2º de Bachillerato

Matemáticas Aplicadas a las Ciencias Sociales II

Contenidos

Cálculo de probabilidades: Distribución normal

1. Origen de la Distribución Normal

Cuando comenzaron los estudios estadísticos modernos, se encontró que la proporción de variables continuas que seguían el modelo de la distribución era tan elevada, que se llegó a pensar que todas las variables aleatorias continuas la seguían.

Imagen de Goldmund100 bajo licencia Creative Commons

Algunos ejemplos de variables asociadas a fenómenos naturales que siguen el modelo de la normal son:

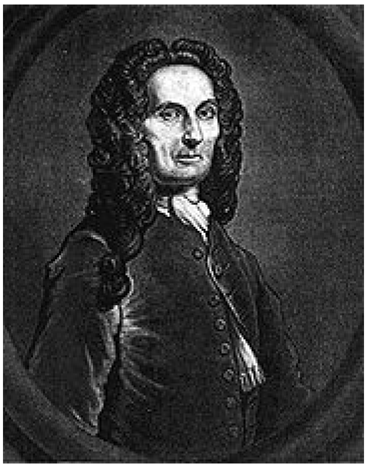
- Caracteres morfológicos de individuos(no sólo personas, sino también animales o plantas) como la estatura, el peso, perímetros, etc.
- Caracteres fisiológicos como el efecto de una dosis de un fármaco o el producido por un abono en un cultivo.
- Caracteres sociológicos como el consumo de cierto producto por un mismo grupo de individuos.
- Caracteres psicológicos como el cociente intelectual o nivel de adaptación a un medio.
- Nivel de ruido o perturbaciones en telecomunicaciones.
- Errores cometidos al medir ciertas magnitudes.
- Etc.

Y en general, cualquier característica que se obtenga como suma de muchos factores independientes se ajustará a un modelo normal.

La distribución Normal se expresa por $N(\mu , \sigma)$, donde el primer parámetro μ, representa la media de la variable aleatoria y σ la desviación típica.

Curiosidad

Antes de continuar, vamos a dar una pequeña pincelada histórica sobre el descubrimiento y evolución de esta importantísima distribución de probabilidad.



Abraham De Moivre. Imagen en Wikimedia Commons bajo licencia Creative Commons

La distribución normal fue presentada por primera vez por Abraham De Moivre en un artículo del año 1733, que fue reimpreso en la segunda edición de su The Doctrine of Chances, de 1738, en el contexto de cierta aproximación de la distribución binomial para grandes valores de n. Su resultado fue ampliado por Laplace en 1812, y en la actualidad el resultado obtenido se llama Teorema de De Moivre-Laplace, teorema que veremos en la siguiente unidad.

Laplace usó la distribución normal en el análisis de errores de experimentos. Gauss, que afirmaba haber usado el método desde 1794, lo justificó rigurosamente en 1809 asumiendo una distribución normal de los errores. El nombre de Gauss se ha asociado a esta distribución porque la usó con profusión cuando analizaba datos astronómicos, llegando a dar la expresión de la función de densidad de estas variables.

Esta distribución también se conoce como ***"distribución de Gauss"*** *o* ***"campana de Gauss"***, y en siguiente apartado verás el porqué. El nombre de "campana" viene de Esprit Jouffret que usó el término "bell surface" (superficie campana) por primera vez en 1872 para una distribución normal.

Algunos autores le atribuyen a Gauss un descubrimiento de esta distribución independiente del de De Moivre. Esta atribución del nombre de la distribución a una persona distinta de su primer descubridor es un claro ejemplo de la Ley de Stigler.

El nombre de "distribución normal" fue otorgado independientemente por Charles S. Peirce, Francis Galton y Wilhelm Lexis hacia 1875.

Galton llegó a decir de esta distribución, *"Si los griegos la hubieran conocido la habrían adorado como a un Dios"*

2. Variables aleatorias continuas

Antes de meternos de lleno en esto de la distribución normal, vamos a ver un poco, al igual que hemos hecho en el tema anterior con las variables discretas, cómo funcionan las variables continuas y cómo se calculan probabilidades en ellas.

Importante

Una variable aleatoria es continua si al realizar el experimento aleatorio, entre cada dos valores, el número de valores que puede tomar es infinito.

Por ejemplo, la altura de una persona, la longitud del dedo índice, el peso de un perro o el caudal de un río son variables aleatorias continuas, pues entre dos posibles valores puede tomar cualquiera que exista. Todo depende de la precisión de nuestro aparato de medida.

Imagen de thombo2 bajo licencia Creative Commons

En una variable aleatoria continua, las probabilidades se calculan de forma un poco distinta de la que hemos visto en las variables discretas. Aquí no tiene sentido el concepto de función de probabilidad, pues al poder tomar la variable infinitos valores, la probabilidad de cada uno en concreto va a ser cero. Lo vemos en el siguiente ejemplo

Ejercicio resuelto

Imagen de jl.cernadas bajo licencia Creative Commons

A la entrada de un parque, hemos realizado una encuesta sobre salud, y entre cosas, a la persona encuestada se le pregunta por su peso. Algunas no han querido responder, pero las que sí, nos han dado estos resultados:

60	75,3	81	86,6	68,1	70,4	69,4	88	94,1	72
77,4	82	57	70,5	65,2	65	81	82,2	63	70
80,5	59	60,5	65	71	73,6	88,4	90	92,5	101

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
87	76	65	69	72	74	98	88,3	60	76,6

¿Con qué probabilidad al elegir una persona pesará 84 kg? ¿Y 71,5? ¿Y 74 kg?

Mostrar retroalimentación

Calcular estas probabilidades es muy fácil, basta buscar el número de casos favorables y aplicar la regla de Laplace. Así, en los dos primeros casos esas probabilidades son cero, pues ninguno de los dos valores, ni 84 ni 71,5 aparecen en la tabla. Si definimos X = Peso de una persona, tenemos:

P(X = 84) = 0

P(X = 71,5) = 0

El tercer valor sí lo tenemos en la tabla, aparece una vez. Así, $P(X=74)=\frac{1}{40}=0{,}025$

Fíjate que no es cero la probabilidad pero casi. Y es que, en estas variables, la probabilidad de que tomen un valor concreto es cero. ¿Qué pasaría si los pesos los damos hasta los gramos, es decir, con tres decimales? Pues sería prácticamente imposible que se repitiera ningún valor y al preguntarnos por cualquier probabilidad sería prácticamente cero.

