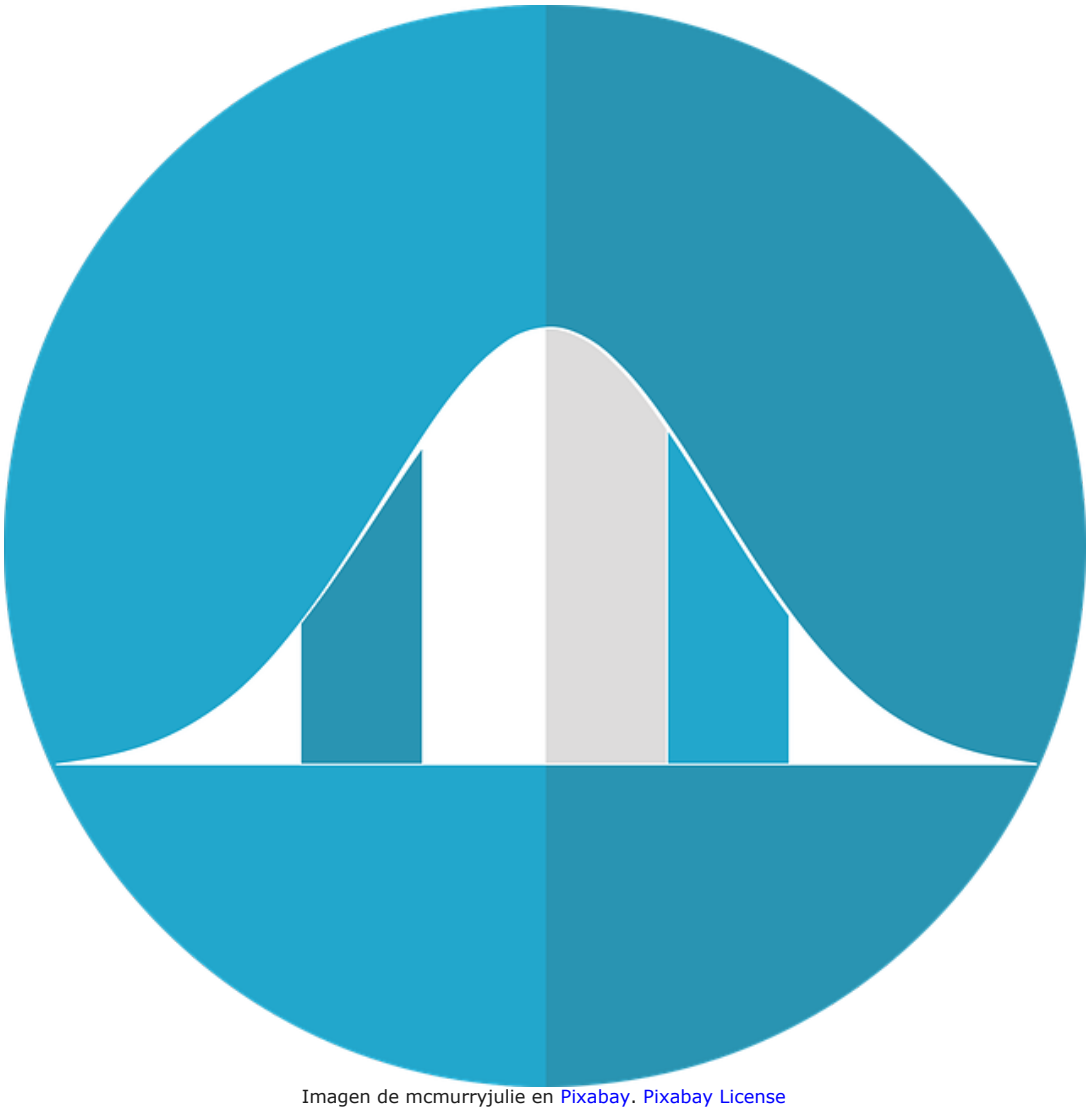


Imagen de mcmurryjulie en Pixabay. Pixabay License

La distribución que tiene por gráfica esa "campana" se denomina **Normal** y a ella se adaptan infinidad de situaciones. Medidas referidas a cualquier tipo de población: alturas, pesos, tallas, volúmenes; órbitas en donde se encuentran los electrones; estudios de errores cometidos en muchos procesos de medición, tanto en la ciencia como en la industria; análisis sobre la utilidad de nuevos fármacos...

De nuevo, las matemáticas modelizan y sirven de esqueleto ideal de multitud de aspectos de la realidad.

2.1. Definición y parámetros

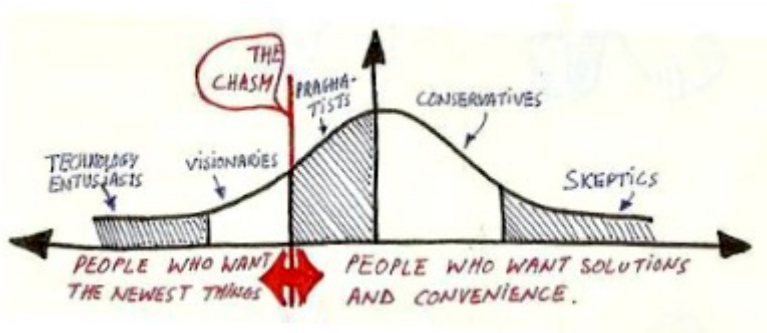


Imagen en Flickr de Lick Galoppin bajo CC

Importante

La función de densidad de una variable **aleatoria Normal** es la curva Normal o campana de Gauss. Su gráfica es una función continua, simétrica, cuyo máximo coincide con la **media de la distribución**, μ .

Para cada valor μ de la media y cada valor σ de la desviación típica, existe una curva Normal y su distribución asociada que llamaremos $N(\mu, \sigma)$.

La expresión analítica de la gráfica de la distribución Normal de media μ y desviación típica σ es:

$$\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

En la siguiente escena de Geogebra puedes trabajar con las propiedades de la gráfica de la Normal que hemos mencionado.

Mostrar retroalimentación

Recuerda que al ser 0 la probabilidad de que X esté en un punto, no tendremos en cuenta los extremos.

