Tema 6. Variables aleatorias continuas. Distribución Normal

Indice

1. Distribuciones de probabilidad continuas.

Una ***variable aleatoria*** se dice ***continua*** cuando puede tomar cualquier valor dentro de un intervalo del conjunto de números reales.

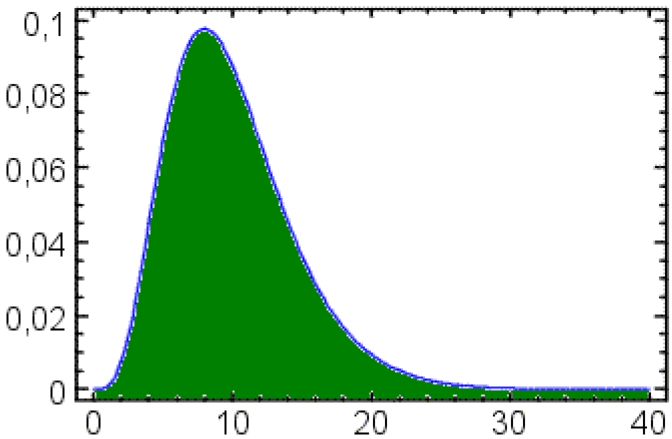
En las distribuciones continuas: la probabilidad de un valor concreto es prácticamente cero, por lo que se estudia las probabilidades asociadas a intervalos: p(a≤X≤b).

Cuando trabajábamos con distribuciones estadísticas continuas agrupábamos los datos en intervalos creados artificialmente, en cambio para las distribuciones de probabilidad continuas se trabaja con la función densidad de probabilidad.

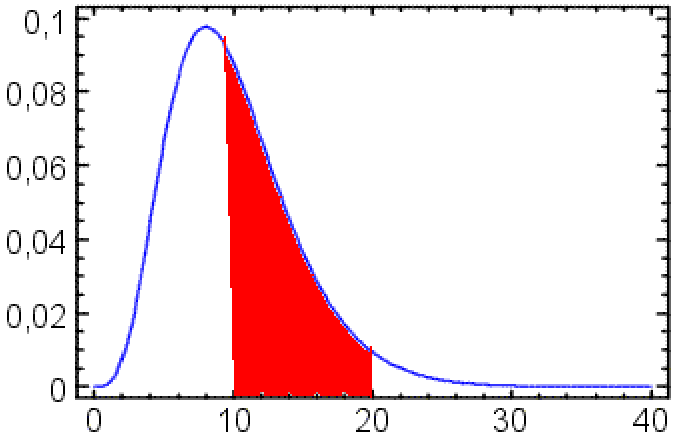
La función densidad de probabilidad, f(x), de una variable continua X, es una función definida positiva (por encima del eje OX) y tal que el área total encerrada entre la curva y el eje OX es uno (la probabilidad de que ocurra cualquier suceso es uno).

La probabilidad de que ocurra el suceso X en el intervalo [a,b], es decir p(a≤X≤b) es igual al área encerrada entre esta curva y el eje OX en este intervalos. Matemáticamente el área se calcula con la integral, p(a≤X≤b)=$\int_a^b f(x)dx$ (que no veremos pues no corresponde al nivel de esta asignatura).

Veamos gráficamente un ejemplo:



P(X tome cualquier valor)=área total de la curva=1



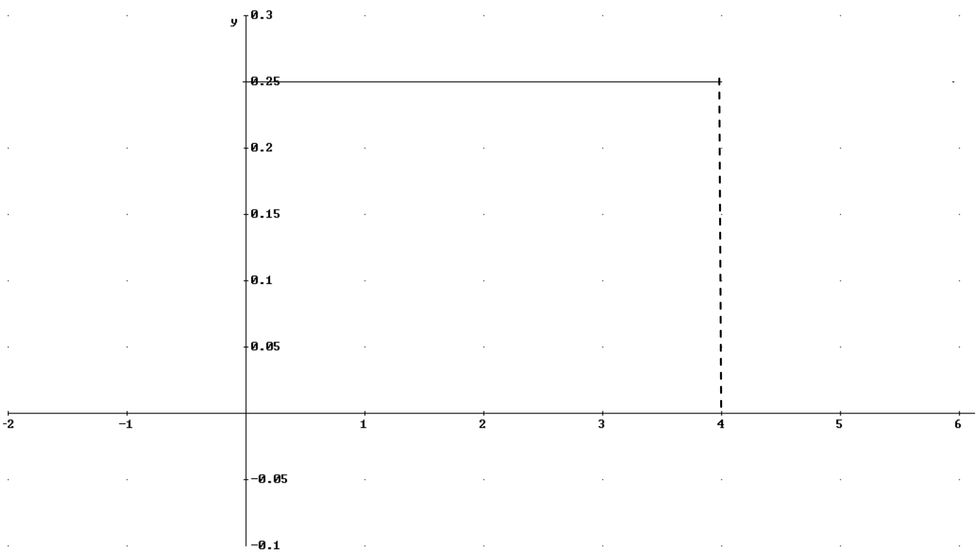
P(10≤x≤20)=área sombreada<1.

