

# La curva normal

Vicente Manzano Arrondo – 2012-2014

## La respuesta está en la historia

Abraham De Moivre nació en Francia en 1667, el mismo año en que su compatriota Jean-Baptiste Denys realizó la primera transfusión de sangre a un humano y otro Jean-Baptiste, esta vez de apellido Poquelin, pero conocido como Molière, escribiera sus últimas obras, ya enfermo.

De Moivre se formó con varios maestros, con los que mostró un excelente intelecto para las matemáticas. Pero aproximadamente con 20 años tuvo que huir de Francia por ser calvinista. Se refugió en Inglaterra. Aterrizó en terreno inglés más o menos cuando Isaac Newton andaba publicando su famoso libro Principia Matemática (que, realmente se llamaba *Philosophiae Naturalis Principia Mathematica*). De Moivre conoció a este famoso inglés, pasado a la posteridad para muchos como la persona más inteligente de la historia. Curiosa sentencia dedicada a quien fuera un puritano escrupuloso y un misógino confeso que expiró orgullosamente virgen tras forzar comisiones para quitarse enemigos de encima y hacer la vida imposible a mucha gente. Por su parte, De Moivre murió en el exilio, tras décadas en el límite continuo de la pobreza, aunque reconocido como un genio matemático por algunos de sus contemporáneos. Tenía nada menos que 87 años. Era 1754, el mismo año en que nació quien sería Luis XVI, el rey guillotinado en la revolución francesa.

Como muchas personas con un intelecto muy desarrollado y algo de tiempo para jugar con él (vivió pobre, pero no nació pobre), De Moivre se dedicó a cosas inútiles sin cuyo desarrollo el conocimiento científico no habría llegado hasta donde hoy se encuentra (esté donde esté ese sitio). Otros franceses aún más famosos que él, como Pascal o Fermat (que murieron ambos cinco años antes de nacer De Moivre), desarrollaban la teoría de las probabilidades con tareas tan serias como el juego de los dados o las partidas de cartas. De Moivre seguía intrigado por estas cuestiones, además de por otras muchas, observando un efecto gráfico curioso. Vamos a reproducirlo, más o menos.

Al lanzar una moneda al aire, hay la misma probabilidad de que aparezca cara que cruz. Hablamos de una probabilidad de  $1/2$  (un medio o 0,5). Si se lanza una moneda una vez, el gráfico de probabilidades para los posibles resultados muestra dos barras con la misma altura: 0,5 para el resultado 1 cara (o 0 cruces), y 0,5 para el resultado 0 caras (es decir, 1 cruz). Si lanzamos la moneda mil veces, cabría esperar una altura de 500 resultados para la cara y otra barra con igualmente 500 resultados para la cruz. Dado que la probabilidad es muy suya, tal vez no sea 500 sino 498, por ejemplo; pero no nos pondremos muy exigentes con ello.

Si se lanza la moneda dos veces, los resultados posibles son: 0, 1 y 2 para el número de caras. Las barras ya no muestran la misma altura. Es natural: lo más probable es que de dos lanzamientos salga una cara y una cruz (probabilidad de 0,5), antes que dos caras (probabilidad de 0,25) o dos cruces (probabilidad de 0,25). Al lanzar la moneda tres veces, los resultados posibles aumentan (0, 1, 2 o 3 caras). Las siguientes representaciones gráficas (agrupadas como figura 1) muestran esa evolución. Al sumar la altura de todas las barras, el resultado ha de ser siempre el mismo: 1, pues representa la totalidad expresada como proporción. Si la expresáramos como porcentaje, el total sería 100%.

La forma que iba tomando la distribución ante los ojos y deducciones del matemático era atractiva (ya ves, la gente se entretiene con cualquier cosa). De Moivre observó que conforme aumenta el número de lanzamientos de monedas y hay más resultados posibles para el número de caras (o de cruces, qué más da), la forma de la

representación gráfica se parecía cada vez más a una campana. Lo de menos era la forma final, lo interesante era la percepción de que se aproximaba a algo. Eso es un aliciente irresistible para un matemático. Y no se resistió.

