¿Qué es una distribución?

Hemos hecho algunos ejercicios para tener una noción de como trabajar con el calculo de probabilidades. Sin embargo hay otra area de las matemáticas que necesitamos usar en probabilidad y esl el **cálculo** y dentro de el; una de los primeros conceptos que se ve es el de **función**.

Precisamente el concepto de función que trabajamos en calculo diferencial, integral o multivariable (vectorial). Nos conduce a que en el area de probabilidades tiene que usarse esa misma palabra **Función** en el mismo sentido matemático y aquí lo que llamamos es que a la función que toma una variable aleatoria y le asigna sus probabilidades, la palabra se le llama **Distribución**.

Es decir; **una distribución de probabilidad** es una función en el sentido matemático del calculo, que recoge una variable aleatoria y a cada uno de sus posibles estados dentro del espacio muestral le asigna una probabilidad

$$\text{X aleatoria} \rightarrow \underbrace{P(X=x)}_{\text{probabilidad de ocurrencia}}$$

$$P = f(X)$$

Entonces recordemos el ejemplo del dado; es un espacio muestral que tiene 6 posibles estados y cada uno de esos estados se le asigna un numero que es $\frac{1}{6}$. En este caso la distribución seria una función constante, que a todos los estados les asigna el mismo valor, solo que en este caso es una función que la llamamos **discreta**.

En general diremos que "x" y "X" es una variable aleatoria donde **P** va a ser la función que a cada una de las ocurrencias o valores posibles de esta variable aleatoria, le va asignar un numero que le llamamos **la probabilidad**, de esta manera entendemos que **P()** es una función que depende de la variable aleatoria y escribimos $P = f(x)$.

Hay una convención que se usa mucho en probabilidad y es que las letras mayúsculas se usan para denotar las variables, mientras que las letras minúsculas se usan para denotar los posibles valores que estas variables aleatorias pueden tomar.

$$X \rightarrow \text{variable aleatoria}$$

$$x \rightarrow \text{valores posibles en el espacio muestran}$$

Como también sucede en calculo las funciones tienen dominios. El dominio viene por todos los valores posibles de la variable aleatoria sobre los cuales la función puede ser calculada.

Estos dominios pueden definir conjuntos:

- discretos
- continuos

Definiendo así funciones que llamamos **discretas** o **continuas**.

Ejemplos de distribuciones **discretas:**

- Lanzar un dado: Poseen números discretos.

Variables aleatorias continuas:

- Temperatura

Entonces decimos que X $\rightarrow$ variable aleatoria

$$X \rightarrow P(X) \rightarrow \text{distribución de probabilidad} = \text{densidad de probabilidad}$$

P(X) puede tener:

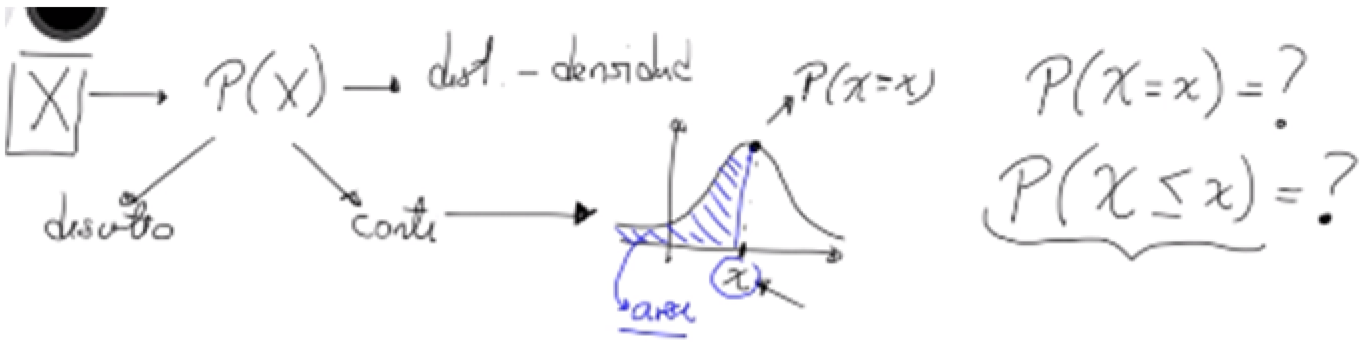
- Forma discreta
- Forma continua

Esto esta delimitado por X; es decir los valores que pueda tomar la variable aleatoria.

Como toda función se puede graficar y ademas se puede realizar operaciones en ellas, como derivación e integración.

¿Qué significa la **Integral de una distribución**? Resulta que existe un concepto que; puedo preguntar ¿cual es la probabilidad de que mi variable tenga un valor particular?, y eso lo hago con la **densidad**, pero también podría preguntar ¿cual es la probabilidad de que mi variable aleatoria tome valores que sean menores o iguales que un valor especifico dado? ¿c´mo se calcula esto?