Ejercicios resueltos

Problema 1: Estudiar ayudándote de la representación gráfica, si las siguientes funciones son funciones de densidad de ciertas variables continuas. En caso afirmativo calcular P(X≤3), p(X≥1), p(1≤X≤2)

a) $f(x)=\begin{cases} \frac{1}{4} \; si \; 0<x<4 \\ 0 \; \; para \; el \; resto \; de \; x \end{cases}$ b) $g(x)=\begin{cases} \frac{1}{8}x-\frac{1}{4} \; si \; 2<x<6 \\ 0 \; \; para \; el \; resto \; de \; x \end{cases}$

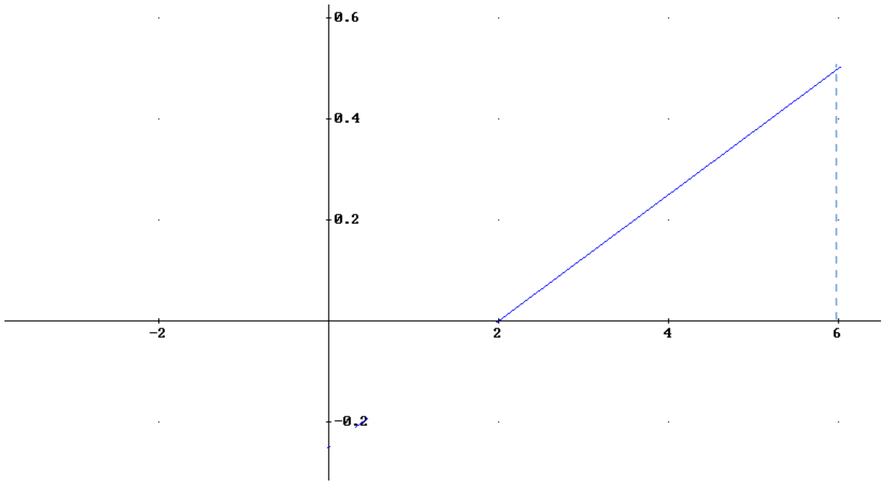
a)



Si es función densidad el área total encerrada ha de ser 1. Veámoslo:

A=$A_{cuadrado}$ =base·altura=4·1/4=1. Luego si es una función densidad de probabilidad.

Probabilidades: P(X≤3)=3·1/4=3/4; p(X≥1)=3·1/4=3/4; p(1≤X≤2)=1·1/4=1/4 b)



Si es función densidad el área total encerrada ha de ser 1. Veámoslo:

A=A$_{triangulo}$ =$\frac{b\cdot h}{2}=\frac{4\cdot 0{,}5}{2}=1$

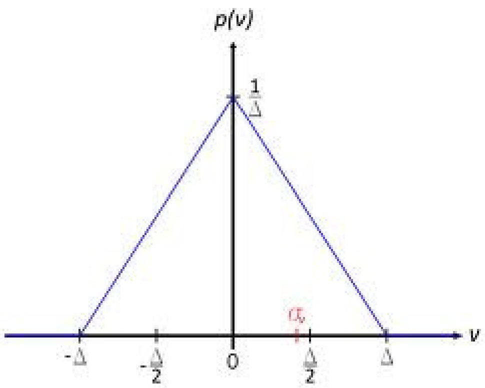
Probabilidades:

p(X≤3)= $\frac{b\cdot h}{2}=\frac{1\cdot 1/8}{2}=\frac{1}{16}$; donde h=f(3)=3/8-1/4=1/8

p(X≥1)= $\frac{b\cdot h}{2}=1$

p(1≤X≤2)=0

Problema 2: Sea la gráfica siguiente una función. Comprobar que independientemente del valor de Δ es una función de densidad. Calcular p(X≥0)