Pero claro, en vista de la muestra, no es igual de probable que una persona de las que acuden al parque pese setenta y tanto o ciento y algo, ¿verdad?

Importante

Si X es una variable aleatoria continua, la probabilidad de que tome un valor concreto es cero.

P[X = a] = 0, para cualquier valor de a.

Por tanto, en variables continuas, las probabilidades se calculan sobre intervalos y serán de esta forma:

- P(a ≤ X ≤ b)
- P(X ≥ a)
- P(X ≤ b)

Además, como el signo "=" en un punto da probabilidad cero, en los intervalos anteriores da igual poner desigualdades estrictas o inclusivas, es decir, da lo mismo escribir ≤ ó < y es lo mismo escribir ≥ ó >.

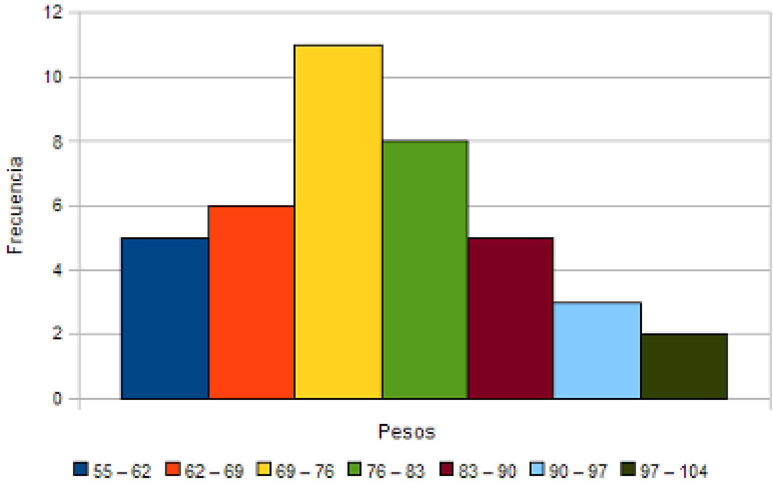
Por ejemplo, si X es una variable aleatoria continua, P(X ≤ 2) = P(X < 2), pues la diferencia entre ambas es que entre el 2 o no, y como ya sabemos que la probabilidad del valor concreto 2 es cero, da igual que lo incluyamos o no.

Ejercicio resuelto

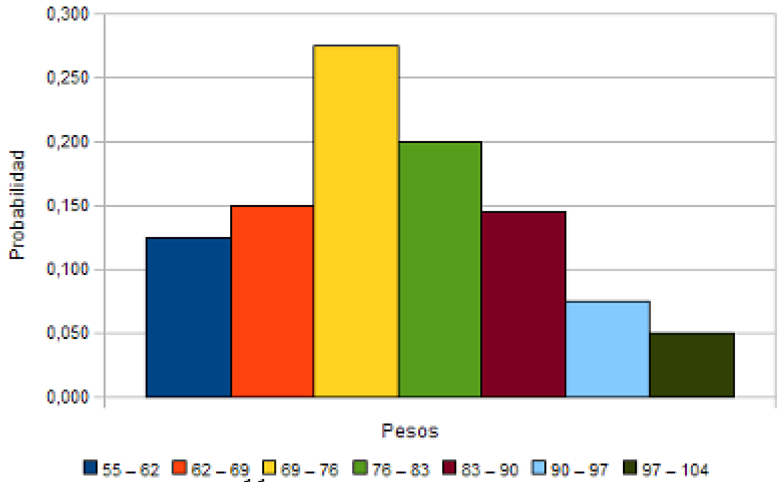
Como puedes pensar, es cierto que no se repite muchos valores en la tabla, pero no es menos cierto que la diferencia entre algunos de ellos es mínima. Está claro que no es lo mismo pesar 70 kilos que 98, pero, ¿hay mucha diferencia entre 70 kilos ó 70,4 kg?

Es por esto por lo que en estas variables los **datos se agrupan en intervalos**. Vamos a hacerlo con los datos de nuestra tabla. Por ejemplo, lo hacemos así:

Intervalo	n.º de casos
55 - 62	5
62 - 69	6
69 - 76	11
76 - 83	8
83 - 90	5
90 - 97	3
97 - 104	2



Ahora sí tiene un poco más de sentido, calcular probabilidades.



Por ejemplo, según esta muestra, la probabilidad de que al elegir una persona al azar, ésta pese entre 69 y 76 kilos es 11/40:

$$P(69 \leq X \leq 76) = \frac{11}{40} = 0{,}275$$

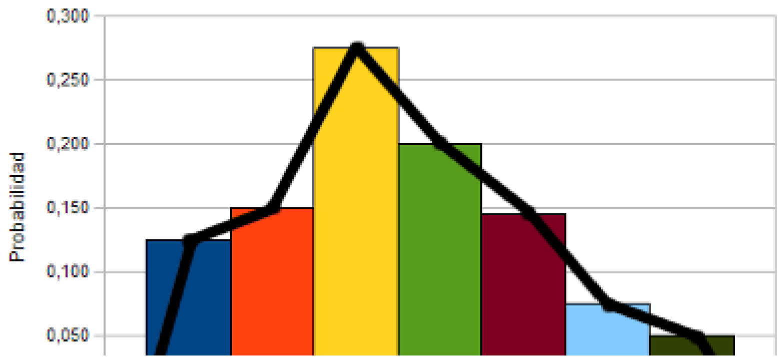
O la probabilidad de que al elegir una persona, ésta pese más de 76 kg es:

$$P(X \geq 76) = \frac{9}{40} = 0{,}225$$

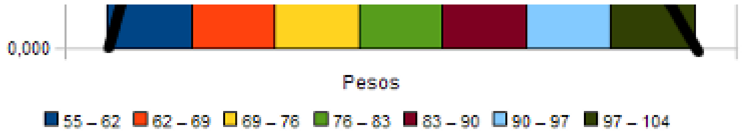
¿Entiendes ahora por qué en las variables continuas las probabilidades siempre se calculan sobre intervalos?

El gráfico de la izquierda nos muestra la distribución de probabilidad de los intervalos, pero, aún no es suficiente, porque ahora mismo podemos calcular sólo la probabilidad de que X esté en alguno de esos intervalos, si queremos la probabilidad de que una persona pese entre 70 y 80 kilos no podemos calcularla. De momento, claro.

Y es que, ahora sí que llegamos ya al equivalente de la función de probabilidad de las variables discretas. ¿Recuerdas el polígono de frecuencias de los temas de estadística del curso pasado? Si no, no pasa nada.



