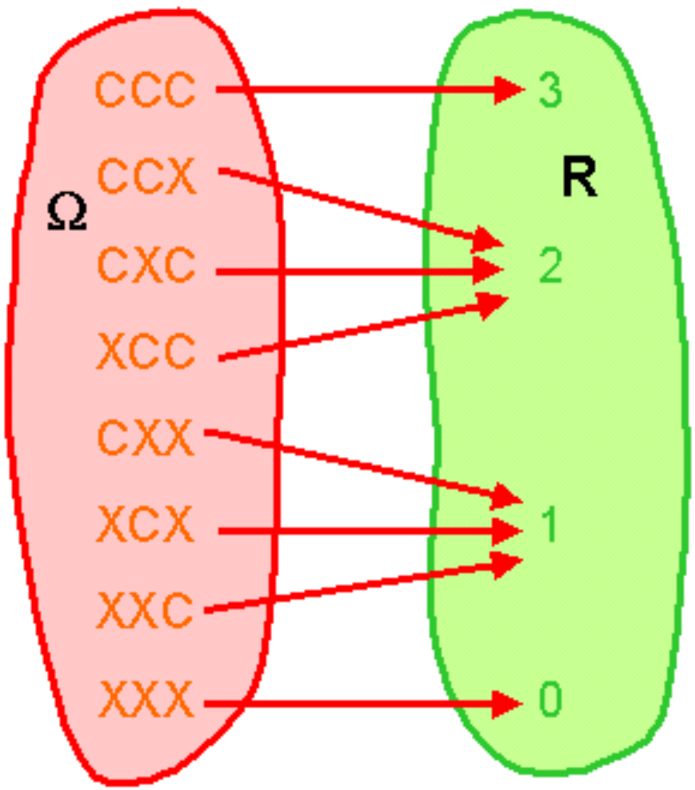
Variables aleatorias

Fenómeno aleatorio: lanzar tres monedas al aire



- La función X = "número de caras" asocia a cada elemento del espacio muestral un número real: es una variable aleatoria.
- Esta variable aleatoria X puede tomar los valores $\{0, 1, 2, 3\}$

- La variable X es discreta porque sólo puede tomar los valores del conjunto $\{0, 1, 2, 3\}$, que son valores discretos.
- Cuando una variable puede tomar, al menos teóricamente, todos los valores en cierto intervalo de la recta real, se dice que es continua.

Distribución de una variable aleatoria

La distribución de probabilidad de una variable aleatoria discreta X asocia a cada uno de los valores que puede tomar X su probabilidad correspondiente.

Si las monedas son perfectas (en el experimento anterior) cada uno de los sucesos elementales del espacio muestral tiene probabilidad 1/8. Por ello la distribución de probabilidad de la variable aleatoria X = "número de caras al tirar 3 monedas" será:

x_i	$P(X = x_i)$
0	1/8
1	3/8
2	3/8
3	1/8

Algunas probabilidades con esta variable:

a) $p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) = 1$
b) $p(1 < X \leq 3) = p(X = 2) + p(X = 3) = 3/8 + 1/8 = 1/2$
c) $p(X > 2) = p(X = 3) = 1/8$
d) $p(X > 1{,}5) = p(X = 2) + p(X = 3) = 3/8 + 1/8 = 1/2$

Parámetros de la distribución de una distribución discreta

μ: media o esperanza. Proporciona cuál es el valor promedio de la distribución.
σ: desviación típica. Dispersión de los valores de X respecto a ese valor promedio

Al realizar un experimento n veces, sean $x_1, x_2, x_3, \ldots, x_k$ los distintos valores de X y sean $F_n(x_i)$, $f_n(x_i)$ las frecuencias absolutas y relativas de x_i. Si n es grande las frecuencias relativas se estabilizan en torno a las probabilidades $p(x_i)$ y se tiene:

- $\overline{x}_n = \frac{1}{n}\sum_{i=1}^{k} x_i F_n(x_i) = \sum_{i=1}^{k} x_i f_n(x_i)$ $\longrightarrow$ $\mu = \sum_{i=1}^{k} x_i\, p(x_i)$
- $S_n = \sqrt{\sum_{i=1}^{k}(x_i - \overline{x})^2 f_n(x_i)}$ $\longrightarrow$ $\sigma = \sqrt{\sum_{i=1}^{k}(x_i - \mu)^2\, p(x_i)} = \sqrt{\sum_{i=1}^{k} x_i^2\, p(x_i) - \mu^2}$

Distribución binomial (I)

Fenómeno aleatorio: lanzar un dado

Éxito: A = "obtener un 6" → $p(A) = \frac{1}{6}$

Fracaso: A^c = "no obtener un 6" → $p(A^c) = \frac{5}{6}$

¿Cuál es la probabilidad del suceso B = "obtener 3 éxitos y 7 fracasos" en ese orden?

$$B = \boxed{A, A, A, A^c, A^c, A^c, A^c, A^c, A^c, A^c} \rightarrow P(B) = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7$$

Formas de obtener 3 éxitos: $= \binom{10}{3}$

Sea X = "número de éxitos en n = 10 pruebas".

$$p(\text{obtener 3 éxitos}) = p(X = 3) = \binom{10}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^7$$

Distribución binomial (II)

Sea p la probabilidad de un suceso A arbitrario. Se dice que X es una binomial de parámetros n y p, se representa **B(n, p)**, si X es el « número de veces que el suceso A se produce» en n pruebas repetidas e independientes del experimento.

La distribución de probabilidad de una variable X binomial de parámetros n y p es

$$p(X = x) = \binom{n}{x} . p^{x} . (1-p)^{n-x}, x = 0, 1, \ldots, n$$

Los parámetros n y p determinan completamente la distribución de la variable. Cada par de valores (n, p) proporciona una variable binomial diferente.

Si X es una binomial B(n, p), entonces $\mu = np$ y $\sigma^2 = np(1-p)$

Si X es una variable B(n, p), entonces para obtener probabilidades para $p > 0{,}5$ basta tener en cuenta que $p(X = x) = p(Y = n - x)$, siendo Y una binomial B(n, 1 - p)

Distribución binomial (III)

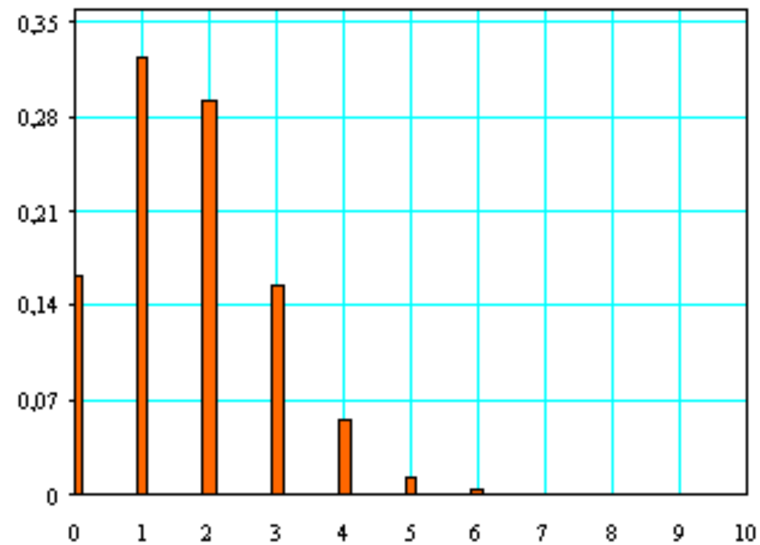
Función de probabilidad de X = "número de seises al tirar un dado 10 veces"

$$p(X = x) = \binom{10}{x} \cdot \left(\frac{1}{6}\right)^{x} \cdot \left(\frac{5}{6}\right)^{10 - x}, \; x = 0, 1, \ldots, 10$$

