

10 Variable aleatoria

10.1. Variable aleatoria

La introducción del concepto de variable aleatoria obedece al hecho de que muchos de los resultados de experimentos aleatorios no son números, y por tanto, es conveniente asociar a cada elemento del espacio muestral un valor numérico con el que poder trabajar.

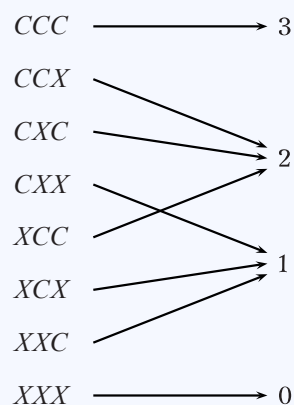
Una *variable aleatoria*, X , es una función que asocia a cada elemento del espacio muestral un número real.

Ejemplo 1

1. En el lanzamiento de tres monedas, el espacio muestral es:

$$E = \{CCC, CCX, CXC, CXX, XCC, XCX, XXC, XXX\}$$

La función que asigna a cada elemento de E el número de caras obtenido es una variable aleatoria:



La variable aleatoria X toma los valores:

$$X = \{0, 1, 2, 3\}$$

2. En el lanzamiento de dos dados, el espacio muestral es:

$$E = \{(1, 1), (1, 2) \dots (6, 6)\}$$

Para este experimento podemos considerar la variable aleatoria "suma de los puntos obtenidos". Esta variable aleatoria toma los valores:

$$X = \{2, 3, 4 \dots 12\}$$

Una vez realizado un experimento aleatorio, se pueden definir múltiples variables aleatorias para el mismo. Por ejemplo, para el lanzamiento de dos dados podríamos considerar la variable aleatoria “número de *unos* obtenidos”; en este caso, la variable toma los valores:

$$X = \{0, 1, 2\}$$

3. En una clínica de maternidad, la función que asigna a cada recién nacido su peso, es una variable aleatoria. También lo sería la función que asocia a cada recién nacido su talla.
4. En una plantación de judías verdes, la función que asigna a cada judía su longitud, es una variable aleatoria.

Observando los ejemplos anteriores se puede llegar a la conclusión de que las variables aleatorias pueden ser:

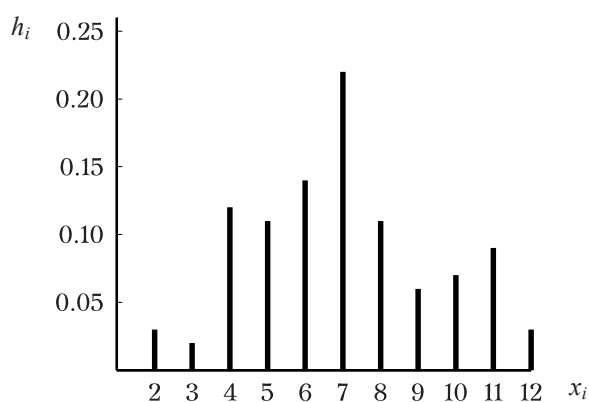
- ☞ *discretas*: sólo pueden tomar valores enteros. Es el caso de los dos primeros ejemplos.
- ☞ *continuas*: pueden tomar, al menos teóricamente, todos los posibles valores de un cierto intervalo de la recta real. Es el caso de los dos últimos ejemplos.

10.2. Distribución de probabilidad discreta

Consideremos el experimento consistente en lanzar reiteradamente dos dados y anotar la suma de los puntos obtenidos en cada lanzamiento. Sea X la variable aleatoria discreta “suma de las puntuaciones”.

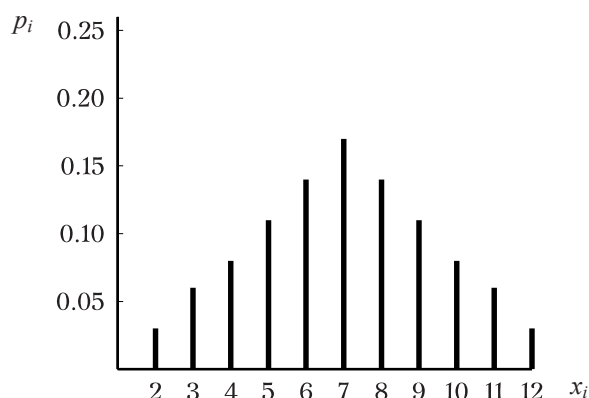
Supongamos que tras 100 lanzamientos, los resultados obtenidos son los que se resumen en la siguiente distribución de frecuencias, y el correspondiente diagrama de barras de frecuencias relativas:

X	f_i	h_i
2	3	0,03
3	2	0,02
4	12	0,12
5	11	0,11
6	14	0,14
7	22	0,22
8	11	0,11
9	6	0,06
10	7	0,07
11	9	0,09
12	3	0,03
$N = 100$		$\sum h_i = 1$



Vamos a representar ahora, en forma de tabla y de diagrama de barras, los resultados esperados a la vista del cálculo de probabilidades:

X	p_i
2	$1/36 = 0,028$
3	$2/36 = 0,056$
4	$3/36 = 0,083$
5	$4/36 = 0,111$
6	$5/36 = 0,139$
7	$6/36 = 0,167$
8	$5/36 = 0,139$
9	$4/36 = 0,111$
10	$3/36 = 0,083$
11	$2/36 = 0,056$
12	$1/36 = 0,028$
$\sum p_i = 1$	



El primer diagrama es experimental, corresponde a una *distribución de frecuencias*; el segundo diagrama es teórico, corresponde a una *distribución de probabilidad*.

Si fuésemos aumentando el número de lanzamientos de los dos dados 1000, 10 000... veces, observaríamos que las frecuencias relativas se aproximan cada vez más a las probabilidades teóricas. Es decir, conforme aumenta el número de pruebas del experimento aleatorio la distribución de frecuencias tiende a convertirse en distribución de probabilidad.

La distribución de probabilidad es, por tanto, una idealización de la distribución de frecuencias, y su utilidad proviene de que puede usarse para predecir, aproximadamente, los resultados de un experimento real.