1. Nos preguntan por la probabilidad de que X esté en el intervalo $(1.7\,,\,2.3)$.

Ese intervalo coincide con $(2-0.3\,,\,2+0.3)$ es decir "la media más/menos una vez la desviación típica". Por tanto la probabilidad será igual a 0.68: $p(1.7 < X \leq 2.3) = 0.68$

2. Tenemos que hallar la probabilidad de que X sea mayor a 2.6 .

$2.6 = 2+2 \cdot 0.3$, es decir, "la media más dos veces la desviación típica". Mirando la gráfica vemos que la probabilidad del intervalo definido por dos veces la desviación típica es igual a 0.954 , por tanto la probabilidad de los dos intervalos extremos que quedan fuera es de 0.046 . Por la simetría de la gráfica, sabemos que esos dos intervalos son iguales, por lo que la probabilidad de cada uno será 0.023 , es decir, $p(2.6<X) = 0.023$

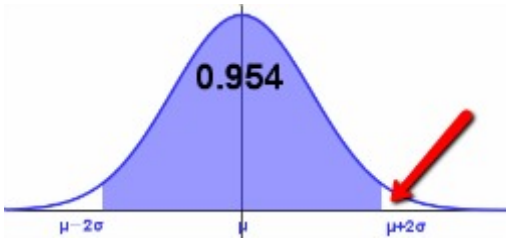


Imagen de elaboración propia

3. Nos piden la probabilidad de que X sea menor que 1.1 .

$1.1 = 2 - 3 \cdot 0.3$, es decir "la media menos tres veces la desviación típica". Razonando del mismo modo que en el apartado anterior, la probabilidad del intervalo definido por tres veces la desviación típica es igual a 0.997 . La probabilidad de los dos intervalos extremos que quedan fuera es de 0.003 , luego, por simetría, la de cada uno será 0.0015 . Tenemos entonces que $p(X \leq 1.1) = 0.0015$

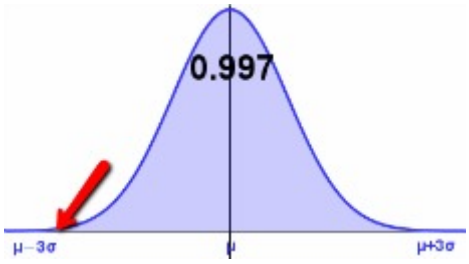


Imagen de elaboración propia

Hemos visto que para calcular la probabilidad de un intervalo tenemos que hallar el área bajo la curva de la Normal en dicho intervalo. La herramienta principal para hallar áreas en matemáticas es la integral definida. Pero el cálculo integral no es motivo de este curso. Nosotros nos ayudaremos de una herramienta más "casera", **las tablas de la distribución Normal**. No te dejes impresionar por el nombre, ya verás que son tablas muy livianas.

2.2. N(0,1)

Imagen en Flickr de decar66 bajo CC

Hasta hace unos años parte del aprendizaje de las matemáticas incorporaba el manejo de tablas para realizar ciertos tipos de cálculos. Por ejemplo, para hallar las razones trigonométricas de un ángulo o conocer el logaritmo decimal de un número.

Tablas llenas de números y números llenos de decimales.

Por suerte, la incorporación de calculadoras científicas, hojas de cálculo y programas de cálculo simbólico han permitido dejar atrás este tipo de trabajo tan rutinario.

Pero, por ahora, aún es necesario ser diestro en el manejo de las **tablas de la Normal** para hallar las probabilidades de experimentos que se ajusten a este tipo de variable aleatoria.

A eso es lo que dedicaremos este apartado, a dar la instrucciones para manejar las tablas de la **Normal estándar**, es decir, la Normal de media 0 y desviación típica 1: $N(0,1)$.

Aún así, y si disponemos de una hoja de cálculo a mano, podremos hallar probabilidades de la Normal sin necesidad de recurrir a las citadas tablas.

En la siguiente imagen se explica cómo se puede hallar con Excel la $p(X \leq 1.4)$ si X es una $N(0,1)$.

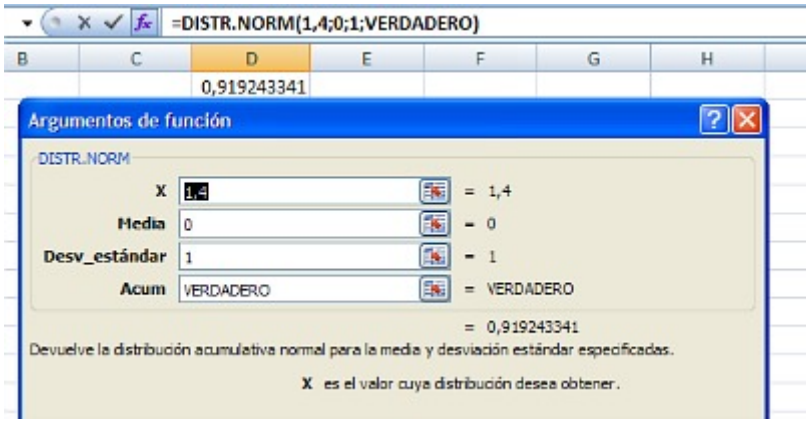


Imagen de elaboración propia

La función que hay que utilizar es **DIST.NORM**, y tendremos que indicar el valor para que el que queremos hallar la probabilidad, la media, la desviación típica de la distribución, e indicar que lo que deseamos es hallar la probabilidad acumulada hasta ese valor.

Importante

Llamaremos **Normal estándar** a la distribución Normal de media 0 y desviación típica 1, $N(0,1)$.

La designaremos por Z.

En la siguiente presentación de Jesús Fernández podemos ver lo pasos que hay que seguir para determinar probabilidades de la Normal estándar haciendo uso de las tablas.

Enviar

Ya sabes, a manejar la tabla y a comprobar los resultados con una hoja de cálculo.

¿Y que ocurre si tenemos que saber la $p(Z>0.85)$, o la $p(0.85 \leq Z \leq 1.42)$, o $p(Z \leq -0.85)$? En la tabla no están esas probabilidades.

Es verdad, no están, pero gracias a la **simetría de la distribución Normal** y a que el área bajo la curva es 1, podemos dar respuesta a todas las probabilidades anteriores con la ayuda de la tabla y razonado un poco.

1. Empecemos por $p(Z>0.85)$:

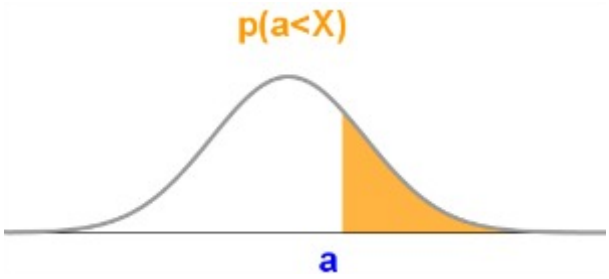


Imagen de elaboración propia

En la imagen podemos ver que si $a=0.85$, el área coloreada es $p(Z>0.85)$ y la zona en blanco es $p(Z \leq 0.85)$.

Por tanto $p(Z>0.85)=1-p(Z \leq 0.85)$.

