Introducción a la Teoría de la Información

Práctico 4: AEP, Información Mutua y Desigualdad de Fano

Año 2025

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

Sean (X_i, Y_i) i.i.d. $\sim p(x, y)$. Formamos el cociente logarítmico de verosimilitud de la hipótesis de que X e Y son independientes frente a la hipótesis de que X e Y son dependientes. ¿Cuál es el límite de

$$\frac{1}{n} \log \frac{p(X^n)p(Y^n)}{p(X^n, Y^n)}?$$

$*$ Problema 2

Sean X_i i.i.d. $\sim p(x)$, donde $x \in \{1, 2, \ldots, m\}$. Sea $\mu = \mathbb{E}[X]$ y $H = -\sum_x p(x) \log p(x)$.
Sea $A_n = \{x^n \in \mathcal{X}^n : \left|-\frac{1}{n} \log p(x^n) - H\right| \leq \epsilon\}$.
Sea $B_n = \{x^n \in \mathcal{X}^n : \left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \leq \epsilon\}$.
(a) ¿Se cumple que $\Pr\{X^n \in A_n\} \to 1$?
(b) ¿Se cumple que $\Pr\{X^n \in A_n \cap B_n\} \to 1$?
(c) Muestre que $|A_n \cap B_n| \leq 2^{n(H+\epsilon)}$ para todo n.
(d) Muestre que $|A_n \cap B_n| \geq \frac{1}{2} 2^{n(H-\epsilon)}$ para n suficientemente grande.

$\star$ Problema 3

Sean $X_1, X_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas, extraídas según la función de masa de probabilidad $p(x)$, donde $x \in \{1, 2, \ldots, m\}$. Por lo tanto, $p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(x_i)$. Sabemos que $-\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) \to H(X)$ en probabilidad. Sea $q(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} q(x_i)$, donde q es otra función de masa de probabilidad en $\{1, 2, \ldots, m\}$.
(a) Evalúe $\lim_{n\to\infty} -\frac{1}{n} \log q(X_1, X_2, \ldots, X_n)$, donde $X_1, X_2, \ldots$ son i.i.d. $\sim p(x)$.
(b) Ahora evalúe el límite del cociente logarítmico de verosimilitud $\frac{1}{n} \log \frac{q(X_1, X_2, \ldots, X_n)}{p(X_1, X_2, \ldots, X_n)}$ cuando $X_1, X_2, \ldots$ son i.i.d. $\sim p(x)$. Notar que las chances a favor de q son exponencialmente pequeñas cuando p es la verdadera distribución de probabilidad.

$\diamondsuit$ Problema 4

Sean X, Y y Z variables aleatorias. Demuestre las siguientes desigualdades y encuentre las condiciones para la igualdad:
(a) $I(X, Y; Z) \geq I(X; Z)$.
(b) $I(X; Z \mid Y) \geq I(Z; Y \mid X) - I(Z; Y) + I(X; Z)$.

$\ast$ Problema 5

Considere una secuencia de n variables aleatorias binarias $X_1, X_2, \ldots, X_n$. Cada secuencia con un número par de 1's tiene una probabilidad de $2^{-(n-1)}$, y cada secuencia con un número impar de 1's tiene una probabilidad de 0.
Encuentre las informaciones mutuas:
$I(X_1; X_2), I(X_2; X_3 \mid X_1), \ldots, I(X_{n-1}; X_n \mid X_1, \ldots, X_{n-2})$.

$\diamondsuit$ Problema 6

Se nos da la siguiente distribución conjunta en (X, Y):

Y / X	a	b	c
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$
2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$

Sea $\hat{X}(Y)$ un estimador para X (basado en Y) y sea $P_e = \Pr\{\hat{X}(Y) \neq X\}$.
(a) Encuentre el estimador de mínima probabilidad de error $\hat{X}(Y)$ y el P_e asociado.
(b) Evalúe la desigualdad de Fano para este problema y compare.

$*$ Problema 7

Se desea identificar un objeto aleatorio $X \sim p(x)$. Se realiza una pregunta $Q \sim r(q)$ de forma aleatoria según $r(q)$. Esto da como resultado una respuesta determinística $A = A(x, q) \in \{a_1, a_2, \ldots\}$. Supongamos que X y Q son independientes. Entonces $I(X; Q, A)$ es la incertidumbre en X eliminada por la pregunta-respuesta (Q, A).

(a) Muestre que $I(X; Q, A) = H(A \mid Q)$. Interprete la igualdad.
(b) Ahora supongamos que se hacen dos preguntas independientes e idénticamente distribuidas $Q_1, Q_2 \sim r(q)$, obteniendo respuestas A_1 y A_2. Muestre que dos preguntas son menos valiosas que el doble de una sola pregunta en el sentido de que $I(X; Q_1, A_1, Q_2, A_2) \leq 2I(X; Q_1, A_1)$.
Guía: Notar que los pares (Q_1, A_1) y (Q_2, A_2) son condicionalmente independientes dado X.

$\star$ Problema 8

¿Cuánta información proporciona la longitud de una secuencia sobre el contenido de la secuencia? Consideramos un proceso $\{X_i\}$ Bernoulli $\left(\frac{1}{2}\right)$. Detenemos el proceso cuando aparece el primer 1. Designemos a este tiempo de detención como N. Por lo tanto, X^N es un elemento del conjunto de todas las secuencias binarias de longitud finita $\{0, 1\}^*$ $= \{0, 1, 00, 01, 10, 11, 000,...\}$.
(a) Encuentre $I(N; X^N)$.
(b) Encuentre $H(X^N \mid N)$.
(c) Encuentre $H(X^N)$.

Ahora consideremos un tiempo de detención diferente. Para esta parte, nuevamente supongamos que $\{X_i\} \sim$ Bernoulli $\left(\frac{1}{2}\right)$ pero detenemos en el tiempo $N = 6$ con probabilidad $\frac{1}{3}$ y detenemos en el tiempo $N = 12$ con probabilidad $\frac{2}{3}$. Supongamos que este tiempo de detención es independiente de la secuencia $X_1 X_2 \cdots X_{12}$.
(d) Encuentre $I(N; X^N)$.
(e) Encuentre $H(X^N \mid N)$.
(f) Encuentre $H(X^N)$.