En la distribución de probabilidad del ejemplo anterior vemos que a cada valor de la variable aleatoria le hacemos corresponder su probabilidad:

Función de probabilidad de una variable aleatoria X es la función que asocia a cada valor x_i de la variable su probabilidad p_i :

$$x_i \longrightarrow p_i = p(X = x_i)$$

Se suele utilizar la siguiente notación para p_i :

$$p(X = x_i) \quad (\text{se lee: probabilidad de que la variable aleatoria tome el valor } x_i)$$

Por ejemplo, para el lanzamiento de los dos dados,

$$p(X = 2) = \frac{1}{36} \quad ; \quad p(X = 7) = \frac{6}{36} \quad ; \quad p(X = 11) = \frac{2}{36} \quad \text{etc.}$$

En una función de probabilidad, debe cumplirse siempre que:

- $0 \leq p_i \leq 1, \quad \forall p_i$
- $\sum p_i = 1$

La función de probabilidad de una variable aleatoria discreta nos proporciona todos los posibles valores de la variable, y la probabilidad de que ocurra dicho valor. Puede venir dada en forma de tabla, de gráfico de barras, o mediante una expresión algebraica.

Ejemplo 2

1. Hallar la función de probabilidad en forma tabular y gráfica de la variable aleatoria “número de cruces obtenidas al lanzar tres monedas”.

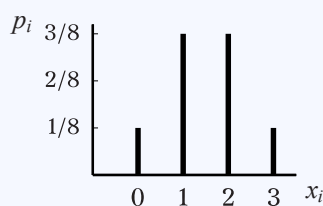
Según el cálculo de probabilidades:

$$p(X=0) = \frac{1}{8}, \quad p(X=1) = \frac{3}{8}, \quad p(X=2) = \frac{3}{8}, \quad p(X=3) = \frac{1}{8}$$

En forma tabular:

x_i	p_i	Observemos que:
0	1/8	$p(X=0) + p(X=1) + p(X=2) + p(X=3) = 1$
1	3/8	y que $0 \leq p_i \leq 1$, para todas las probabilidades de la
2	3/8	tabla.
3	1/8	

En forma gráfica:



2. Una variable aleatoria X tiene la distribución de probabilidad dada por la siguiente tabla. Halla el valor de c .

x_i	0	1	2	3	4
$p(X=x_i)$	1/5	1/6	3/10	1/5	c

$$\text{Como: } \sum p(X=x_i) = 1 \Rightarrow \frac{1}{5} + \frac{1}{6} + \frac{3}{10} + \frac{1}{5} + c = 1 \Rightarrow c = \frac{4}{30} = \frac{2}{15}$$

3. Demuestra que la función $P(x) = \frac{x^2+1}{34}$, para $x = 1, 2, 3, 4$, es una función de probabilidad. Escríbela en forma tabular.

$$P(1) = \frac{1^2+1}{34} = \frac{2}{34}$$

$$P(2) = \frac{2^2+1}{34} = \frac{5}{34}$$

$$P(3) = \frac{3^2+1}{34} = \frac{10}{34}$$

$$P(4) = \frac{4^2+1}{34} = \frac{17}{34}$$

Todas las $P(i)$ cumplen: $0 \leq P(i) \leq 1$

$$\text{Además: } \sum P(i) = \frac{2}{34} + \frac{5}{34} + \frac{10}{34} + \frac{17}{34} = 1$$

Luego $P(x)$ es una función de probabilidad. En forma de tabla:

x_i	1	2	3	4
p_i	2/34	5/34	10/34	17/34

Mediante la función de probabilidad podemos conocer la probabilidad de que la variable X tome un valor concreto x_i . Pero también interesa conocer la probabilidad de que la variable tome valores menores o iguales que el de x_i , es decir, interesa conocer $p(X \leq x_i)$.

Si ordenamos los valores $x_1, x_2, \dots, x_n$ de menor a mayor podemos hallar el valor $p(X \leq x_i)$ del

siguiente modo:

$$p(X \leq x_i) = p(X = x_1) + p(X = x_2) + \cdots + p(X = x_i)$$

Se llama *función de distribución* de una variable aleatoria X a la función que asocia a cada valor x_i de la variable su probabilidad acumulada:

$$x_i \longrightarrow p(X \leq x_i)$$

Ejemplo 3

Sea X la variable aleatoria “nº obtenido al lanzar un dado”. La función de probabilidad y la de distribución serán:

X	$p(X = x_i)$	$p(X \leq x_i)$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6

Así: $p(X \leq 3) = \frac{3}{6}$, $p(X \leq 5) = \frac{5}{6}$, etc.

10.3. Parámetros de una variable aleatoria discreta

De modo análogo a como se procede en las distribuciones de frecuencias en las que la información se resume mediante el cálculo de una serie de parámetros, en las distribuciones de probabilidad también interesa conocerlos.

10.3.1. Media

Recordemos que en las distribuciones de frecuencias la media se calcula:

$$\bar{x} = \frac{\sum x_i f_i}{N} = \sum x_i h_i$$

Por analogía, se calcula la media de una distribución de probabilidad del siguiente modo:

$$E(X) = \mu = \sum x_i p_i$$

A esta media se la llama también *esperanza matemática*, y se la representa por $E(X)$. Este nombre proviene de los juegos de azar (origen del cálculo de probabilidades), ya que la media representa la ganancia esperada por un jugador, en promedio, cuando efectúa un gran número de apuestas, y es el parámetro utilizado para medir si un juego es equitativo o no.

10.3.2. Varianza y desviación típica

En una distribución de frecuencias la varianza se calcula del siguiente modo:

$$s^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{N} = \sum (x_i - \bar{x})^2 h_i \quad \text{o también} \quad s^2 = \frac{\sum x_i^2 f_i}{N} - \bar{x}^2 = \sum x_i^2 h_i - \bar{x}^2$$

Para una distribución de probabilidad, la fórmula de la varianza será:

$$Var(X) = \sigma^2 = \sum (x_i - \mu)^2 p_i \quad \text{o bien} \quad Var(X) = \sigma^2 = \sum x_i^2 p_i - \mu^2 = E(X^2) - [E(X)]^2$$

La raíz cuadrada positiva de la varianza es la desviación típica, que se representa por σ o $SD(X)$.

