Teoría de la información

Funciones

Las funciones tienen datos de entrada y datos de salida. Entre la entrada y la salida existe una transformación, que es lo que llamamos "función". La función $f(x) = y$ se denota por

$$f\colon X \mapsto Y$$

la función f mapea valores de X (input) a Y (output)

Por ejemplo, la función $y = 5x + 2$ nos dice que por cada valor de x, y valdrá 5 veces x más 2. Si $x = (1, 2, 3)$, la función transforma a x en la colección de valores

$$y = (1 * 5 + 2, 2 * 5 + 2, 3 * 5 + 2)$$

En una tabla es más fácil de visualizar

Operador Σ

Σ - Suma de i elementos hasta n, donde i el índice de sumación

$$\sum_{i=m}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

Donde m es el límite inferior de la suma, y n el límite superior de la suma.

Ejemplo: sumar los valores de x del 1 al 4

$$x = \{1_{[1]}, 8_{[2]}, 3_{[3]}, 5_{[4]}\}$$

$$\sum_{i=2}^{4} = 8 + 3 + 5 = 10$$

En corchete coloco *el orden* (el índice) del elemento

Conjuntos (brevísima)

Un conjunto es una colección de distintos objetos. A es un subconjunto de B si todo elemento de A es también incluido en B, que se simboliza como $A \subset B$.

- El conjunto vacío, denotado $\emptyset$, es el conjunto que no contiene nada.
- Denotamos por $|S|$ a la cardinalidad (maomeno *el número de sus elementos*) de S. Por ejemplo, $|\emptyset| = 0$.
- $x \in A$ es "x es un miembro del conjunto A". En probabilidad, los eventos se tratan como conjuntos de valores. Si lanzas una moneda 10 veces, el evento "veces que cae águila" es un conjunto de valores.

Probabilidad

Si todos los resultados son igualmente posibles, la probabilidad de A

$$P_{\text{naïve}}(A) = \frac{\text{veces que sale } A}{\text{total de resultados (i.e., } A \cup A^c)} = \frac{|A|}{|S|}$$

Donde por $|\cdot|$ entendemos la *cardinalidad*, o el número de elementos.

Probabilidad - Variables y espacios:

Variable aleatoria: es una función que mapea los resultados de un experimento aleatorio al conjunto de los números reales (comúnmente). Se suele representar con letra mayúscula (e.g., X).

Existen variables aleatorias discretas (sus valores son infinitos contables), y continuas (sus valores son infinitos no contables).

Espacio muestral: el conjunto de todos los resultados posibles. Se suele representar con Ω. De este conjunto la X mapea a los reales: $X : \Omega \to \mathbb{R}$. Es decir, a cada elemento de Ω asigna un número real, $X(\omega)$.

Evento: subconjunto de Ω, usualmente representado por una vocal mayúscula, e.g., A. Si lanzamos una moneda dos veces, $\Omega = \{HH, HT, TT, TH\}$. El evento "la primera moneda cae H' ' es $A = \{HH, HT\}$.

Operador esperanza $\mathsf{E}[X]$

Valor esperado - (a.k.a. *media*, *esperanza*, o *promedio*) es una suma ponderada de los posibles resultados de nuestra variable aleatoria. Matemáticamente, si $x_1, x_2, x_3, \ldots$ son todas distintos posibles valores que X puede tomar, el valor esperado de X es

$$\mathsf{E}[X] = \sum_i x_i p(X = x_i), \text{ si } x \text{ es discreta}$$

$$\mathsf{E}[X] = \int_{\mathbb{R}} x f(x) dx, \text{ si } x \text{ es continua}$$

Por brevedad, podemos simplemente escribir $p(x_i)$ para el caso discreto. Para el caso continuo, $f(x)$ denota la función de densidad de probabilidad.

Si tenemos n datos, todos con la misma probabilidad de ser tomados por X, entonces la esperanza es simplemente la media aritmética:

$\mathsf{E}[X] = \frac{1}{n}\sum_i x_i$

Con $p(x_1) = p(x_2) = ... = p(x_n) = 1/n$. Por ejemplo, si $X = (1, 5, 9, 10)$, todos con probabilida $p(x) = 1/4$,

$\mathsf{E}[X] = \frac{1}{n}\sum_i x_i = \frac{1}{4} \times (1 + 5 + 9 + 10)$

Propiedades de los logaritmos

Los logaritmos solo están definidos para los números reales

$$\log_b(x) = a, \forall x \in \mathbb{R} > 0$$

Que se lee "para todos los x del conjunto $\mathbb{R}$ mayores que 0".

Un logaritmo se puede definir como el valor al que hay que elevar la base b para obtener x.

Propiedad 1 $log_b(x \times y) = \log_b(x) + \log_b(y)$

En palabras, esto significa que el logaritmo base b del producto de dos números es igual a la suma de los logaritmos de esos números.

Propiedad 2 $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$

Esta propiedad es simplemente la operación inversa de la **Propiedad 1**. Esta propiedad permite expresar las razones (o proporciones) en términos de diferencias.

Si, por ejemplo, x es mayor a y, el rango de valores que puede tomar x/y va desde 1 a infinito. Por otro lado, si $y \geq x$, el rango de valores está entre 0 y 1. Por lo tanto, las razones no son funciones simétricas.

Los logaritmos sí. Si $x > y$, $\log(x/y) > 1$, si $x < y$, $\log(x/y) < 0$. Si $x = y$, $\log(x/y) = 1$.

Propiedad 3 $\log_b(x^a) = a\log_b(x)$

Esto se sigue de la **Propiedad 1**. Supongamos que $a = 3$, entonces $\log_b(x^a) = \log_b(x^3) = \log_b(x \times x \times x)$, lo que es lo mismo a escribir $\log_b(x) + \log_b(x) + \log_b(x)$, y dado que $\log_b(x)$ se repite 3 veces, la expresión $\log_b(x) + \log_b(x) + \log_b(x)$ es igual a $3\log_b(x)$. La expresión es generalizable a números fraccionarios, como 1/2.

Propiedad 3 $\log_b(1) = 0$

Que es lo inverso a la **Propiedad 2.2** de los exponentes, donde $b^0 = 1$.