

# PROBABILIDAD

## ÍNDICE

1. Sucesos aleatorios
2. Definición de probabilidad
3. Probabilidad condicionada
4. Teorema de la Bayes
5. Variable aleatoria
6. Función de probabilidad
7. Función de distribución
8. Media y varianza
9. Distribución binomial
10. Distribución normal

## SUCESOS ALEATORIOS

**Experimento aleatorio** : es aquel que se caracteriza porque al repetirlo bajo análogas condiciones jamás se puede predecir el resultado que se va a obtener . En caso contrario se llama experimento determinista .

**Espacio muestral E** : ( de un experimento aleatorio ) es el conjunto de todos los resultados posibles del experimento .

**Suceso de un experimento aleatorio** : es un subconjunto del espacio muestral . Puede haber los siguientes tipos :

- suceso elemental
- suceso compuesto ( de varios sucesos elementales )
- suceso seguro
- suceso imposible
- suceso contrario

### Operaciones con sucesos :

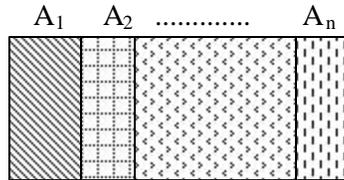
- **Unión de sucesos** : la unión de dos sucesos A y B es el suceso que se realiza cuando se realiza A ó B
- **Intersección de sucesos** : la intersección de A y B es el suceso que se realiza cuando se realizan simultáneamente los sucesos A y B . Cuando es imposible que los sucesos se realicen simultáneamente se dice que son **incompatibles** . Si . . . B . . . entoncesson incompatibles. En caso contrario se dice que son compatibles .

### Propiedades :

	Unión	Intersección
Asociativa	$(A \cup B) \cup C = A \cup (B \cup C)$	$(A \cap B) \cap C = A \cap (B \cap C)$
Conmutativa	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Idempotente	$A \cup A = A$	$A \cap A = A$
Simplificativa	$A \cup (B \cap A) = A$ $A \cap (B \cup A) = A$	
Distributiva	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	
Suceso contrario	$A \cup \bar{A} = E$	$A \cap \bar{A} = \Phi$

**Sistema completo de sucesos :** Se dice que un conjunto de sucesos  $A_1, A_2, \dots$  constituyen un sistema completo cuando se verifica :

- $A_1 \cup A_2 \cup \dots = E$
- $A_1, A_2, \dots$  son incompatibles 2 a 2.



## PROBABILIDAD

**Ley de los grandes números :** La frecuencia relativa de un suceso tiende a estabilizarse en torno a un número, a medida que el número de pruebas del experimento crece indefinidamente. Este número lo llamaremos **probabilidad de un suceso**.

**Definición clásica de probabilidad : (regla de Laplace)**

$$p(A) = \frac{\text{n}^\circ \text{ de casos favorables}}{\text{n}^\circ \text{ de casos posibles}}$$

( para aplicar esta definición se supone que los sucesos elementales son equiprobables )

**Definición axiomática de probabilidad : ( Kolmogorov )** Se llama probabilidad a una ley que asocia a cada suceso A un número real que cumple los siguientes axiomas :

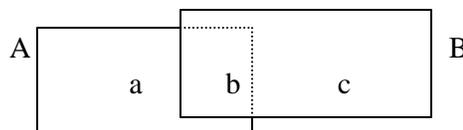
1. La probabilidad de un suceso cualquiera del espacio de sucesos siempre es positiva, es decir  $p(A) \geq 0$
2. La probabilidad del suceso seguro es 1, es decir,  $p(E) = 1$
3. La probabilidad de la unión de sucesos incompatibles es igual a la suma de probabilidades de cada uno de ellos, o sea,  $p(A \cup B) = p(A) + p(B)$

**Consecuencias de los axiomas :**

- $p(\bar{A}) = 1 - P(A)$
- $p(\Phi) = 0$
- $0 \leq p(A) \leq 1$
- Si  $A \subset B$   $p(A) \leq p(B)$
- Si los sucesos son compatibles :  $p(A \cup B) = p(A) + p(B) - p(A \cap B)$   
Para el caso de tres sucesos compatibles sería :  
 $p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C)$