Tabla de valores de B(10, 1/6)

x	p(X = x)
0	0,161505583
1	0,323011166
2	0,290710049
3	0,155045360
4	0,054265876
5	0,013023810
6	0,002170635
7	0,000248073
8	0,000018605
9	0,000000827
10	0,000000017

Gráfica de la función de probabilidad



Distribuciones de variables continuas (I)

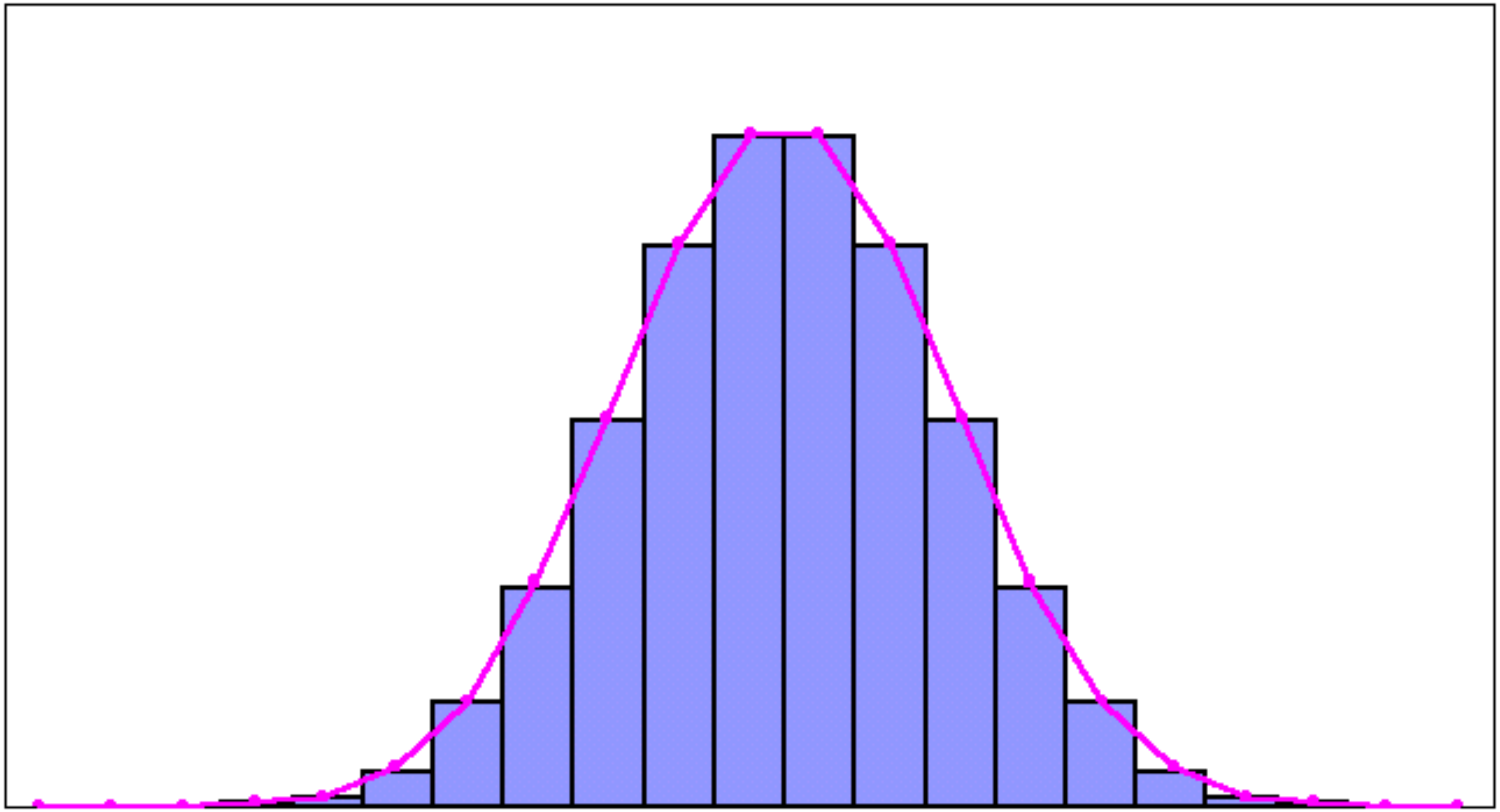
Para una variable aleatoria continua X, sobre una población muy grande, con una muestra de tamaño n construimos el histograma en escala de densidad. En estos histogramas:

- Los intervalos de clase son de igual amplitud.
- La superficie encerrada por cada rectángulo es igual a la frecuencia relativa correspondiente al intervalo.
- Por tanto la suma de las áreas de todos los rectángulos es 1.

Al hacer que n (tamaño de la muestra) crezca, y la amplitud de los intervalos disminuya observamos que la curva se aproxima a una curva f que se denomina función de densidad de la variable X. Observemos una simulación de dicho proceso:

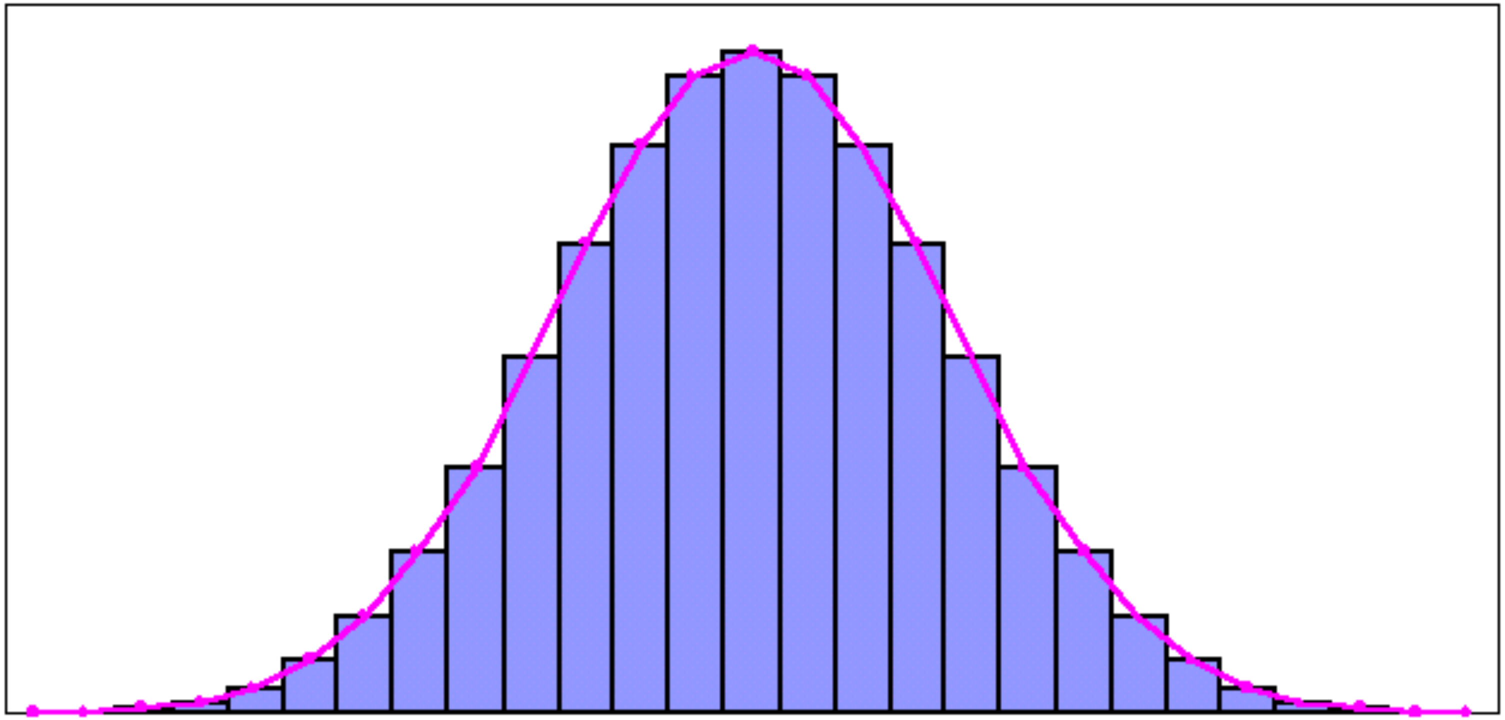
Distribuciones de variables continuas (II)