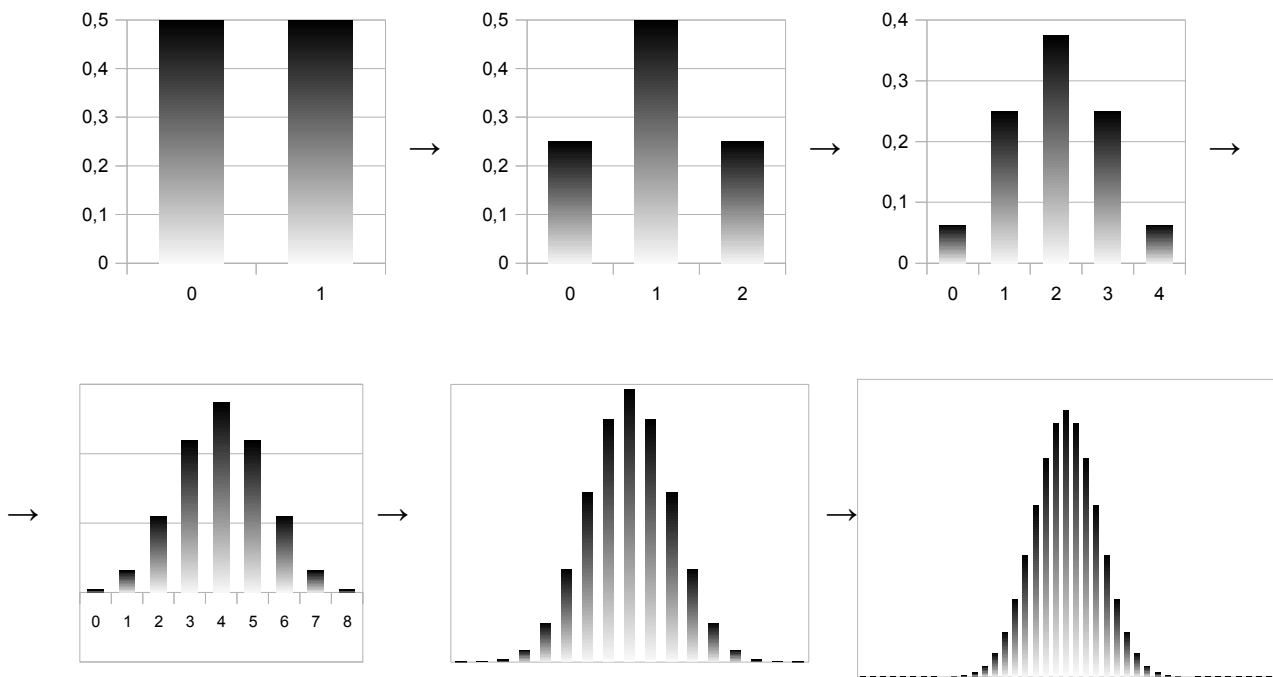


Figura 1. Evolución hacia la normal del lanzamiento de una moneda.

De Moivre se planteó identificar la función matemática a la que se aproxima la distribución de probabilidades de  $n$  lanzamientos de una moneda. Y la encontró. Imagina la de días que se llevó el hombre en esta tarea, en una época sin ordenadores ni calculadoras de bolsillo, donde la gente escribía mojando una pluma de ave en un frasco de cristal con tinta de calamar. El resultado es, para muchas personas, una función matemática de extrema belleza. Para otras, tiene la misma hermosura que un caracol comiendo perejil. Sea como fuere, esta curva inició con De Moivre una historia que llega hasta ti. La función que encontró el calvinista exiliado suministra la altura de la curva para cada valor del eje horizontal, que es:

$$f(x) = \frac{1}{S \sqrt{e^{-\frac{(X_i - \bar{X})^2}{S^2}} 2\pi}} = \frac{1}{\sqrt{e^{-Z_i^2} 2\pi}} = (e^{-Z_i^2} 2\pi)^{-\frac{1}{2}}$$

Para quien no maneje con soltura las expresiones matemáticas, la fórmula anterior puede generarle cierto estupor. Observa, no obstante, que todos sus elementos deben resultarte familiares.  $S$  es la desviación tipo de la variable  $X$ , de la que puedes ver también la representación del valor concreto ( $X_i$ ) y la media aritmética ( $\bar{X}$ ).  $e$  es el llamado número natural, cuyo valor tiene una cantidad infinita de dígitos decimales (2,71828182845904...). Y  $\pi$  es el número *pi*, cuyo valor, también provisto de infinitos dígitos decimales, viene a ser 3,14159265358979... A la derecha observas la misma función, pero expresada para puntuaciones tipo en lugar de puntuaciones originales. En ese caso, sabes que  $S=1$  (por lo que desaparece) y el exponente de  $e$  se simplifica mucho.

Al representar gráficamente la función que encontró De Moivre, se obtiene lo que puedes observar en la figura 2.

Casi un siglo después, a inicios del XIX, un ya famoso matemático y físico, Gauss, lanzó al estrellato la curva normal. Este genio del siglo utilizó la curva casi hasta en la sopa, especialmente para estudiar los errores de medición en astronomía, un asunto que había cautivado a Galileo, quien por cierto casi desarrolla la curva normal un siglo antes que De Moivre. Tal fue el papel de Gauss en la fama y utilización de esta función, que es normal referirse a ella como *campana de Gauss*. Antes que este alemán de lujo, otro francés, temido por los estudiantes de probabilidad desde hace tres siglos, Laplace, formalizó varios desarrollos a partir de la curva normal, por lo que también se la conoce como curva de Laplace-Gauss. En fin, para todos los gustos.



Figura 2. Curva normal.

### Por qué nos ha dado por la normal

Eso, poco más o menos, es lo que les escuchaba a los vecinos cuando yo era pequeño. Ella le gritaba “¿Tú cuándo vas a ser normal?”. Él le respondía “¡Y dale con lo normal!”.