Área=$\frac{b\cdot h}{2}=\frac{2\Delta\cdot\frac{1}{\Delta}}{2}=1$

p(X≥0)=1/2

Ejercicios propuestos:

Problema 3: Calcular el valor de a para que la función $f(x)=\begin{cases} a \;\; si -2<x<4 \\ 0 \;\; para\ el\ resto\ de\ x \end{cases}$ sea una función de densidad para la variable aleatoria continua X. Calcular para ese valor de a las probabilidades p(x≤2), p(x≥0) y p(-1≤x≤1).

Problema 4: comprobar que la función $f(x)=\begin{cases} 2x \quad si\ 0<x<1 \\ 0 \;\; para\ el\ resto\ de\ x \end{cases}$ es una función densidad de probabilidad para la variable continua X y calcular p(x≤2), p(x≥0) y p(-1≤x≤1)

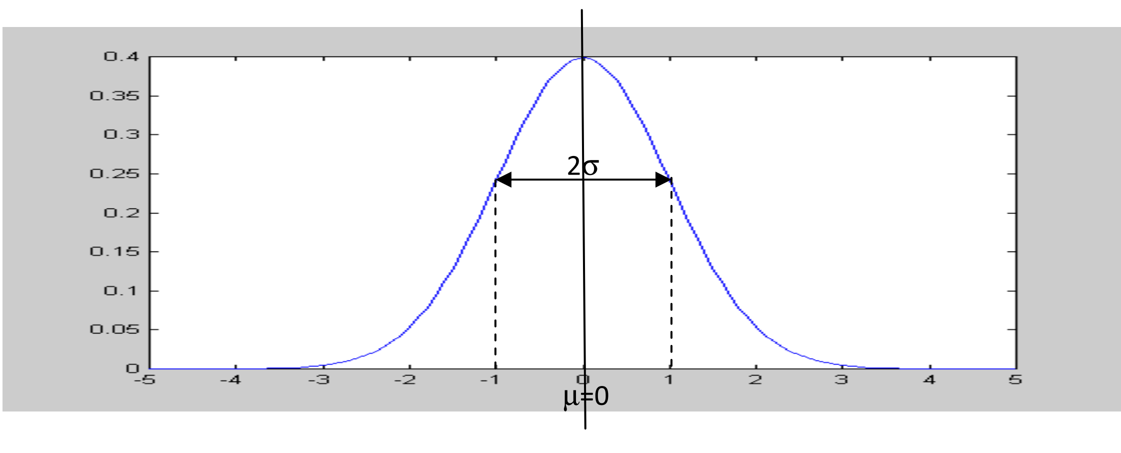
2. Distribución Normal

El calificativo normal de esta distribución se debe a que es típica de muchos experimentos y observaciones, especialmente en fenómenos de la naturaleza.

2.1. Distribución Normal estándar N(0,1).

Como veremos más adelante cualquier distribución normal con media μ y desviación típica σ puede asociarse a una distribución normal con media 0 y desviación típica 1, mucho más sencilla de trabajar. Esta distribución normal con media 0 y desviación típica 1 se llama ***distribución normal estándar*** y se denota como ***N(0,1).***

La variable aleatoria normal que se rige por el comportamiento N(0,1) se designa como Z y se llama ***variable normal estándar*** o ***tipificada***. La función distribución relacionada con esta distribución es la ***Gaussina*** (en honor a Gauss) centrada y de anchura unidad: $\phi(z)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{z^2}{2}}$ cuya media es 0 y desviación típica 1:



La ventaja de la distribución tipificada frente a las demás distribuciones normales es que las probabilidades P(Z≤a) se encuentran baremados en la tabla de la distribución normal:

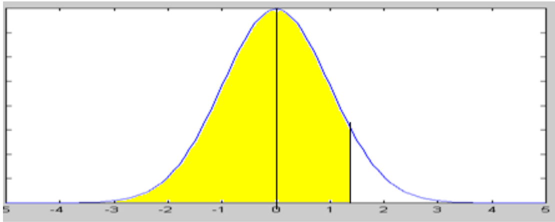
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.7985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