Buscamos en la tabla de la $N(0,1)$, y tenemos que $p(Z \leq 0.85)=0.8023$. Realizando la operación anterior tenemos que:

$$p(Z>0.85)=1-0.8023=0.1977$$

2. Veamos ahora $p(0.85 \leq Z \leq 1.42)$:

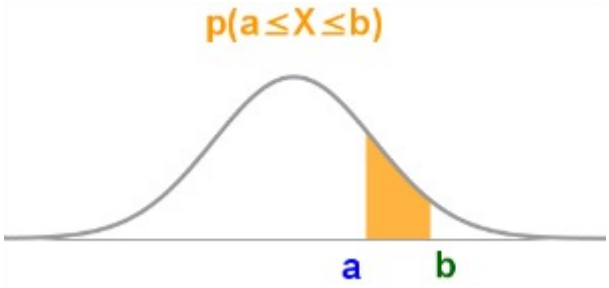


Imagen de elaboración propia

En la imagen anterior, si $a=0.85$ y $b=1.42$, la $p(0.85 \leq Z \leq 1.42)$ coincidiría con el área coloreada. Para hallar su valor lo que haremos es restar a la $p(Z \leq b)$ la $p(Z \leq a)$. Es decir, restar al área hasta b el área hasta a. Buscamos ambas probabilidades en la tabla y tendremos que:

$$p(0.85 \leq Z \leq 1.42)=p(Z \leq 1.42)-p(Z \leq 0.85)=0.9222-0.8023=0.1199$$

3. Por último, hallemos $p(Z \leq -0.85)$:

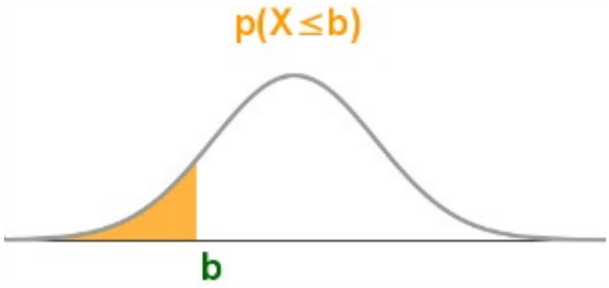


Imagen de elaboración propia

Si $b=-0.85$, la $p(Z\leq -0.85)$ es el área coloreada. Por simetría de la función Normal se tiene que $p(Z\leq -0.85)=p(Z\geq 0.85)$. Es decir, el área coloreada coincide con la del primer ejemplo que hemos visto. Por tanto:

$$p(Z\leq -0.85)=p(Z\geq 0.85)=1-p(Z<0.85)=1-0.8023=0.1977$$

Comprueba lo aprendido

Calcula las siguientes probabilidades:

1. p(Z < 0.98) = ☐
2. p(Z ≥ 0.98) = ☐
3. p(Z < -0.98) = ☐
4. p(Z > -0.98) = ☐

Enviar

La primera probabilidad se halla mirando en las tablas. La segunda y la tercera se calculan aplicando lo explicado más arriba.

Para la última, volvemos a trabajar con la simetría de la distribución Normal.

Imágenes de elaboración propia

Si consideramos $a=-098$ y $b=0.98$, observamos que $p(Z>-0.98)=p(Z<0.98)$, y esta segunda probabilidad es la misma que hemos hallado en el primer apartado de este ejercicio, 0.8365.

2.3. Tipificación

¿Y qué hacemos para calcular probabilidades del resto de distribuciones Normales, es decir con media distinta de cero y desviación típica distinta de 1? ¿Hay tablas para todas?

No, por suerte solo hay una tabla de la distribución Normal, la de la estándar. Por lo que para conocer probabilidades de otras Normales tenemos que recurrir a herramientas tecnológicas como las hojas de cálculo o aprender a **tipificar**.

¿Tipificar? Sí, vamos a realizar un **cambio de variable** de cualquier Normal a la estándar que no va a modificar la probabilidad. A continuación, hallaremos la probabilidad de la estándar, y por tanto la de la Normal inicial.

Imagen en Flickr de HckySo bajo CC

¿Un cambio de variable que mantiene la probabilidad? Ese hecho no es tan extraño. ¿Recuerdas que la probabilidad de que una Normal estuviera en el intervalo $(\mu-\sigma\,,\mu+\sigma)$ era 0.683 y este valor no dependía del valor de la media ni de la desviación típica? Pues lo mismo ocurre con otro tipos de intervalos.

En las siguientes imágenes aparecen tres Normales diferentes, cada una con su media y desviación típica. Pero las probabilidades de las tres áreas coinciden porque las veces que se han sumado las desviaciones típicas a las medias son iguales en los tres casos:

Imágenes de elaboración propia

Importante

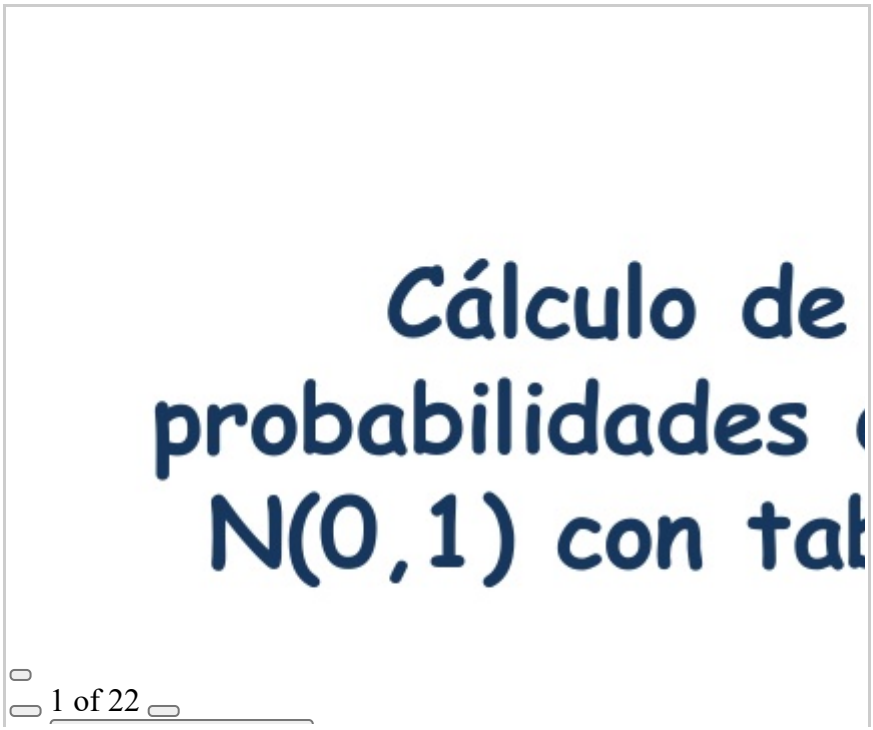
- Si a una variable aleatoria le restamos un número, la media μ de la variable disminuye también en ese número.
- Del mismo modo, si la dividimos por un número, la desviación típica σ de la variable también queda dividida por ese número.