**Probabilidad condicionada  $p(A/B)$  :** Se llama probabilidad del suceso A condicionado por B a la probabilidad de que se cumpla A una vez que se ha verificado el B.

$$p(A/B) = \frac{p(A \cap B)}{p(B)}$$



$$p(A \cap B) = \frac{b}{a+b+c} \quad p(B) = \frac{b+c}{a+b+c} \quad p(A/B) = \frac{b}{b+c}$$

Otra forma de ver la fórmula es :

$$p(A \cap B) = p(B) \cdot p(A/B) = p(A) \cdot p(B/A) = p(B \cap A)$$

Generalizando :  $p(A \cap B \cap C) = p(A) \cdot p(B/A) \cdot p(C/A \cap B)$

Ejemplo :

	Hombres	Mujeres	
Fuman	70	40	110
No Fuman	20	30	50
	90	70	160

$$p(H) = 90/160 \quad p(M) = 70/160 \quad p(F) = 110/160 \quad p(NF) = 50/160$$

$$p(H/NF) = 20/50 \quad p(H/F) = 70/110 \quad p(M/NF) = 30/50 \quad p(M/F) = 40/110$$

$$p(H \cap F) = 70/160 = p(F) \cdot p(H/F) = (110/160) \cdot (70/110)$$

Lo mismo se podría hacer con color de ojos ( marrones y azules ) y color de pelo ( rubio y castaño ) .

**Sucesos independientes :** dos sucesos A y B se dice que son independientes si  $p(A) = p(A/B)$  . En caso contrario ,  $p(A) \neq p(A/B)$  , se dice que son dependientes .

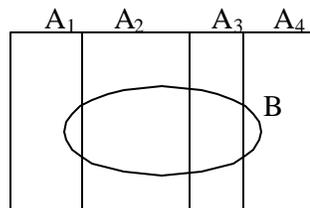
**Probabilidad de la intersección o probabilidad compuesta :**

- Si los sucesos son dependientes  $p(A \cap B) = p(A) \cdot p(B/A) = p(B) \cdot p(A/B)$
- Si los sucesos son independientes  $p(A \cap B) = p(A) \cdot p(B)$

Ejemplo : si al extraer dos cartas de una baraja lo hacemos con devolución tendremos dos sucesos independientes ,  $p(A \cap B) = p(A) \cdot p(B)$  pero si lo hacemos sin devolución ahora si son dependientes  $p(A \cap B) = p(A) \cdot p(B/A)$  .

**Teorema de la probabilidad total :** sea un sistema completo de sucesos y sea un suceso B tal que  $p(B/A_i)$  son conocidas , entonces :

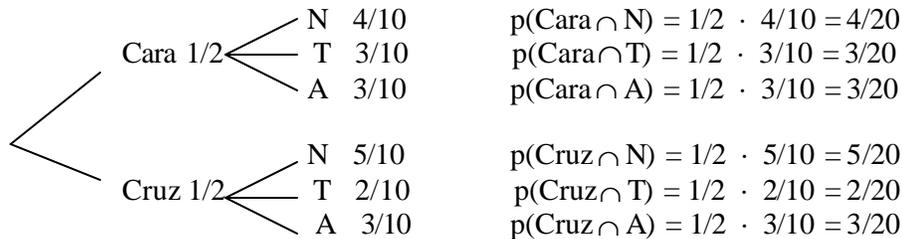
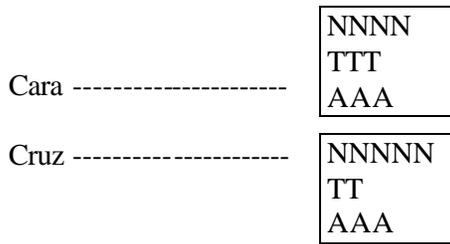
$$p(B) = p(B \cap A_1) + p(B \cap A_2) + \dots = \sum p(B \cap A_i)$$



**Teorema de Bayes :** sea un sistema completo de sucesos y sea un suceso B tal que  $p(B/A_i)$  son conocidas , entonces :

$$p(A_i / B) = \frac{p(A_i) \cdot p(B / A_i)}{\sum p(A_i) \cdot p(B / A_i)}$$

Ejemplo importante : Se va a realizar el siguiente experimento , se tira una moneda , si sale cara se saca una bola de una urna en la que hay 4 bolas negras , 3 turquesa y 3 amarillas , si sale cruz se saca una bola de otra urna en la que hay 5 bolas negras , 2 turquesa y 3 amarillas .