Ejemplo 4

El tutor de una clase de 1º de Bachillerato ha llegado a la conclusión de que el número de entrevistas anuales con los padres de cada alumno sigue la distribución de probabilidad dada en la tabla:

nº entrevistas	1	2	3	4
probabilidad	1/2	1/3	1/10	1/15

- a) Calcula la media y la desviación típica de la variable aleatoria “nº de entrevistas por alumno”.
- b) Si tiene un total de 30 alumnos, halla el nº de entrevistas que espera realizar en total durante todo el curso.

a)	x_i	p_i	$x_i p_i$	$x_i^2 p_i$	$\mu = \sum x_i p_i = \frac{26}{15} = 1,73$
	1	1/2	1/2	1/2	
	2	1/3	2/3	4/3	
	3	1/10	3/10	9/10	
	4	1/15	4/15	16/15	
			$\Sigma = 26/15$	$\Sigma = 19/5$	$\sigma^2 = \sum x_i^2 p_i - \mu^2 = \frac{19}{5} - \left(\frac{26}{15}\right)^2 =$
					$= \frac{179}{225}; \quad \sigma = \sqrt{\frac{179}{225}} = 0,89$

- b) Si en promedio hace 1,73 entrevistas por alumno cada curso, para un total de 30 alumnos:

$$30 \cdot 1,73 = 51,9 \approx 52 \text{ entrevistas.}$$

Ejemplo 5

En un juego de azar, el jugador acciona una ruleta dividida en cuatro partes iguales numeradas del 1 al 4, y gana la cantidad de dinero que se muestra en la tabla:

nº	1	2	3	4
€	1	2	5	8

Halla:

- a) La ganancia esperada en promedio por un jugador.
- b) Suponiendo que cada jugada cueste 5€, ¿recomendarías a alguien jugar?

- a) Sea X la variable aleatoria “ganancia obtenida en un juego”. Como todos los números son igualmente probables, la distribución de probabilidad será:

x_i	p_i	$x_i p_i$
1	1/4	1/4
2	1/4	2/4
5	1/4	5/4
8	1/4	8/4
		$\Sigma = 16/4 = 4$

$$E(X) = \mu = \sum x_i p_i = 4$$

La ganancia en promedio es de 4 €.

- b) No recomendaría jugar porque el juego cuesta 5 €, mientras que la ganancia esperada es de 4 €. Para que un juego sea equitativo la apuesta debe ser igual a la esperanza o media del juego (no hay ventaja ni para la banca ni para el jugador).

10.4. La distribución binomial

10.4.1. Introducción

La distribución binomial es un caso particular de distribución de probabilidad de variable aleatoria discreta. Fue enunciada por Jakob Bernoulli hacia el año 1700, y por sus aplicaciones, es posiblemente la más importante de las distribuciones discretas de probabilidad.

Esta distribución corresponde a la realización de un experimento aleatorio que cumpla las siguientes condiciones:

- ☞ al realizar el experimento sólo se consideran dos posibles resultados: el suceso A , llamado *éxito*, o su contrario $\bar{A}$, llamado *fracaso*.
- ☞ al repetir el experimento, el resultado obtenido es independiente de los resultados obtenidos anteriormente.
- ☞ la probabilidad del suceso A es constante, es decir, no varía de una prueba del experimento a otra. Si llamamos p a la probabilidad de A , $p(A) = p$, entonces $p(\bar{A}) = 1 - p = q$.
- ☞ en cada experimento se realizan n pruebas idénticas.

La variable aleatoria que expresa el número de éxitos obtenidos en la realización de n pruebas del experimento se llama *variable aleatoria binomial*. Esta variable se representa:

$$X \longrightarrow B(n, p) \quad , \text{ donde:}$$

- n = número de pruebas del experimento.
- p = probabilidad de éxito.

La variable binomial puede tomar los valores $0, 1, 2, \dots, n$, por tanto, es una variable discreta; n y p se llaman *parámetros de la distribución*.

Ejemplo 6

Son distribuciones binomiales las siguientes:

1. Lanzamos diez veces una moneda y contamos el número de caras. Llamamos:

cara = éxito $\Rightarrow p = 0,5$

cruz = fracaso $\Rightarrow q = 0,5$

Como se hacen 10 lanzamientos: $n = 10$. El número de caras en diez lanzamientos sigue una binomial $B(10; 0,5)$.

2. De una bolsa con 3 bolas blancas y 7 negras se extraen 6 bolas, sucesivamente y con devolución, y se anota el número de bolas blancas. Llamamos:

blanca = éxito $\Rightarrow p = 0,3$

negra = fracaso $\Rightarrow q = 0,7$

El número de bolas blancas obtenidas en las seis extracciones sigue una binomial $B(6; 0,3)$.

Observemos que no se trataría de una distribución binomial si las extracciones fuesen sin devolución, ya que la probabilidad de extraer bola blanca (éxito) no permanecería constante de una extracción a la siguiente.

3. Un profesor ha observado a lo largo de los años que aprueba al 65 % de sus estudiantes. Tomamos una muestra formada por 20 estudiantes y contamos el número de aprobados. Llamamos:

aprobado = éxito $\Rightarrow p = 0,65$

suspenso = fracaso $\Rightarrow q = 0,35$

El número de aprobados en la muestra sigue una binomial $B(20; 0,65)$.

10.4.2. Función de probabilidad de la variable binomial

La variable aleatoria binomial, como toda variable aleatoria discreta, tiene asociada una función de probabilidad.

La función de probabilidad de la variable binomial nos da la probabilidad de obtener r éxitos en n pruebas del experimento aleatorio:

$$p(X = r)$$

Ejemplo 7

Sea la binomial correspondiente al número de caras obtenidas al lanzar 10 veces una moneda, $B(10; 0,5)$; la variable X (número de caras) podrá tomar los valores:

$X = \{0, 1, 2, \dots, 10\}$.

La función de probabilidad nos da las probabilidades $p(X = 0)$, $p(X = 1)$, $\dots$, $p(X = 10)$.

Para obtener la función de probabilidad de una variable aleatoria binomial $B(n, p)$, se deberá calcular $p(X = r)$, donde $r = 0, 1, 2, \dots, n$.

