Info Theory HW 2

24.04.02

> For a continuous random variable X, given EX, DX, proof that $N(EX, DX)$ is a distribution of X which has the Maximum Entropy in parametric family.

Let $Y \sim N(EX, DX)$. Denote the p.d.f. of X as f, EX as μ, and DX as σ^2. We prove that $D(X \parallel Y) = h(Y) - h(X)$.

$$\begin{aligned}
D(X \parallel Y) &= \int_{-\infty}^{\infty} f(x) \log_2 (f(x)\sqrt{2\pi\sigma^2} e^{\frac{(x-\mu)^2}{2\sigma^2}}) \,\mathrm{d}x \\
&= -\int_{-\infty}^{\infty} f(x) \log_2 \frac{1}{f(x)} \,\mathrm{d}x + \frac{1}{2}\log_2(2\pi\sigma^2) \int_{-\infty}^{\infty} f(x) \,\mathrm{d}x + \frac{\log_2 e}{2\sigma^2} \int_{-\infty}^{\infty} (x-\mu)^2 f(x) \,\mathrm{d}x \\
&= -h(X) + \frac{1}{2}\log_2(2\pi\sigma^2) + \frac{1}{2}\log_2 e \\
&= \frac{1}{2}\log_2(2\pi\sigma^2 e) - h(X).
\end{aligned}$$

We have proved in the previous homework that $h(Y) = \frac{1}{2}\log_2(2\pi\sigma^2 e)$, so $D(X \parallel Y) = h(Y) - h(X)$.

Knowing $D(X \parallel Y) \geq 0$ for any X, Y, we conclude that Y is the distribution with the maximum entropy.

> For a discrete and non-negative integer random variable X, given EX. Show the distribution of X that has the Maximum Entropy and prove it.

Denote $p_k = \Pr(X = k)$ and $\mu = EX$.

Let Y be such a distribution that $\Pr(Y = k) = \frac{e^{-\lambda k}}{\sum_{i=0}^{\infty} e^{-\lambda i}}$ and choose λ to let $EY = \mu$.

Now we prove this Y is the distribution with the maximum entropy.

Denote $M = \sum_{i=0}^{\infty} e^{-\lambda i}$. We first calculate the entropy of Y.

$$\begin{aligned}
H(Y) &= \sum_{k=0}^{\infty} \frac{1}{M} e^{-\lambda k} \log_2 (M e^{\lambda k}) \\
&= \log_2 M \sum_{k=0}^{\infty} \frac{1}{M} e^{-\lambda k} + \lambda \log_2 e \sum_{k=0}^{\infty} \frac{k}{M} e^{-\lambda k} \\
&= \log_2 M + \lambda\mu \log_2 e.
\end{aligned}$$

We next calculate $D(X \parallel Y)$.

$$\begin{aligned}
D(X \parallel Y) &= \sum_{k=0}^{\infty} p_k \log_2 (p_k M e^{\lambda k}) \\
&= -\sum_{k=0}^{\infty} p_k \log_2 p_k + \log_2 M \sum_{k=0}^{\infty} p_k + \lambda \log_2 e \sum_{k=0}^{\infty} k p_k \\
&= -H(X) + \log_2 M + \lambda\mu \log_2 e \\
&= H(Y) - H(X).
\end{aligned}$$

So, identical to the argument in the last problem, we can conclude that Y is the distribution with the maximum entropy.