Tª de la probabilidad total :  $p(N) = p(\text{Cara} \cap N) + p(\text{Cruz} \cap N) = 4/20 + 5/20 = 9/20$

Tª de Bayes :  $p(\text{Cara}/N) = \frac{p(\text{Cara} \cap N)}{p(\text{Cara} \cap N) + p(\text{Cruz} \cap N)}$  que no es ni más ni menos que casos favorables entre casos posibles .

### DISTRIBUCIONES DISCRETAS : DISTRIBUCIÓN BINOMIAL

**Variable aleatoria X :** es toda ley que asocia a cada elemento del espacio muestral un número real . Esto permite sustituir los resultados de una prueba o experimento por números y los sucesos por partes del conjunto de los números reales .

Las variables aleatorias pueden ser discretas o continuas .

Por ejemplo en el experimento aleatorio de lanzar tres monedas el espacio muestral es  $E = [ CCC , CCX , CXC , XCC , CXX , XCX , XXC , CCC ]$  . Supongamos que a cada suceso le asignamos un número real igual al número de caras obtenidas . Esta ley o función que acabamos de construir la llamamos variable aleatoria ( discreta ) que representa el nº de caras obtenidas en el lanzamiento de tres monedas .

Consideremos el experimento que consiste en elegir al azar 100 judías de una plantación y medimos su longitud . La ley que asocia a cada judía su longitud es una variable aleatoria ( continua ) .

Por ejemplo al lanzar un dado podemos tener la variable aleatoria  $x_i$  que asocia a cada suceso el nº que tiene en la parte de arriba .

Por ejemplo al lanzar dos dados podemos tener la variable aleatoria  $x_i$  que asocia a cada suceso el producto de los dos números que tiene en la parte de arriba .

**Función de probabilidad :** ( de una variable aleatoria ) es la ley que asocia a cada valor de la variable aleatoria  $x_i$  su probabilidad  $p_i = p( X = x_i )$ .

**Función de distribución F(x) :** ( de una variable aleatoria ) es la ley que asocia a cada valor de la variable aleatoria , la probabilidad acumulada de este valor .

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

**Media de una variable aleatoria discreta :**  $\mu = \sum x_i \cdot p_i$

**Varianza de una variable aleatoria discreta :**  $\sigma^2 = \sum (x_i - \bar{x})^2 \cdot p_i$

Ejemplo : en una bolsa hay bolas numeradas : 9 bolas con un 1 , 5 con un 2 y 6 con un 3 . Sacamos una bola y vemos que número tienen .

La función de probabilidad es :

$x_i$	1	2	3
$p_i$	9/20	5/20	6/20

La función de distribución es :

$x_i$	1	2	3
$p_i$	9/20	14/20	20/20

La media es  $1 \cdot (9/20) + 2 \cdot (5/20) + 3 \cdot (6/20) = 1'85$

La varianza es  $(1-1'85)^2 \cdot 9/20 + (2-1'85)^2 \cdot 5/20 + (3-1'85)^2 \cdot 6/20 = 0'72$

**Distribución binomial :** Una variable aleatoria es binomial si cumple las siguientes características :

1. Los elementos de la población se clasifican en dos categorías , éxito o fracaso .
2. El resultado obtenido en cada prueba es independiente de los resultados anteriores
3. La probabilidad de éxito y fracaso es siempre constante

Ejemplos : fumadores de una población , nº de aprobados de la clase , días de lluvia a lo largo de un año , nº de caras al tirar una moneda , etc .