Supongamos que al realizar n pruebas del experimento aleatorio obtenemos r veces el suceso A (r éxitos) y $n - r$ veces el suceso $\bar{A}$ ($n - r$ fracasos). Como obtener éxito en una cualquiera de las pruebas es independiente de los éxitos obtenidos en las pruebas anteriores, los sucesos son independientes, y la probabilidad de su intersección es el producto de sus probabilidades:

$$\begin{aligned} p(\underbrace{A \cap A \cap \dots \cap A}_{r \text{ veces}} \cap \underbrace{\bar{A} \cap \bar{A} \cap \dots \cap \bar{A}}_{n-r \text{ veces}}) &= \underbrace{p(A) \cdot p(A) \cdots p(A)}_{r \text{ veces}} \cdot \underbrace{p(\bar{A}) \cdot p(\bar{A}) \cdots p(\bar{A})}_{n-r \text{ veces}} = \\ &= \underbrace{p \cdot p \cdot p \cdots p}_{r \text{ veces}} \cdot \underbrace{q \cdot q \cdot q \cdots q}_{n-r \text{ veces}} = p^r q^{n-r} \end{aligned}$$

Pero la ordenación anterior no es la única posible para obtener r éxitos y $n - r$ fracasos. Hay tantas ordenaciones posibles como permutaciones con repetición de n elementos entre los que hay r repetidos y $n - r$ repetidos:

$$PR_n^{r, n-r} = \frac{n!}{r!(n-r)!} = \binom{n}{r}$$

Luego:

$$p(X = r) = \binom{n}{r} p^r q^{n-r} \quad \text{Función de probabilidad de la distribución binomial}$$

Ejemplo 8

Un examen de tipo test consta de 10 preguntas con cuatro respuestas cada una, siendo sólo una de ellas correcta. Un alumno contesta todas las preguntas al azar. Calcula la probabilidad de:

- Acertar exactamente cuatro preguntas.
- No acertar ninguna.
- Acertar todas.
- Acertar al menos ocho.
- Acertar a lo sumo tres.

Consideremos:

- éxito: contestar bien una pregunta $\Rightarrow p = 0,25$
- fracaso: contestar mal una pregunta $\Rightarrow q = 0,75$

El número de preguntas bien contestadas del total de 10 es una variable aleatoria que sigue una distribución binomial $B(10; 0,25)$.

$$\text{a)} \quad p(X = 4) = \binom{10}{4} \cdot 0,25^4 \cdot 0,75^6 = 0,1460 = 14,60\%$$

$$\text{b)} \quad p(X = 0) = \binom{10}{0} \cdot 0,25^0 \cdot 0,75^{10} = 0,0563 = 5,63\%$$

$$\text{c)} \quad p(X = 10) = \binom{10}{10} \cdot 0,25^{10} \cdot 0,75^0 = 9,53 \cdot 10^{-7} \simeq 0$$

$$\begin{aligned} \text{d)} \quad p(X \geq 8) &= p(X = 8) + p(X = 9) + p(X = 10) = \binom{10}{8} \cdot 0,25^8 \cdot 0,75^2 + \\ &+ \binom{10}{9} \cdot 0,25^9 \cdot 0,75 + \binom{10}{10} \cdot 0,25^{10} = 0,0005 = 0,05\% \end{aligned}$$

$$\begin{aligned} \text{e)} \quad p(X \leq 3) &= p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) = \\ &= \binom{10}{0} \cdot 0,75^{10} + \binom{10}{1} \cdot 0,25 \cdot 0,75^9 + \binom{10}{2} \cdot 0,25^2 \cdot 0,75^8 + \binom{10}{3} \cdot 0,25^3 \cdot 0,75^7 = \\ &= 0,7759 = 77,59\% \end{aligned}$$

10.4.3. Media y varianza de una distribución binomial

El cálculo de la media, varianza y desviación típica de una distribución binomial es inmediato una vez conocidos los parámetros n y p .

☞ Supongamos el caso $n = 1$:

x_i	p_i	$x_i p_i$	$x_i^2 p_i$	
0	q	0	0	$\mu = p$
1	p	p	p	$\sigma^2 = p - p^2 = p(1 - p) = pq$
	$\Sigma = p$		$\Sigma = p$	

☞ Caso $n = 2$:

x_i	p_i	$x_i p_i$	$x_i^2 p_i$	
0	q^2	0	0	$\mu = 2pq + 2p^2 = 2p(q + p) = 2p$
1	$2pq$	$2pq$	$2pq$	$\sigma^2 = 4p^2 + 2pq - 4p^2 = 2pq$
2	p^2	$2p^2$	$4p^2$	

Etc.

Procediendo por inducción se llega a:

$$\mu = np \quad ; \quad \sigma^2 = npq \quad ; \quad \sigma = \sqrt{npq}$$

Ejemplo 9

En una urna hay cuatro bolas blancas y tres negras. Se extraen tres bolas sucesivamente retornándolas después de cada extracción. Sea la variable aleatoria X = “número de bolas blancas extraídas”. Halla la media y la desviación típica.

Se trata de una $B(3; \frac{4}{7})$ $\mu = np = 3 \cdot \frac{4}{7} = 1,71$

$\sigma^2 = npq = 3 \cdot \frac{4}{7} \cdot \frac{3}{7} = \frac{36}{49}$ $\sigma = \sqrt{\frac{36}{49}} = \frac{6}{7} = 0,857$

10.4.4. Probabilidades binomiales y calculadora gráfica

Las siguientes instrucciones son válidas para la calculadora CASIO fx-9750G PLUS y similares.

☞ Menú **STAT**.

☞ **F5 (DIST) ⇒ F5 (BINM) ⇒** hay dos posibilidades:

* **Bpd** (*binomial probability distribution*): se usa para calcular probabilidades puntuales, del tipo $p(X = r)$. Introducir los datos del siguiente modo:

Data: Variable x: n° de éxitos (r) Numtrial: n° pruebas del experimento (n) p: probabilidad de éxito (p)	}	$\Rightarrow$ F1 (CALC)
---	---	--------------------------------

* **Bcd** (*binomial cumulative distribution*): se usa para calcular probabilidades acumuladas, del tipo $p(X \leq r)$. Los datos se introducen como en el caso anterior.

Para calcular, por ejemplo, $p(X \geq 8)$, hemos de tener en cuenta que:

$$p(X \geq 8) = 1 - p(X < 8) = 1 - p(X \leq 7)$$

Por tanto, hallamos $p(X \leq 7)$ con la calculadora, y el resultado lo restamos de 1.