n = 100



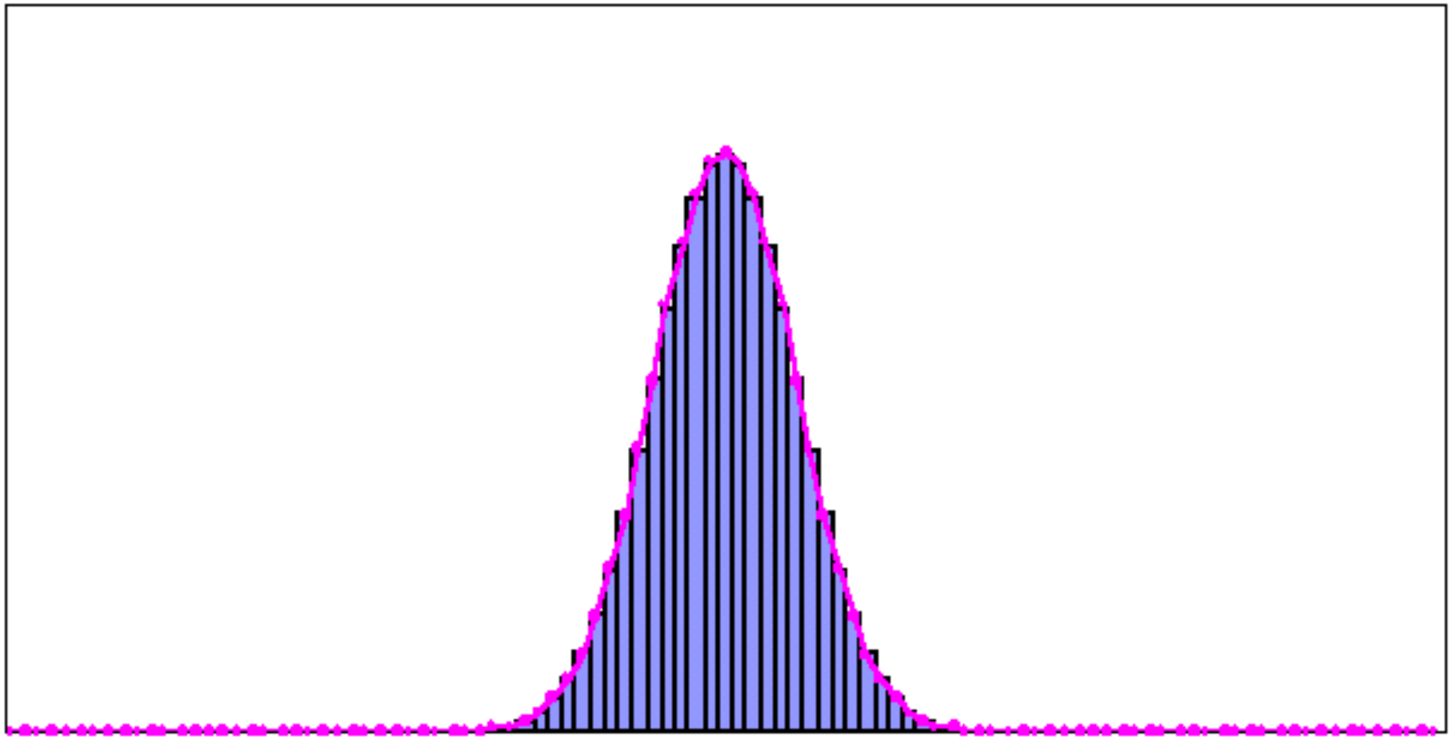
Distribuciones de variables continuas (III)

n = 1500

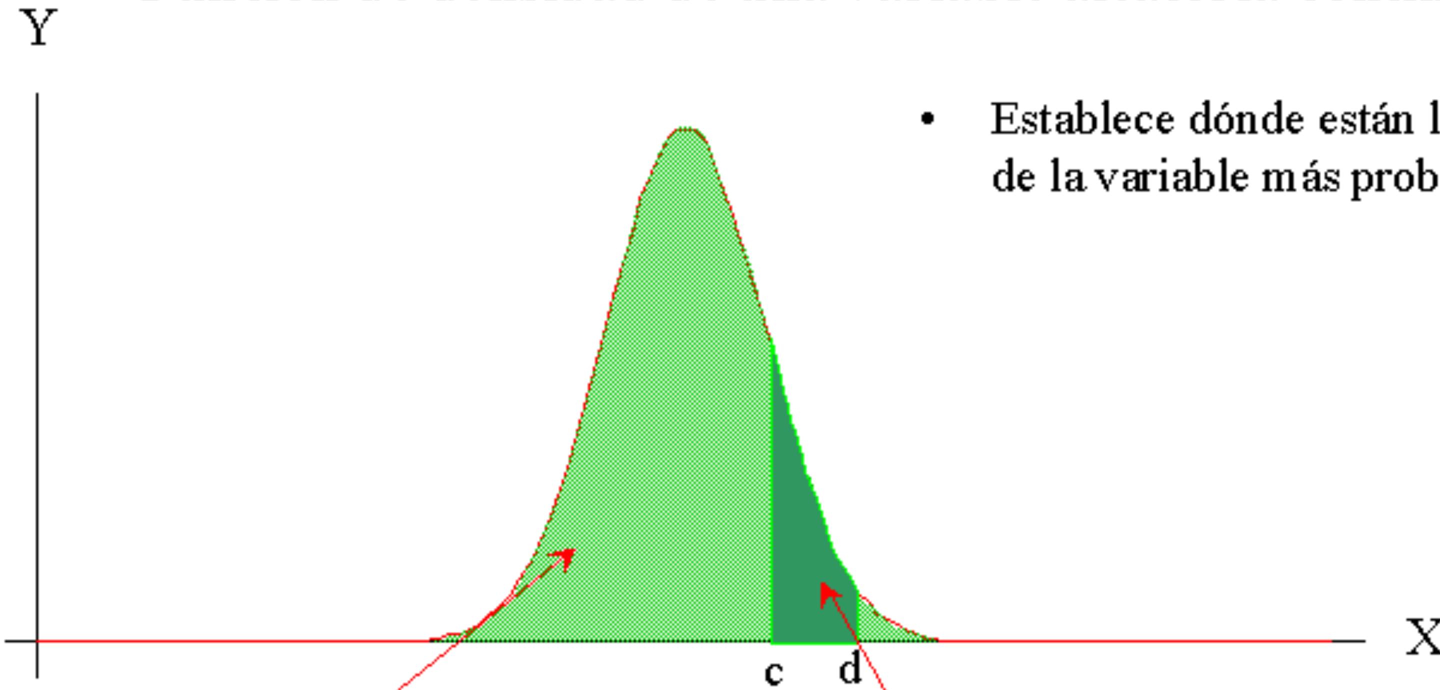


Distribuciones de variables continuas (IV)

n = 3000



Función de densidad de una variable aleatoria continua



- Establece dónde están los valores de la variable más probables.