- **Función de probabilidad**  $p(X = r) = \binom{n}{r} p^r q^{n-r}$  donde p es la probabilidad de éxito , q la probabilidad de fracaso , n el numero total de pruebas y r el número de éxitos .

- **Función de distribución**  $p(X \leq x) = \sum_{r=0}^{r=x} \binom{n}{r} p^r q^{n-r}$

- **Media**  $\mu = n \cdot p$

- **Varianza**  $\sigma^2 = n \cdot p \cdot q$

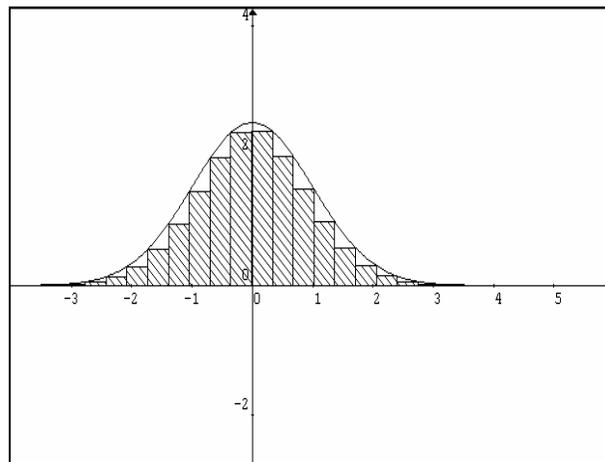
Ejemplo : Se lanza una moneda 11 veces :

- ¿Cuál es la probabilidad de obtener 5 caras ?
- ¿Cuál es la probabilidad de obtener 5 o menos caras ?
- ¿Cuántas caras se obtienen por término medio ?
- ¿Cuál es la desviación típica ?

## DISTRIBUCIONES CONTINUAS : DISTRIBUCIÓN NORMAL

**Función de densidad  $f(x)$**  : cuando en un histograma de frecuencias relativas de una variable continua aumentamos el nº de clases y por lo tanto su amplitud es más pequeña vemos que el polígono de frecuencias relativas se acerca a una función  $f(x)$  que llamaremos función de densidad que cumple las siguientes propiedades :

- $f(x) \geq 0$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$  el área encerrada bajo la curva de la función es igual a la unidad .
- $\int_a^b f(x)dx = p(a \leq X \leq b)$  área bajo la curva correspondiente a ese intervalo .



**Función de distribución  $F(x) = p(X \leq x)$**  : cuando en un histograma de frecuencias relativas acumuladas de una variable continua aumentamos el nº de clases y por lo tanto su amplitud es más pequeña vemos que el polígono de frecuencias relativas acumuladas se acerca a una función  $F(x)$  que llamaremos función de distribución que cumple las siguientes propiedades :

- $F(a) = \int_{-\infty}^a f(x)dx = p(-\infty \leq X \leq a)$  por lo tanto :

$$p(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a)$$

- $F(x)$  es nula para todo valor de  $x$  anterior al menor valor de la variable aleatoria y es igual a la unidad para todo valor posterior al mayor valor de la variable aleatoria . Si es continua se dice que  $F(-\infty)=0$  y  $F(+\infty)=1$
- Por ser una probabilidad  $0 \leq F(x) \leq 1$  .
- Es una función creciente .

**Media de una variable aleatoria continua :**  $\mu = \int_a^b x \cdot f(x) dx$

**Varianza de una variable aleatoria continua :**  $\sigma^2 = \int_a^b (x - \bar{x})^2 f(x) dx$

**Distribución normal :** una variable aleatoria es normal si se rige según las leyes del azar . La mayoría de las distribuciones más importantes son normales . Por ejemplo la distribución de los pesos de los individuos de cualquier especie , la estatura de una población , Tª del mes de agosto a lo largo de 100 años , la longitud de los tornillos que salen de una fábrica , etc .

