Índice general

1. Criptosistemas simétricos o de clave privada **1**

1.1. Entropía 1

Capítulo 1

Criptosistemas simétricos o de clave privada

1.1. Entropía

Teorema 1.1. *Una función continua definida sobre el conjunto de funciones de distribución de longitud n que cumpla las condiciones:*

(1) $H\left(\frac{1}{n},\dots,\frac{1}{n}\right) < H\left(\frac{1}{n+1},\dots,\frac{1}{n+1}\right)$

(2) $H\left(\frac{1}{n},\dots,\frac{1}{n}\right) = H\left(\frac{k_1}{n},\dots,\frac{k_m}{n}\right) + \sum_{i=1,k_i\neq 0}^{m} \frac{k_i}{n} H\left(\frac{1}{k_i},\dots,\frac{1}{k_i}\right)$ *siempre que* $\sum_{i=1}^{m} k_i = n$

es de la forma

$$H(p_1,\dots,p_n) = \sum_{i=1,p_i\neq 0}^{n} p_i \log_b\left(\frac{1}{p_i}\right) = -\sum_{i=1,p_i\neq 0}^{n} p_i \log_b p_i$$

para algún $b > 1$.

Demostración. Si $m|n$, entonces

$$\begin{aligned} H\left(\frac{1}{n},\dots,\frac{1}{n}\right) &= H\left(\frac{m}{n},\dots,\frac{m}{n}\right) + \sum_{i=1}^{n/m} \frac{m}{n} H\left(\frac{1}{m},\dots,\frac{1}{m}\right) = \\ &= H\left(\frac{m}{n},\dots,\frac{m}{n}\right) + H\left(\frac{1}{m},\dots,\frac{1}{m}\right) \end{aligned}$$

En particular, si $n = m^s$, entonces

$$H\left(\frac{1}{m^s},\dots,\frac{1}{m^s}\right) = H\left(\frac{1}{m^{s-1}},\dots,\frac{1}{m^{s-1}}\right) + H\left(\frac{1}{m},\dots,\frac{1}{m}\right)$$

Sea $g(n) = H\left(\frac{1}{n},\dots,\frac{1}{n}\right)$, entonces

$$g(m^s) = g(m^{s-1}) + g(m)$$

y por inducción sobre s, se obtiene que

$$g(m^s) = sg(m)$$

La condición (1), implica que g es estrictamente creciente y, por tanto, para todo $m > 1$ tenemos $g(m^s) < g(m^{s+1})$, es decir, $sg(m) < (s+1)g(m)$. Por tanto $g(m)$ es positivo.

Sean n, k y m enteros mayores a 1 y sea s

$$s = \text{máx}\left\{j \in \mathbb{Z} : j \geq 0, m^j \leq n^k\right\}$$

entonces $m^s \leq n^k < m^{s+1}$. Como g es estrictamente creceinte, $g(m^s) \leq g(n^k) \leq g(m^{s+1})$, o equivalentemente

$$sg(m) \leq kg(n) \leq (s+1)g(m)$$

Como log también es una función creciente también tenemos

$$s\log(m) \leq k\log(n) \leq (s+1)\log(m)$$

Por tanto,

$$\frac{s}{k} \leq \frac{g(n)}{g(m)} \leq \frac{s+1}{k} \quad \text{y} \quad \frac{s}{k} \leq \frac{\log(n)}{\log(m)} \leq \frac{s+1}{k}$$

luego

$$\left|\frac{g(n)}{g(m)} - \frac{\log(n)}{\log(m)}\right| \leq \frac{1}{k}$$

Como k es arbitrario,

$$\frac{g(n)}{g(m)} = \frac{\log(n)}{\log(m)}$$

es decir,

$$\frac{g(n)}{\log(n)} = \frac{g(m)}{\log(m)} = C$$

Luego $g(n) = C\log(n)$ para algún número postivo C. Por tanto, si elegimos una base b adecuada, tendremos que $g(n) = \log_b n$.

Supongamos ahora que $(p_1, \ldots, p_k)$ es una distribución de probabilidad formada por números racionales. Poniéndolos con común denominador podemos suponer que $p_i = \dfrac{b_i}{n}$ y, de la propiedad (2) tenemos,

$$H(p_1, \ldots, p_k) = H\left(\frac{b_1}{n}, \ldots, \frac{b_k}{n}\right) = g(n) - \sum_{i=1, b_i \neq 0}^{k} \frac{b_i}{n} g(b_i) = \log_b n - \sum_{i=1, b_i \neq 0}^{k} \frac{b_i}{n} \log_b b_i =$$

$$= \sum_{i=1, b_i \neq 0}^{k} \frac{b_i}{n} \log_b \frac{n}{b_i} = \sum_{i=1, p_i \neq 0}^{k} p_i \log_b \frac{1}{p_i}$$

Como H es continua, entonces

$$H(p_1, \ldots, p_k) = \sum_{i=1, p_i \neq 0}^{k} p_i \log_b \frac{1}{p_i}$$

para toda $k-$ upla $(p_1, \ldots, p_k)$ de números reales en el dominio de H. □

Definición 1.1. Sea b un número real mayor que 1. Se llama *entropía* en base b de una distribución de probabilidad $P = (p_1, \ldots, p_k)$ a

$$H_b(p_1, \ldots, p_k) = \sum_{i=1}^{k} p_i \log_b \frac{1}{p_i}$$

La entropía de una variable aleatoria discreta es la entropía de su distribución de probabilidad.