La curva normal aparece con frecuencia y es difícil de ver, depende del contexto. Existen muchas características que parecen comportarse según una ley normal, es decir, cuya representación gráfica se asemeja a la curva del señor De Moivre. Un ejemplo clásico es la altura. Si medimos la altura de una población numerosa, encontraremos que existe una gran aglomeración en torno a una altura *media* o característica. Conforme nos alejamos de ella, la frecuencia disminuye rápidamente, hasta que esa disminución desacelera, es decir, sigue disminuyendo pero con más suavidad, perdiéndose en los extremos. Otras muchas características no siguen muy bien una función normal, pero se aproximan en el sentido de que muestran cierta simetría y agolpamiento central con dispersión en los extremos.

No obstante, en sentido estricto, hay pocas variables que sigan realmente una ley normal. Lo que sí ocurre más o menos según De Moivre - Laplace - Gauss es que muchas distribuciones *indirectas* son normales. Ocurre con las distribuciones muestrales, lo que comprobaremos en el monográfico siguiente sobre la estimación estadística. Recuerda que una distribución muestral es un conjunto de datos donde cada uno de ellos proviene de medir un mismo estadístico pero en muestras diferentes. Por ejemplo, la

distribución muestral de medias es el resultado de calcular la media aritmética en una infinidad de muestras (imagina, por ejemplo, un millón de muestras, de cada una de las cuales se ha calculado la media aritmética obteniendo un millón de resultados). Pues bien, la forma con que se distribuyen esas medias es habitualmente normal. Lo mismo ocurre con la distribución muestral de proporciones, etc. La facilidad con que los estadísticos se distribuyen según una ley normal es una circunstancia que ha desarrollado la estadística muchísimo. Y lo veremos en otra unidad, si es que quieres que nos encontremos en ella. El modo en que De Moivre llegó a formalizar la curva normal es un buen ejemplo de lo que estamos hablando: una distribución originalmente binomial (tipo éxito/fracaso o cara/cruz) se aproxima a la normal conforme aumenta  $n$ , es decir, el número de veces que se lanza la moneda al aire y se cuenta si ha salido cara o cruz.

Podemos pensar que la curva normal se llama así porque es así, es decir, normal, habitual o frecuente. También nos vale pensar en que el nombre indica que la curva sirve para normalizar o estandarizar determinados procedimientos en estadística. Y también es cierto. Pero lo que va a ser normal es que vamos a hartarnos de utilizarla. Así que vete acostumbrando.

### **Cosas tan curiosas como importantes**

La curva normal tiene algunas características importantes. Veamos algunas de ellas.

1. La curva puede variar de posición a lo largo del eje horizontal, es decir, puede estar más hacia la izquierda o más hacia la derecha. Esa posición se representa bien por la media aritmética. Esto le pasa a todos los conjuntos de datos. Lo peculiar de la curva normal es que la media aritmética es una de las dos únicas características que definen la función.
2. Manteniendo la misma escala, la curva puede ser más ancha o más estrecha, según la desviación tipo de la variable que sigue ese comportamiento normal. En esto tampoco hay nada de particular. Lo relevante es que se trata de la segunda característica de la curva.

Ya no tiene más; es decir, conociendo la media y la desviación tipo, podemos representarla sin necesidad de más información.

Las características 1 y 2 permiten concluir que si sabemos que un conjunto de datos se distribuye según una ley normal y conocemos su media aritmética y su desviación tipo, entonces conocemos todos los datos. Por ejemplo, si la variable  $X$  se distribuye según una ley normal (o “es normal” o “sigue una distribución normal”) y tiene el valor 90 como media aritmética y 20 de desviación tipo, entonces sabemos que un 15% de la población tiene puntuaciones comprendidas entre  $X_i = 70$  y  $X_i = 80$ , que en términos estandarizados se expresarían respectivamente como  $Z_i = -1$  y  $Z_i = -0,5$ .

3. La curva es simétrica. En otras palabras: según un eje de simetría vertical, una de las dos mitades es un reflejo exacto de la otra. Como es simétrica, la media y la mediana coinciden (a ambos lados se encuentra el mismo número de datos y el mismo peso). Como solo tiene una moda, le pasa como a todas las distribuciones simétricas unimodales: la moda se encuentra necesariamente en el centro, por lo que también coincide con la mediana y con la media.
4. Muestra agolpamiento en el centro y dispersión hacia los extremos.
5. El modo en que se dispersa desde el centro es acelerada, es decir, disminuye con rapidez, hasta que llega a un punto de inflexión en que se desacelera. Ese punto

es el valor que se encuentra a 1 desviación tipo a ambos lados de la media. (obviamente, es a *ambos lados* puesto que la curva es simétrica).

6. Aproximadamente el 95% de los datos (la gran mayoría) se encuentra, como mucho, a 2 desviaciones tipo de la media. El 99% (la inmensa mayoría) llega a poco más de 2,5 desviaciones tipo de distancia.

Lo habitual es que la curva normal se utilice estandarizada, es decir, que las puntuaciones originales se traduzcan a puntuaciones tipo. Es lo que he hecho en el punto 6 del esquema: que algo esté a 2 desviaciones tipo de la media es que muestra una puntuación tipo de valor 2 (si está por encima de la media) o de valor -2 (si está por debajo).