Una consecuencia de las dos propiedades anteriores es la siguiente:

Si queremos transformar una variable Normal $X\sim N(\mu\,,\sigma)$, en otra variable Z Normal estándar $N(0\,,1)$, le restamos μ a X y luego la dividimos entre σ.

Es decir, si $X\sim N(\mu\,,\sigma)$, entonces $Z=\frac{X-\mu}{\sigma}\sim N(0\,,1)$

A este proceso, que mantiene las probabilidades, le llamaremos **tipificar** una variable aleatoria continua.



Haciendo clic en la siguiente imagen podrás descargarte una tabla de la distribución Normal estándar. Es conveniente que la imprimas para poder realizar los ejercicios que se plantean a continuación.

z_0	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	z_0
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359	0,0
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753	0,1
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141	0,2
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517	0,3
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879	0,4
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224	0,5
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549	0,6
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852	0,7
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133	0,8
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389	0,9
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621	1,0
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830	1,1
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015	1,2
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177	1,3
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319	1,4

Imagen de elaboración propia

Comprueba lo aprendido

Haciendo uso de las tablas de la Normal estándar, determina las siguientes probabilidades:

1. $p(Z \leq 0.79) =$ ☐
2. $p(Z \leq 4.5) =$ ☐
3. $p(Z \leq 3) =$ ☐
4. $p(Z < 0.02) =$ ☐
5. $p(Z \leq 1.77) =$ ☐
6. $p(Z \leq$ ☐ $) = 0.7190$
7. $p(Z \leq$ ☐ $) = 0.8643$

Mostrar retroalimentación

Llamemos X = "precio del kg de gambas frescas (en euros)"

Sabemos que $X \sim N(16.5\,,\,2)$

Pero esta variable X hay que tipificarla y convertirla en una nueva variable $Z \sim N(0\,,\,1)$, es decir, en una Normal estándar.

Para ello, restamos a X el valor de la media (en este caso $\mu = 16.5$), así tendremos una nueva variable con media $\mu = 0$ (16.5-16.5=0) y luego dividimos entre la desviación típica (en este caso $\sigma = 2$) y obtendremos otra nueva variable con desviación típica $\sigma = 1$ (2/2=1).

Así, a partir de X, obtendremos la variable $Z = \frac{X-16.5}{2}$ que sigue una distribución $N(0\,,\,1)$.

Por tanto, para dar respuesta a la cuestión hacemos esta transformación tanto a la variable como al valor del que queremos determinar su probabilidad. Vamos estableciendo una cadena de igualdades (haciendo uso de las propiedades vistas en el apartado anterior) hasta llegar a la probabilidad buscada.

$$p(X<15) = p\left(\frac{X-16.5}{2} < \frac{15-16.5}{2}\right) = p(Z<-0.75) = p(Z>0.75) =$$
$$= 1-p(Z \leq 0.75) = 1-0.7734 = 0.2266$$

La $p(Z \leq 0.75)$ la hemos obtenido de la tabla de la $N(0\,,\,1)$.

Por tanto, la probabilidad de que un kilogramo de gambas frescas nos cueste menos de 15 euros es 0.2266.

(Interpretación: Hay un 22.66 % de probabilidad que 1 kg de gambas nos cueste menos de 15 €)

b) ¿Y la probabilidad de que nos cueste entre 15 y 17 euros?

Mostrar retroalimentación

Seguiremos el mismo esquema empleado en la anterior cuestión.
Llamemos X = "Precio del kg de gambas frescas (en euros)"

Sabemos que $X \sim N(16.5\,,\,2)$ y, por tanto, $Z = \frac{X-16.5}{2} \sim N(0\,,\,1)$

Por tanto, para dar respuesta a la cuestión hacemos esta transformación tanto a la variable como a los valores entre los que queremos determinar su probabilidad. Vamos estableciendo una cadena de igualdades (haciendo uso de las propiedades vistas en el apartado anterior) hasta llegar a la probabilidad buscada.

$$p(15<X<17) = p(X<17) - p(X<15)$$

$$p(X<17) = p\left(\frac{X-16.5}{2} < \frac{17-16.5}{2}\right) = p(Z<0.25) = 0.5987$$

$$p(X<15) = p\left(\frac{X-16.5}{2} < \frac{15-16.5}{2}\right) = p(Z<-0.75) = p(Z>0.75) =$$
$$= 1-p(Z \leq 0.75) = 1-0.7734 = 0.2266$$

Las probabilidades $p(Z<0.25)$ y $p(Z<0.75)$ las hemos obtenido de la tabla de la $N(0\,,\,1)$.

Por tanto:

$$p(15<X<17) = p(X<17) - p(X<15) = 0.5987 - 0.226 = 0.3721$$

De lo anterior deducimos que la probabilidad de que un kilogramo de gambas frescas nos cueste entre 15 y 17 € vale 0.3721.

(Interpretación: Hay un 37.21 % de probabilidad de que 1 kg de gambas nos cueste entre 15 y 17 €)

c) ¿Y la probabilidad de que un kilogramo nos cueste más de 17 euros?

Mostrar retroalimentación

Llamemos X = "Precio del kg de gambas frescas (en euros)"

Sabemos que $X \sim N(16.5\,,\,2)$, y por tanto, $Z = \frac{X-16.5}{2} \sim N(0\,,\,1)$
Por tanto, para dar respuesta a la cuestión hacemos esta transformación tanto a la variable como a los valores entre los que queremos determinar su probabilidad. Vamos estableciendo una cadena de igualdades (haciendo uso de las propiedades vistas en el apartado anterior) hasta llegar a la probabilidad buscada.

$$p(X>17) = p(\frac{X-16.5}{2} > \frac{17-16.5}{2}) = p(Z>0.25) = 1-p(Z \leq 0.25) =$$
$$= 1-0.5987 = 0.4013$$

La $p(Z<0{,}25)$ la hemos obtenido de la tabla de la $N(0\,,\,1)$.

La probabilidad de que un kilogramo de gambas frescas nos cueste más de 17 € vale $0{,}4013$.

(Interpretación: Hay un 40,13 % de probabilidad que 1 kg de gambas nos cueste más de 17 €)

Ejercicio resuelto

Curso 2009/2010

En un colegio se estudia la distribución de la nota de Matemáticas de sus estudiantes, resultando ser una Normal de media 7.2 y desviación típica 1.2. Se elige al azar un estudiante de ese colegio, ¿cuál será la probabilidad de que su nota en esta asignatura sea mayor que 7.5?

Mostrar retroalimentación

Sea X = "nota de la asignatura de Matemáticas". Sabemos que $X \sim N(7.2\,,\,1.2)$

Para resolver el problema que se nos plantea tenemos que hallar la $p(X>7.5)$.

Hagamos el cambio de variable $Z = \frac{X-7.2}{1.2}$, y tendremos que Z es la Normal estándar.