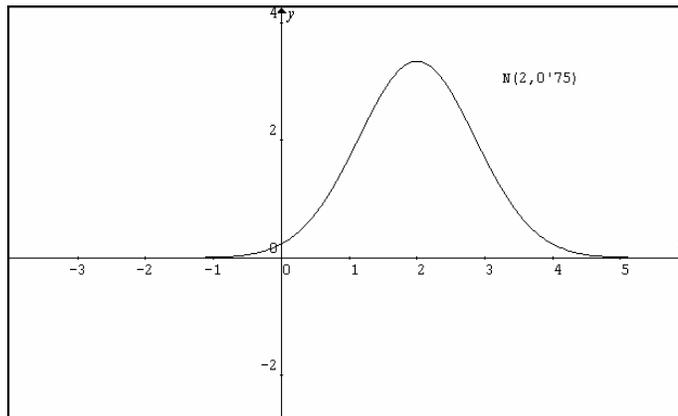
No todas las distribuciones son normales por ejemplo si clasificamos según el nivel de renta a los ciudadanos españoles son muy pocos los que poseen niveles de rentas altas y en cambio son muchos los que poseen niveles de rentas bajas , por tanto la distribución no sería simétrica y en consecuencia no se adapta al modelo normal .

**Función de densidad :** una variable continua X sigue una distribución normal de media  $\mu$  y desviación típica  $\sigma$  , y se designa por  $N(\mu, \sigma)$  , si cumple que

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

Podríamos comprobar que :

$$\int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \mu \qquad \int_{-\infty}^{+\infty} (x - \mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \sigma^2$$



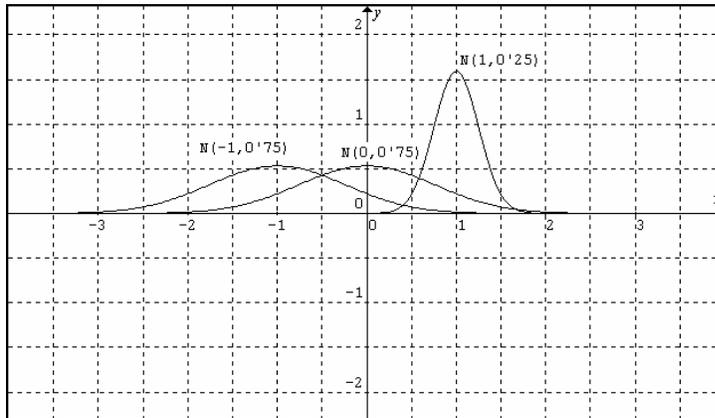
Para calcular los máximos y mínimos deberíamos hacer :

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

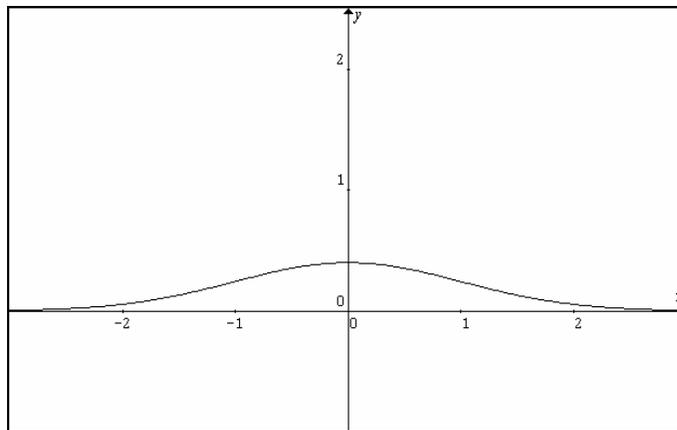
$f'(x) = -\frac{x-\mu}{\sigma} f(x)$  , puesto que  $f(x)$  nunca puede valer 0 entonces , si  $x = \mu$   $f'(x) = 0$  por lo que será un posible máximo o mínimo .

$f''(x) = -\frac{1}{\sigma^2} \left[ 1 - \left( \frac{x-\mu}{\sigma} \right)^2 \right] f(x)$  luego  $f''(\mu) < 0$  por lo que es hay un máximo en el punto  $\left( \mu, \frac{1}{\sigma\sqrt{2\pi}} \right)$

Conviene observar que cuando la desviación típica es elevada aumenta la dispersión y se hace menos puntiaguda la función ya que disminuye la altura del máximo . Por el contrario para valores pequeños de  $\sigma$  obtenemos una gráfica menos abierta y más alta .



