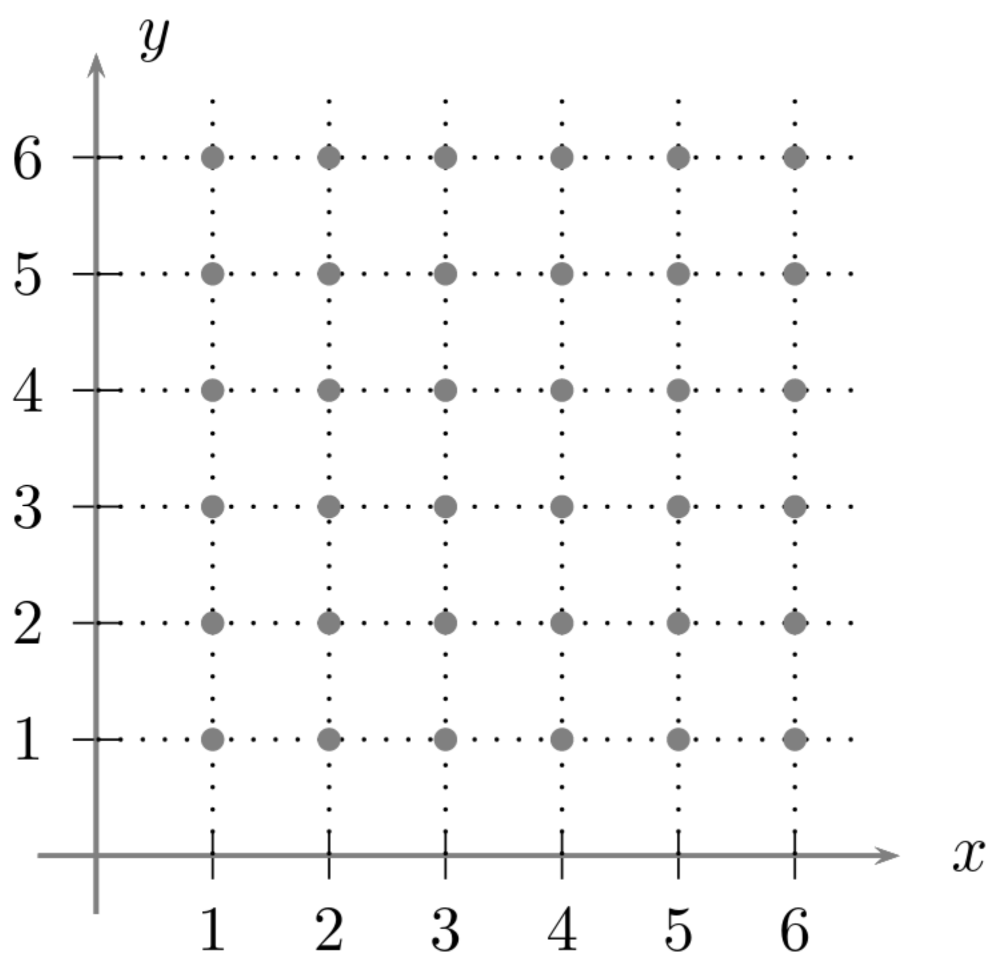
BABY NOTAS DE PROBABILIDAD

JORGE DÁVILA

August 31, 2022

1. Variables aleatorías, Eventos y Probabilidad

Experimento 1: Consideremos el experimento aleatorio que consiste en lanzar dos dados equilibrados. Considerando el conjunto de resultados de dicho experimento Ω como la siguiente figura. Calcular la probabilidad de los siguientes eventos.



a) $A = \{(x,y) : x \leq y\}$.

b) $B = \{(x,y) : x + y = 6\}$.

c) $C = \{(x,y) : x + y \leq 4\}$.

d) $D = \{(x,y) : x - y = 3\}$

Para mentes curiosas

e) $E = \{(x,y) : \max\{x,y\} = 3\}$.

