Introducción a la Teoría de la Información

Práctico 1: Entropía y Divergencia

Año 2025

Cada ejercicio tiene un símbolo que indica su dificultad de acuerdo a la siguiente escala: $\diamondsuit$ básica, $\star$ media, $*$ avanzada, y ‡ difícil.

$\diamondsuit$ Ejercicio 1

Sea $p(x, y)$ dada por

	Y	0	1
X			
0		1/3	1/3
1		0	1/3

Encuentre:
(a) $H(X)$, $H(Y)$.
(b) $H(X|Y)$, $H(Y|X)$.
(c) $H(X, Y)$.

$\diamondsuit$ Ejercicio 2

Sea la variable aleatoria X con tres posibles valores $\{a, b, c\}$. Considera dos distribuciones para esta variable aleatoria:

Símbolo	$p(x)$	$q(x)$
a	$\frac{1}{2}$	$\frac{1}{3}$
b	$\frac{1}{4}$	$\frac{1}{3}$
c	$\frac{1}{4}$	$\frac{1}{3}$

(a) Calcule $H(p)$, $H(q)$, $D(p||q)$, y $D(q||p)$. Verifique que en este caso, $D(p||q) \neq D(q||p)$.

(b) Dé un ejemplo de dos distribuciones p y q en un alfabeto binario tal que $D(p||q) = D(q||p)$ (que no sea el caso trivial $p = q$).

♢ Ejercicio 3

El *tiro desde los once pasos* es la pena máxima en el fútbol; se desea analizar con herramientas de teoría de la información las opciones del guardameta. El arco de fútbol se divide en cinco zonas $Z = \{z_1, z_2, z_3, z_4, z_5\}$ hacia donde el golero puede arrojarse.

Según el perfil del ejecutante (zurdo o diestro), el arquero conoce una distribución de probabilidad $p_I(Z)$ o $p_D(Z)$ de las zonas como se muestra en la figura.

(a) Calcule la incertidumbre que tiene el cuidavallas para cada uno de los perfiles de ejecutante.

(b) Si la probabilidad de que un ejecutante sea zurdo es igual a la de que sea diestro, calcule la incertidumbre media que tiene el cuidavallas *dado* el perfile de ejecutante.

♢ Ejercicio 4

¿Cuál es el valor mínimo de $H(p_1, \ldots, p_n) = H(\mathbf{p})$, donde $\mathbf{p}$ varía sobre el conjunto de vectores de probabilidad n-dimensionales? Encuentre todos los vectores $\mathbf{p}$ que alcanzan este mínimo.

♢ Ejercicio 5

Sea X una variable aleatoria discreta. Demuestre que la entropía de una función de X es menor o igual que la entropía de X justificando los siguientes pasos:

$$\begin{aligned} H(X, g(X)) &\overset{(a)}{=} H(X) + H(g(X) \mid X) \\ &\overset{(b)}{=} H(X); \\ H(X, g(X)) &\overset{(c)}{=} H(g(X)) + H(X \mid g(X)) \\ &\overset{(d)}{\geq} H(g(X)). \end{aligned}$$

Por lo tanto $H(g(X)) \leq H(X)$.

♢ Ejercicio 6

Demuestra que si $H(Y|X) = 0$, entonces Y es una función de X. Es decir, para todos los x con $p(x) > 0$, hay solo un valor posible de y con $p(x, y) > 0$.

$\star$ Ejercicio 7

Una moneda justa es lanzada hasta que se obtenga la primera cara. Sea X el número de tiradas necesario.

(a) Encuentre la entropía $H(X)$ en bits. Las siguientes expresiones pueden ser útiles:

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \quad \text{y} \quad \sum_{n=0}^{\infty} n \cdot r^n = \frac{r}{(1-r)^2}.$$

(b) Una variable aleatoria X se obtiene según esta distribución. Encuentre una secuencia "eficiente" de preguntas sí-no para determinar X, de la forma: "¿Está X contenida en el conjunto S?". Compare $H(X)$ con el número esperado de preguntas requeridas para determinar X.

$\star$ Ejercicio 8

Sean X_1 y X_2 variables aleatorias discretas con distribuciones de probabilidad p_1 y p_2, respectivamente, sobre los alfabetos $\mathcal{X}_1 = \{1, 2, ..., m\}$ y $\mathcal{X}_2 = \{m+1, ..., n\}$. Sea

$$X = \begin{cases} X_1 \text{ con probabilidad } \alpha, \\ X_2 \text{ con probabilidad } 1-\alpha. \end{cases}$$

(a) Encuentre $H(X)$ en términos de $H(X_1)$, $H(X_2)$, y α.