Fíjate en el gráfico de la izquierda. Hemos unido los puntos medios de los intervalos mediante rectas (esto era el polígono de frecuencias.) Pues bien, si hacemos que los intervalos tengan una amplitud muy pequeña, esa línea poligonal va



esa línea poligonal va cogiendo cada vez una forma más redondeada, el de función, y se va a llamar **función de densidad**.

Esta función es la que marca las probabilidades en las variables aleatorias continuas, y es que cualquier **probabilidad** de un intervalo se calcula como el **área encerrada bajo la función de densidad en ese intervalo**.

Así, para una variable aleatoria continua, será necesario conocer su función de densidad para calcular cualquier probabilidad.

Ojo, la función de densidad siempre tiene que ser positiva y el área total encerrada por ella será 1, pues como siempre, la probabilidad de todo el espacio muestral es 1.

Comprueba lo aprendido

Contesta verdadero o falso a las siguientes cuestiones. Repasa lo comentado anteriormente y las propiedades generales de la probabilidad.

1) Si X es una variable aleatoria continua, $P(X = 4) = 0$

Sugerencia

Verdadero Falso

Verdadero
Las probabilidades de valores concretos en distribuciones continuas son siempre nulas

2) Si X es una variable aleatoria continua, $P(X > 4) = 0$

Sugerencia

Verdadero Falso

Falso
No tiene porqué. Dependerá de los valores de X

3) El área máxima bajo la función de densidad puede ser mayor que 1.

Sugerencia

Verdadero Falso

Falso
La probabilidad máxima es 1, y es la de todo el espacio muestral.

4) Si X es una variable aleatoria continua, $P(X < 2) = P(X \leq 2)$

Sugerencia

Verdadero Falso

Verdadero

5) En una variable continua, la $P(X > 3) = P(X \leq 3)$

Sugerencia

Verdadero Falso

Falso
No tiene porqué y seguramente no lo sea.

3. Distribución Normal

Imagen de Arenamontanus bajo licencia Creative Commons

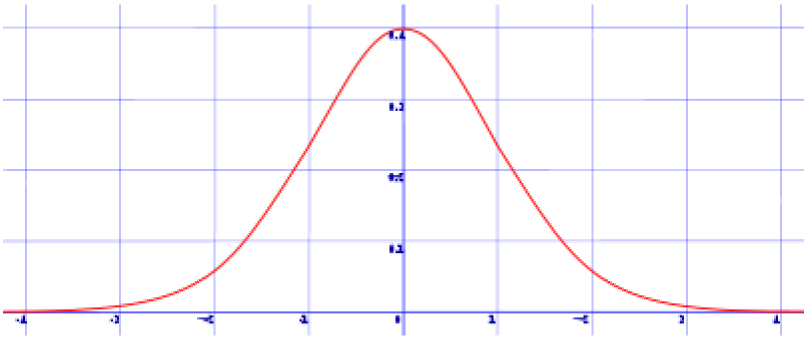
Ahora sí, vamos ya de lleno con la Distribución Normal.

Acabamos de decir que para calcular probabilidades en una variable continua tenemos que conocer la función de densidad de ésta. Pues bien, si una variable aleatoria continua sigue una distribución normal, debemos conocer antes de nada cuáles son la **media μ** y la **desviación típica σ**. Estos dos valores son los parámetros que **caracterizan** y distinguen cada distribución normal (al igual que "n" y "p" en la binomial.)

Conocidos estos dos parámetros, si X sigue una distribución normal de parámetros μ y σ, **X~N(μ , σ)**, la función de densidad de X será:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R},$$

Y la representación gráfica, aplicando todo lo que vimos en el tema de representación de funciones (dominio, cortes con ejes, extremos relativos,...) será similar a esta:



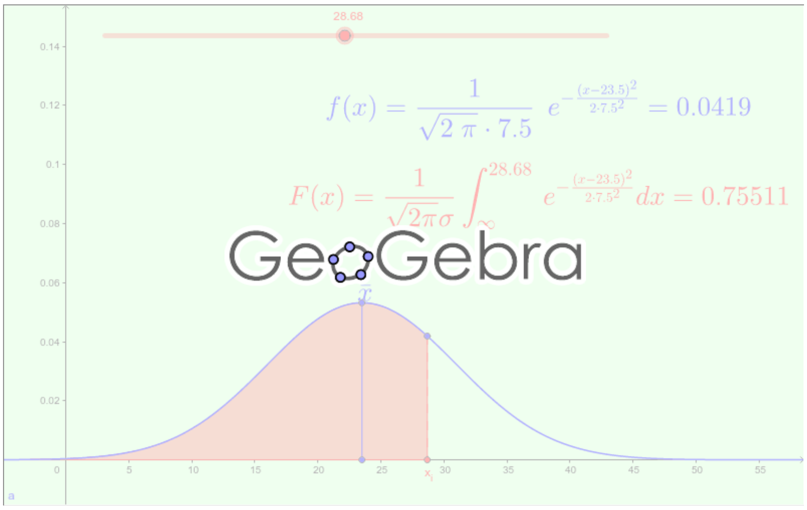
Aquí hemos representado la que tiene de media 0 y desviación típica 1, pero la gráfica de esta función para cualquier valor de μ y σ siempre tiene el mismo aspecto, como una campana, de ahí, que también a la Distribución Normal se le llame *"Campana de Gauss."* ¿Pillas ahora el título del tema?

¿Te has recuperado ya del susto? Bueno, no te preocupes que tampoco es que tengamos que utilizar ese pedazo de función para lo que nos interesa.

Ahora ya podríamos calcular la probabilidad de cualquier intervalo en cualquier distribución normal calculando el área que hay bajo esa función de densidad dentro del intervalo en cuestión. Pero eso no es que sea un juego de niños precisamente.

En Matemáticas, esto se hace usando el concepto de integral, algo similar a la operación contraria de la derivada, pero para esta función es tan sumamente complicado que se utilizan otras herramientas para calcular estas probabilidades, aunque en el fondo de esos cálculos están las integrales.

En la siguiente escena de geogebra realizada por Jorge Cotera, aparece representada la función de densidad de una distribución normal. En la cual puedes ir deslizando el valor de la variable x, y podrás ver sombreada el área de la región que corresponde con el valor que toma la función de probabilidad.