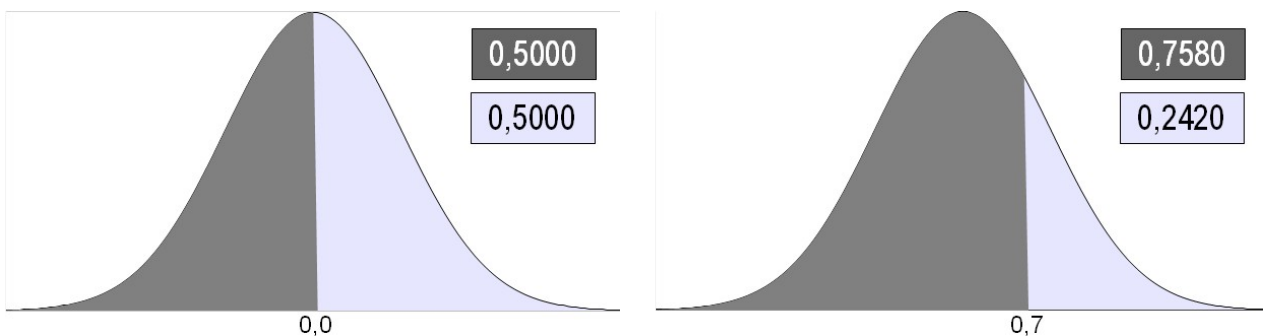
Recordemos que las puntuaciones tipo tienen de media 0 y desviación tipo 1. Esto simplifica bastante la fórmula de De Moivre, tal y como hemos visto más arriba. Como resulta que una curva normal está caracterizada por su media y su desviación tipo y en la estandarizada estos valores son siempre los mismos y de cuantías interesantes, entonces cualquier conjunto de datos que sea normal puede ser representado por la curva estandarizada, sin necesidad de manejar infinitas curvas, una para cada una de las infinitas combinaciones posibles de valores para la media y la desviación tipo.

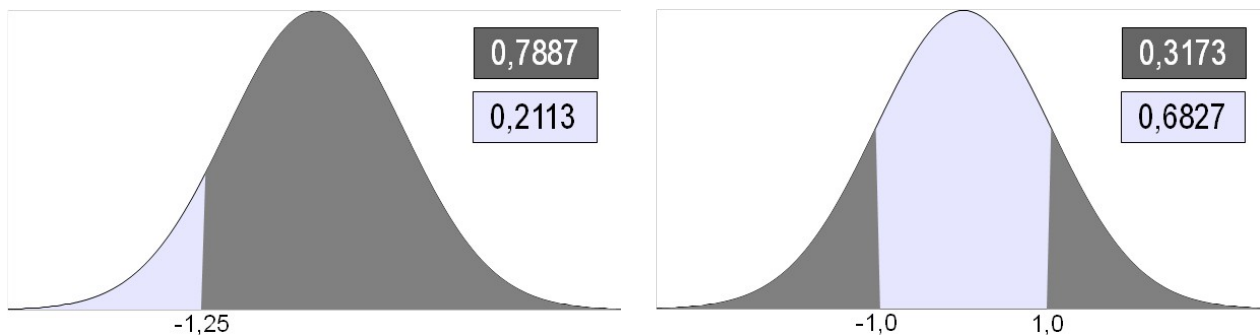
Para muchos menesteres se utilizan tablas de la curva normal estandarizada, que asocian puntuaciones tipo con probabilidades. Por ello, estas tablas permiten traducir de puntuaciones tipo a probabilidades o de probabilidades a puntuaciones tipo. Por ejemplo, utilizando una tabla podemos saber que el 95% de los datos en una curva normal se encuentran alejados de la media en no más de 1,96 desviaciones tipo. En otras palabras, en una curva normal, el 95% central de los datos se encuentra entre las puntuaciones estandarizadas de  $Z = -1,96$  y  $Z = 1,96$ . Dicho también de otro modo: la probabilidad de encontrar en una curva normal datos que se alejen de la media en no más de 1,96 desviaciones tipo es del 95% o de 0,95.

Para afianzar estas ideas, veamos algunas áreas bajo la curva normal, asociadas a valores concretos de puntuaciones tipo. Para entender bien los gráficos, recordemos que las probabilidades se miden como proporciones o tantos por uno, es decir, como porciones de la unidad. Así, por ejemplo:

1. Algo imposible de que ocurra tiene la probabilidad 0.
2. Algo que ocurre seguro tiene la probabilidad 1.
3. Algo que ocurre la mitad de las veces, tiene de probabilidad 0,5.
4. Algo que ocurre un 75% de las ocasiones, tiene una probabilidad de valor 0,75.

En las siguientes representaciones gráficas de áreas bajo la curva normal, se muestra el eje horizontal en escala de puntuaciones típicas. Cada una de las áreas coloreadas representa una probabilidad, una proporción o una porción de área que se expresa en el recuadro de su mismo color.