Realicemos las operaciones necesarias con el cambio anterior:

$$p(X>7.5) = p(\frac{X-7.2}{1.2} > \frac{7.5-7.2}{1.2}) = p(Z>0.25)$$

Como el área debajo de la gráfica de la Normal es 1, tenemos que

$p(Z>0.25)=1-p(Z\leq 0.25)$

Buscamos en la tabla, y obtenemos

$p(X>7.5)=p(Z>0.25)=1-p(Z\leq 0.25)=1-0.5987=0.4016$

Por lo que la probabilidad de que la nota sea mayor que 7.5 es de un 40.16 por ciento.

Ejercicio resuelto

Curso 2010/2011

El peso de las manzanas que se producen en una huerta sigue una ley Normal con una media de 150 gramos y desviación típica de 20 gramos.

a) ¿Qué porcentaje de estas manzanas tendrá un peso inferior a 115 gramos?

b) Halla la probabilidad de que una manzana, elegida al azar en este huerto, tenga un peso que se encuentre entre 165 y 220 gramos?

Mostrar retroalimentación

Sea X = "peso de las manzanas"

$X\sim N(150\,,20)$, y se nos plantean dos cuestiones $p(X<115)$ y $p(165<X<220)$.

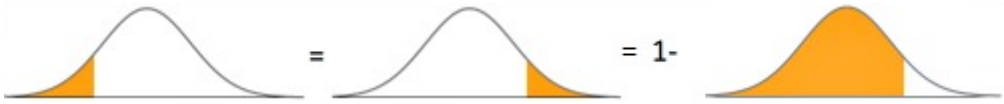
Realizamos el cambio de variable $Z=\frac{X-150}{20}$, por tanto Z es $N(0\,,1)$.

a) Operamos:

$$p(Z<115)=p(\frac{X-150}{20}<\frac{115-150}{20})=p(Z<-1.75)$$

Ahora bien, esa probabilidad no viene en las tablas, pero gracias a las propiedades de esta distribución, tenemos que:

$$p(Z<-1.75)=p(Z>1.75)=1-p(Z<1.75)$$



En la imagen se explica gráficamente lo que se ha hecho.

Buscamos en las tablas, y obtenemos $p(Z<1.75)=0.9599$.

Por tanto $p(X<115)=1-p(Z<1.75)=1-0.9599=0.0401$.

Luego el porcentaje de manzanas que pesa menos de 115 gramos es del 4 %.

b) Resolvamos la segunda cuestión que se nos plantea: $p(165<X<220)$.

Haciendo el cambio de variable, tenemos que:

$$p(165<X<220)=p(\frac{165-150}{20}<\frac{X-150}{20}<\frac{220-150}{20})=p(0.75<Z<3.5)$$.

Aplicando las propiedades de la probabilidad, tenemos que:

$p(0.75<Z<3.5)=p(Z<3.5)-p(Z<0.75)=0.99977-0.7734=0.2263$.

Lo que quiere decir que el 22.63 % de las manzanas pesan entre 165 y 220 gramos.

Ejercicio resuelto

Curso 2011/2012

La duración de un tipo de pilas alcalinas sigue una distribución Normal de media 55 horas y una desviación típica de 6 horas.

a) Calcula la probabilidad de que una pila elegida al azar dure más de 53 horas.

b) Halla la probabilidad de que una pila elegida al azar dure entre 56 y 58 horas.

Mostrar retroalimentación

Sea X = "duración de las pilas alcalinas"

Estamos ante una variable $X \sim N(55,6)$.

En primer lugar realizamos el cambio de variables para obtener una Normal estándar:

$Z=\frac{X-55}{6} \sim N(0,1)$.

a) Veamos en primer lugar $p(X>53)$.

Tenemos que $p(X>53)=p(\frac{X-53}{6}>\frac{55-53}{6})=p(Z>-0.33)$.

Gracias a la simetría de la distribución Normal tenemos que:

$$p(Z>-0.33)=p(Z\leq 0.33)$$

Por último, y con la ayuda de las tablas, obtenemos que

$p(X<53)=p(Z\leq 0.33)=0.6304$, lo que quiere decir que el 63 % de las pilas dura más de 53 horas.

b) Veamos ahora la $p(56<X<58)$.

Tipificando tenemos que:

$p(56<X<58)=p(\frac{56-55}{6}<\frac{X-55}{6}<\frac{58-55}{6})=p(0.17<Z<0.5)$, y aplicando las propiedades de la probabilidad:

$$p(0.17<Z<0.5)=p(Z<0.5)-p(Z<0.17)$$

Como ejemplo, vamos a calcular la $p(X \leq 4.5)$ sabiendo que X es una $N(3, 1.2)$.

En primer lugar, si hacemos el cambio indicado arriba, tenemos que $Z = \frac{X-3}{1.2}$ es una $N(0, 1)$. Vamos a operar:

$$p(X \leq 4.5) = p(\frac{X-3}{1.2} \leq \frac{4.5-3}{1.2}) = p(Z \leq 1.25)$$

Las tablas de la Normal estándar nos dicen que la $p(Z \leq 1.25) = 0.8944$, y por tanto $p(X \leq 4.5) = 0.8944$.

A continuación puedes ver un vídeo del canal de juanmemol donde se resuelve un problema en contexto usando la tipificación de una distribución Normal:

Problema sobre distribución normal

Ejercicio resuelto

Imagen en Flickr de Diego López bajo CC

El precio del kilogramo de gambas frescas en el mercado de abastos de Huelva sigue una distribución Normal de media 16.5 euros y desviación típica 2 euros.

Responde a las siguientes cuestiones:

a) ¿Cuál es la probabilidad de que un kilogramo de gambas nos cueste menos de 15 euros?

3. Apéndice

El último tema del curso está dedicado al estudio de las variables aleatorias continuas, y más concretamente a la distribución Normal.

La Normal y la famosa gráfica de su función de densidad, la Campana de Gauss, modeliza infinidad de situaciones de la realidad: sociales, físicas, médicas, biológicas.

Para determinar la probabilidad de esta distribución es necesario, en primer lugar, tipificarla y posteriormente hacer uso de las tablas de la Normal estándar. Estos dos procesos son los que se desarrollan en la última parte del tema.

Importante

En el siguiente mapa conceptual se trazan las líneas fundamentales de los conceptos que se han tratado a lo largo del tema.