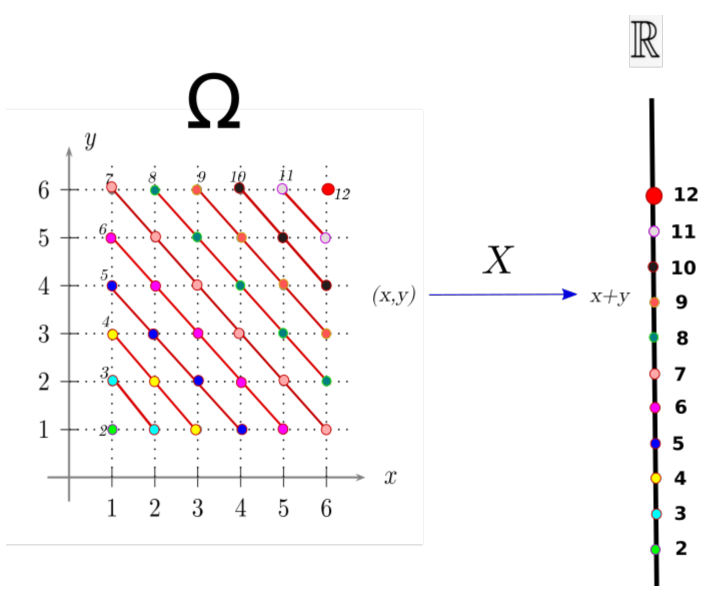
f) $F = \{(x,y) : \max\{x,y\} \leq 3\}$.

Notemos que los anteriores eventos, se pueden ver como imagenes inversas de una variables aleatoria, veamos esto.

Definamos $X : \Omega \to \mathbb{R}$ por

$$X(x,y) = x + y.$$

En este caso se tiene que $B = X^{-1}(6)$ y $C = X^{-1}\big([4,\infty)\big)$. Sigamos aprendiendo más sobre esta varieble aleatoria. Se tiene que el conjunto de valores que toma X son $\{2,3,4,\ldots,12\}$.¿Como se ven las imagenes inversas de estos valores?



Veamos un gráfica de frecuencia de los valores de X

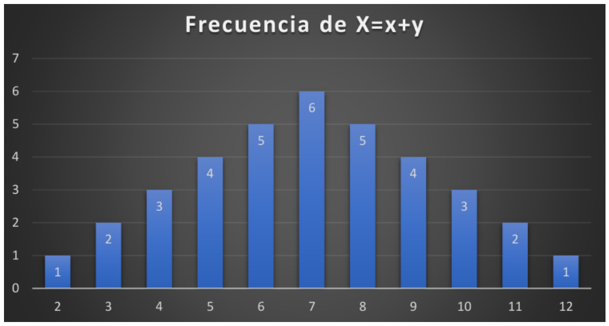


FIGURE 1. gráfica de frecuencias X=x+y

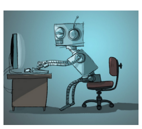
Notemos que

$$P_c(B) = P_c(X^{-1}(6)) = \frac{\#B}{\#\Omega} = \frac{\text{No. de veces que X=6}}{\#\Omega} = \frac{5}{36}.$$

La ideas clave esta en que la probabilidad de B, se determina a partir de la tabla de frecuencias de la variable aleatoria X.

$$P_c(B) = \frac{\text{No. de veces que X=6}}{\#\Omega}. \tag{1.1}$$

La fracción anterior se llama **frecuencia relativa**.