10.5. Distribución de probabilidad continua

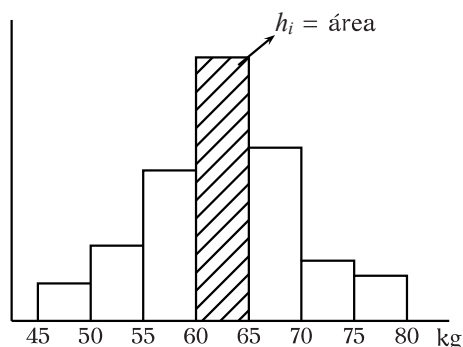
Dada una variable aleatoria continua, carece de sentido asignar a cada valor x_i de la variable su probabilidad p_i , ya que al ser X continua puede tomar los infinitos valores pertenecientes a un cierto intervalo de la recta real.

Como no podemos definir la función de probabilidad para una variable continua del modo en que lo hemos hecho para variables discretas, es necesario introducir un nuevo concepto matemático que caracterice a dicho tipo de variables.

Comenzamos con un ejemplo: sea la variable continua X que mide los pesos (en kg) de 50 alumnos de un Instituto. Los valores que toma dicha variable los agrupamos en intervalos o clases, por ejemplo:

[45; 50); [50; 55); [55; 60) ... [75; 80)

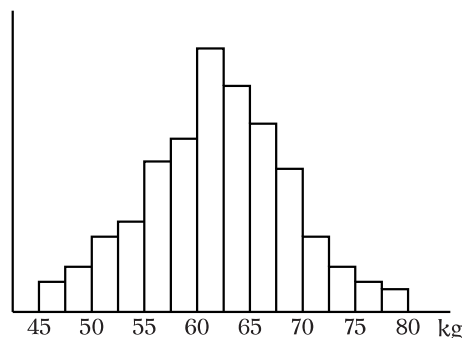
Una vez hecho el recuento de alumnos cuyo peso pertenece a cada intervalo, se pueden calcular las frecuencias relativas h_i , y representar estos datos en forma de histograma de frecuencias relativas, obteniendo un resultado del siguiente tipo:



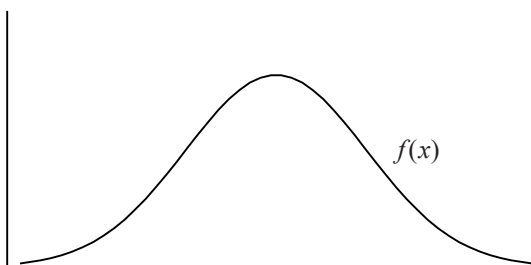
Con un sencillo cambio en la escala del eje vertical se puede conseguir que sea el área de cada rectángulo, y no su altura, la que represente la frecuencia relativa de cada clase.

Los histogramas contruidos de este modo se llaman *histogramas de densidad*. En ellos **la suma de las áreas de todos los rectángulos es igual a 1**.

Supongamos ahora que tenemos datos sobre el peso de más alumnos (100, por ejemplo), y que construimos un histograma de densidad con los intervalos de clase más pequeños (por ejemplo, reduciendo la amplitud a la mitad). Obtendríamos una representación similar a la de la derecha:



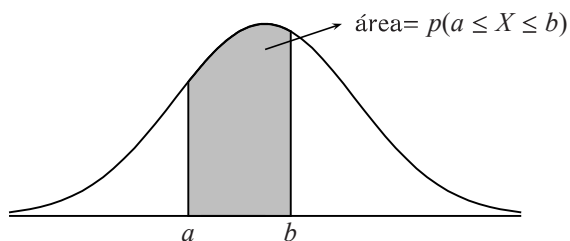
Si continuáramos con este proceso de reducir el tamaño del intervalo y aumentar el tamaño de la muestra de forma indefinida, obtendríamos un número infinitamente grande de intervalos infinitamente pequeños; la gráfica escalonada se transformaría entonces en una curva continua, y las frecuencias relativas de las clases se aproximarían a las probabilidades respectivas:



Esta curva corresponde a una función real de variable real $f(x)$, y recibe el nombre de *función de densidad de la variable aleatoria continua*. (Nota: aunque aquí se ha dibujado una curva simétrica, no tiene por qué serlo.)

Una función de densidad $f(x)$ debe cumplir las siguientes condiciones:

- ☞ $f(x) \geq 0$ en todo su dominio (las probabilidades son siempre positivas).
- ☞ el área total encerrada bajo la curva de $f(x)$ es igual a 1.
- ☞ la probabilidad de que la variable aleatoria tome valores en un intervalo determinado $[a, b]$ es el área encerrada bajo la curva en dicho intervalo:



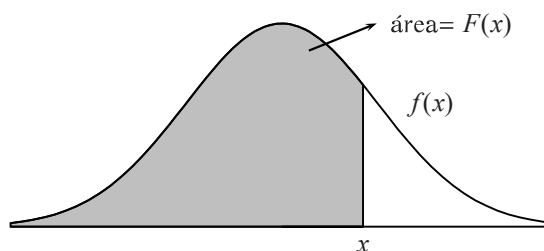
- ☞ la probabilidad de que la variable X tome un valor puntual, $p(X = a)$, es siempre cero, como corresponde al área de una región cuya base es un punto.
Por ejemplo, para los pesos de los alumnos no tiene sentido plantearse cuál es la probabilidad de que un alumno pese exactamente 63,5847 kg, pero sí tiene sentido preguntarse por la probabilidad de que el peso esté comprendido entre 63 y 64 kg.

Por tanto, para una variable continua, las probabilidades se obtienen calculando áreas bajo la función de densidad. El cálculo del área puede ser complicado, sobre todo cuando no se dispone de la herramienta adecuada, que es el cálculo integral. Para facilitarlo, se define una nueva función, representada por $F(x)$ y llamada *función de distribución*:

La función de distribución $F(x)$ de una variable aleatoria continua proporciona el área encerrada bajo la gráfica de la función de densidad $f(x)$, desde $-\infty$ hasta x , es decir:

$$F(x) = p(X \leq x)$$

Gráficamente:



10.6. La distribución normal

10.6.1. Introducción

La distribución normal es un caso particular de distribución de probabilidad de variable continua, sin duda el más importante, ya que hay multitud de fenómenos que se ajustan a ella.

La distribución normal la presentan aquellas variables que dependen de muchos factores; dicha distribución acumula muchos individuos en los valores centrales (próximos a la media), pero el número de éstos va decreciendo según se aleja la variable de la media, tanto por encima como por debajo, de forma simétrica.