(b) Interprete cada término del resultado.

$*$ Ejercicio 9

Sean X e Y variables aleatorias que toman valores en $\{x_1, \ldots, x_r\}$ y $\{y_1, \ldots y_s\}$, respectivamente, y sea $Z = X + Y$.

(a) Demuestre que $H(Z|X) = H(Y|X)$. Argumente que si X e Y son independientes, entonces $H(Y) \leq H(Z)$ y $H(X) \leq H(Z)$. Por lo tanto, la suma de variables aleatorias independientes agrega incertidumbre.

(b) Proporcione un ejemplo de variables aleatorias (necesariamente dependientes) en las que $H(X) > H(Z)$ y $H(Y) > H(Z)$.

(c) ¿Bajo qué condiciones $H(Z) = H(X) + H(Y)$?

$\star$ Ejercicio 10

Sea $X \in \mathcal{X}$, $X \sim p$, una variable aleatoria, donde $\mathcal{X}$ es el conjunto de naturales entre 0 y $2n-1$, inclusive, para cierta constante positiva n. Sean

$$Q = X \,\mathbf{div}\, 2\,, \quad R = X \,\mathbf{mod}\, 2$$

el cociente entero y el resto de dividir X por 2. Establecer desigualdades y especificar qué condiciones debe cumplir p para que se de la igualdad en los siguientes casos:

- $H(X)$ vs. $H(Q,R)$
- $H(X)$ vs. $H(Q)$
- $H(R)$ vs. 1

$\star$ Ejercicio 11

Se tiene un vector de probabilidad $\mathbf{p} = (p_1, p_2, \dots, p_i, \dots, p_j, \dots p_n)$.

(a) Compare su entropía con la de $\mathbf{p}' = (p_1, \dots, \frac{p_i+p_j}{2}, \dots, \frac{p_i+p_j}{2}, \dots, p_n)$. Justifique la desigualdad.

(c) Más en general, compare $H(\mathbf{p})$ con $H(\mathbf{p}'')$, siendo
$\mathbf{p}'' = (p_1, \dots, \lambda p_i + (1-\lambda)p_j, \dots, (1-\lambda)p_i + \lambda p_j, ..., p_n)$ con $\lambda \in (0,1)$.

$*$ Ejercicio 12

Sea $X_1, X_2, ..., X_n$ variables aleatorias que denotan los resultados de lanzamientos independientes de una moneda cargada, con $P(X_i = 1) = p$, y $P(X_i = 0) = 1-p$, donde p es desconocido. Denotamos $\{0,1\}^* = \{\Lambda, 0, 1, 00, 01, ...\}$ al conjunto de todas las secuencias binarias de longitud finita, donde Λ es la cadena nula.

Deseamos obtener una secuencia $Z_1, Z_2, ..., Z_K$ de lanzamientos justos de moneda a partir de $X_1, X_2, ..., X_n$. Con este fin, sea $f : X^n \to \{0,1\}^*$ una función tal que $f(X_1, X_2, \dots, X_n) = (Z_1, Z_2, \dots, Z_K)$, donde $Z_i \sim Bernoulli(1/2)$, y K puede depender de $X_1, \dots, X_n$.

Para que la secuencia $Z_1, Z_2, \dots$ parezca ser de lanzamientos justos de moneda, la asignación f debe tener la propiedad de que todas las 2^k secuencias $(Z_1, Z_2, ..., Z_k)$ de una longitud k dada tengan la misma probabilidad (eventualmente 0), para $k = 1, 2, ...$ Por ejemplo, para $n = 2$, la asignación $f(01) = 0$, $f(10) = 1$, $f(00) = f(11) = \Lambda$, tiene la propiedad de que $P(Z_1 = 1|K = 1) = P(Z_1 = 0|K = 1) = 1/2$. Justifique las siguientes desigualdades:

$$\begin{aligned}
nH(p) &\overset{\text{(a)}}{=} H(X_1, \dots, X_n) \\
&\overset{\text{(b)}}{\geq} H(Z_1, Z_2, \dots, Z_K, K) \\
&\overset{\text{(c)}}{=} H(K) + H(Z_1, \dots, Z_K \mid K) \\
&\overset{\text{(d)}}{=} H(K) + E[K] \\
&\overset{\text{(e)}}{\geq} E[K].
\end{aligned}$$

Por lo tanto, no se pueden obtener más de $nH(p)$ lanzamientos justos de moneda a partir de $X_1, \dots, X_n$, en promedio. Muestre un buen mapeo f para secuencias de longitud 4.