Notación:

$$P(X = a) := P(X^{-1}(a)).$$

$$P(X \leq a) := P(X^{-1}([a, \infty)).$$

$$P(a \leq X < b) := P(X^{-1}([a, b)).$$

Para el caso del evento C tenemos $P_c(C) = \dfrac{\text{No. de veces que } X \leq 4}{\#\Omega}$

$C = \{(x, y) : x + y = 2\} \cup \{(x, y) : x + y = 3\} \cup \{(x, y) : x + y = 4\}$

Por lo tanto

$$P_c(C) = \frac{\text{No. de veces que X=2}}{\#\Omega} + \frac{\text{No. de veces que X=3}}{\#\Omega} + \frac{\text{No. de veces que X=4}}{\#\Omega}$$

Usando la notacion anterior

$$P_c(C) = P(X \leq 4)) = P(X = 2) + P(X = 3) + P(X = 4) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36}.$$

1.1. **Variables aleatorias & espacios de probabilidad**

Dado un espacio de probabilidad $(\mathcal{A}, \Omega, P)$ podemos pensar a la variable aleatoria definida por

$$X : (\mathcal{A}, \Omega) \to (\mathcal{B}(\mathbb{R}), \mathbb{R})$$

Lo anterior quiere decir que X es una función $X : \Omega \to \mathbb{R}$ que además cumple para todo $B \in \mathcal{B}(\mathbb{R})$ se tiene $X^{-1}(B) \in \mathcal{A}$, que no es otra cosa que nuestra definición de variable aleatoria.

Donde $\mathcal{B}(\mathbb{R})$ es llamada la σ-álgebra de Borel, y la vamos a imaginar compuesta por todos los intervalos de $\mathbb{R}$, por cuestiones de simplicidad.

La importancia de una variable aletoria, reducción de la complejidad de $(\mathcal{A}, \Omega, P)$

Podemos decir que una variable aleatoria X es una herramienta que nos permite hacer dos cosas muy importantes.

Lo primero que nos permite hacer una variable aleatoria X es **reducir el conjunto de eventos**, hacemos esto considerando solo a los eventos $E \in \mathcal{A} \subseteq \mathcal{P}(\Omega)$ que son imagenes inversas de dicha variable, es decir, $E = X^{-1}(B)$ para algún elemento $B \in \mathcal{B}(\mathbb{R})$. Desde un punto de vista práctico estos son los eventos en que estamos interesados cuando analizamos una característica de una población dada por la variable aleatoria X. Recordemos que $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ puede ser muy complejo.

Lo Segundo que nos permite hacer la variable aleatoria Xes definir una probabildad en $(\mathcal{B}(\mathbb{R}), \mathbb{R})$, que denotaremos por P_X de la siguiente forma, esto es básicamente la idea clave 1.1

$$P_X(B) := P(X^{-1}(B)) \text{ Para todo } B \in \mathcal{B}(\mathbb{R}). \tag{1.2}$$

 Notemos que lo anterior transforma nuestro espacio de probabilidad original $(\mathcal{A}, \Omega, P)$ en el epacio de probabilidad $(\mathcal{B}(\mathbb{R}), \mathbb{R}, P_X)$, gracias a nuestra variable aleatoria X.

"Orden en el Caos"

$$(\mathcal{A}, \Omega, P) \rightsquigarrow (\mathcal{B}(\mathbb{R}), \mathbb{R}, P_X).$$

En resumen: Dada la variable aleatoria X, podemos olvidarnos del espacio de probabilidad $(\mathcal{A}, \Omega, P)$, y trabajar sobre el espacio $(\mathcal{B}(\mathbb{R}), \mathbb{R}, P_X)$, mucho más facil de trabajar he imaginar.

Regresando al problema de construir espacios de probabilidad para fenómenos aleatorio: Hemos visto que en general es difícil construir la función de probabilidad P, es decir, tenemos $(\mathcal{A}, \Omega, ?)$. ¿Que ocurre si tenemos una variable aleatoria X y un espacio de probabilidad $(\mathcal{B}(\mathbb{R}), \mathbb{R}, \widetilde{P})$?, ¿Nos permite X definir una probabilidad en $(\mathcal{A}, \Omega)$?.