La función de densidad de una variable normal es una curva simétrica llamada *campana de Gauss*, en honor al matemático que la estudió inicialmente (ver los gráficos anteriores).

Muchos fenómenos de la vida real siguen este modelo de probabilidad: estaturas, pesos, cociente intelectual, vida media de los aparatos, gasto en productos de consumo, errores en las mediciones de magnitudes, etc. Esta ubicuidad en fenómenos tan dispares puede justificarse gracias al llamado *teorema central del límite*, que afirma que aunque en un fenómeno intervengan diferentes causas que no siguen una distribución normal, su presencia simultánea hace que la distribución tienda a la normal.

No todas las distribuciones de variable continua son normales. Por ejemplo: nivel de renta de los españoles, distancia del impacto al centro de la diana al realizar el disparo un tirador profesional, ... son ejemplos de distribuciones asimétricas.

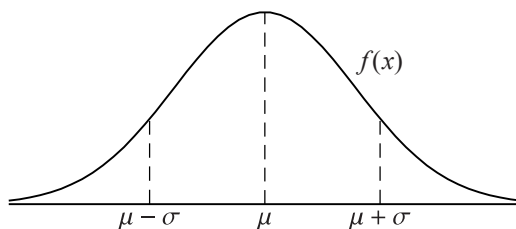
10.6.2. Función de densidad de la distribución normal

La función de densidad de una variable aleatoria con distribución normal tiene la siguiente expresión:

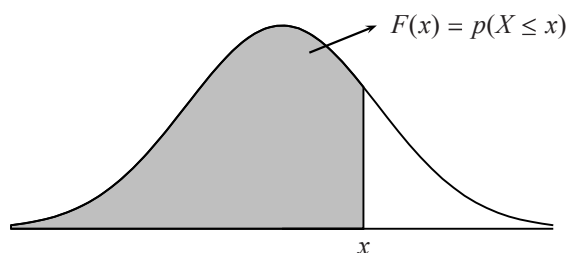
$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2}$$

donde: μ : media de la distribución.
 σ : desviación típica.

Esta expresión analítica es compleja (¡no es preciso aprenderla!), pero su representación gráfica, la campana de Gauss, nos permite descubrir muchas de sus propiedades:



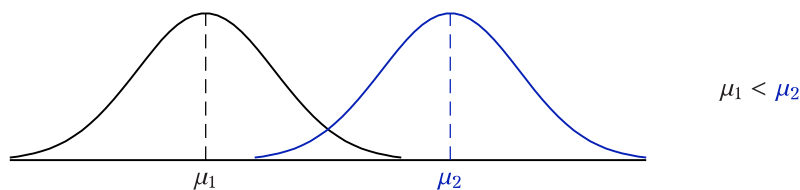
- ☞ es una función definida para cualquier valor de x ; su dominio, por tanto es $\mathbb{R}$.
- ☞ es una función simétrica respecto de la recta $x = \mu$, y alcanza el máximo en dicho valor.
- ☞ el eje de abscisas es una asíntota horizontal.
- ☞ la función posee dos puntos de inflexión en $\mu - \sigma$ y $\mu + \sigma$.
- ☞ el área encerrada entre la curva $f(x)$ y el eje de abscisas es 1. Por ser simétrica, el eje $x = \mu$ deja un área igual a 0,5 a la derecha, y otra igual a 0,5 a la izquierda.
- ☞ la función de distribución correspondiente, $F(x)$, nos da las áreas comprendidas entre la curva y el eje OX , desde $-\infty$ hasta cada valor de x . Este área coincide con $p(X \leq x)$:



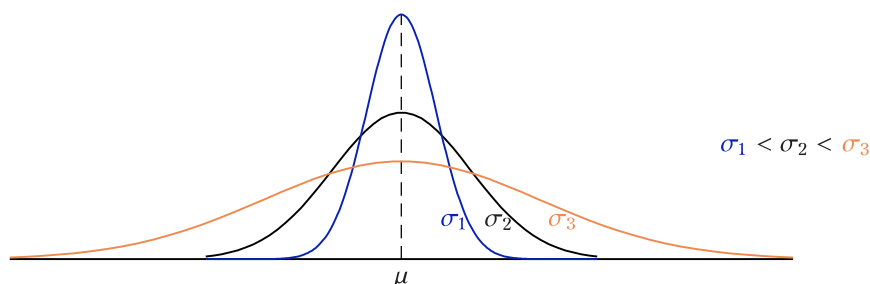
Una variable aleatoria que sigue una distribución normal de media μ y desviación típica σ se representa por $N(\mu, \sigma)$.
 μ y σ son los parámetros de la distribución.

Para cada valor de μ y σ tendremos una función de densidad distinta; estos valores influyen sobre la forma de la campana de Gauss:

- ☞ la media desplaza la curva a izquierda o derecha:



- ☞ la desviación típica estrecha o ensancha la curva: cuando σ es alta, la dispersión es alta y la gráfica menos estilizada y más abierta; cuando σ disminuye, la dispersión es más pequeña, y la gráfica más estilizada (datos más concentrados en torno a la media). Hay que tener en cuenta que el área bajo la campana se mantiene constante e igual a 1:



10.6.3. Distribución normal estándar

Para calcular probabilidades en una distribución normal, es necesario calcular el área bajo la curva de la función de densidad entre dos valores cualesquiera. Como este cálculo no es trivial, se han elaborado tablas para la función de distribución, $F(x) = p(X \leq x)$. El problema es que existen infinitas distribuciones normales, tantas como pares (μ, σ) distintos hay.

Ahora bien, todas ellas tienen propiedades comunes, y se puede reducir una cualquiera de ellas a cualquier otra, sólo cambiando la variable convenientemente. Por tanto, bastará con tener las tablas de una única distribución normal.

Se ha tabulado la distribución normal más sencilla, que es la que tiene media $\mu = 0$ y desviación típica $\sigma = 1$, es decir, la $N(0, 1)$. Ésta se llama *distribución normal estándar*.

La variable aleatoria de la $N(0, 1)$ se representa por Z (en lugar de X), y se llama *variable normal tipificada*.