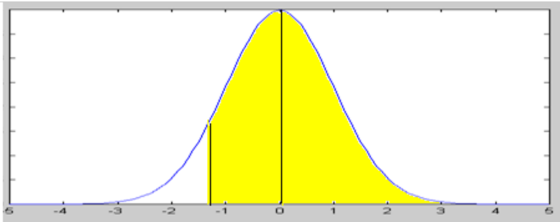
Ejemplo: P(z≤1,32)= 0.9066 (fila 1.3, columna 0,002)

2.1.1 Utilización de la tabla de la distribución N(0,1)

En este apartado calcularemos todos los posibles casos de probabilidades que se puedan presentar a la hora de calcular probabilidades de una distribución N(0,1). Los valores numéricos $P(Z\leq a)$ se pueden encontrar en la tabla anterior. Nota: para entender los resultados hay que tener dos consideraciones:

1. $P(-\infty\leq Z\leq\infty)$=área total=1
2. La función es simétrica: $P(Z\leq a)=P(Z\geq -a)$



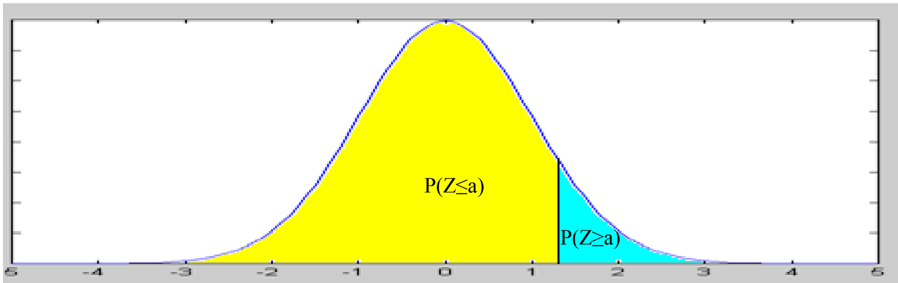


A partir de estas dos propiedades es fácil ver los valores de las siguientes probabilidades:

1. $P(Z\geq a)$

Es fácil a partir del dibujo y de la primera propiedad que la propiedades $P(Z\geq a)$ se puede obtener a partir de $P(Z\leq a)$ como:

$$P(Z\geq a)=1-P(Z\leq a)$$

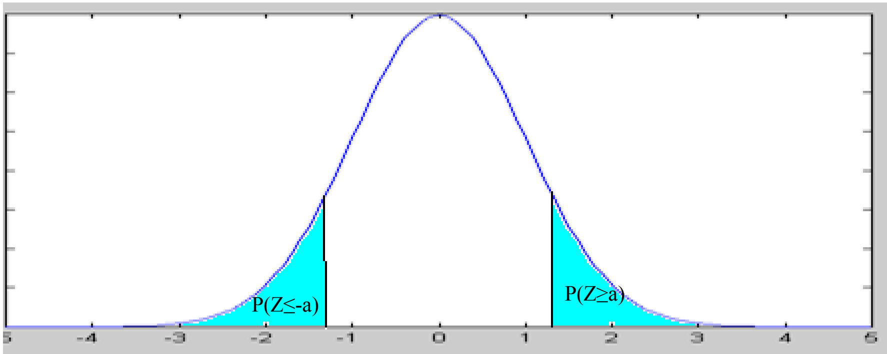


Ejemplo: $P(Z\geq 1.34)=1-P(Z\leq 1.34)=1-0.9099=0.0901$

2. **P(Z≤ -a)**

Es fácil a partir del dibujo y de la primera propiedad que la propiedades P(Z≥a) se puede obtener a partir de P(Z≤a) como:

$$P(Z\leq -a)=p(Z\geq a)=1-p(Z\leq a)$$

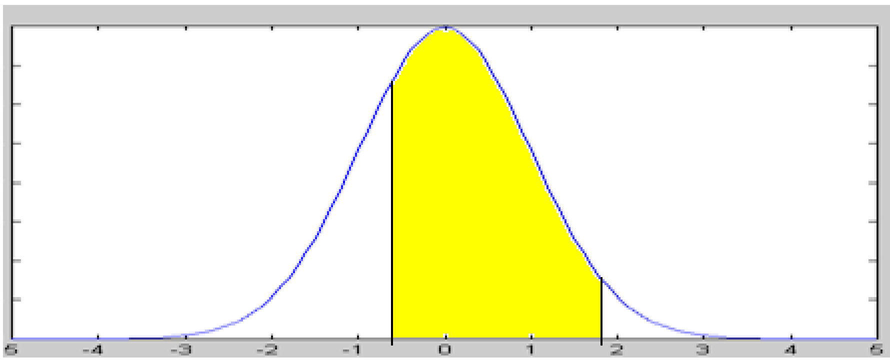


Ejemplo: $P(Z\leq -0.31)=P(Z\geq 0.31)=1-P(z\leq 0.31)=1-0.6217=0.3783$

3. **P(a≤Z≤b)**

El área comprendida entre a y b, será el área desde -∞ hasta b menos el área desde b a -∞. Es decir:

$$P(a\leq Z\leq b)=P(Z\leq b)-P(Z\leq a)$$



Ejemplo: $P(0.47\leq Z\leq 2.13)=P(Z\leq 2.13)-P(Z\leq 0.47)=0.9834-0.6808=0.3026$

Ejercicios resueltos:

1. En la distribución normal N(0,1) calcular el valor de k en los siguientes casos:
   a) $P(Z\leq k)=0.7673$:
   Basta con ver el valor en la tabla cuya probabilidad sea la más próxima a 0.7673, k será el valor que nos marque la fila y columna con dicha probabilidad: k=0.73
   b) $P(Z\leq k)=0.9761$
   De manera análoga al caso anterior k=1.98
   c) $P(Z\geq k)=0.1075$
   En la tabla solo tenemos las probabilidades para valores de Z menores o iguales que k, por lo que tendremos que aplicar la relación (1):
   $P(Z\geq k)=1-P(z\leq k)=0.1075 \rightarrow P(Z\leq k)=1-0.1075=0.8925 \rightarrow k=1.24$
   d) $P(-k\leq Z\leq k)=0.516$.
   Aplicando las propiedades (3) y (2):
   $P(-k\leq Z\leq k)=P(Z\leq k)-P(Z\leq -k)=P(Z\leq k)-(1-P(Z\leq k))=2\cdot P(Z\leq k)-1=0.516$
   $\rightarrow P(Z\leq k)=\frac{(1+0.516)}{2}=0.758 \rightarrow k=0,70$
2. Sea Z una variable aleatoria que sigue una distribución normal N(0,1). Hallar las siguientes probabilidades:
   a. $P(Z\geq 0.32)$ **b.** $P(Z\leq 0)$ **c.** $P(Z\geq -2.3)$ **d.** $P(-0.51\leq Z\leq 1.1)$

   Soluciones:

   a. $P(Z\geq 0.32)=1-P(Z\leq 0.32)=1-0.6255=0.3745$
   b. $P(Z\leq 0)=0.5$
   c. $P(Z\geq -2.3)= P(Z\leq 2.3)=0.9893$
   d. $P(-0.51\leq Z\leq 1.1)=P(Z\leq 1.1)-P(Z\leq -0.51)=0.8665-(1-P(Z\leq 0.51))=0.8665-(1-0.595)=0.4615$

Ejercicios propuestos:

1. Si Z es una variable aleatoria que sigue la distribución N(0,1) hallar el valor de a:

a) $P(Z\leq a)=0.7673$ **b)** $P(0\leq Z\leq a)=0.4115$ **c)** $P(Z\geq a)=0.994$ **d)** $P(Z\leq 2+a)=0.9974$

2. Calcular las siguientes probabilidades sabiendo que Z sigue distribución N(0,1)

a. $P(Z\geq 2.12)$ **b.** $P(Z\leq 0.45)$ **c.** $P(Z\leq -2.5)$ **d.** $P(-2.51\leq Z\leq 1.23)$