Reflexiona

Manipula la siguiente escena y contesta a las siguientes cuestiones:

1. ¿Qué efecto produce sobre la función de densidad aumentar la media? ¿Y disminuirla?
2. ¿Qué efecto produce sobre la función de densidad aumentar la desviación típica de la Normal? ¿Y disminuirla?
3. Si X es N(0,1), ¿Cuál es la probabilidad de que X esté entre 2,1 y 3,5?
4. Si X ~ N(0,1), calcula P(-3<X<0)
5. Si X ~ N(0,1), calcula P(X ≥ -1) (Usa el zoom y aleja el punto Q)
6. Si X ~ N(1,5 ; 0,6), calcula P(-2<X<2)
7. Si X ~ N(-0,5 ; 1,5), calcula P(-2<X<2)
8. Si X ~ N(8 , 2), calcula P(0≤X≤10) (Usa el zoom para ver la función)
9. Si X ~ N(8 , 2), calcula P(X≥0)
10. Si X ~ N(3 ; 0,5), calcula P(X<3)
11. Si X ~ N(-2; 0,6), calcula P(X<-2)

Mostrar retroalimentación

Las soluciones te tienen que salir más o menos, decimal arriba decimal abajo, estas:

1. Al aumentar la media la función se desplaza a la derecha y al disminuirla a la izquierda.
2. Al aumentar la desviación típica, la función de densidad se achata, mientras que al disminuirla se hace más puntiaguda.
3. 0,0172
4. 0,4988
5. 0,8429
6. 0,7988
7. 0,7940
8. 0,8436
9. 1
10. 0,5
11. 0,5

Curiosidad

Si te fijas bien en la gráfica de esta función, puedes ver que:

El dominio es el conjunto de todos los números reales.

Respecto a la media, la función es simétrica.

Tiene asíntota horizontal $y = 0$, tanto cuando x se va a $-\infty$ como a $+\infty$.

La función es creciente hasta que x llega al valor de la media y decreciente a partir de él. Por tanto, en $x = \mu$ la función alcanza su máximo.

Otra propiedad curiosa que cumple esta distribución es que las tres medidas de centralización coinciden, es decir, moda, mediana y media coinciden en el valor de μ.

Imagen de kevindooley bajo licencia Creative Commons

3.1. Distribución Normal estándar


Imagen de Arkangel bajo licencia Creative Commons

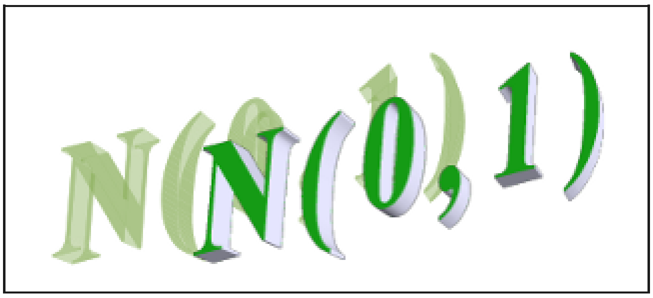
Antes de seguir, necesitamos aprender a calcular probabilidades con la distribución normal, y es que, la escena que acabas de ver en el apartado anterior está muy bien, pero con ella lo que pretendemos es que entiendas el significado de las probabilidades que estás calculando. Además, tiene sus limitaciones, pues no puedes darle ni a la media ni a la desviación típica todos los valores que quieras, y habrá situaciones en la que necesitemos más.

El cálculo de estas probabilidades se hace a partir de tablas, pero claro, ¡imagínate que hay una tabla para cada distribución normal! ¡Necesitaríamos un tocho impresionante!

Con las técnicas que vamos a aprender en los dos próximos apartados, verás que con un folio basta, pero antes que nada, necesitamos manejar con soltura "*ese folio*".

¡Verás qué fácil es!

Ejercicio resuelto



La distribución normal más elemental y la que se denomina estándar es la que tiene de media 0 y de desviación típica 1, es decir, la normal N(0,1) y las variables aleatorias que siguen esta distribución se suelen representar con la letra Z para distinguirla. Bien, pues vamos a familiarizarnos con esta distribución y con la tabla donde se calculan las probabilidades. En el enlace, te puedes descargar la tabla de valores de la distribución Normal N(0,1).

Lo primero que tienes que observar es que con esta tabla se calculan probabilidades del tipo $P(Z \leq a)$, donde *Z* representa una variable aleatoria que sigue una N(0,1) y *"a"* el valor extremo. Es decir, el área de la zona dibujada en azul.

Por ejemplo, vamos a calcular la probabilidad de que si Z es una variable N(0,1), un valor tomado al azar sea menor que 1,37.

Para calcularla, en la columna donde pone z_0(la primera o la última) hemos de buscar la fila en la que aparece el valor entero y la primera cifra decimal, es decir, 1,3. Una vez encontrado, nos vamos a las columnas y buscamos en el encabezamiento la segunda cifra decimal, o sea, donde aparezca 0,07.

El lugar donde se cruzan esas dos búsquedas me indica la probabilidad. En nuestro caso, aparece, 0,9147.

Luego $P(Z \leq 1{,}37) = 0{,}9147$.

Si te ha quedado alguna duda, viendo la siguiente presentación seguro que se te aclaran.

Ve pulsando sobre la presentación para avanzar:

Haz clic para habilitar Adobe Flash Player

Comprueba lo aprendido

Blanco

Calcula las siguientes probabilidades a partir de la tabla.

Usa la coma (,) para escribir los decimales.

$P(Z \leq 0{,}95) =$ ______

$P(Z \leq 3{,}14) =$ ______

$P(Z \leq 1) =$ ______

$P(Z < 2{,}92) =$ ______

$P(Z \leq 3{,}07) =$ ______

Enviar

Repasa la presentación anterior si algún resultado no te sale bien.

Curiosidad

Fíjate que la tabla llega hasta un valor máximo de 3,99 y ya ese valor tiene una probabilidad de casi 1. O sea, que para cualquier valor mayor que éste, se puede considerar que la probabilidad de que sea menor que él es 1, o lo que es lo mismo, es **seguro** que en una distribución normal N(0,1) el valor de la variable es menor o igual que 4.

Por ejemplo, $P(Z \leq 4) = 1$ ó $P(Z \leq 6,24) = 1$.

3.2. Cálculo de probabilidades en una Distribución Normal

Imagen de Hermenpaca bajo licencia Creative Commons

Todavía nos falta un poquito para poder calcular cualquier probabilidad, y es que, con esa tabla podemos calcular probabilidades del tipo P(Z < a) y además con "a" un número positivo, pero, ¿Y si me interesa algo del tipo Z>a o Z< -a o que Z esté entre dos valores?