Sea $\mathcal{A}' = \{E \in \mathcal{A} : E = X^{-1}(B) \text{ para algún } B \in \mathcal{B}(\mathbb{R})\}$ y definamos la probabilidad en $(\mathcal{A}' \subseteq \mathcal{A}, \Omega)$ por

$$P(A) := \widetilde{P}(B).$$

2.

Funciones de Distribución de probabilidad

Def: Dada una variable aleatoria X en $(\mathcal{A}, \Omega, P)$, decimos que es:

1) **Discreta**, si X toma un conjunto finito o númerable de valores.

2) **Continua** cuando toma cualquier valor de un intervalor de la forma $(a, b), (a, \infty), (-\infty, a), (-\infty, \infty)$.

Consideremos en primer lugar el caso donde la variable aleatoria X es discreta y su conjunto de resultados $S = \{x_1, x_1, \dots, x_n, \dots\}$, es finito o infinito númerable.

Def: Sea X variable aleatoria discrete. Una **función de distribución de probabilidad** asociada a X, es una función $f_X : S \to \mathbb{R}$ que cumple:

1)

$$f_X(x) > 0 \text{ para todo } x \in S.$$

2)

$$\sum_{i=1}^{n,\infty} f_X(x_i) = 1$$

si S es finito o infinito.

Ejemplo básico: Consideremos el caso de lanzar dos dados y la variable aleatoria $X(x, y) = x + y$. Tenemos en este caso que $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ y definamos la siguiente función

(2.1)

$$f_X(x) := P(X^{-1}(\{x_i\})) = P_X(\{x_i\}) = \frac{\text{No. de veces que } X = x_i}{\#\Omega} \text{ para toda } x_i \in S$$

Asi, f_X no es otra cosa que la frecuencia relativa de los valores en S. Claramente tenemos $f(x_i) > 0$ y además se tiene que $\sum_{i=1}^{12} f_X(x_i) = 1$, ver fig. 1. Por lo tanto, la función de frecuencia no es otra cosa que una función de distribución de probabilidad.

Llegamos a uno de los puntos más importantes de todo este hermoso viaje

La igualdad anterior (2.1) nos guarda muchos regalos. ¿Que relación hay entre f_X y P_X?, ¡bueno!, en algún sentido son lo mismo, veamos esto.

Definamos primero la llamada **función indicadora** de un conjunto A. Dado $A \subseteq \mathbb{R}$ se define su función indicadora de la siguiente forma

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Dado $A \subset \mathbb{R}$ y una función de distribución f_X se tiene

$$P_X(A) = \sum_{x_i \in S} 1_A(x_i) f_X(x_i) \tag{2.2}$$

Segunda ideas clave: Tenemos que f_X define a P_X y viceversa. Por esta raźon trabajamos de ahora en adelante con funciones de distribución de probabilidad de variables aleatoria.

Ejemplo: Nuestro ya conocido experimento. Consideremos un experimento de lanzar dos dados

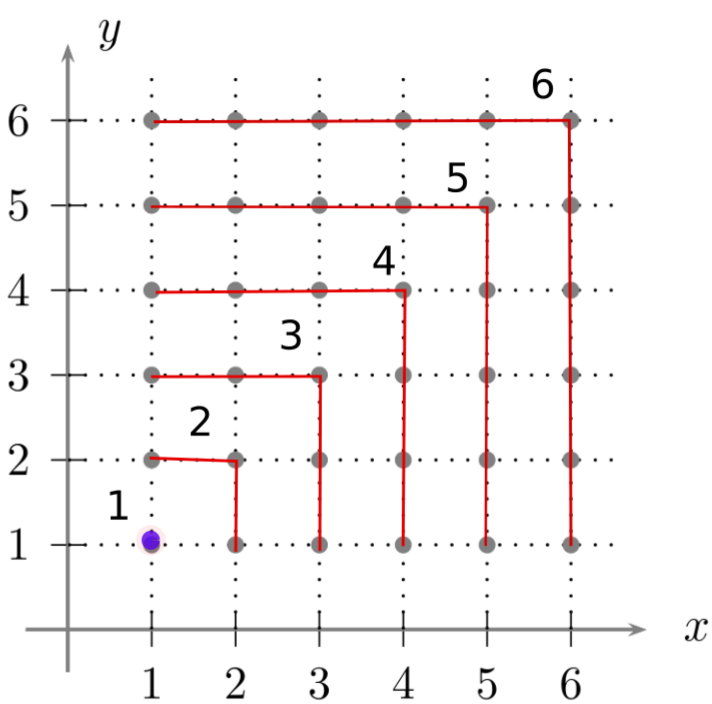
1) Espacio de resultados $\Omega = \{(x, y) : x, y = 1, 2, 3, 4, 5, 6\}$.