2.2. Tipificación de la variable

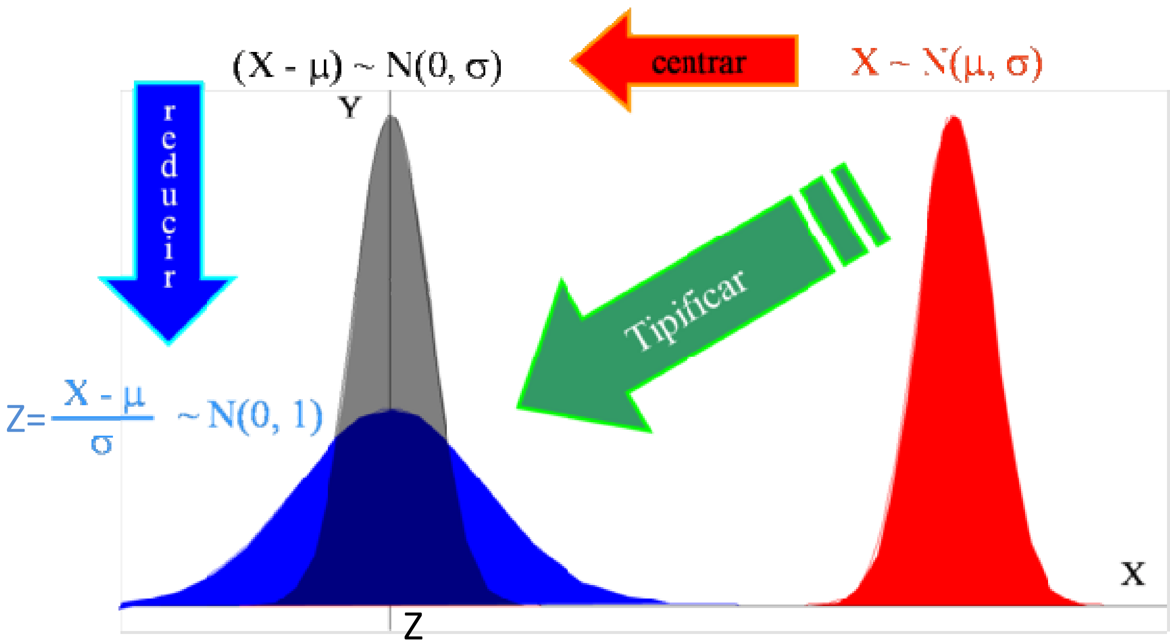
Hemos visto en el apartado anterior que la distribución normal N(0,1) se encuentra tabulada y resulta sencillo calcular las distintas probabilidades.

En la práctica las distribuciones normales que se utilizan en la vida real rara vez coincide con esta distribución normal cuya media $\mu=0$, y desviación típica es $\sigma=1$, N(0,1). Por tanto es necesario relacionar cualquier normal con media μ y desviación σ que pueda tomar cualquier valor. La transformación de la normal $N(\mu,\sigma)$ a la normal N(0,1) se denomina ***tipificación.***

Tipificación: es el paso de una variable X que sigue una distribución normal $N(\mu,\sigma)$ a otra variable Z que sigue la distribución N(0,1) y que sabemos calcular su probabilidad al estar tabuladas sus probabilidades. La transformación consiste en:

1. Traslado, es decir trasladar la media μ a cero ($\mu=0$)
2. Escalado, que consiste en hacer el ancho de la Gaussiana como si tuviera $\sigma=1$.

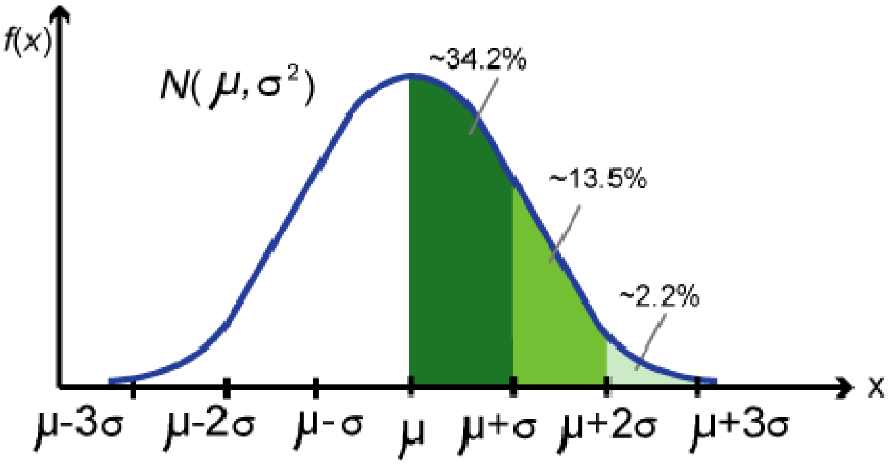
Las dos transformaciones anteriores se consiguen con el siguiente cambio de variable: —— . De esta forma podemos pasar de la variable X que sigue $N(\mu,\sigma)$ a la variable Z que sigue N(0,1).