Bien, pues vamos a ver cómo resolver estas cuestiones. Empezamos por la más fácil: probabilidad de Z mayor que un número positivo:

P(Z > a)

si te fijas, Z > a es lo contrario de Z ≤ a. Por tanto, esta probabilidad se calcula aplicando la propiedad de los sucesos complementarios $P(\overline{A})=1-P(A)$

Luego P(Z > a) = 1 - P(Z ≤ a)

Vamos ahora con la probabilidad de que una distribución sea menor que un número negativo:
P(Z < -a)
P(Z ≤ -a) = 1 - P(Z ≤ a)

Fíjate en el siguiente ejemplo:

Supongamos que queremos calcular P{ Z ≤ -1,53}. Dicha probabilidad está representada por el área sombreada en la siguiente figura 1.

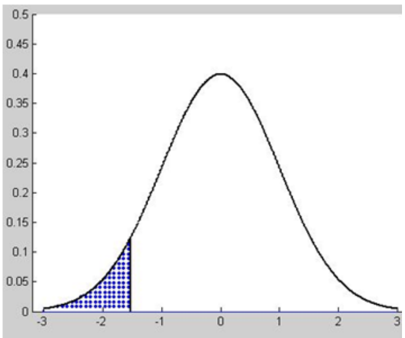


Figura 1

El número -1,53 no figura en la tabla, pero eso no nos impide calcular la probabilidad en cuestión. Simplemante hay que tener en cuenta que, por la simetría de la campana de Gauss se tiene:

P{ Z ≤-1,53}= P{ Z >1,53}

La probabilidad que figura en el segundo miembro de la ecuación está respresentada en el área sombreada en la figura 2:

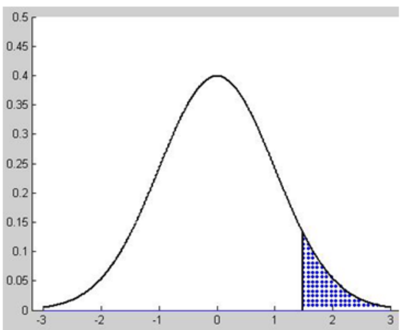


Figura 2

Dicha probabilidad es la complementaria de la probabilidad P{ Z ≤ 1,53}, representada en la figura 3.

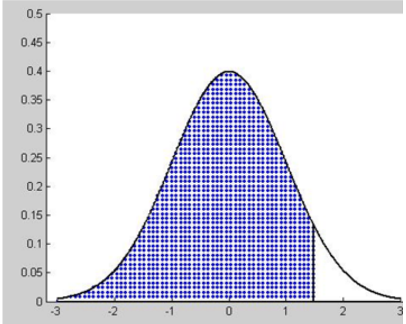


Figura 3

Es decir: P{ Z ≤ 1,53}+ P{ Z > 1,53}= 1. Para hallar P{ Z ≤ 1,53} simplemente vamos a la tabla y consultamos:

Tabla de la distribución nor

z	0.00	0.01	0.02	0.03	
0.0	0.5000	0.5040	0.5080	0.5120	0.5
0.1	0.5398	0.5438	0.5478	0.5517	0.5
0.2	0.5793	0.5832	0.5871	0.5910	0.5
0.3	0.6179	0.6217	0.6255	0.6293	0.6
0.4	0.6554	0.6591	0.6628	0.6664	0.6
0.5	0.6915	0.6950	0.6985	0.7019	0.7
0.6	0.7257	0.7291	0.7324	0.7357	0.7
0.7	0.7580	0.7611	0.7642	0.7673	0.7
0.8	0.7881	0.7910	0.7939	0.7967	0.7
0.9	0.8159	0.8186	0.8212	0.8238	0.82
1.0	0.8413	0.8438	0.8461	0.8485	0.85
1.1	0.8643	0.8665	0.8686	0.8708	0.87
1.2	0.8849	0.8869	0.8888	0.8907	0.89
1.3	0.9032	0.9049	0.9066	0.9082	0.90
1.4	0.9192	0.9207	0.9222	0.9236	0.925
1.5	0.9332	0.9345	0.9357	0.9370	0.93
1.6	0.9452	0.9463	0.9474	0.9484	0.94
1.7	[illegible]	[illegible]	0.9573	[illegible]	[illegible]

De aquí obtenemos P{ Z ≤ 1,53} = 0,9370 y, por lo tanto:

P{ Z ≤ -1,53}= P{ Z > 1,53} = 1 - P{ Z ≤ 1,53}= 1- 0,9370 = 0,0630

Y ahora la de Z > que un número negativo:

P(Z > -a)

Hacemos lo mismo que antes, aplicando la propiedad de sucesos complementarios, esta probabilidad es 1 - P(Z ≤ -a) que es la que acabamos de calcular, luego:

P(Z > -a) = 1 - P(Z ≤ -a) = 1 - [1 - P(Z ≤ a)] = 1 - 1 + P(Z ≤ a) = P(Z ≤ a).

Y para terminar la de un intervalo:

P(a ≤ Z ≤ b)

Se cumple que:

P(a ≤ Z ≤ b) = P(Z ≤ b) - P(Z ≤ a)

Fíjate en el siguiente ejemplo:

Supongamos que queremos calcular P{0,41 < Z ≤ 1,62}. Esta probabilidad está representada por el área sombreada en la figura 1.

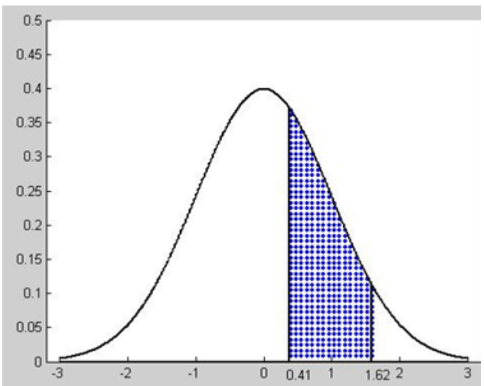


Figura 1

Dicha probabilidad se puede calcular como

$$P\{ 0,41 < Z \leq 1,62\}.= P\{ Z \leq 1,62\}- P\{Z \leq 0,41\}$$

El minuendo y el sustraendo están representados por las áreas sombreadas en las siguientes figuras 2 y 3, respectivamente.