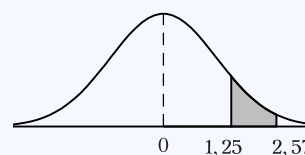
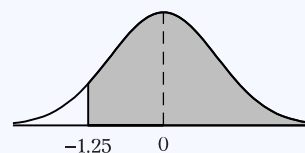
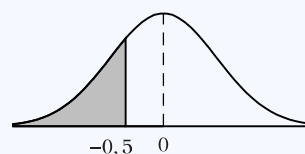
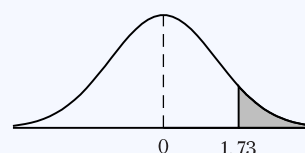
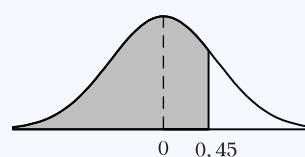
Las tablas de la distribución $N(0, 1)$ exponen los valores de las áreas bajo la campana desde $-\infty$ hasta k , es decir las $p(Z \leq k)$. Estos valores se encuentran en el cruce de la fila que contiene las unidades y décimas de k con la columna de sus centésimas (ver tabla de la distribución normal estándar).

La información proporcionada por la tabla, junto con la simetría de la curva, y el hecho de que el área total bajo la campana sea igual a 1, es suficiente para calcular cualquier probabilidad.

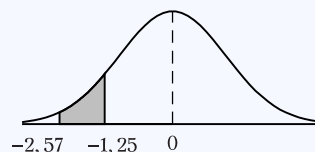
Ejemplo 10

Si Z sigue una distribución $N(0, 1)$:

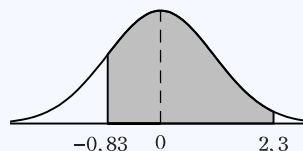
- $p(Z \leq 0,45) = 0,6736$
(lectura directa de la tabla)
- $p(Z \geq 1,73) = 1 - p(Z \leq 1,73) =$
 $= 1 - 0,9582 = 0,0418$
- $p(Z \leq -0,5) = p(Z \geq 0,5) =$
 $= 1 - p(Z \leq 0,5) = 1 - 0,6915 = 0,3085$
- $p(Z \geq -1,25) = p(Z \leq 1,25) =$
 $= 0,8944$
- $p(1,25 \leq Z \leq 2,57) =$
 $= p(Z \leq 2,57) - p(Z \leq 1,25) =$
 $= 0,9949 - 0,8944 = 0,1005$



$$\begin{aligned} \bullet \quad & p(-2,57 \leq Z \leq -1,25) = \\ & = p(1,25 \leq Z \leq 2,57) = 0,1005 \end{aligned}$$



$$\begin{aligned} \bullet \quad & p(-0,83 \leq Z \leq 2,3) = \\ & = p(Z \leq 2,3) - p(Z \leq -0,83) = \\ & = p(Z \leq 2,3) - (1 - p(Z \leq 0,83)) = \\ & = 0,9893 - 1 + 0,7967 = 0,7860 \end{aligned}$$



10.6.4. Tipificación de la variable

En los problemas reales la distribución normal no será una $N(0,1)$, sino $N(\mu, \sigma)$. Para calcular las probabilidades en una $N(\mu, \sigma)$ hay que transformar la variable X en otra variable Z que siga una $N(0,1)$. Esta transformación se conoce con el nombre de *tipificación de la variable*.

Para llevar a cabo esta transformación hay que realizar dos pasos:

- 1º) Centrar, es decir, trasladar la campana de Gauss hasta que la media μ sea igual a 0. Ello se consigue restando $X - \mu$.
- 2º) Reducir la desviación típica a 1. Esto equivale a contraer o dilatar la campana de Gauss para que coincida con la de la normal estándar. Para ello, basta con dividir entre σ : $\frac{X - \mu}{\sigma}$.

En resumen:

$$\begin{array}{ccc} N(\mu, \sigma) & \text{-----} \rightarrow & N(0,1) \\ X & \xrightarrow{\text{tipificación}} & Z = \frac{X - \mu}{\sigma} \end{array}$$

Los valores de la variable tipificada lo que miden en realidad es **cuántas desviaciones típicas se alejan los valores de la variable respecto de la media**.

Si X es $N(\mu, \sigma)$, con la tipificación la probabilidad $p(X \leq k)$ se calcula así:

$$p(X \leq k) = p\left(\frac{X - \mu}{\sigma} \leq \frac{k - \mu}{\sigma}\right) = p\left(Z \leq \frac{k - \mu}{\sigma}\right)$$

donde Z es $N(0,1)$. El valor de esta última probabilidad se mira en la tabla.

Ejemplo 11

Se fabrican unas pilas alcalinas cuya duración en horas sigue una distribución normal de media 60 h, y desviación típica 5 h. Si se elige una pila al azar, qué probabilidad hay de que dure:

- a) Menos de 50 h.
- b) Entre 52 y 65 h.
- c) Examinadas 200 pilas, ¿cuántas durarán más de 68 h?

- a) Sea X = número de horas de duración de una pila: $X \rightarrow N(60; 5)$

$$p(X < 50) = p\left(\frac{X - 60}{5} < \frac{50 - 60}{5}\right) = p(Z < -2) = p(Z > 2) =$$

$$= 1 - p(Z \leq 2) = 1 - 0,9772 = 0,0228$$

- b) $p(52 \leq X \leq 65) = p\left(\frac{52 - 60}{5} \leq \frac{X - 60}{5} \leq \frac{65 - 60}{5}\right) =$
- $$= p(-1,6 \leq Z \leq 1) = P(Z \leq 1) - p(Z \leq -1,6) =$$
- $$= p(Z \leq 1) - p(Z \geq 1,6) = p(Z \leq 1) - [1 - p(Z \leq 1,6)] =$$
- $$= 0,8413 - 1 + 0,9452 = 0,7865$$

- c) $p(X \geq 68) = p\left(\frac{X - 60}{5} \geq \frac{68 - 60}{5}\right) = p(Z \geq 1,6) = 1 - p(Z \leq 1,6) =$
- $$= 1 - 0,9452 = 0,0548 = 5,48\%$$

Para un total de 200 pilas: $200 \cdot 0,0548 = 10,96 \approx 11$ pilas.