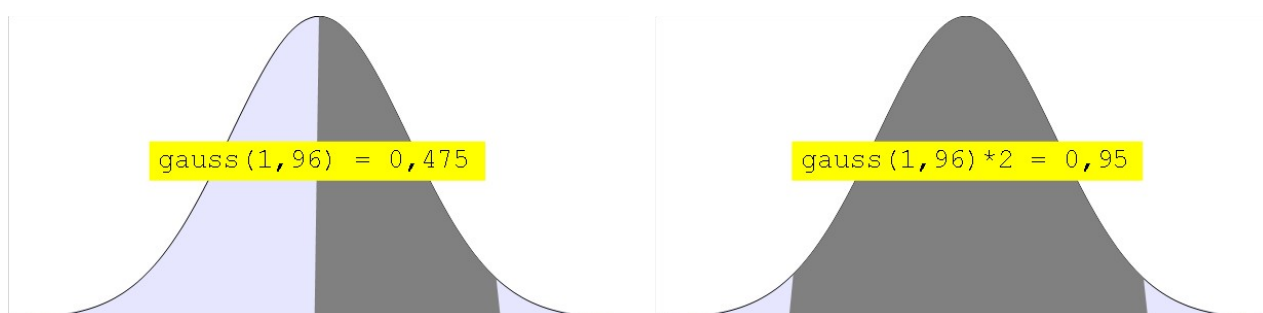


## Recursos y tablas

No solo para afianzar estos conocimientos teóricos sino también porque necesitaremos calcular probabilidades (o frecuencias o áreas) a partir de puntuaciones directas o típicas de una curva normal, o hacer el camino inverso... conviene ponerse manos a la obra y realizar algunos de esos cálculos. Inicialmente, este cometido exige calcular áreas bajo la curva normal aplicando integrales definidas a la fórmula de De Moivre. No te preocupes. Aunque hay personas que disfrutan con ello, supongo que no es tu caso. Lo habitual es acudir a tablas impresas y, cada vez con mayor frecuencia, a programas de ordenador. Para ejemplificar ambos procedimientos, vamos a acudir a una tabla que verás al final de este documento y a la hoja informatizada de cálculo de Libre Office Calc (de libre distribución, muy similar al software propietario de Microsoft Excel). Empecemos por este último recurso. Si no vas a utilizarlo, puedes saltar el apartado.

### *Pasar de puntuaciones a proporciones con Calc*

La función `gauss(X)` devuelve el área bajo la curva normal que se encuentra entre la media aritmética (centro de la distribución) y  $X$  desviaciones tipo hacia uno de los lados. Da igual cuál, ya que recuerda que estamos ante una curva simétrica. Por ejemplo: `gauss(1,96) = 0,475`. En las tres figuras siguientes, observamos cuatro tipos de áreas diferentes calculadas a partir de la misma función `gauss(X)`.





La siguiente tabla muestra un conjunto de datos ( $X_i$ ) que provienen de una distribución normal de media 50 y desviación tipo 10. Por ejemplo,  $X_i = 52$  se encuentra a 2 unidades por encima de la media. Como  $S = 10$ , esas dos unidades se estandarizan como  $Z_i = 0,2$  ( $Z_i = [52-50]/10$ ). El objetivo de la tabla es calcular la probabilidad de encontrar datos en esa distribución que se alejan de la media tanto o menos que cada valor  $X_i$ . Por eso, el resultado se denomina % centrado, como podría llamarse área centrada, entre otras posibilidades. Esa columna se expresa en porcentajes. Para utilizar la función de cálculo, hay que introducir como argumento distancias estandarizadas, no valores o puntuaciones directas. Por eso he creado la columna  $Z_i$ . Lo que hace es traducir el valor  $X_i$  a una distancia estandarizada, siguiendo la expresión que ya conocemos:

$$Z_i = \frac{X_i - \bar{X}}{S} = \frac{X_i - 50}{10}$$

Otro ejemplo: un 66% de los datos de ese conjunto que sigue una ley normal se alejan de la media (50) en no más de 9 unidades (pues  $59-50=9$ ). Para eso, imagina que el valor 59 se encuentra en la celda A5, entonces primero se aplica la fórmula  $= (A5-50) / 10$  para obtener  $Z_i$  (por ejemplo, en la casilla B5). Acto seguido calculamos  $=\text{gauss}(B5) * 200$  (es decir,  $*2*100$  para que no solo se muestre el área centrada que se corresponde con el doble de  $\text{gauss}()$  sino también para que se exprese en tantos por ciento en lugar de tantos por uno o proporción).

Características			
Media=			50
Desv. tipo=			10
$X_i$	$Z_i$	% centrado	
52	0,24	19	
57	0,65	48	
59	0,95	66	
50	0,03	03	
48	-0,20	16	
62	1,23	78	
61	1,06	71	
43	-0,70	52	
54	0,40	31	
52	0,24	19	
56	0,65	48	
46	-0,40	31	
66	1,63	90	
53	0,32	25	
55	0,45	35	

## Pasar de puntuaciones a proporciones con la tabla

Hay muchas posibilidades para construir una tabla de la curva normal tipificada. Y hay muchas situaciones en las que podemos necesitarla. En los intervalos de confianza (lo veremos en otra unidad), se manejan áreas centradas. En las pruebas de significación de la hipótesis nula (otra unidad más), se utilizan áreas extremas que utilizan la misma proporción en ambos extremos (prueba de dos colas) o solo en un extremo (pruebas de una cola). Y en diversas situaciones hace falta otro tipo de áreas, como las que se inician en el extremo izquierdo y superan la media (probabilidad acumulada).