Bueno recordando los conceptos de calculo integral sabemos que la integral serian los valores debajo de la curva

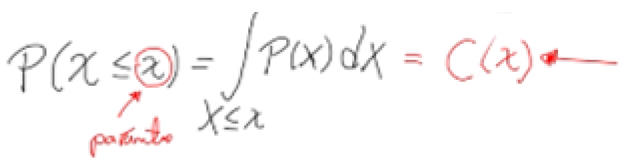


Entonces yo podría decir **¿cual es la probabilidad de que mi variable aleatoria tome valores que sean menores o iguales que un valor especifico? es**

$$P(X \leq x) = \int_{X \leq x} P(X) dX$$

Esta es una integral y es un area bajo la curva y representa también una probabilidad.

En general decimos que cuando "x" (equis minúscula) mo es un valor numérico sino que es un valor cualquiera dado; es decir puede considerarse tipo **parámetro**, entonces toda la integral determina una nueva función que llamamos **La distribución acumulada** C(x).

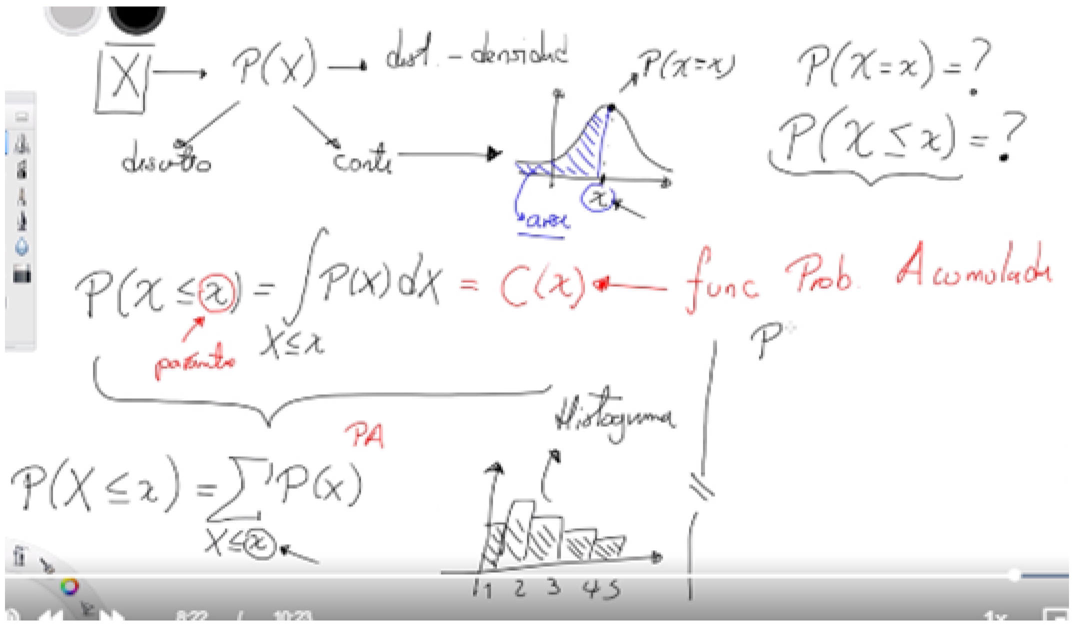


Entonces la distribución acumulada representa la probabilidad de que mi variable aleatoria tome valores menores o iguales que esa X (equis mayúscula) dada. Así la llamaremos **función de probabilidad acumulada**. Entonces la función de probabilidad acumulada es la integral de la función densidad de probabilidad y sirve para responder ese tipo de preguntas donde no estoy preguntando por valores particulares de la probabilidad. Por ejemplo; ¿cual es la probabilidad de que mi dado caiga en 2? sino ¿cual es la probabilidad de que al tirar un dado el resultado sea numero igual o menor que 2?. Ahí esta la diferencia.

Claro este mismo razonamiento funciona para variables discretas:
Solo que en el caso de funciones discretas la gráfica ya no está dada por una curva suave sino que estaría mas dado por un histograma. Entonces esto es que la **Probabilidad de cada evento** es la frecuencia con que ocurre ese evento representado como una barra dentro del histograma.

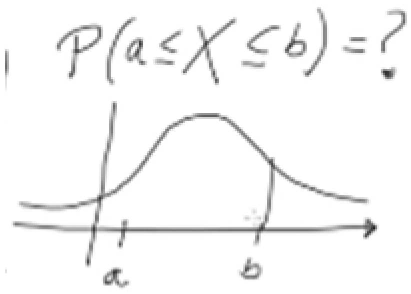
Entonces en este sentido, cuando yo quiero calcular la **probabilidad acumulada** de una función discreta y a su vez quiero responder preguntas iguales a este tipo de preguntas de; ¿cual es la probabilidad de que mi variable tome valores iguales o menores que un cierto valor? Ya no se hacen con integrales sino que se hacen con sumas discretas.

$$P(X \leq x) = \sum_{X \leq x} P(X) = \text{Probabilidad acumulada para funciones discretas}$$



También podemos tener la pregunta de:

¿Cual es la probabilidad de que mi variable tome valores entre 2 valores umbrales? $P(a \leq X \leq b)$ Consideremos el caso de una variable continua en el que tengo un intervalo con a y b.



Entonces gráficamente una probabilidad de ese estilo vendría dada por un area bajo la curva. ¿cómo podríamos escribir una expresión matemática para eso?

$$P(a \leq X \leq b) = \int_a^b P(X)dX = p(b) - p(a) = C(X)$$