- Está siempre por encima del eje OX
- El área encerrada bajo la función de densidad es 1

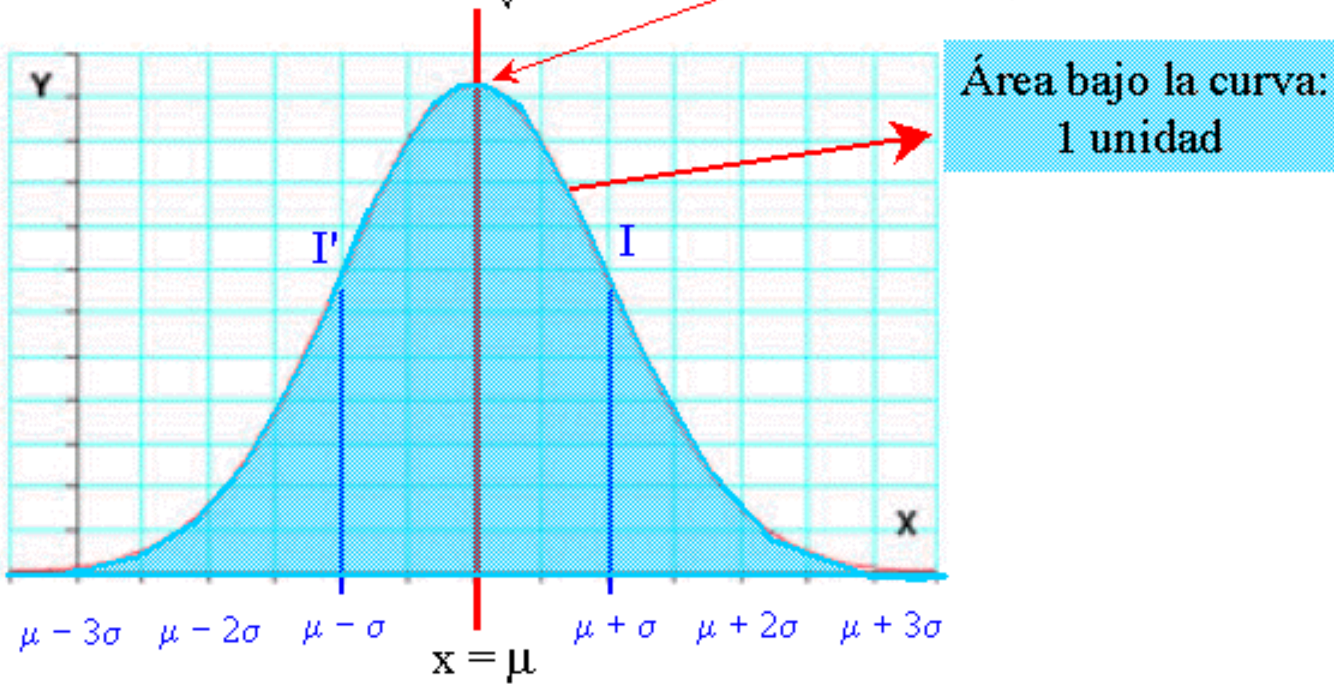
$$p(c < X < d) = p(c \leq X \leq d) = \int_c^d f(x)\,dx$$

Media y varianza de una variable aleatoria continua

Variable estadística	Variable aleatoria discreta	Variable aleatoria continua
$\overline{x} = \sum_{i=1}^{n} x_i f(x_i)$	$\mu = \sum_{i=1}^{n} x_i p(x_i)$	$\mu = \int_a^b x \,.\, f(x)\, dx$
$s^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 f(x_i)$	$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 p(x_i)$	$\sigma^2 = \int_a^b (x - \mu)^2 \,.\, f(x) \,.\, dx$

Función de densidad de la N(μ, σ)

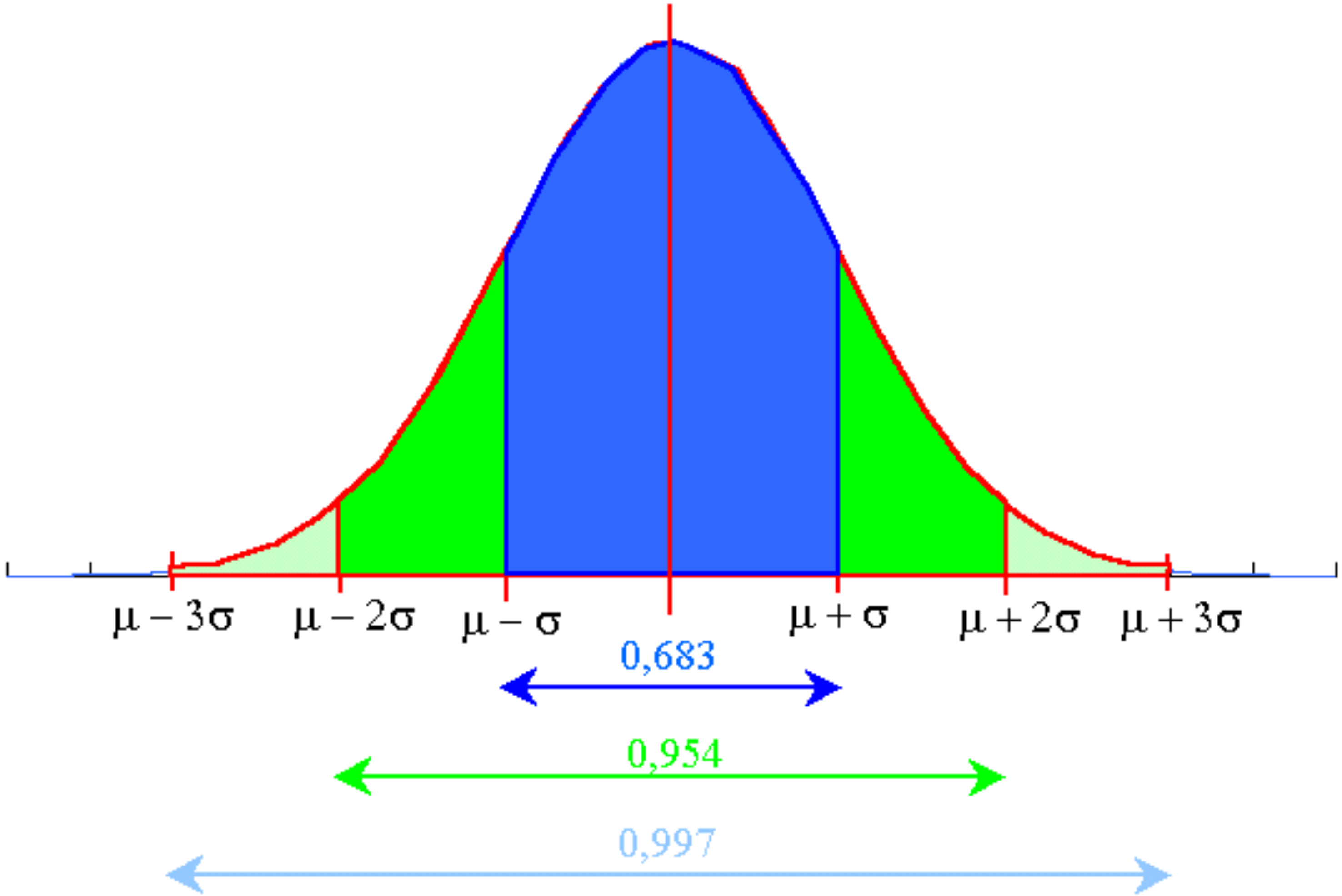
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