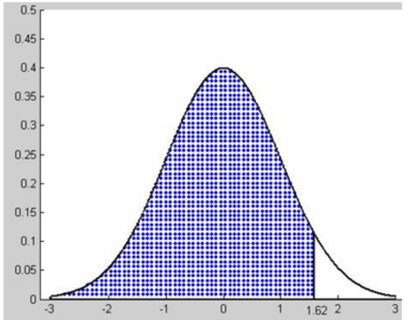


Figura 2

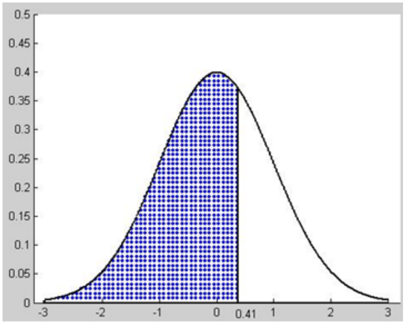


Figura 3

La busca en la tabla nos da los valores:

$$P\{ Z \leq 1,62\} = 0,9474, \quad \text{y} \quad P\{Z \leq 0,41\} = 0,6591 .$$

Por lo tanto:

$$P\{ 0,41 < Z \leq 1,62\}= 0,9474 - 0,6591= 0,2883$$

Importante

Cálculo de probabilidades en una distribución Normal.

Si "**a**" es un número positivo y **Z** sigue una distribución **N(0,1)**:

- $P(Z < a) \rightarrow$ A partir de la tabla de probabilidades de una distribución N(0,1)
- $P(Z > a) = 1 - P(Z < a)$
- $P(Z < -a) = 1 - P(Z < a)$
- $P(Z > -a) = P (Z < a)$
- $P(a < Z < b) = P(Z < b) - P(Z < a)$

Imagen de Dave Dugdale bajo licencia Creative Common

Ejercicio resuelto

Imagen de Fartese bajolicencia Creative Commons

Uno de los juegos favoritos de Gonzalo y Blanca cuando van al casino es la ruleta, pero tras mucho tiempo jugando a ella, María José, la profesora de la Universidad de Nagora, ha analizado sus ganancias en cientos de euros y ha llegado a la conclusión de que siguen una distribución normal N(0,1), por lo que no es demasiado rentable, pues como término medio esperan ganar 0 €. ¡Bueno, al menos no pierde!

Pero aprovechando esto, vamos a plantear algunas cuestiones sobre las posibilidades que tienen Blanca y Gonzalo uno de los días que van al casino.

Por ejemplo, ¿con qué probabilidad ganan algo? ¿Con qué probabilidad ganan menos de 250 €? ¿Y más de 125? ¿Con qué probabilidad pierden más de 234 €? ¿Y menos de 301 €?

¿Es muy probable que un día cualquiera su ganancia esté entre 94 y 338 €?

Mostrar retroalimentación

Imagen de justj0000lie bajo licencia Creative Commons

Empezamos. como tenemos una distribución normal de media cero y desviación típica 1, vamos a llamar Z a nuestra variable, aunque podíamos haberla llamado X perfectamente. Entonces:

Z=Dinero ganado por Gonzalo y Blanca (en cientos de €) $\rightarrow$ Z~N(0,1)

En cada caso, tenemos que observar como cuál delos cinco apartados del "Importante" es, y una vez que hemos hecho el cambio pertinente, buscar en la tabla la probabilidad de Z menor que el número positivo.

P(Ganar algo) = P(Z > 0) → Estamos en el segundo caso, así que:

P(Z > 0)= 1 - P(Z ≤ 0) = 1 - 0,5 = 0,5.

O sea, que tiene las mismas probabilidades de ganar algo de dinero que de perderlo.

P(Ganar menos de 250 €) = P(Z < 2,50) *(recuerda que la variable se expresa en cientos de euros, no en euros)*→ Primer caso, por tanto, miramos directamente en la tabla.

P(Z < 2,50) = 0,9938

P(Ganar más de 125 €) = P(Z > 1,25) → 2º caso, luego 1 menos probabilidad Z <.

P(Z > 1,25) = 1 - P(Z ≤ 1,25) = 1 - 0,8944 = 0,1056

P(perder más de 234 €). Fíjate que perder dinero es como ganar dinero negativo (Z expresa ganancia), y si pierdes más de 234 euros quiere decir que has perdido 235, 236,... ó lo que es lo mismo que ganas -235, -236,...

Luego P(perder más de 234) = P(ganar menos de -234 €) = P(Z < -2,34). Estamos entonces en el tercer caso.

Entonces, P(Z < -2,34) = 1 - P(Z < 2,34) = 1 - 0,9904 = 0,0096. Es, por tanto, muy poco probable que pierda tanto dinero.

P(perder menos de 301 €). Aplicando lo mismo de antes, perder menos de 301 es ganar más de -301 €, luego tenemos que calcular la probabilidad P(Z > -3,01). 4º caso pues (Z > negativo = Z < positivo)

$P(Z > -3{,}01) = P(Z < 3{,}01) = 0{,}99865$ mirando en la tabla. Es por tanto casi seguro que va a perder menos de 301 €.

P(Ganar entre 94 y 338) $= P(0{,}94 < Z < 3{,}38)$ Estamos en el caso de un intervalo, así la probabilidad de que sea menor que el segundo menos la probabilidad de que sea menor que el primero:

$P(0{,}94 \leq Z \leq 3{,}38) = P(Z \leq 3{,}38) - P(Z \leq 0{,}94) = 0{,}99964 - 0{,}8264 = 0{,}17324$, así que no, no es demasiado probable.

Comprueba lo aprendido Blanco

Ahora te toca a ti. Calcula las siguientes probabilidades usando las propiedades vistas arriba y la tabla de probabilidades de la distribución Normal N(0,1) y como siempre separa la parte entera de la decimal con ",":

Z sigue una distribución N(0,1). Calcula:

1. $P(Z \geq 0{,}32) =$
2. $P(Z \leq 0) =$
3. $P(Z > 0{,}7) =$
4. $P(-0{,}51 \leq Z \leq 0{,}51) =$
5. $P(Z > -2{,}63) =$

Enviar

1. $P(Z \geq 0{,}32) = 1 - P(Z < 0{,}32) = 1 - 0{,}6255 = 0{,}3745$
2. $P(Z \leq 0) = 0{,}5$
3. $P(Z > 0{,}7) = 1 - P(Z < 0{,}7) = 1 - 0{,}7580 = 0{,}242$
4. $P(-0{,}51 \leq Z \leq 0{,}51) = P(Z \leq 0{,}51) - P(Z < -0{,}51) = P(Z \leq 0{,}51) - [1 - P(Z < 0{,}51)] = 0{,}695 - (1 - 0{,}695) = 0{,}39$
5. $P(Z > -2{,}63) = P(Z < 2{,}63) = 0{,}9957$