10.6.5. Significado de la desviación típica

Dada una distribución normal $N(\mu, \sigma)$, es interesante conocer cuál es el porcentaje de individuos que se encuentra en el intervalo $(\mu - \sigma, \mu + \sigma)$, es decir, que no se alejan de la media más de 1 desviación típica:

$$p(\mu - \sigma \leq X \leq \mu + \sigma) = p\left(\frac{\mu - \sigma - \mu}{\sigma} \leq Z \leq \frac{\mu + \sigma - \mu}{\sigma}\right) = p(-1 \leq Z \leq 1) =$$

$$= p(Z \leq 1) - p(Z \leq -1) = p(Z \leq 1) - [1 - p(Z \leq 1)] =$$

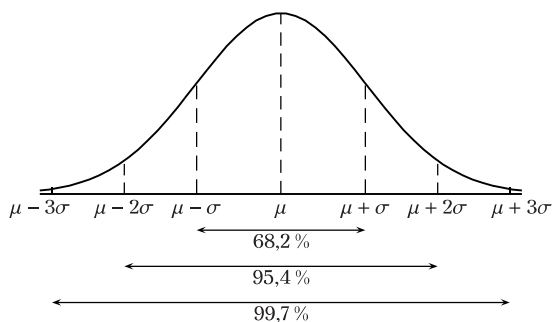
$$= 0,8413 - 1 + 0,8413 = 0,6826 = 68,26\%$$

El 68,26 % de los individuos se encuentran en el intervalo $(\mu - \sigma, \mu + \sigma)$.

De forma similar se puede calcular que:

☞ el 95,4 % de los individuos están en el intervalo $(\mu - 2\sigma, \mu + 2\sigma)$.

☞ el 99,7 % de los individuos se encuentran en el intervalo $(\mu - 3\sigma, \mu + 3\sigma)$.



Observando los porcentajes anteriores, se comprende fácilmente por qué las tablas de la distribución $N(0,1)$ se han construido para valores de la variable menores que 3,59. La probabilidad, para valores superiores a 3,59 (o inferiores a $-3,59$) es prácticamente cero.

10.6.6. Cálculo de percentiles

La tabla normal se usa, generalmente, para calcular la probabilidad según el valor de Z ; pero también puede utilizarse en sentido inverso: conocida la probabilidad, hallar el valor de Z que le corresponde.

Esto es especialmente útil para el cálculo de percentiles (percentil P_α es el valor de la variable que deja por debajo un $\alpha\%$ de los individuos).

Ejemplo 12

Aplicado un test a un grupo de 400 personas, se ha obtenido una distribución normal de media 60 puntos y desviación típica 5 puntos. Halla el percentil 67.

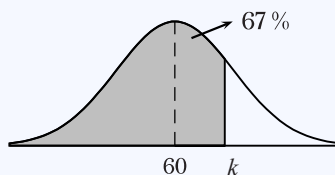
Sea X = puntuación en el test: $X \rightarrow N(60; 5)$

Si al percentil P_{67} lo llamamos k :

$$p(X \leq k) = 0,67 \Rightarrow p\left(Z \leq \frac{k - 60}{5}\right) = 0,67 \Rightarrow$$

$$\frac{k - 60}{5} = 0,4399 \Rightarrow k = 60 + 5 \cdot 0,4399 =$$

$$= 62,1995 \approx 62,2 \Rightarrow P_{67} = 62,2 \text{ puntos}$$



Ejemplo 13

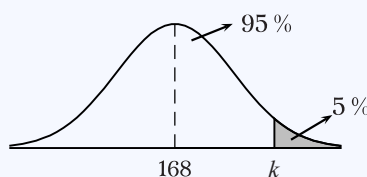
Se sabe que los soldados de un cuartel poseen una estatura que se distribuye normalmente con media 168 cm y desviación típica 8 cm. Si se quiere seleccionar el 5% de los soldados más altos, ¿a partir de qué altura ha de hacerse?

Sea X = estatura de los soldados: $X \rightarrow N(168; 8)$

$$p(X \leq k) = 0,95 \Rightarrow p\left(Z \leq \frac{k - 168}{8}\right) = 0,95 \Rightarrow$$

$$\frac{k - 168}{8} = 1,6449 \Rightarrow k = 168 + 8 \cdot 1,6449 =$$

$$= 181,1592 \approx 181 \text{ cm}$$



Ejemplo 14

Un test dio una puntuación que sigue una ley normal con media 100 y desviación típica 15. Halla el rango intercuartil.

Se trata de hallar un intervalo centrado en la media que contenga al 50 % de la población.

Sea X = puntuación en el test: $X \rightarrow N(100; 15)$

$$p(X \leq k_1) = 0,75 \Rightarrow p\left(Z \leq \frac{k_1 - 100}{15}\right) = 0,75 \Rightarrow \frac{k_1 - 100}{15} = 0,6745 \Rightarrow$$

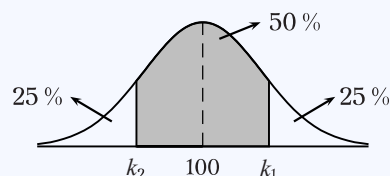
$$k_1 = 100 + 15 \cdot 0,6745 = 110,1175 \simeq 110 \text{ puntos}$$

Por simetría:

$$k_2 = 100 - 15 \cdot 0,6745 = 89,8825 \simeq 90 \text{ puntos.}$$

En el intervalo $[90, 110]$ se encuentra el 50 % central de la población:

$$\text{Rango intercuartil} = 110 - 90 = 20$$

**10.6.7. Aproximación de la binomial por la normal**

(estudiar por apuntes de clase)

10.6.8. Probabilidades normales y calculadora gráfica

Las siguientes instrucciones son válidas para la calculadora CASIO fx-9750G PLUS y similares.

☞ Menú **STAT**.

☞ **F5 (DIST) ⇒ F1 (NORM)**

* Para calcular probabilidades: **F2 (Ncd)**. Introducir los datos del siguiente modo:

$$\left. \begin{array}{l} \text{límite inferior} \\ \text{límite superior} \\ \sigma \\ \mu \end{array} \right\} \Rightarrow \text{F1 (CALC)}$$

* Para calcular percentiles: **F3 (InvN)**. Introducir los datos del siguiente modo:

$$\left. \begin{array}{l} \text{área (probabilidad)} \\ \sigma \\ \mu \end{array} \right\} \Rightarrow \text{F1 (CALC)}$$

Nota: tomar $-10^{99} \simeq -\infty$ y $10^{99} \simeq +\infty$.