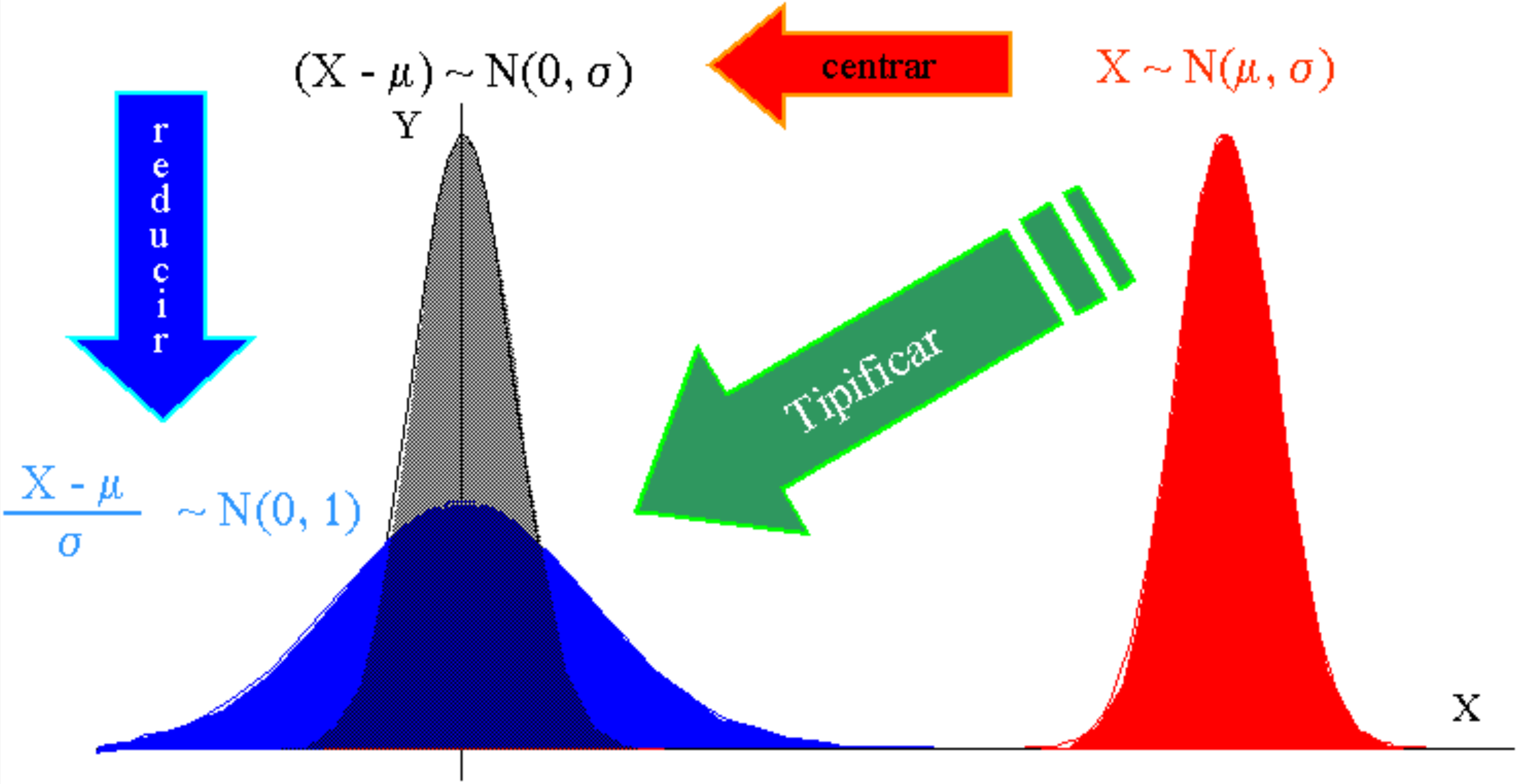
Campo de existencia = (- ∞, + ∞)

Creciente | Decreciente

Probabilidades de una variable N(μ, σ)



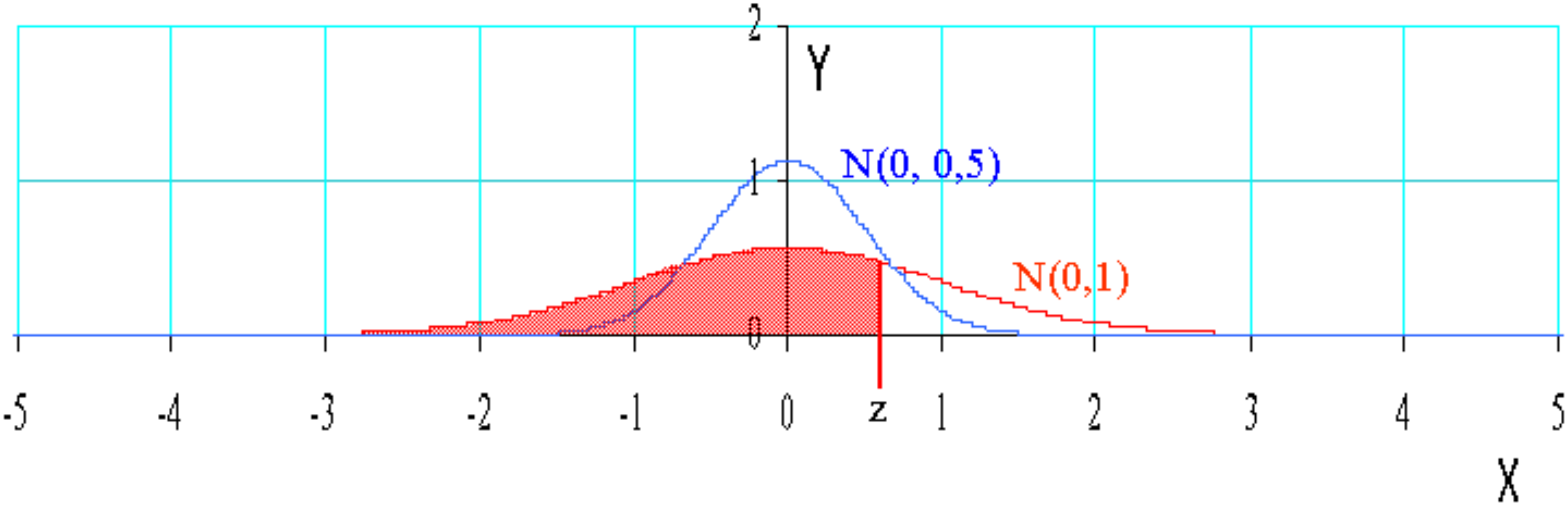
Tipificación de una variable $N(\mu, \sigma)$



Si X es una $N(\mu, \sigma)$ entonces $p\,(a \le X \le b) = p\,(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma})$ en donde Z es una $N(0, 1)$

Función de densidad de la N(0, 1)

Esta N(0, 1) se parece mucho a la verdadera (igual tamaño de unidades en ambos ejes)



Función de distribución de la N(0, 1)

- Tabulada: proporciona la superficie sombreada
- Esta superficie equivale a la probabilidad acumulada hasta z