3.3. Tipificación de una variable aleatoria que sigue una Distribución Normal

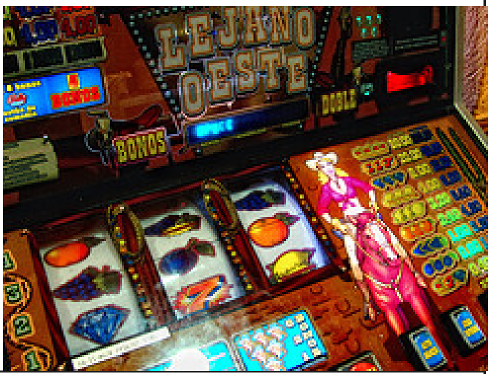


Imagen de Andres Rueda bajo licencia Creative Commons

Ahora ya sí que podemos calcular cualquier probabilidad normal.

Bueno, casi, y es que, la media de una distribución no tiene porqué ser siempre cero, ¿no? Por ejemplo, las ganancias mensuales de uno de los casinos que frecuentan Blanca y Gonzalo son aproximadamente de 18.500 € de media, la vida útil de las máquinas tragaperras del casino, como la que tienes a la izquierda es de 8,4 años, el número medio de clientes que acuden diariamente al bar La Tertulia, es de 254, y así podríamos poner un montón de ejemplos. Bueno, y tampoco la desviación típica va a ser siempre uno.¡Seguro que va tener distintos valores!

Entonces, ¿qué hacemos? ¿No habíamos dicho que con un folio íbamos a tener suficiente?

Pues así es querido alumno o querida alumna. Tan solo hemos de hacer una pequeña operación para llevarnos cualquier distribución Normal a una N(0 , 1) y ese proceso se llama TIPIFICAR. Cuando tengamos esto, ya si te prometemos que no nos hace falta nada más, que podremos calcular la probabilidad de que una televisión dure 20 años , que una bombilla se funda en un par de meses o de que tengamos que ir al médico porque nuestro nivel de colesterol no esté muy allá.

Importante

Tipificar una variable aleatoria continua que sigue una distribución normal de parámetros μ y σ, consiste en convertirla en una Normal de parámetros 0 y 1. Para ello, se le resta a la variable el valor de la media y se divide todo por el valor de la desviación típica. Es decir:

$$Si\ \ X \sim N(\mu,\sigma) \longrightarrow Z = \frac{X-\mu}{\sigma} \sim N(0{,}1)$$

Por ejemplo, supón que X~N(14 ; 3,1) y necesitas calcular la probabilidad de que X sea menor que 11. Procederíamos así:

$$P(X<11) = P\left(\frac{x-14}{3{,}1} < \frac{11-14}{3{,}1}\right) = P(Z<-0{,}9677)$$

Y eso ya lo calcularíamos como en el apartado anterior, pues Z es ya una distribución N(0,1).

De todas formas, no te preocupes que con los siguientes ejemplos, seguro que lo vas a entender.

Ejercicio resuelto

En este primer ejemplo, vamos a aplicar la distribución normal a la fabricación de coches. Fíjate cómo resuelve la situación en el siguiente vídeo:

Mostrar retroalimentación

X~N(36; 1,4) y nos pide calcular la probabilidad de que X sea menor que 38,8. El cálculo se ha hecho así:

$$P(X \leq 38{,}8) = P(Z \leq \frac{38{,}8-36}{1{,}4}) = P(Z \leq 2) = 0{,}9772$$

Luego la probabilidad de que el asiento le venga bien al conductor es del 97,72%

Segundo ejemplo. Hay que vigilar un poquito el peso...

Imagen de Bocadorada bajo licencia Creative Commons.

Y tercer ejemplo, como la salud preocupa a todo el mundo, ahora hablamos de uno de los problemas que tiene buena parte de nuestra población en la sedentaria sociedad actual; el colesterol.

La distribución normal tiene muchísimas aplicaciones en la medicina pues muchos indicadores se distribuyen según un modelo normal. En este ejemplo, vemos que el nivel de colesterol en sangre de una persona se distribuye según una normal.

El enunciado nos dice que en una ciudad de 300.000 habitantes el nivel de colesterol en sangre sigue una distribución normal de media 192 y varianza 144. Nos pide calcular la probabilidad de que una persona elegida al azar tenga un nivel de colesterol comprendido entre 186 y 200, el número de personas con nivel de colesterol elevado, considerándose colesterol elevado si es superior a 235 y por último, nos pide determinar un nivel de colesterol de forma que el 84,2% de la población tenga un nivel superior a ese.

En la presentación que sigue está la solución de estas cuestiones.

Mostrar retroalimentación

Ve pulsando sobre la animación para ir avanzando

Curiosidad

Fíjate que en este vídeo. Se aplica la distribución normal al tiempo de ejecución de una obra. **¡Si es que vale para todo!**

Aquí, al final, la probabilidad la calculan de forma un poco distinta. Nosotros nos iríamos a la tabla directamente con el valor que nos sale al tipificar. Comprueba que es el mismo en la tabla.

Comprueba lo aprendido

ltiple

Hemos dicho al principio de este apartado que el número de clientes diarios del bar La Tertulia es de 254. Vamos a suponer que X=n.º de clientes diarios, sigue una distribución Normal y además, que la desviación típica es 5,4.

Lo mismo te has dado cuenta de que el número de clientes es una variable discreta, no puede haber 4 clientes y medio. Si es así, enhorabuena, esto marcha muy bien. Lo que ocurre es que a pesar de ser discreto, cuando los valores son muy grandes puede aproximarse a un modelo normal. en la próxima unidad veremos por qué es así.

Bueno, independientemente de esto, calcula las siguientes probabilidades:

1) $P(X \leq 260)$

Sugerencia

- 0,8438
- 0,8686
- 0,8461
- 0,8665

No es correcto, busca bien en la tabla.

No es correcto, redondea bien.

No has tipificado bien.

Muy bien

Solution

1. Incorrecto
2. Incorrecto
3. Incorrecto
4. Opción correcta

2) La probabilidad de que un día acudan menos de 240 personas

Sugerencia

- 0,8212
- 0,9952
- 0,0048
- 0,1788

Tipifica bien y aplica bien las reglas de la N(0,1)

No es correcto, revisa cómo se calcula la probabilidad de un número negativo.