Teniendo en cuenta todas las situaciones posibles, lo más usual es utilizar una tabla que parta de áreas definidas entre la media aritmética y un valor de  $Z$  especificado. A partir de esa tabla, con determinados cálculos, se llega a responder a todas las necesidades. La lógica es exactamente la misma que has visto en el subapartado anterior, que recurre a la función  $\text{gauss}(Z)$  de Calc:

1. La tabla suministra el área ( $A$ ) que se encuentra entre la media y el valor de  $Z$  que se introduce o se utiliza como referente.
2. Si interesa un área diferente, habrá que hacer operaciones:
  - a) Área centrada, es decir, entre  $-Z$  y  $+Z$ :  $2A$ .
  - b) Área extrema, es decir, inferior a  $-Z$  y superior a  $+Z$ :  $1-2A$ .
  - c) Área acumulada, es decir, desde  $-\infty$  hasta  $Z$ :
    - Si  $Z > 0$ , entonces el área es  $0,5+A$ .
    - Si  $Z < 0$ , entonces el área es  $0,5-A$ .

Veamos un ejemplo para cada caso, acudiendo a la tabla de áreas respecto a la media que tienes en el Anexo. Observa que el contenido de cada casilla o celda es la probabilidad de encontrar valores en la curva normal estandarizada que se encuentren entre 0 y el valor de  $Z$  compuesto entre la cabecera de la fila (hasta las décimas) y la cabecera de la columna (centésimas). Por ejemplo, una  $Z=1,35$  se encuentra en la fila 1,30 y la columna 0,05 señalando la celda con valor 0,411. En otras palabras, la probabilidad de encontrar una distancia estandarizada de valor comprendido entre 0 y 1,35 en una distribución normal es del 41,1%.

1. ¿Cuál es el porcentaje de datos en una curva normal que superan a la media en no más de 1,37 desviaciones tipo? Puntuaciones tipo implicadas:  $0 \leq Z \leq 1,37$  (fila 1,30 y columna 0,07). Es precisamente lo que genera la tabla. Luego, el área es  $A=0,415$ . Respuesta: 41,5%.
2. ¿Qué porcentaje de datos se aleja de la media en no más de 1,96 desviaciones tipo? Puntuaciones tipo implicadas:  $-1,96 \leq Z \leq 1,96$ . La tabla solo suministra una mitad (desde 0 hasta 1,96). Luego, el área es  $2A=2 \cdot 0,475=0,95$ . Respuesta: 95%.
3. ¿Qué porcentaje de datos se aleja de la media en, por lo menos, 1,5 desviaciones tipo? Puntuaciones tipo implicadas:  $1,5 \leq |Z|$  (es decir, el valor absoluto de la distancia estandarizada supera a 1,5). El área es  $1-2A=1-2 \cdot 0,433=0,134$ . O bien,  $2(0,5-A)=2(0,5-0,433)=0,134$ . La respuesta es 13,4%.
4. Para áreas acumuladas, dos opciones:
  - a) ¿Cuántos datos, en una curva normal, tienen una puntuación tipo igual o inferior a 1,28?  $-\infty \leq Z$ . Área implicada:  $0,5+0,4=0,9$ . Respuesta: 90%.
  - b) ¿Cuántos datos, en una curva normal, tienen una puntuación tipo igual o inferior a -0,80?  $-\infty \leq Z$ . Área implicada:  $0,5-0,288=0,212$ . Respuesta: 21,2%.



### Pasar de proporciones a puntuaciones con Calc

La función es de nombre algo enrevesado, pero claro: `=distr.norm.estand.inv(p)`. Devuelve la puntuación tipificada que se corresponde con el valor de probabilidad acumulada  $p$ . Por ejemplo, la probabilidad acumulada 0,5 se corresponde con la puntuación tipo  $Z=0$ . Según lo que se desee calcular, lo que hay que hacer es:

Se quiere calcular la Z que se corresponde con	Lo que hay que hacer es:
el área acumulada $p$	<code>=distr.norm.estand.inv(p)</code>
el área centrada $p$	<code>=distr.norm.estand.inv(p/2+0,5)</code>
el área de $p/2$ a cada extremo	<code>=distr.norm.estand.inv(1-p/2)</code>
el área $p$ del extremo inferior	<code>=distr.norm.estand.inv(p)</code>
el área $p$ del extremo superior	<code>=distr.norm.estand.inv(1-p)</code>

Para comprender el contenido de la tabla anterior hay que pensar en términos gráficos, manteniendo en mente la imagen de la curva normal. Pensemos, por ejemplo, en cómo encontrar el argumento  $p$  de la expresión `=distr.norm.estand.inv(p)`, cuando lo que queremos calcular es el valor de  $Z$  que acota un área centrada y no un área acumulada. Hay varios razonamientos que llevan al mismo punto. Así, si  $p$  es un área centrada que tiene como límite superior a  $+Z$ , entonces esa  $+Z$  es la misma que:

- Deja por debajo de sí a  $p$  más el extremo inferior. Lo que  $p$  deja sin tocar es  $1-p$ , repartido por igual en ambos extremos. Luego, el inferior vale  $(1-p)/2$ . Al sumarlo a  $p$ , entonces:  $p+(1-p)/2 = p/2+0,5$ .
- Deja por debajo de sí a toda la curva menos el extremo superior. Como ese extremo vale  $(1-p)/2$ , entonces el área acumulada es  $1-(1-p)/2=p/2+0,5$ .
- Deja por debajo de sí a la mitad del área centrada más toda la mitad de la curva, es decir, a  $p/2+0,5$ .