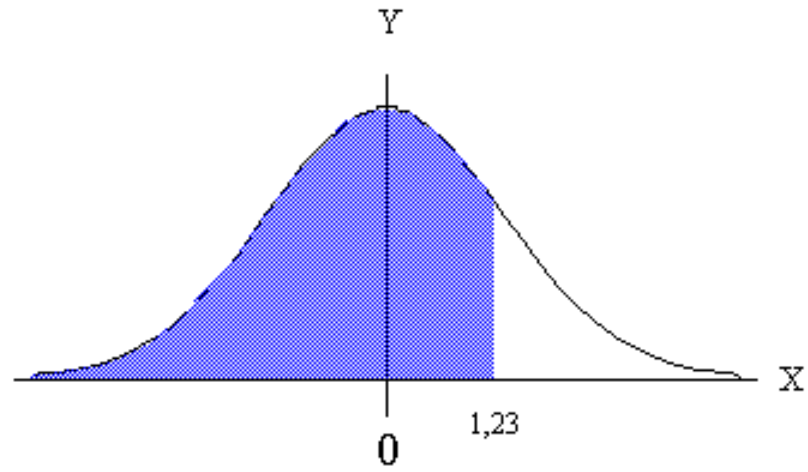
Tablas de la normal N(0, 1)

$$F(z) = p(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}t^2} dt$$

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

Cálculo de probabilidades con variables normales (I)

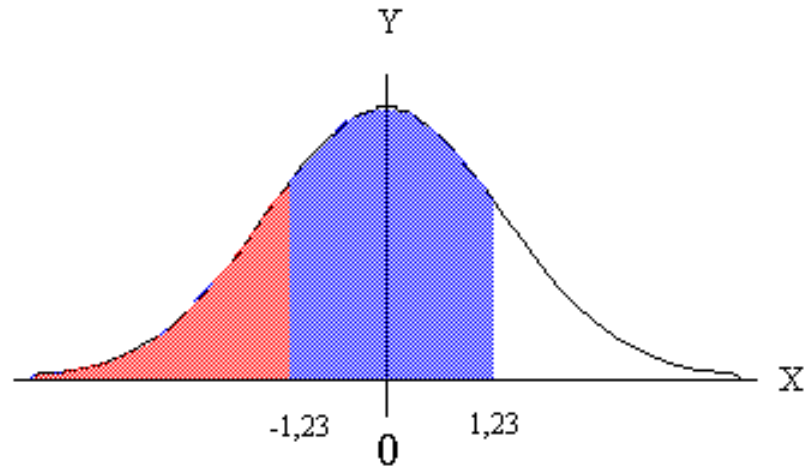
x	0,00	0,01	0,02	0,03
0,0	0,5000	0,5040	0,5080	[illegible]
0,1	0,5398	0,5438	0,5478	[illegible]
0,2	0,5793	0,5832	0,5871	[illegible]
0,3	0,6179	0,6217	0,6255	[illegible]
0,4	0,6554	0,6591	0,6628	[illegible]
0,5	0,6915	0,6950	0,6985	[illegible]
0,6	0,7257	0,7291	0,7324	[illegible]
0,7	0,7580	0,7611	0,7642	[illegible]
0,8	0,7881	0,7910	0,7939	[illegible]
0,9	0,8159	0,8186	0,8212	[illegible]
1,0	0,8413	0,8438	0,8461	[illegible]
1,1	0,8643	0,8665	0,8686	[illegible]
1,2	[illegible]	[illegible]	[illegible]	0,8907
1,3	0,9032	0,9049	0,9066	0,9082



$$p(Z \leq 1{,}23) = 0{,}8907$$

Cálculo de probabilidades con variables normales (II)

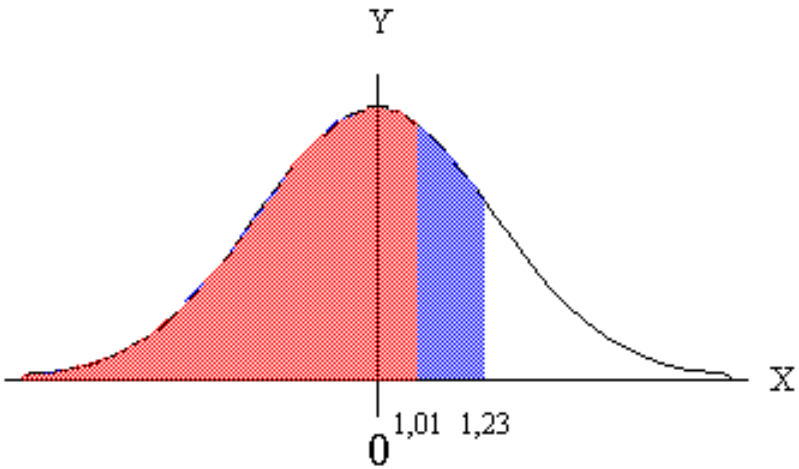
x	0,00	0,01	0,02	0,03
0,0	0,5000	0,5040	0,5080	0,5120
0,1	0,5398	0,5438	0,5478	0,5517
0,2	0,5793	0,5832	0,5871	0,5910
0,3	0,6179	0,6217	0,6255	0,6293
0,4	0,6554	0,6591	0,6628	0,6664
0,5	0,6915	0,6950	0,6985	0,7019
0,6	0,7257	0,7291	0,7324	0,7357
0,7	0,7580	0,7611	0,7642	0,7673
0,8	0,7881	0,7910	0,7939	0,7967
0,9	0,8159	0,8186	0,8212	0,8238
1,0	0,8413	0,8438	0,8461	0,8485
1,1	0,8643	0,8665	0,8686	0,8708
1,2	[illegible]	[illegible]	[illegible]	0,8907
1,3	0,9032	0,9049	0,9066	0,9082



$$p(Z \leq -1{,}23) = 1 - p(Z \leq 1{,}23) = 1 - 0{,}8907 = 0{,}1093$$

Cálculo de probabilidades con variables normales (III)

x	0,00	0,01	0,02	0,03
0,0	0,5000	0,5040	0,5080	0,5120
0,1	0,5398	0,5438	0,5478	0,5517
0,2	0,5793	0,5832	0,5871	0,5910
0,3	0,6179	0,6217	0,6255	0,6293
0,4	0,6554	0,6591	0,6628	0,6664
0,5	0,6915	0,6950	0,6985	0,7019
0,6	0,7257	0,7291	0,7324	0,7357
0,7	0,7580	0,7611	0,7642	0,7673
0,8	0,7881	0,7910	0,7939	0,7967
0,9	0,8159	0,8186	0,8212	0,8238
1,0	[illegible]	$\boxed{0,8438}$	0,8461	0,8485
1,1	0,8643	0,8665	0,8686	0,8708
1,2	[illegible]	[illegible]	[illegible]	$\boxed{0,8907}$
1,3	0,9032	0,9049	0,9066	0,9082

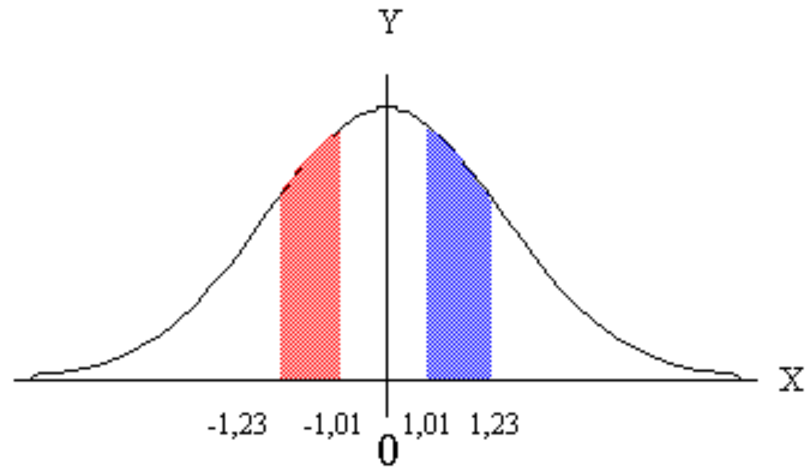


$$p(1{,}01 \leq Z \leq 1{,}23) = p(Z \leq 1{,}23) - p(Z \leq 1{,}01) =$$

$$= 0{,}8907 - 0{,}8438 = 0{,}1469$$

Cálculo de probabilidades con variables normales (IV)

x	0,00	0,01	0,02	0,03
0,0	0,5000	0,5040	0,5080	0,5120
0,1	0,5398	0,5438	0,5478	0,5517
0,2	0,5793	0,5832	0,5871	0,5910
0,3	0,6179	0,6217	0,6255	0,6293
0,4	0,6554	0,6591	0,6628	0,6664
0,5	0,6915	0,6950	0,6985	0,7019
0,6	0,7257	0,7291	0,7324	0,7357
0,7	0,7580	0,7611	0,7642	0,7673
0,8	0,7881	0,7910	0,7939	0,7967
0,9	0,8159	0,8186	0,8212	0,8238
1,0	[illegible]	0,8438	0,8461	0,8485
1,1	0,8643	0,8665	0,8686	0,8708
1,2	[illegible]	[illegible]	[illegible]	0,8907
1,3	0,9032	0,9049	0,9066	0,9082