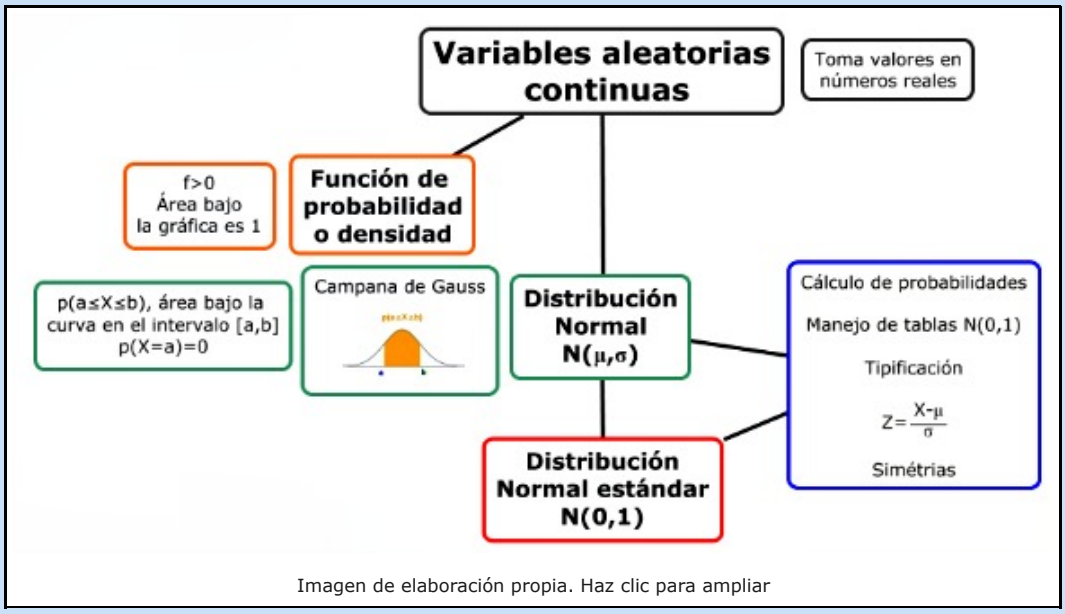
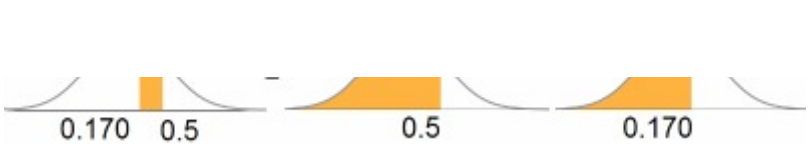


Imagen de elaboración propia. Haz clic para ampliar



Haciendo uso de las tablas:

$$p(0.17<Z<0.5) = p(Z<0.5) - p(Z<0.17) = 0.6915 - 0.5675 = 0.124$$.

Lo que implica que el 12 % de las pilas duran entre 56 y 58 horas.

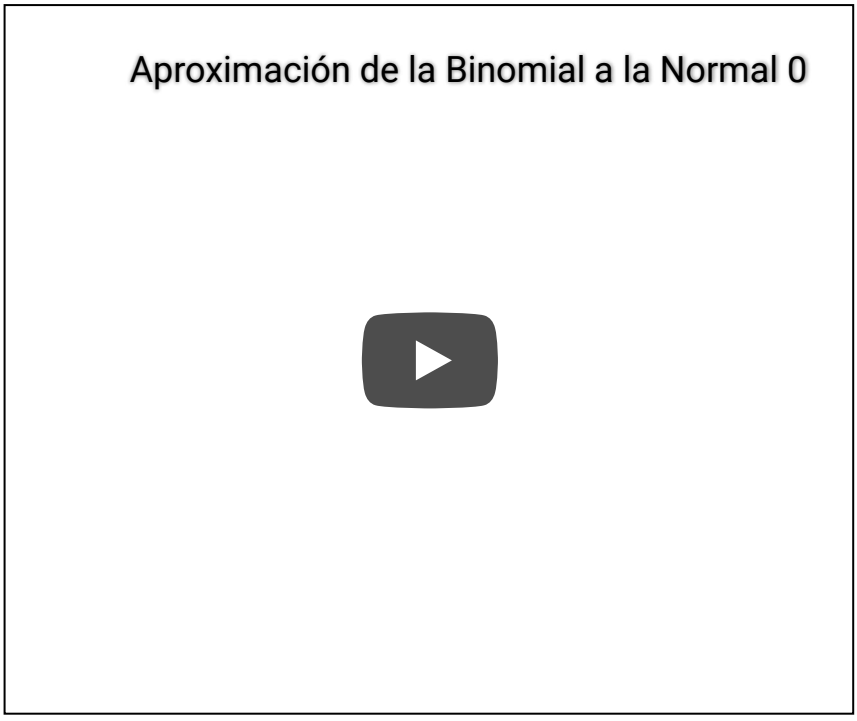
Ya hemos mencionado anteriormente que una distribución Binomial con un número de repeticiones del experimento grande, tiende a hacerse continua. Incluso vimos al inicio del tema, que en ese caso, los diagramas de barras de las probabilidades se van aproximando a la campana de Gauss.

En general, una binomial $B(n\,,\,p)$ se parece a una **Normal** cuanto **mayor es el producto** $n \cdot p$ (o $n \cdot q$ si q es mayor que p).

De hecho si $n \cdot p$ y $n \cdot q$ son ambos mayores que 3, la aproximación es buena, y si superan 5, es perfecta.

En ese caso la Binomial se transforma en una Normal de media $\mu = n \cdot p$ y desviación típica $\sigma = \sqrt{n \cdot p \cdot q}$

En el siguiente vídeo se explica este proceso con más detalle:

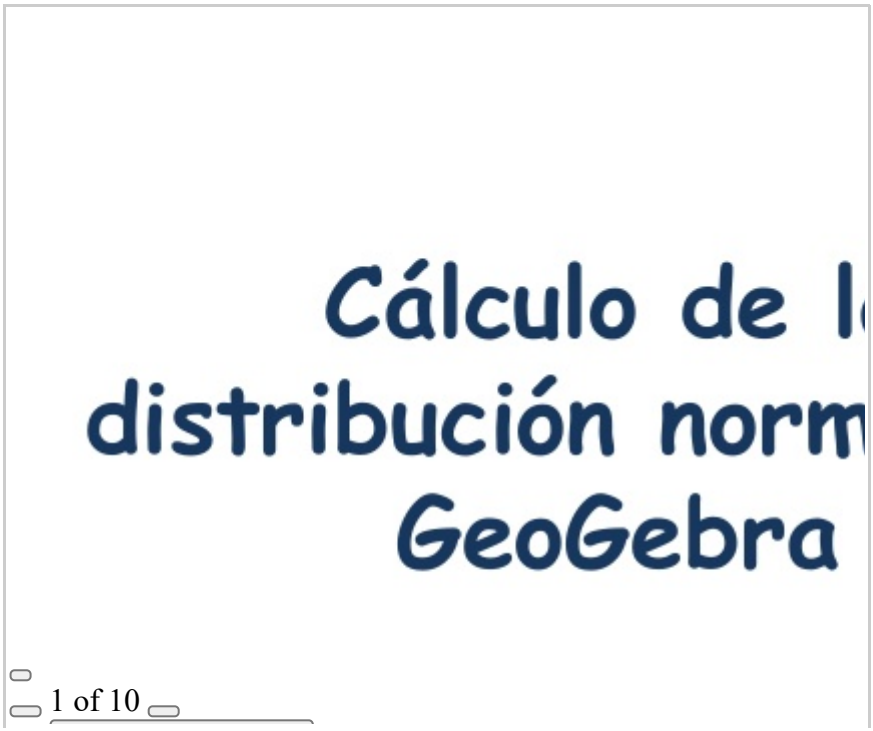


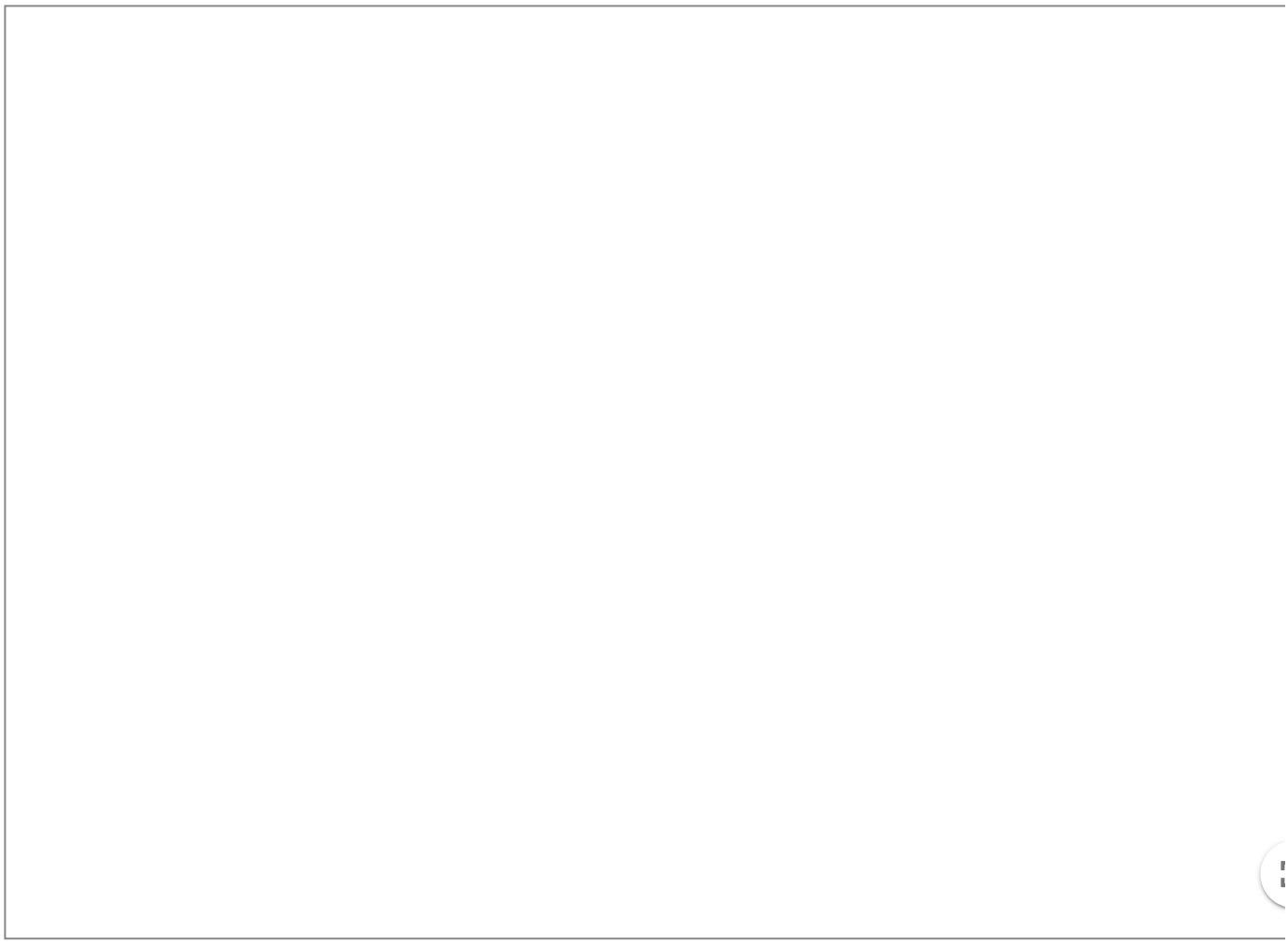
3.1. Recursos

Por fortuna, en la actualidad existen una amplia variedad de recursos informáticos que nos facilitan el cálculo de probabilidades de la distribución Normal y que nos ahorran tanto el proceso de tipificación como el de tener que utilizar las tablas.

Ya hemos visto en el desarrollo del tema que las hojas de cálculo incorporan funciones para determinar dichas probabilidades.

Como no podía ser de otra manera, con GeoGebra también es posible conocer probabilidades de la Normal de forma fácil, rápida y visual. En la siguiente presentación puedes verlo:





Recurso de Ignacio López alojado en Geogebra. Licencia CC

Como el área bajo la curva de la Normal es 1, y la gráfica es simétrica respecto de la media, podemos afirmar que $p(X \leq \mu) = p(X \geq \mu) = 0.5$

Pero la distribución Normal guarda algunas sorpresas a la hora de la asignación de probabilidades a algunos intervalos.

Como podemos observar en este enlace, independientemente de la media μ y desviación típica σ, la probabilidad de cada uno de los intervalos $(\mu-\sigma,\mu+\sigma)$ $(\mu-2\sigma,\mu+2\sigma)$ $(\mu-3\sigma,\mu+3\sigma)$ es **0.683**, **0.954** y **0.997** respectivamente.

Esto nos ayuda a entender cómo se concentra o dispersa la distribución Normal en función de su desviación típica, y a calcular sin mucho esfuerzo algunas probabilidades.

Ejercicio resuelto

La altura X de unos tornillos siguen una distribución Normal de media 2 centímetros y desviación típica 0.3 centímetros, $X \sim N(\mu, \sigma)$

Determina las siguientes probabilidades.

1. $p(1.7 < X \leq 2.3)$
2. $p(2.6 < X)$
3. $p(X \leq 1.1)$

Aviso Legal

El presente texto (en adelante, el "**Aviso Legal**") regula el acceso y el uso de los contenidos desde los que se enlaza. La utilización de estos contenidos atribuye la condición de usuario del mismo (en adelante, el "**Usuario**") e implica la aceptación plena y sin reservas de todas y cada una de las disposiciones incluidas en este Aviso Legal publicado en el momento de acceso al sitio web. Tal y como se explica más adelante, la autoría de estos materiales corresponde a un trabajo de la **Comunidad Autónoma Andaluza, Consejería de Educación y Deporte (en adelante Consejería de Educación y Deporte**).

