Javier Sáez
77448344F
franciscojavier.saez@estudiante.uam.es

Ejercicios
Teoría de la Información

2021-10-22

Problema 1
Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ para ciertos $\mu, \sigma^2 \in \mathbb{R}$. Calcula su entropía.

Recordamos que la función de densidad de una variable aleatoria que sigue una distribución gaussiana de parámetros μ, σ^2 es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

y que la entropía de una variable aleatoria continua X en general viene dada por la expresión:

$$H(X) = -\mathbb{E}\left[\log X\right] = \int_{-\infty}^{\infty} f_X(x)\left(-\log f_X(x)\right) dx.$$

Podemos resolver esta integral de forma muy sencilla si sustituimos la función de densidad f en la parte del logaritmo pero no en la parte que aparece sola. Tenemos que

$$-\log f(x) = \log\frac{1}{f(x)} = \log\left((2\pi\sigma^2)^{\frac{1}{2}} \exp\left(\frac{1}{2\sigma^2}(x-\mu)^2\right)\right) = \frac{1}{2}\log\left(2\pi\sigma^2\right) + \frac{1}{2\sigma^2}(x-\mu)^2$$

Con esto se tiene que, en nuestro caso:

$$H(X) = \frac{1}{2}\log\left(2\pi\sigma^2\right)\underbrace{\int_{-\infty}^{\infty} f_X(x)dx}_{(1)} + \frac{1}{2\sigma^2}\underbrace{\int_{-\infty}^{\infty} f_X(x)(x-\mu)^2 dx}_{(2)}$$

Donde vemos que en (1) estamos integrando en todo el dominio posible la función de densidad de la normal, por lo que tenemos que eso es igual a 1. Además, en (2) tenemos la definición de la varianza $\nu = \sigma^2$ de la distribución normal de media μ y desviación típica σ.

$$H(X) = \frac{1}{2}\log\left(2\pi\sigma^2\right) + \frac{\sigma^2}{2\sigma^2} = \frac{1}{2}\left(\log\left(2\pi\sigma^2\right) + 1\right) = \frac{1}{2}\left(\log\left(2\pi\sigma^2\right) + \log(e)\right) = \frac{1}{2}\log\left(2\pi e\sigma^2\right)$$

□

Problema 2

Sean X, Y dos variables aleatorias. Probar que la información mutua entre ambas es simétrica, es decir, $MI(X;Y) = MI(Y;X)$.

Este ejercicio es una consecuencia prácticamente trivial de que la distribución conjunta de $P(X,Y)$ es la misma que la de $P(Y,X)$. Lo vemos formalmente. Recordamos que

$$MI(X;Y) = \sum_{x,y} P_{XY}(x,y)\log\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)} \tag{1}$$

Ahora, basta cambiar el orden de las sumatorias y aplicar que $P(X,Y) = P(Y,X)$ para ver que:

$$MI(X;Y) = \sum_{x,y} P_{XY}(x,y)\log\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)} = \sum_{y,x} P_{YX}(y,x)\log\frac{P_{YX}(y,x)}{P_Y(y)P_X(x))} = MI(Y;X)$$

como queríamos ver.
En el caso continuo, la demostración es análoga, pues podemos aprovecharnos de Teorema de Fubini para cambiar las integrales y obtener el mismo resultado.

□

Problema 3

Sean X, Y dos variables aleatorias. Demostrar que, si $H(X)$ es la entropía de X, y $H(X,Y)$ es la entropía conjunta de X e Y, entonces:

$$MI(X;Y) = H(X) + H(Y) - H(X,Y).$$

Lo primero que hacemos es recordar la definición de entropía y entropía conjunta para variables discretas. Tenemos que

$$H(X) = -\mathbb{E}\left[\log X\right] \tag{2}$$
$$H(X,Y) = -\mathbb{E}\left[\log P_{XY}(x,y)\right] \tag{3}$$