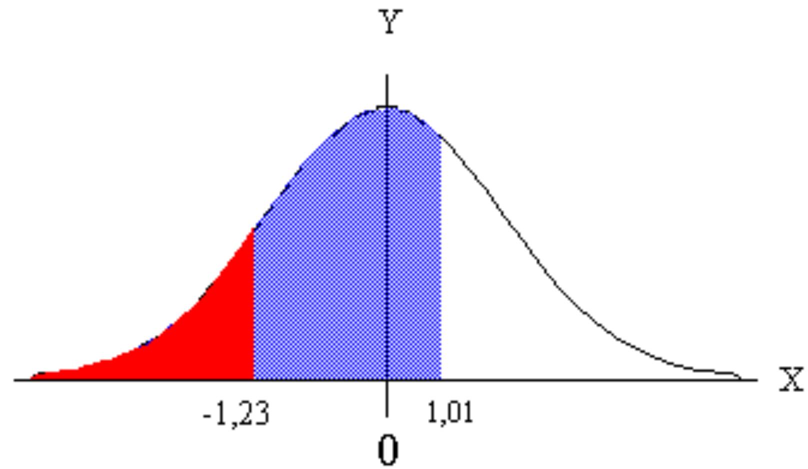


$$p(-1{,}23 \le Z \le -1{,}01) = p(1{,}01 \le Z \le 1{,}23) =$$

$$= p(Z \le 1{,}23) - p(Z \le 1{,}01) = 0{,}8907 - 0{,}8438 = 0{,}1469$$

Cálculo de probabilidades con variables normales (V)

x	0,00	0,01	0,02	0,03
0,0	0,5000	0,5040	0,5080	0,5120
0,1	0,5398	0,5438	0,5478	0,5517
0,2	0,5793	0,5832	0,5871	0,5910
0,3	0,6179	0,6217	0,6255	0,6293
0,4	0,6554	0,6591	0,6628	0,6664
0,5	0,6915	0,6950	0,6985	0,7019
0,6	0,7257	0,7291	0,7324	0,7357
0,7	0,7580	0,7611	0,7642	0,7673
0,8	0,7881	0,7910	0,7939	0,7967
0,9	0,8159	0,8186	0,8212	0,82[illegible]8
1,0	[illegible]	0,8438	0,8461	0,8485
1,1	0,8643	0,8665	0,8686	0,8708
1,2	[illegible]	[illegible]	[illegible]	0,8907
1,3	0,9032	0,9049	0,9066	0,9082



$$p(-1{,}23 \leq Z \leq 1{,}01) = p(Z \leq 1{,}01) - p(Z \leq -1{,}23) =$$

$$= p(Z \leq 1{,}01) - (1 - p(Z \leq 1{,}23)) = 0{,}8907 - 1 + 0{,}8438 = 0{,}7345$$

Intervalo que corresponde a una probabilidad fijada

x	0,00	0,01	0,02	0,03
0,0	0,5000	0,5040	0,5080	0,5120
0,1	0,5398	0,5438	0,5478	0,5517
0,2	0,5793	0,5832	0,5871	0,5910
0,3	0,6179	0,6217	0,6255	0,6293
0,4	0,6554	0,6591	0,6628	0,6664
0,5	[illegible]	[illegible]	0,6985	0,7019
0,6	0,7257	0,7291	0,7324	0,7357
0,7	0,7580	0,7611	0,7642	0,7673
0,8	[illegible]	[illegible]	[illegible]	0,7967
0,9	0,8159	0,8186	0,8212	0,8238
1,0	0,8413	0,8438	0,8461	0,8485
1,1	0,8643	0,8665	0,8686	0,8708
1,2	0,8849	0,8869	0,8888	0,8907
1,3	0,9032	0,9049	0,9066	0,9082

- ¿Qué valor de z corresponde a $p(Z \leq z) = 0{,}6985$?

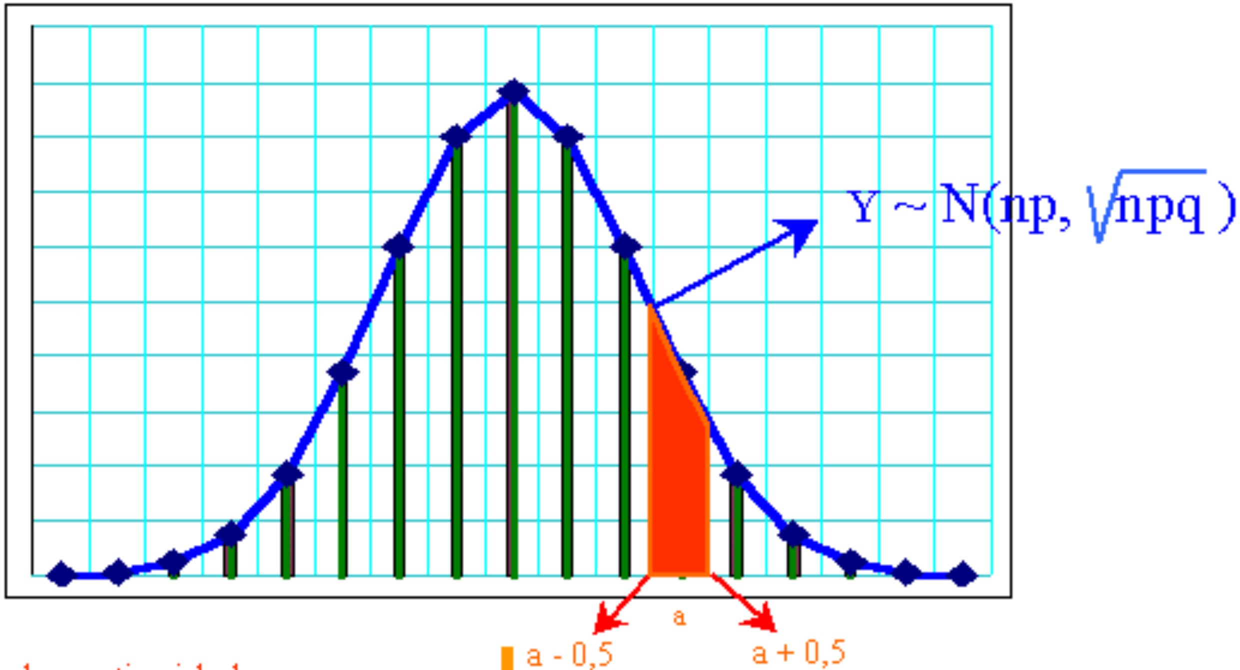
- Observamos que z = 0,52

- ¿Qué valor de z corresponde a $p(Z \leq z) = 0{,}2033$?