Con el fin de mejorar las prestaciones de los contenidos ofrecidos, la Consejería de Educación y Deporte se reserva el derecho, en cualquier momento, de forma unilateral y sin previa notificación al usuario, a modificar, ampliar o suspender temporalmente la presentación, configuración, especificaciones técnicas y servicios del sitio web que da soporte a los contenidos educativos objeto del presente Aviso Legal. En consecuencia, se recomienda al Usuario que lea atentamente el presente Aviso Legal en el momento que acceda al referido sitio web, ya que dicho Aviso puede ser modificado en cualquier momento, de conformidad con lo expuesto anteriormente.

Régimen de Propiedad Intelectual e Industrial sobre los contenidos del

Curiosidad

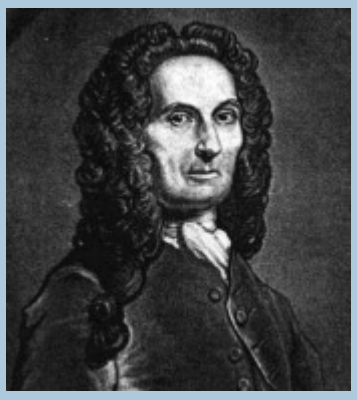


Imagen en Wikimedia Commons de Searobin bajo Dominio Público

A la hora de trabajar con la distribución Normal hemos podido comprobar que la Campana de Gauss juega un papel destacado.

Pero no fue Gauss el primero en construir esa gráfica.

Fue el matemático francés Abraham De Moivre, que vivió entre los siglos XVII y XVIII.

Moivre profundizó en los trabajos de Jacobo Bernouilli relacionados con el lanzamiento de monedas y el estudio de la distribución Binomial. Observó que los cálculos se complicaban a la hora de trabajar con monedas sesgadas y aumentar el número de lanzamientos. La dificultad para hallar los números combinatorios le llevó a definir una función a la que llamó "**función error**" o "**distribución Normal**".

La primitiva distribución Normal creada por Moivre, estaba expresada del siguiente modo:

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

Pero su trabajo cayó en el olvido ya que su aplicación la limitó tan solo a resolver problemas sobre juegos.

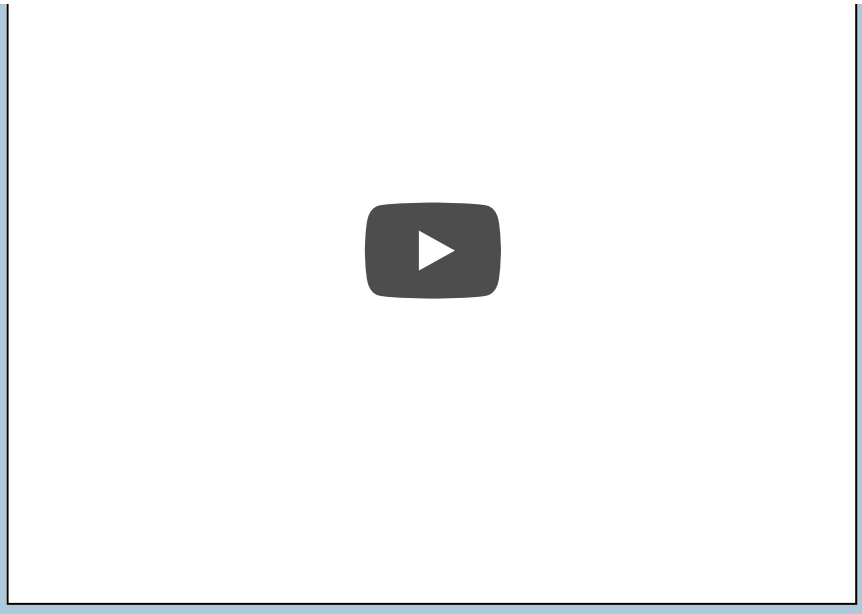
Hubo que esperar algo más de medio siglo hasta que Gauss y Laplace volvieron a retomar la idea de Moivre para evaluar la distribución de errores en distintos tipos de mediciones. Por ejemplo, en estudios astronómicos.

Posteriormente, se demostró que la "curva de los errores" se adaptaba a una amplia gama de situaciones, como ya hemos comentado a lo largo del tema.

Curiosidad

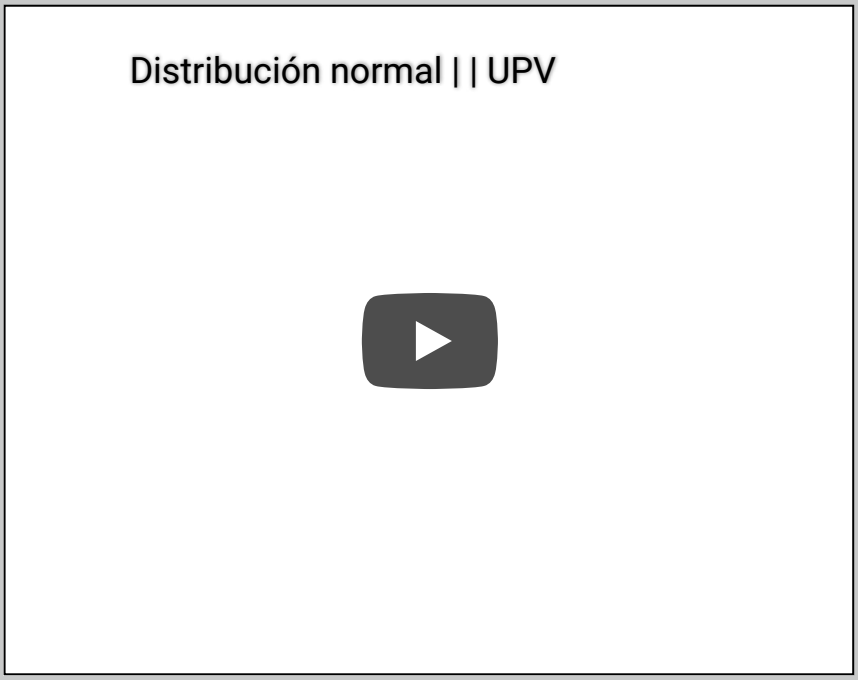
Para terminar, veamos un simpático vídeo donde se explica de forma agradable y amena una aplicación de la distribución Normal.

Distribución Normal.wmv



Para saber más

En el siguiente vídeo se resumen con claridad los aspectos más importantes relacionados con la distribución Normal:



Para saber más

Terminamos con el siguiente enlace en el que podrás repasar y ampliar muchos de los contenidos tratados en el tema:

- Distribución Normal .