La base b en la que se calcule la función de entropía sólo implica un cambio de escala debido a la igualdad $\log_b x = \log_{b'} x \cdot \log_b b'$ que implica

$$H_b(X) = H_{b'}(X) \cdot \log_b b'$$

Proposición 1.1. *Sea $(p_1, \ldots, p_n, q_1, \ldots, q_m)$ una distribución de probabilidad. Si $a = \sum_{i=1}^n p_i$, con $0 < a < 1$ entonces*

$$H(p_1, \ldots, p_n, q_1, \ldots, q_m) = H(a, 1-a) + aH\left(\frac{p_1}{a}, \ldots, \frac{p_n}{a}\right) + (1-a)H\left(\frac{q_1}{1-a}, \ldots, \frac{q_m}{1-a}\right)$$

Demostración.

$$\begin{aligned}
H(p_1, \ldots, p_n, q_1, \ldots, q_m) &= \sum_{i=1}^n p_i \log\frac{1}{p_i} + \sum_{i=1}^m q_i \log\frac{1}{q_i} = \sum_{i=1}^n p_i \log\frac{a}{ap_i} + \sum_{i=1}^m q_i \log\frac{1-a}{(1-a)q_i} = \\
&= \sum_{i=1}^n p_i \left(\log\frac{a}{p_i} + \log\frac{1}{a}\right) + \sum_{i=1}^m q_i \left(\log\frac{1-a}{q_i} + \log\frac{1}{1-a}\right) = \\
&= \sum_{i=1}^n p_i \log\frac{a}{p_i} + \sum_{i=1}^n p_i \log\frac{1}{a} + \sum_{i=1}^m q_i \log\frac{1-a}{q_i} + \sum_{i=1}^m q_i \log\frac{1}{1-a} = \\
&= a\log\frac{1}{a} + (1-a)\log\frac{1}{1-a} + \sum_{i=1}^n p_i \log\frac{a}{p_i} + \sum_{i=1}^m q_i \log\frac{1-a}{q_i} = \\
&= H(a, 1-a) + a\sum_{i=1}^n \frac{p_i}{a} \log\frac{a}{p_i} + (1-a)\sum_{i=1}^m \frac{q_i}{1-a} \log\frac{1-a}{q_i} = \\
&= H(a, 1-a) + aH\left(\frac{p_1}{a}, \ldots, \frac{p_n}{a}\right) + (1-a)H\left(\frac{q_1}{1-a}, \ldots, \frac{q_m}{1-a}\right)
\end{aligned}$$

□

Vamos ahora a ver cual es el rango de la función de entropía. Más concretamente vamos a demostrar el siguiente.

Teorema 1.2. *Sea X una variable aleatoria discreta con n sucesos posibles. Entonces*

$$0 \leq H_b(X) \leq \log_b n$$

Además $H_b(X) = 0$ precisamente si $P(X = x) = 1$ para algún suceso x y $H_b(X) = \log_b n$ si y sólo si la distribución de probabilidad de X es uniforme.

Para demostrar el Teorema 1.2 necesitaremos dos lemas. El primero es bien conocido:

Lema 1.1. *Para todo número real positivo x se verifica $\log x \leq x - 1$ y la igualdad se verifica precisamente si $x = 1$.*

El segundo es un poco más complicado:

Lema 1.2. *Sea $P = (p_1, \ldots, p_n)$ una distribución de probabilidad y $Q = (q_1, \ldots, q_n) \in \mathbb{R}^n$ con $0 \leq q_i \leq 1$ y $\sum_{i=1}^n q_i \leq 1$. Entonces*

$$\sum_{i=1, p_i \neq 0}^n p_i \log\frac{1}{p_i} \leq \sum_{i=1, q_i \neq 0} p_i \log\frac{1}{q_i}$$

Además la igualdad se verifica precisamente si $p_i = q_i$ para todo i.

Demostración. Del lema 1.1 se tiene que si $p \neq 0$ y $q \neq 0$ entonces

$$\log \frac{q}{p} \leq \frac{q}{p} - 1$$

y, por tanto,

$$p \log \frac{1}{p} \leq p \log \frac{1}{q} + q - p$$

Puesto que $\sum_{i=1}^{n} q_i \leq 1 = \sum_{i=1}^{n} p_i$, se tiene

$$\sum_{i=1, p_i \neq 0}^{n} p_i \log \frac{1}{p_i} \leq \sum_{i=1, p_i \neq 0, q_i \neq 0}^{n} \left(p_i \log \frac{1}{q_i} + q_i - p_i \right) \leq \sum_{i=1, q_i \neq 0}^{n} p_i \log \frac{1}{q_i}$$

Supongamos que se da la igualdad, esto es,

$$\sum_{i=1, p_i \neq 0}^{n} p_i \log \frac{1}{p_i} = \sum_{i=1, q_i \neq 0} p_i \log \frac{1}{q_i}$$

Entonces,

$$p_i \log \frac{1}{p_i} = p_i \log \frac{1}{q_i} + q_i - p_i$$

para todo i con $p_i \neq 0$ y $q_i \neq 0$, o equivalentemente,

$$\log \frac{q_i}{p_i} = \frac{q_i}{p_i} - 1$$

Pero del lema 1.1, esto equivale a que $p_i = q_i$ para todo i.

$\square$