### Pasar de proporciones a puntuaciones con la tabla

La primera tabla del anexo traduce directamente de puntuaciones tipo a probabilidades o áreas hasta la media. La segunda tabla lleva de áreas acumuladas a puntuaciones tipo. La tercera, de áreas centradas a puntuaciones tipo. Así que tenemos cubiertas muchas situaciones. Realmente bastaría con cualquiera de las tres tablas, siendo la primera la que suministra información más precisa (llega a tres dígitos decimales para las áreas o probabilidades) y llega a ser la más habitual. Utilizando esta tabla, necesitaremos realizar varias operaciones según sea la situación que haya que resolver. Precisamente para abreviar por el camino y afrontar situaciones muy habituales, cuentas con las otras dos tablas, menos extendidas pero muy útiles.

Veamos un ejemplo.

Las puntuaciones de un test de memoria se distribuyen según una ley normal de media 100 y desviación tipo 20. Queremos saber qué puntuaciones obtiene el 90% de la población.

Para responder, suponemos que nos referimos al 90% *central*, es decir, qué puntuaciones en memoria tiene el 90% menos alejado de la media o, en otros términos,

más representativo (el 10% restante aglutina los valores de memoria más raros, sean por exceso o por defecto).

Para resolverlo, lo más fácil es acudir a la tercera tabla, que ya maneja áreas centradas. El 90% se corresponde con la fila 0,90 y la columna ,00. El valor de la casilla es  $Z=1,645$ . Luego, el 90% de la población suministra valores de memoria comprendidos entre  $-1,645$  y  $+1,645$  en número de desviaciones tipo que se alejan de la media. Hay que traducirlo a puntuaciones directas para terminar convenientemente la tarea:

$$Z_i = \frac{X_i - \bar{X}}{S} \Rightarrow X_i = \bar{X} \pm Z_i S = 100 \pm 1,645 \cdot 20 = 100 \pm 32,9 = \{ 67,1 ; 132,9 \}$$

Luego, el 90% de la población cuenta con puntuaciones en memoria comprendidas entre 67,1 y 132,9.

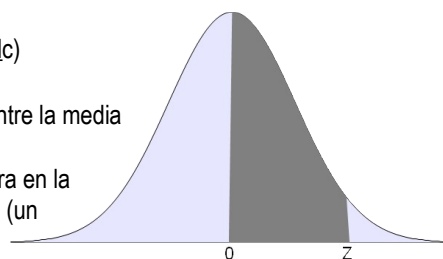
Si se consulta la segunda tabla, que suministra áreas acumuladas, hay que considerar que un 90% central está acotado por la misma  $Z$  que acota el 95% acumulado (el 90% central más el 5% inferior). Piensa que el 90% central deja un 10% en los extremos, que se reparten equitativamente en un 5% inferior y un 5% superior. El 95% acumulado se corresponde con la fila 0,90 y la columna ,05, cuya celda apuntada es  $Z=1,645$ . El resto del proceso es idéntico al anterior hasta llegar a 67,1 y 132,9.

También podemos partir de la primera tabla. En esta, las puntuaciones tipo no se encuentran en las celdas sino en las cabeceras de filas y columnas. Las celdas de esta tabla, como hemos visto, expresan áreas desde la media hasta  $Z$ . Pues bien, un área centrada de valor 90% es lo mismo que dos áreas del 45% hasta la media. Luego, hay que buscar la celda que muestre el valor 0,450. No existe, pero sí vemos dos celdas contiguas con valores 0,449 y 0,451, que se corresponden respectivamente con las respectivas  $Z$  de valores 1,64 y 1,65, así que podemos tomar el valor intermedio:  $Z=1,645$ . El resto del proceso hasta llegar 67,1 y 132,9 es idéntico a los dos procedimientos previos.