- Observamos que $p(Z \leq -z) = 1 - 0{,}2033 = 0{,}7967$.

- Entonces - z = 0,7967 y por tanto z = -0,7967

Cálculo de probabilidades en la binomial mediante la normal (I)



X es B(n, p)

corrección de continuidad

Y es $N(np, \sqrt{npq})$

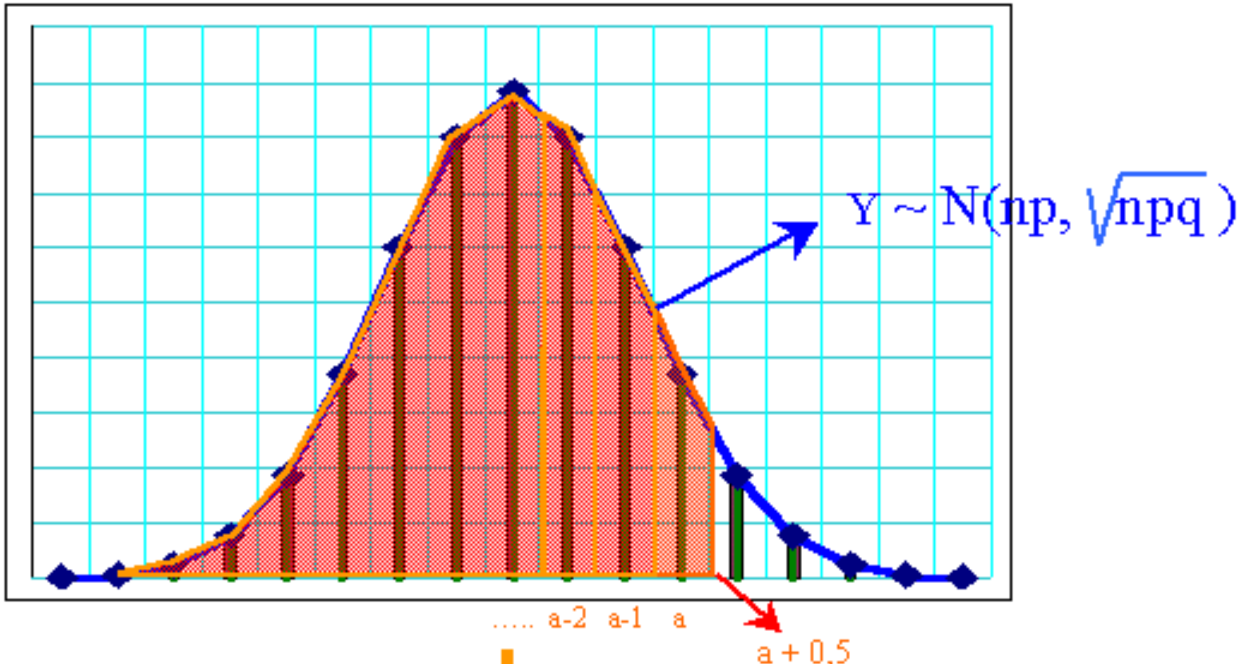
$$Z = \frac{Y - np}{\sqrt{npq}} \text{ es } N(0, 1)$$

$$p(X = a) = p(a - 0{,}5 \le Y \le a + 0{,}5) =$$

$$= p\left(\frac{a - 0{,}5 - np}{\sqrt{npq}} \le \frac{X' - np}{\sqrt{npq}} \le \frac{a + 0{,}5 - np}{\sqrt{npq}}\right) =$$

$$= p\left(\frac{a - 0{,}5 - np}{\sqrt{npq}} \le Z < \frac{a + 0{,}5 - np}{\sqrt{npq}}\right)$$

Cálculo de probabilidades en la binomial mediante la normal (II)



X es B(n, p)

corrección de continuidad

Y es $N(np, \sqrt{npq})$

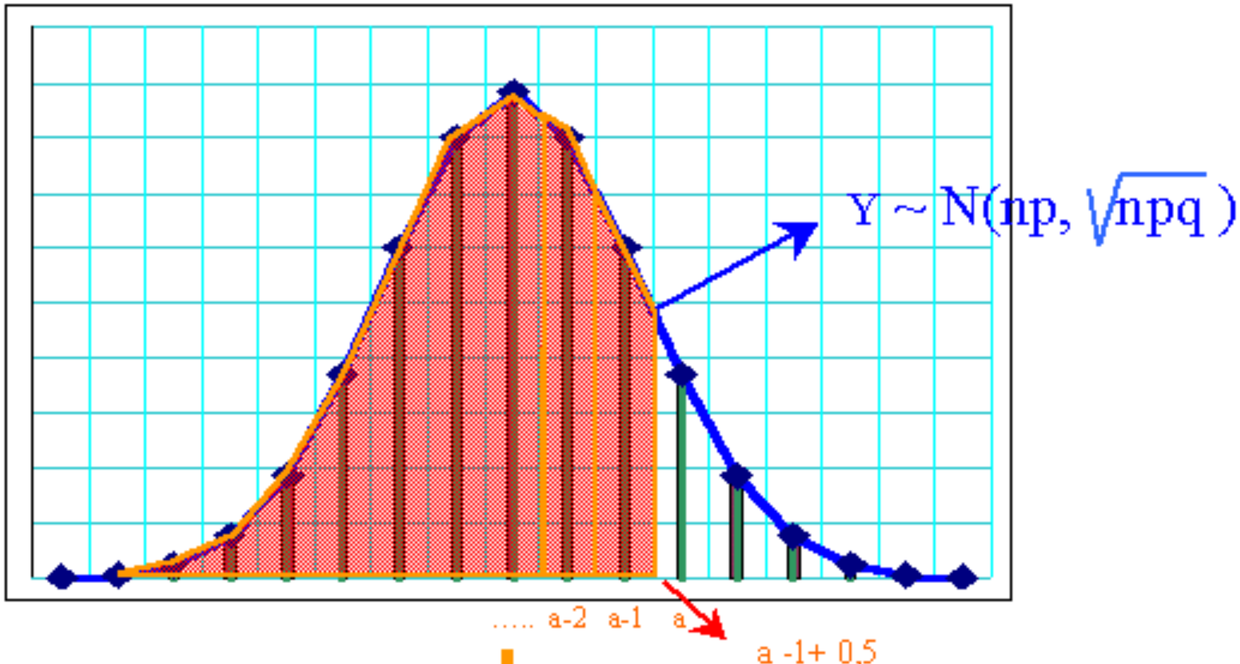
$$Z = \frac{Y - np}{\sqrt{npq}} \text{ es } N(0, 1)$$

$$p(X \leq a) = p(Y \leq a + 0{,}5) =$$

$$= p\left(\frac{Y - np}{\sqrt{npq}} \leq \frac{a + 0{,}5 - np}{\sqrt{npq}}\right) =$$

$$= p\left(Z \leq \frac{a + 0{,}5 - np}{\sqrt{npq}}\right)$$

Cálculo de probabilidades en la binomial mediante la normal (III)



X es B(n, p)

corrección de continuidad

Y es $N(np, \sqrt{npq})$

$$Z = \frac{Y - np}{\sqrt{npq}} \sim N(0, 1)$$

$$p(X < a) = p(X \leq a-1) = p(Y \leq a - 1 + 0{,}5) =$$

$$= p\left(\frac{Y - np}{\sqrt{npq}} < \frac{a - 0{,}5 - np}{\sqrt{npq}}\right) =$$

$$= p\left(Z < \frac{a - 0{,}5 - np}{\sqrt{npq}}\right)$$