2) σ-álgebra $\mathcal{A} = \mathcal{P}(\Omega)$.

3) $P = P_c$, probabilidad Clásica.

4) Variable aleatoria discreta $X : \Omega \to \mathbb{R}$ con $X(x, y) = \max\{x, y\}$.

Los valores que toma X son $S = \{1, 2, 3, 4, 5, 6\}$.



Definamos la función de distribución de probabilidad f_X por

$f_X(1) = P_X(X=1) = P(X^{-1}(1)) = P(\{(1,1)\}) = \dfrac{1}{36}$.

$f_X(2) = P_X(X=2) = P(X^{-1}(2)) = P(\{(1,2),(2,1),(2,2)\}) = \dfrac{3}{36}$.

$f_X(3) = P_X(X=3) = P(X^{-1}(3)) = P(\{(1,3),(3,1),(2,3),(3,2),(3,3)\}) = \dfrac{5}{36}$.

$f_X(4) = P_X(X=4) = P(\{(1,4),(4,1),(2,4),(4,2),(3,4),(4,3),(4,4)\}) = \dfrac{7}{36}$.

De forma similar se tiene $f_X(5) = P(X=5) = \dfrac{9}{36}$ y $f_X(6) = P(X=6) = \dfrac{11}{36}$. Veamos su tabla de frecuencias.

X	1	2	3	4	5	6
f_X	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Table 1. Tabla de f_X

Notemos que f_X se define a partir de P, veamos lo inverso, dada f_X definamos P_X. Para ello consideremos una variable aleatoria X cuyos valores fueron $S = \{2,3,11\}$ y función de distribución dada por la siguiente tabla.

X	2	3	11
f_X	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Definamos $P(X=2) = f_X(2), P(X=3) = f_X(3)$ y $P(X=11) = f_X(11)$. Si tememos $X^{-1}(0,-\infty), X^{-1}[2,-\infty)$ y $X^{-1}(5,-\infty)$, ¿Cuáles son sus probabilidad?. Tenemos que $X^{-1}(0,-\infty) = \emptyset, X^{-1}[2,-\infty) = \{2\}$ y $X^{-1}(5,-\infty) = \{2,3\}$.

Otra manera de definir f_X sería

X	2	3	11
f_X	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Si tenemos el intervalo $(0,5)$ ¿Cuál sería su probabilidad ?.

Usemos nuestra notación de función indicadora.

$$P_X(0,5) = \sum_{x\in\{2,3,11\}} 1_{(0,5)}(x) f_X(x) = 1_{(0,5)}(2) f_X(2) + 1_{(0,5)}(3) f_X(3) + 1_{(0,5)}(11) f_X(11)$$

$$1_{(0,5)}(2) f_X(2) + 1_{(0,5)}(3) f_X(3) = f_X(2) + f_X(3) = \frac{1}{4} + \frac{1}{2}.$$