El resultado que queremos obtener se sigue de operar en la definición de información mutua (1). Lo primero que haremos es separar el logaritmo usando las propiedades de logaritmo del producto y del cociente:

$$\begin{aligned}
MI(X;Y) &= \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x)P_Y(y)} \\
&= \sum_{x,y} P_{XY}(x,y) \left(\log P_{XY}(x,y) - \log P_X(x) - \log P_Y(y)\right) \\
&= \underbrace{\sum_{x,y} P_{XY}(x,y) \log P_{XY}(x,y)}_{-H(X,Y)} - \sum_{x,y} P_{XY}(x,y) \log P_X(x) - \sum_{x,y} P_{XY}(x,y) \log P_Y(y)
\end{aligned} \tag{4}$$

Ya hemos obtenido uno de los términos que queríamos. Ahora, basta ver que

$$\sum_x \sum_y P_{XY}(x,y) \log P_X(x) = \sum_x \log P_X(x) \sum_y P_{XY}(x,y) \overset{(*)}{=} \sum_x P_X(x) \log P_X(x) = -H(X)$$

donde en $(*)$ hemos usado que estamos sumando para un x fijo sobre todos los valores de y, luego nos queda simplemente la probabilidad de ese x en la distribución marginal P_X. Así, sustiyuyendo esto en la ecuación (4), obtemos lo que queríamos:

$$MI(X;Y) = -H(X,Y) - (-H(X)) - (-H(Y)) = H(X) + H(Y) - H(X,Y)$$

□

Problema 4

Sean X, Y dos variables aleatorias. Con la notación del ejercicio anterior, demostrar que:

- $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$
- $MI(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$
- $MI(X,X) = H(X)$

Recordamos primero la definición de entropía condicionada:

$$H(X|Y) = -\mathbb{E}_{P_{XY}}\left[\log \frac{P_{XY}(x,y)}{P_X(x)}\right] = -\sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x)} = -\sum_x P_X(x) \sum_y P_{Y|X}(y|x) \log P_{Y|X}(y|x) \tag{5}$$

Vamos a usar esta definición y las que se han dado en ejercicios anteriores para probar estas igualdades.

- $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.

  Usando la definición de la entropía conjunta dada en la ecuación (3), y aplicando la definición de probabilidad condicionada $P_{XY}(x,y) = P_{Y|X}(y|x)P_X(x)$ en varias ocasiones obtenemos:

$$\begin{aligned}
H(X,Y) &= -\sum_{x,y} P_{XY}(x,y)\log P_{XY}(x,y) \\
&= -\sum_{x,y} P_{XY}(x,y)\log\big(P_{Y|X}(y|x)P_X(x)\big) \\
&= -\sum_{x,y} P_{XY}(x,y)\big(\log P_{Y|X}(y|x) + \log P_X(x)\big) \\
&= -\sum_{x,y} P_{XY}(x,y)\log P_{Y|X}(y|x) - \sum_{x,y} P_{XY}(x,y)\log P_X(x) \\
&= -\sum_{x} P_X(x)\log P_X(x) - \sum_{x,y} P_{XY}(x,y)\log P_{Y|X}(y|x) \\
&= H(X) - \sum_{x} P_X(x)\sum_{y} P_{Y|X}(y|x)\log P_{Y|X}(y|x) \\
&= H(X) + H(Y|X)
\end{aligned}$$

Del mismo modo pero usando la otra descomposición de la probabilidad conjunta, obtenemos la segunda igualdad.

- $MI(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$.

  Para esta, podemos usar que hemos probado en el ejercicio anterior que $MI(X;Y) = H(X) + H(Y) - H(X,Y)$ y que en el apartado anterior hemos probado que $H(X,Y) = H(Y) + H(X|Y)$ para obtener que

  $$MI(X;Y) = H(X) + H(Y) - H(X,Y) = H(X) + H(Y) - (H(Y) + H(X|Y)) = H(X) - H(X|Y),$$

  y, del mismo modo podemos probar la otra igualdad usando que $H(X,Y) = H(X) + H(Y|X)$.

- $MI(X;X) = H(X)$.
  Este es consecuencia directa del anterior y de que

  $$H(X|X) = -E_{P_X}\left[\log\frac{P_X(x)}{P_X(x)}\right] = -E_{P_X}[\log 1] = 0.$$

quedando así probadas todas las igualdades $\square$