## Anexos: tabla de la distribución normal estandarizada

*Tabla de la distribución normal estandarizada respecto a la media*

- Cabecera de filas: unidades y décimas de la puntuación tipo ( $Z = u,dc$ )
- Cabecera de columnas: centésimas de la puntuación tipo ( $Z = u,dc$ )
- Celdas: proporción de área bajo la curva normal que se encuentra entre la media y la puntuación tipo señalada por la fila y la columna.
- Ejemplo: la proporción entre la media y  $Z=1,53$  es la que se encuentra en la celda de la fila 1,50 y la columna ,03 ( $1,50+.03=1,53$ ), es decir 0,437 (un 43,7% del área)



	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,00	,000	,004	,008	,012	,016	,020	,024	,028	,032	,036
0,10	,040	,044	,048	,052	,056	,060	,064	,067	,071	,075
0,20	,079	,083	,087	,091	,095	,099	,103	,106	,110	,114
0,30	,118	,122	,126	,129	,133	,137	,141	,144	,148	,152
0,40	,155	,159	,163	,166	,170	,174	,177	,181	,184	,188
0,50	,191	,195	,198	,202	,205	,209	,212	,216	,219	,222
0,60	,226	,229	,232	,236	,239	,242	,245	,249	,252	,255
0,70	,258	,261	,264	,267	,270	,273	,276	,279	,282	,285
0,80	,288	,291	,294	,297	,300	,302	,305	,308	,311	,313
0,90	,316	,319	,321	,324	,326	,329	,331	,334	,336	,339
1,00	,341	,344	,346	,348	,351	,353	,355	,358	,360	,362
1,10	,364	,367	,369	,371	,373	,375	,377	,379	,381	,383
1,20	,385	,387	,389	,391	,393	,394	,396	,398	,400	,401
1,30	,403	,405	,407	,408	,410	,411	,413	,415	,416	,418
1,40	,419	,421	,422	,424	,425	,426	,428	,429	,431	,432
1,50	,433	,434	,436	,437	,438	,439	,441	,442	,443	,444
1,60	,445	,446	,447	,448	,449	,451	,452	,453	,454	,454
1,70	,455	,456	,457	,458	,459	,460	,461	,462	,462	,463
1,80	,464	,465	,466	,466	,467	,468	,469	,469	,470	,471
1,90	,471	,472	,473	,473	,474	,474	,475	,476	,476	,477
2,00	,477	,478	,478	,479	,479	,480	,480	,481	,481	,482
2,10	,482	,483	,483	,483	,484	,484	,485	,485	,485	,486
2,20	,486	,486	,487	,487	,487	,488	,488	,488	,489	,489
2,30	,489	,490	,490	,490	,490	,491	,491	,491	,491	,492
2,40	,492	,492	,492	,492	,493	,493	,493	,493	,493	,494
2,50	,494	,494	,494	,494	,494	,495	,495	,495	,495	,495
2,60	,495	,495	,496	,496	,496	,496	,496	,496	,496	,496
2,70	,497	,497	,497	,497	,497	,497	,497	,497	,497	,497
2,80	,497	,498	,498	,498	,498	,498	,498	,498	,498	,498
2,90	,498	,498	,498	,498	,498	,498	,498	,499	,499	,499
3,00	,499	,499	,499	,499	,499	,499	,499	,499	,499	,499
3,10	,499	,499	,499	,499	,499	,499	,499	,499	,499	,499
3,20	,499	,499	,499	,499	,499	,499	,499	,499	,499	,499
3,30	,500	,500	,500	,500	,500	,500	,500	,500	,500	,500

*Tablas de puntuaciones tipo correspondientes a una curva normal*

La suma de las cabeceras de filas y columnas se corresponde con el área, proporción o probabilidad.

Puntuaciones estandarizadas para probabilidades acumuladas										
	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,00		-2,326	-2,054	-1,881	-1,751	-1,645	-1,555	-1,476	-1,405	-1,341
0,10	-1,282	-1,227	-1,175	-1,126	-1,080	-1,036	-,994	-,954	-,915	-,878
0,20	-,842	-,806	-,772	-,739	-,706	-,674	-,643	-,613	-,583	-,553
0,30	-,524	-,496	-,468	-,440	-,412	-,385	-,358	-,332	-,305	-,279
0,40	-,253	-,228	-,202	-,176	-,151	-,126	-,100	-,075	-,050	-,025
0,50	,000	,025	,050	,075	,100	,126	,151	,176	,202	,228
0,60	,253	,279	,305	,332	,358	,385	,412	,440	,468	,496
0,70	,524	,553	,583	,613	,643	,674	,706	,739	,772	,806
0,80	,842	,878	,915	,954	,994	1,036	1,080	1,126	1,175	1,227
0,90	1,282	1,341	1,405	1,476	1,555	1,645	1,751	1,881	2,054	2,326

Puntuaciones estandarizadas para probabilidades centradas (int. de confianza)										
	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,00		,013	,025	,038	,050	,063	,075	,088	,100	,113
0,10	,126	,138	,151	,164	,176	,189	,202	,215	,228	,240
0,20	,253	,266	,279	,292	,305	,319	,332	,345	,358	,372
0,30	,385	,399	,412	,426	,440	,454	,468	,482	,496	,510
0,40	,524	,539	,553	,568	,583	,598	,613	,628	,643	,659
0,50	,674	,690	,706	,722	,739	,755	,772	,789	,806	,824
0,60	,842	,860	,878	,896	,915	,935	,954	,974	,994	1,015
0,70	1,036	1,058	1,080	1,103	1,126	1,150	1,175	1,200	1,227	1,254
0,80	1,282	1,311	1,341	1,372	1,405	1,440	1,476	1,514	1,555	1,598
0,90	1,645	1,695	1,751	1,812	1,881	1,960	2,054	2,170	2,326	2,576