Otro ejemplo. Si tenemos el intervalo $(3,12]$, entonces su probabilidad es

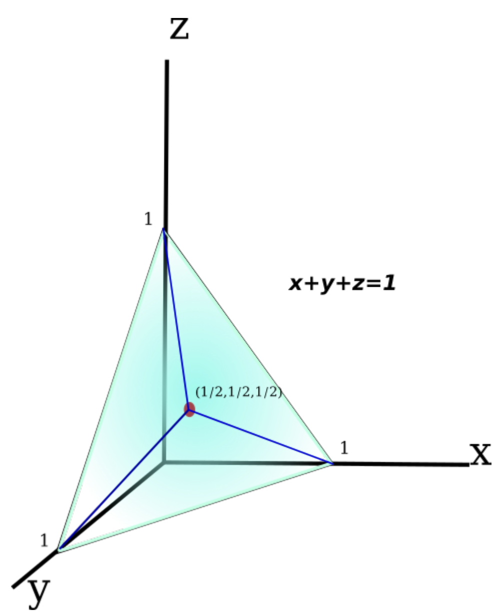
$$P_X(3,12] = \sum_{x\in\{2,3,11\}} 1_{(3,12]}(x) f_X(x) = 1_{(3,12]}(2)f_X(2)+1_{(3,12]}(3)f_X(3)+1_{(3,12]}(11)f_X(11)$$

$$= 1_{(3,12]}(11)f_X(11) = f_X(11) = \frac{1}{4}.$$

El poder definir varias funciones de distribución de probabilidad f_X ya lo habíamos notado al considerar tener tres números $x, y, z > 0$ tales que $x + y + z = 1$ y definir en general la f_X por

X	2	3	11
f_X	x	y	z

Cualquier punto dentro del triángulo de la figura siguiente, nos define una función de distribución de probabilidad f_X.



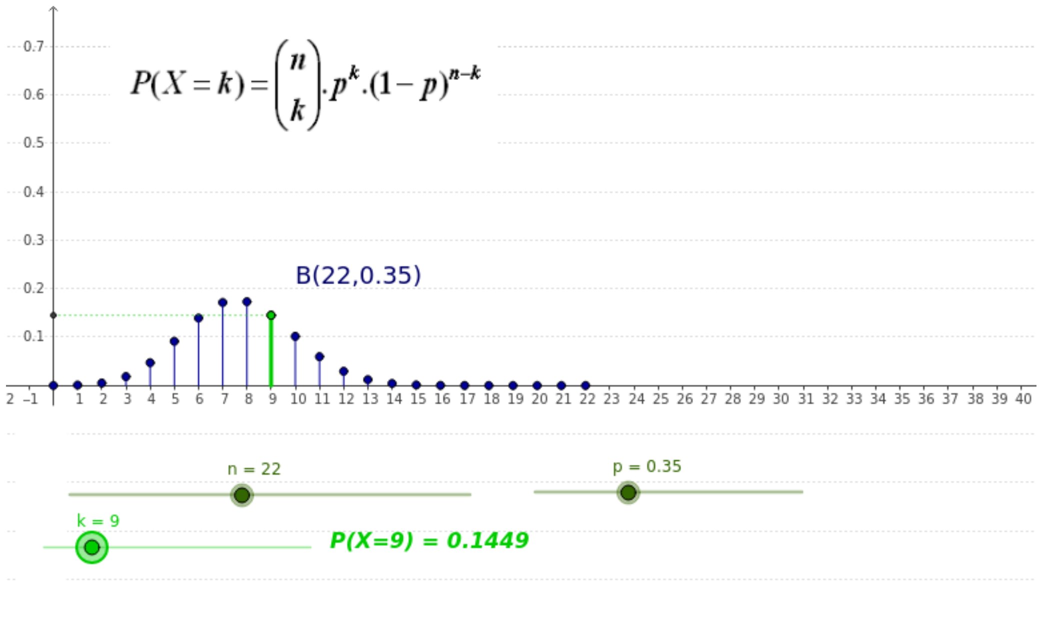
Ejemplo básico: Función de densidad Binomial. Sea $S = \{0, 1, 2, \ldots, n\}$ y $p \in [0, 1]$. Definamos

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k \in S.$$

Notemos que f depende de los parámetros n y $,p$. Se dice que n es el número de veces que se realiza un experiento con dos posibles resultados y p es la probabilidad de que salga uno de esos resultados.