Cuando  $\mu=0$  y  $\sigma=1$  ,  $N(0,1)$  se dice que tenemos una distribución normal reducida , estandar o simplificada .



**Función de distribución :**  $F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = p(X \leq x)$

**Distribución Normal Estándar  $N(0,1)$  :** La distribución  $N(0,1)$  se encuentra tabulada , lo cual permite un cálculo rápido de las probabilidades asociadas a esta distribución . Pero en general la media no suele ser 0 , ni la varianza 1 , por lo que se hace una transformación que se llama **tipificación de la variable** , que consiste en hacer el siguiente cambio de variable :

$$Z = \frac{x - \mu}{\sigma}$$

a partir del cual obtenemos una variable Z que si es N(0,1) y que por lo tanto podemos calcular sus probabilidades .

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} \sigma \cdot dz$$

Ejemplo : si tenemos N(2,4) y queremos calcular p(x<7) entonces :

$$p(x<7) = p\left(\frac{x-2}{4} < \frac{7-2}{4}\right) = p(z < 5/4) = 0'1056$$

**Manejo de tablas :** pueden presentarse los siguientes casos :

$$p(z<1'45) = 0'9265$$

$$p(z<-1'45) = 0'0735$$

$$p(1'25<z<2'57) = 0'1005$$

$$p(-2'57<z<-1'25) = 0'1005$$

$$p(-0'53<z<2'46) = 0'695$$

**Utilización conjunta de  $\mu$  y  $\sigma$  :**

En  $(\mu \pm \sigma)$  está el 68'26% de los datos ya que :

$$p(\mu - \sigma < X < \mu + \sigma) = p\left(\frac{\mu - \sigma - \mu}{\sigma} < Z < \frac{\mu + \sigma - \mu}{\sigma}\right) = p(-1 < Z < 1) = 0.6826$$

Análogamente se puede comprobar que en  $(\mu \pm 2\sigma)$  está el 95'4% de los datos y en  $(\mu \pm 3\sigma)$  está el 99'7% .

Ejemplo : El C.I. de los 5600 alumnos de una provincia se distribuyen N(112,6) .

Calcular aproximadamente cuántos de ellos tienen :

- a) más de 112 .....2800 alumnos.....la mitad de los alumnos
- b) entre 106 y 118 .....3823 alumnos .....este es el caso :  $(\mu \pm \sigma)$
- c) entre 106 y 112 .....1911 alumnos
- d) menos de 100 .....128 alumnos
- e) más de 130 .....7 alumnos
- f) entre 118 y 124 .....761 alumnos

( ojo hay que multiplicar % obtenido en la tabla por 5600/100 , que sale de una regla de tres )

**Aproximación normal para la binomial :**

Cuando los valores a calcular para la binomial superan a los de las tablas para obtener un resultado aproximado se utiliza la distribución normal , es decir , la variable

$$y = \frac{x - np}{\sqrt{npq}} \text{ obedece a una distribución } N(0,1)$$

El resultado es tanto más fiable cuanto mayor es el tamaño de la muestra n y cuanto más cerca está p de 0'5 .

Ejemplo : Se ha comprobado que la probabilidad de tener un individuo los ojo marrones es 0'6 . Sea X la variable aleatoria que representa el n° de individuos que tienen los ojos marrones de un grupo de 1100 . Calcular p(X>680) y p(X=680)

$$p(X>680) = 1 - p(X<680) = 1 - p\left(Y < \frac{680 - 110 \cdot 0'6}{\sqrt{1100 \cdot 0'6 \cdot 0'4}}\right) = 1 - p(Y < 1'23) = 0'1093$$

p(X = 680) = p(679'5 < X < 680'5) se debe hacer así puesto que en una variable continua no tiene sentido calcular probabilidades de valores puntuales .