2.3. Interpretación de la desviación típica, s, en la distribución Normal

En este apartado no queremos más que hacer ver el significado de σ, tal que se cumple:

1. $P(\mu-\sigma \leq X \leq \mu+\sigma)=0,684$
2. $P(\mu-2\sigma \leq X \leq \mu+2\sigma)=0,954$
3. $P(\mu-3\sigma \leq X \leq \mu+3\sigma)=0,998$



Ejercicios resueltos:

1. La temperatura T durante el mes de Agosto está distribuida normalmente con una media μ=36º y desviación típica σ=6º. Calcular la probabilidad de que la temperatura de un día de Agosto a) este entre 40º y 44º, b) sea menor que 32º.

Solución: T $\Rightarrow$ N(36,6), por tanto $Z=\frac{T-36}{6} \Rightarrow N(0,1)$

a. $P(40\leq T\leq 44)=P(\frac{40-36}{6} \leq Z \leq \frac{44-36}{6})=P(0.67\leq Z\leq 1.33)=P(Z\leq 1.33)-P(Z\leq 0.67)=0.16$

b. $P(T\leq 32)=P(Z \leq \frac{32-36}{6})=P(Z\leq -0.67)=1-P(Z\leq 0.67)=1-0.7486=0.2514$

2. Supongamos que las estaturas de los 800 estudiantes de bachillerato en España distribuidas con una media de 168 cm y desviación típica de 10cm. Hallar el número de alumnos con las siguientes estaturas: a) Entre 165 y 178; b) mayor o igual que 183cm.

Solución: E=estarura $\Rightarrow$ N(168,10), por tanto $Z=\frac{E-168}{10} \Rightarrow N(0,1)$

a. $P(165\leq E\leq 178)=P(\frac{165-168}{10} \leq Z \leq \frac{178-168}{10})=P(-0.3\leq Z\leq 1)=P(Z\leq 1)-P(Z\leq -0.3)=$

$=0.8413-(1-0.6179)=0.4594$

Luego el 45.94% de los alumnos en esta altura: 45.94% de 800=367.36≈367

b. $P(E\geq 183)=P(Z\geq \frac{183-168}{10})$ $=P(Z\geq 1.5)=1-P(Z\leq 1.5)=1-0.9332=0.0668$ →6,68%

Luego los alumnos con esta altura son el 6,68% de 800=53,44≈53 alumnos.

Ejercicios propuestos:

1. El diámetro de los tornillos producidos por una fábrica están distribuidos normalmente con una media de 0.58 cm y una desviación típica de 0,05cm. Se considera defectuoso un tornillos si su diámetro es menor que 0.46 o mayor de 0.65. Hallar el porcentaje de tornillos defectuosos de esta fábrica. Si produce 3000 tronilloos a la hora ¿Cuántos tornillos serán fabricados defectuosos en dos horas?

2. Las puntuaciones de un examen están distribuidas normalmente con una media de 7.6 y una desviación típica de 1.5. El 15% de los estudiantes, los mejores estudiantes, reciben sobresaliente y el 10%, los peores reciben insuficiente.

a) Hallar la puntuación mínima para tener un sobresaliente.

b) Calcular la puntuación mínima para no suspender.

3. Se sabe que la distribución de los cocientes intelectuales (C.I.), cociente entre la edad mental y la edad cronológica, de los reclutas de un reemplazo siguen la distribución normal de media 0.9 y desviación típica de 0,4. Para un batallón compuesto por 1000 reclutas, calcular el número de reclutas con a) C.I. menor que 1, b) menor que 0.2, c) superior a 1.4, d) entre 0.8 y 1.3.

4. Una confitura puede ser calificada de almíbar si contiene entre 420 y 520 gramos de azúcar por kilo. Un fabricante analiza 200 botes de confitura de 1kg y encuentra que el valor medio de azúcar es de 465g con una desviación típica de 30g. Sabiendo que el contenido de azúcar se distribuye normalmente calcular el porcentaje de botes que no deberían llevar la mención de almíbar.