Muy bien

No es correcto. Repasa cómo se hace la tipificación.

Solution

1. Incorrecto
2. Incorrecto
3. Opción correcta
4. Incorrecto

3) Manolo, el dueño del bar, hombre previsor donde los haya, siempre tiene reservas de existencias para atender a unas 280 personas que consuman lo habitual en su bar. ¿Puede estar tranquilo o es probable que algún día vayan más de 280 personas?

Sugerencia

- Sí, puede estar tranquilo porque $P(X > 280) = 0$
- Al contrario, es seguro que cualquier día le van a entrar más de 280 personas, pues la $P(X>280) = 1$.
- $P(X>280) = 0,8599$, así que no debe estar nada de tranquilo.
- Así así, pues la probabilidad de que un día acudan más de 280 personas es del 14,01%.

Correcto

Repasa las reglas de la N(0,1)

Repasa cómo se tipifica y cómo se aplican las reglas de la normal.

Repasa cómo se tipifica

Solution

1. Opción correcta
2. Incorrecto
3. Incorrecto
4. Incorrecto

Para saber más

Para terminar, en este vídeo te muestran algunas propiedades de la distribución y las relaciones que indican los valores de la media y la desviación típica:

4. Ejemplo de Selectividad

Ejercicio resuelto

Selectividad

En selectividad suele aparecer de forma habitual al menos un ejercicio en el que aparece la Distribución Normal, pero siempre unido con la realización de muestreos estadísticos, que aprenderás en la Unidad 5, o junto a lo que se llaman los "Contrastes de hipótesis", que verás en la Unidad 6.

En cualquier caso hay que saber tipificar la variable normal y consultar valores en la tabla de la N(0, 1), como has aprendido en la presente unidad.

JUNIO 2009

El cociente intelectual de los alumnos de un centro educativo se distribuye según una ley Normal de media 110 y desviación típica 15. Se extrae una muestra aleatoria simple de 25 alumnos.
a) (1'5 puntos) ¿Cuál es la probabilidad de que la media del cociente intelectual de los alumnos de esa muestra sea superior a 113?
b) (0'5 puntos) Razone cómo se vería afectada la respuesta a la pregunta anterior si el tamaño de la muestra aumentase.

Ejercicio tomado de http://www.iesayala.com/selectividadmatematicas/

Mostrar retroalimentación

a) Tenemos una variable aleatoria X que sigue una normal N(μ, σ), de la que nos dan la media μ=110 y la desviación típica σ = 15.

Sin embargo, al decirnos que utilizamos una muestra de 25 alumnos, la variable que hay que analizar es lo que se llama distribución muestral de las medias, que se denota por $\overline{X}$, que es una variable normal $N(\mu,\frac{\sigma}{\sqrt{n}})$ con n=25 en nuestro caso (el número de elementos de la muestra), luego la distribución normal de la que hay que estudiar la probabilidad y tipificarla será:

$$\overline{X} \sim N(110,\frac{15}{\sqrt{25}}) = N(110,\frac{15}{5}) = N(110,3)$$

Conociendo esto, veamos la resolución:

Datos: X sigue una normal N(110, 15), luego $\mu = 110 = \overline{x}$ y $\sigma = 15$; n = 25

Me están pidiendo p($\overline{X}$ >113) = { tipifico $Z = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$ } = p($Z > \frac{113-110}{15/\sqrt{25}}$) = p(Z > 1) = 1 - p(Z ≤ 1) = {mirando en las tablas de la N(0'1) } = 1 – 0'8413 = 0'1547.

b)
Razone cómo se vería afectada la respuesta a la pregunta anterior si el tamaño de la muestra aumentase.

Si nos fijamos en la expresión "{ tipifico $Z = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$ } = p(Z > 1)", y aumentase el tamaño de "n", el número "1"de p(Z < "1") aumentaría, y aumentaría p(Z < "nº nuevo), con lo cual como estamos calculando la probabilidad del contrario, **la probabilidad que me están pidiendo disminuiría**.

Resumen

Importante

Una variable aleatoria es continua si al realizar el experimento aleatorio, entre cada dos valores, el número de valores que puede tomar es infinito.

Por ejemplo, la altura de una persona, la longitud del dedo índice, el peso de un perro o el caudal de un río son variables aleatorias continuas, pues entre dos posibles valores puede tomar cualquiera que exista. Todo depende de la precisión de nuestro aparato de medida.

Imagen de thombo2 bajo licencia Creative Commons

Importante

Si X es una variable aleatoria continua, la probabilidad de que tome un valor concreto es cero.

P[X = a] = 0, para cualquier valor de a.

Por tanto, en variables continuas, las probabilidades se calculan sobre intervalos y serán de esta forma:

- $P(a \leq X \leq b)$
- $P(X \geq a)$
- $P(X \leq b)$

Además, como el signo "=" en un punto da probabilidad cero, en los intervalos anteriores da igual poner desigualdades estrictas o inclusivas, es decir, da lo mismo escribir ≤ ó < y es lo mismo escribir ≥ ó >.

Por ejemplo, si X es una variable aleatoria continua, $P(X \leq 2) = P(X < 2)$, pues la diferencia entre ambas es que entre el 2 o no, y como ya sabemos que la probabilidad del valor concreto 2 es cero, da igual que lo incluyamos o no.

Importante

Cálculo de probabilidades en una distribución Normal

distribución Normal.

Si "**a**" es un número positivo y **Z** sigue una distribución **N(0,1)**:

- $P(Z < a)$ → A partir de la tabla de probabilidades de una distribución N(0,1)
- $P(Z > a) = 1 - P(Z < a)$
- $P(Z < -a) = 1 - P(Z < a)$
- $P(Z > -a) = P(Z < a)$
- $P(a < Z < b) = P(Z < b) - P(Z < a)$

Imagen de Dave Dugdale bajo licencia Creative Common

Importante

Tipificar una variable aleatoria continua que sigue una distribución normal de parámetros μ y σ, consiste en convertirla en una Normal de parámetros 0 y 1. Para ello, se le resta a la variable el valor de la media y se divide todo por el valor de la desviación típica. Es decir:

$$Si\ \ X \sim N(\mu,\sigma) \longrightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

Aviso Legal

AVISO DEL SERVIDOR

Por motivos de seguridad esta página web solo está accesible mediante acceso seguro (https):

https://www.juntadeandalucia.es/Aviso_Legal_Andalucia_v04.htm

Por favor, actualice sus marcadores. Gracias.