Introducción a la Teoría de la Información

Práctico 1: Entropía y Divergencia

Año 2024

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

$\diamondsuit$ Problema 1

¿Cuál es el valor mínimo de $\mathrm{H}(\mathrm{p}_1, \ldots, \mathrm{p}_n) = \mathrm{H}(\mathrm{p})$? Donde p varía sobre el conjunto de vectores de probabilidad n-dimensionales. Encuentra todos los p que alcanzan este mínimo.

$\diamondsuit$ Problema 2

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

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

Y entonces $H(g(X)) \leq H(X)$.

$\diamondsuit$ Problema 3

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 $p(x, y) > 0$.

$\star$ Problema 4

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 $H(X) \leq H(Z)$. Así, 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 $H(Z) = H(X) + H(Y)$?

$*$ Problema 5

Sean X1 y X2 variables aleatorias discretas con funciones de distribución de probabilidad $p1()$ y $p2()$ respectivamente, sobre los alfabetos $X1 = \{1, 2, \ldots, m\}$ y $X2 = \{m + 1, \ldots, n\}$. 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 $H(X1)$, $H(X2)$, y α.

(b) Interpretar cada término del resultado.

$\star$ Problema 6

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

(a) Comparar su entropía con la del conjunto $\mathbf{p'} = (p_1, ..., \frac{p_i+p_j}{2}, ..., \frac{p_i+p_j}{2}, ..., p_n)$.

(b) Justificar la desigualdad.

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

(d) ¿La entropía es cóncava o convexa? Vincular esta propiedad con el resultado de (c).

⋆ Problema 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. Encuentra una secuencia “eficiente” de preguntas de sí-no de la forma: “¿Está X contenida en el conjunto S?” Compara $H(X)$ con el número esperado de preguntas requeridas para determinar X.

♢ Problema 8

Sea la variable aleatoria X con tres posibles valores a, b, c. Considera dos distribuciones en 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) De 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 $p = q$).

♢ Problema 9

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) $H(X, Y)$.

(d) $H(Y) - H(Y|X)$.

(e) $I(X; Y)$.

(f) Dibuje un diagrama de Venn para las cantidades en las partes (a) a (e).

⋆ Problema 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

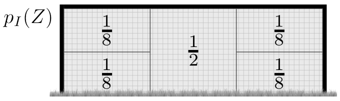
$$Q = X \,\mathbf{div}\, 2\,, \quad R = X \,\mathbf{mod}\, 2$$

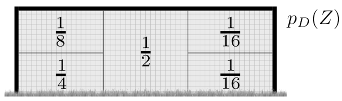
el cociente entero y el resto de dividir X por 2. Establecer desigualdades y especificar qué condiciones debe cumplir p para que se de la igualdad en los siguientes casos:

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

♢ Problema 11

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.

Hallar la información media que tiene el cuidavallas si sabe el perfil del ejecutante, tanto para el caso que sea zurdo o diestro.

⋆ Problema 12

Hay un grupo de 16 personas y se sabe que una de ellas (y una sola) tiene COVID. Se quiere ubicar quién es.

(a) Dar un modelo de la fuente de información. ¿Cuántos bits de información implica saber quién es el infectado?

(b) La primera estrategia que se le ocurre a alguien es ir haciendo los test de PCR uno por uno. ¿Cuántos análisis en media requiere encontrar al infectado de esa manera?

(c) Proponer una estrategia mejor a partir de la idea de mezclar las muestras. ¿Cuál es la estrategia óptima y por qué se puede afirmar que lo es?

(d) ¿Cómo cambia el problema si también pudiera no haber ninguna persona infectada? ¿Cuántos análisis son el mínimo ahora?

$*$ Problema 13

Sea $X_1, X_2, ..., X_n$ variables aleatorias que denotan los resultados de lanzamientos independientes de una moneda cargada. Así, $\Pr\{X_i = 1\}$ = p, $\Pr\{X_i = 0\}$ = 1-p, donde p es desconocido. 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 $f : X^n \to \{0,1\}^*$ (donde $\{0,1\}^* = \{\Lambda, 0, 1, 00, 01, ...\}$ es el conjunto de todas las secuencias binarias de longitud finita) una asignación $f(X_1, X_2, ..., X_n) = (Z_1, Z_2, ..., Z_K)$, donde $Z_i \sim Bernoulli(1/2)$, y K puede depender de $X_1, ..., X_n$.

Para que la secuencia $Z_1, Z_2, ...$ parezca ser de lanzamientos justos de moneda, la asignación f de lanzamientos de moneda cargada a lanzamientos justos 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 $f(01) = 0$, $f(10) = 1$, $f(00) = f(11) = \Lambda$ (la cadena nula) tiene la propiedad de que $\Pr\{Z_1 = 1 | \mathrm{K} = 1\} = \Pr\{Z_1 = 0 | \mathrm{K} = 1\} = 1/2$. Dé razones para las siguientes desigualdades:

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

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

$\star$ Problema 14

Dado $X \sim p(x)$, $x = 1, 2, ..., m$. Se nos da un conjunto $S \subseteq \{1, 2, ..., m\}$. Preguntamos si $X \in S$ y recibimos la respuesta:

$$Y = \begin{cases} 1 & \text{if } X \in S \\ 0 & \text{if } X \notin S \end{cases}$$

Supongamos que $Pr\{X \in S\} = \alpha$.

(a) Halle la disminución en la incertidumbre $H(X) - H(X|Y)$. Aparentemente, cualquier conjunto S con un dado α es tan bueno como cualquier otro.

(b) Relacione el resultado anterior con las mejores estrategias cuando se tiene que averiguar algo con preguntas de respuesta binaria.