Se dice que una variable aleatoria X se distribuye de forma binomial si su funcion de densidad es f y son llamdos experimentos binomiales donde se quiere saber la probabilidad de obtener x exitos en un experiemento que se repite n veces ($k \leq n$).

Un link para ver la gráfica de la distribucion binomial
`https://www.geogebra.org/m/MfxBWg5K`



número de experimentos n=22, número de exitos k=9

El viaje por Ω finito es muy interesante, ¿Que pasa en los casos donde $\#\Omega = \#\mathbb{N} = \aleph_0$ o $\#\Omega = \#\mathbb{R} = \mathfrak{c}$?.

Vayamos el caso donde $\#\Omega = \#\mathbb{R} = \mathfrak{c}$, que sera nuestro caso para datos continuos y variables aleatorias continuas.

Experimento 3: Consideremos un experimento donde se toman la temperatura de una placa rectangular y se tienen los siguiente datos:

1) Espacio de resultados $\Omega = \{(x, y) : x, y = [1, 6]\}$.

2) σ-álgebra $\mathcal{A} \subseteq \mathcal{P}(\Omega)$.

3) $P = P_g$, probabilidad Geométrica.

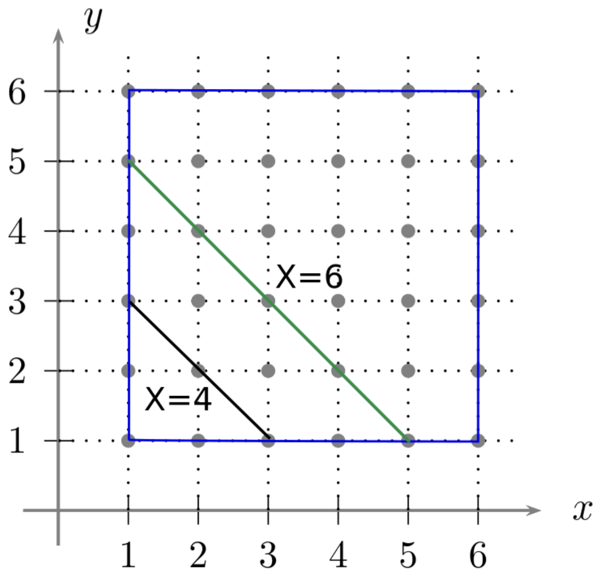
4) Variable aleatoria continua $X : \Omega \to \mathbb{R}$ con $X(x, y) = x + y$.

Los valores que toma X están en el intervalo $[2, 12]$.

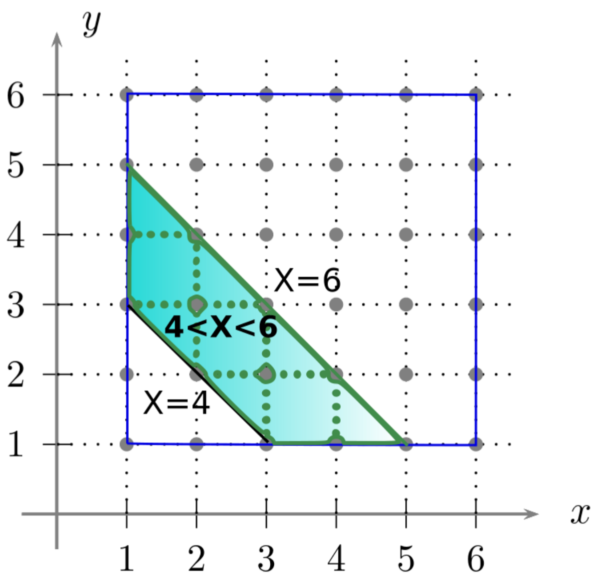
¿Que diferencia hay entre el Experimento 1 y el 2?. Tomemos el evento $B = \{(x, y) : x + y = 6\}$., en este caso se tiene

$$P(B) = P_g(B) = P(X^{-1}(6)) = P(X = 6) = \frac{Area(B)}{Area(\Omega)} = \frac{0}{25} = 0.$$

Veamos la siguiente figura, nos ayudaremos de la figura del Experimento 1, para ver el cambio de lo discreto a lo continuo. Tenemos que B es línea verde, y como no tiene área su probabilidad es cero. En el caso de tener variables aleatorias continuas es común considerar eventos de la forma $B = X^{-1}(a, b)$ y considerar sus probabilidades, $P(B) = (X^{-1}(a, b)) = P(a < X < b)$.



Consideremos el evento $C = \{(x, y) : 4 < x + y < 6\} = X^{-1}((4, 6))$, tenemos que su probabilidad es el área de la figura siguiente , dividida entre el área del rectángulo



$$P_g(C) = \frac{\frac{4*4}{2}}{25} - \frac{\frac{2*2}{2}}{25}.$$

2.1. funciones de distribución de probabilidad para X continuas

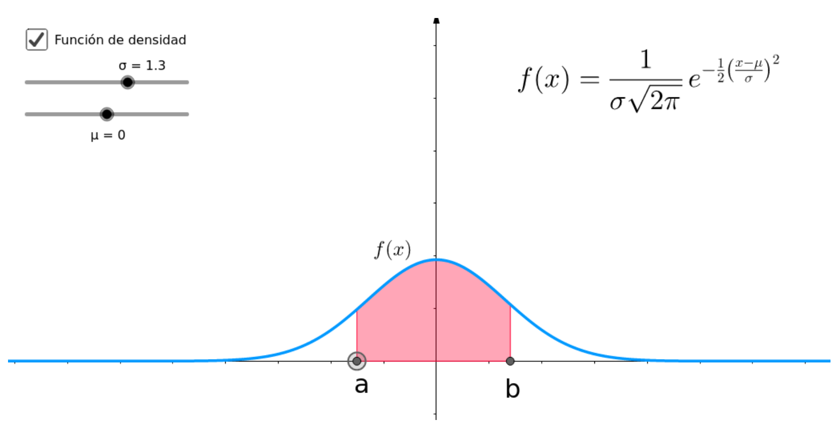
Def: Sea X variable aleatoria continua. Una **función de distribución de probabilidad** asociada a X, es una función $f_X : \mathbb{R} \to \mathbb{R}$ que cumple:

1)
$$f_X(x) \geq 0 \text{ para todo } x \in \mathbb{R}.$$

2)
$$\int_{\mathbb{R}} f_X(x)dx = 1$$
.

Ejemplos:
1) $1_{[0,1]} : \mathbb{R} \to \mathbb{R}$, función indicadora.

Tal vez el ejemplo más importante, la distribución normal con media μ y desviación estandar σ.
2)
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$
.
Veamos su gráfica para $\mu = 0$ y $\sigma = 1.3$



link: `https://www.geogebra.org/m/v38xjmfz`

La idea al igual que el caso discrerto es construir una función de probabilidad dada la variable aletoria continua X y una función de distribución de probabilidad f_X.

$$P_X(B) =:= \int_{-\infty}^{\infty} 1_B f_X(x)dx \text{ para todo } B \in \mathcal{B}(\mathbb{R}).$$

Si $B = (a, b)$ tenemos

$$P(a < X < b) := P_X(B) = \int_{-\infty}^{\infty} 1_{(a,b)} f_X(x)dx = \int_a^b f_X(x)dx$$

References

Email address: `jorge.davila@cimat.mx`