Info theory Recitation 2

Kanad Sarkar

2) **AEP and mutual information**. Let (X_i, Y_i) be i.i.d. $\sim p(x, y)$. We form the log likelihood ratio of the hypothesis that X and Y are independent vs. the hypothesis that X and Y are dependent. What is the limit of

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

$$\begin{aligned}
\frac{1}{n}\log\frac{p(X^n)p(Y^n)}{p(X^n,Y^n)} &= \frac{1}{n}\log\prod_{i=1}^{n}\frac{p(X_i)p(Y_i)}{p(X_i,Y_i)} \\
&= \frac{1}{n}\sum_{i=i}^{n}\log\frac{p(X_i)p(Y_i)}{p(X_i,Y_i)} \\
&\to E(\log\frac{p(X_i)p(Y_i)}{p(X_i,Y_i)}) \\
&= -I(X;Y)
\end{aligned}$$

Thus, $\frac{p(X^n)p(Y^n)}{p(X^n,Y^n)} \to 2^{-nI(X;Y)}$, which will converge to 1 if X and Y are indeed independent.

3) **Piece of cake**

A cake is sliced roughly in half, the largest piece being chosen each time, the other pieces discarded. We will assume that a random cut creates pieces of proportions:

$$P = \begin{cases} (\frac{2}{3}, \frac{1}{3}) & \text{w.p. } \frac{3}{4} \\ (\frac{2}{5}, \frac{3}{5}) & \text{w.p. } \frac{1}{4} \end{cases}$$

Thus, for example, the first cut (and choice of largest piece) may result in a piece of size $\frac{3}{5}$. Cutting and choosing from this piece might reduce it to size $(\frac{3}{5})(\frac{2}{3})$ at time 2, and so on.

How large, to first order in the exponent, is the piece of cake after n cuts?

Let C_i be the fraction of the piece of cake that is cut at the ith cut, and let T_n be the fraction of cake left after n cuts. Then we have $T_n = C_1 C_2 \dots C_n = \prod_{i=1}^{n} C_i$. Hence, as in Question 2 of Homework Set #3,

$$\begin{aligned}\lim \frac{1}{n} \log T_n &= \lim \frac{1}{n} \sum_{i=1}^{n} \log C_i \\ &= E[\log C_1] \\ &= \frac{3}{4} \log \frac{2}{3} + \frac{1}{4} \log \frac{3}{5}.\end{aligned}$$

10) **Random box size.** An n-dimensional rectangular box with sides $X_1, X_2, X_3, \ldots, X_n$ is to be constructed. The volume is $V_n = \prod_{i=1}^{n} X_i$. The edge length l of a n-cube with the same volume as the random box is $l = V_n^{1/n}$. Let $X_1, X_2, \ldots$ be i.i.d. uniform random variables over the unit interval $[0, 1]$. Find $\lim_{n\to\infty} V_n^{1/n}$, and compare to $(EV_n)^{\frac{1}{n}}$. Clearly the expected edge length does not capture the idea of the volume of the box. The geometric mean, rather than the arithmetic mean, characterizes the behavior of products.

$$\int \log_a x \, dx = x \log_a x - \frac{x}{\ln a} = \frac{x}{\ln a}(\ln x - 1)$$

10) *Random box size*. The volume $V_n = \prod_{i=1}^n X_i$ is a random variable, since the X_i are random variables uniformly distributed on $[0, 1]$. V_n tends to 0 as $n \to \infty$. However

$$\log_e V_n^{\frac{1}{n}} = \frac{1}{n} \log_e V_n = \frac{1}{n} \sum \log_e X_i \to E(\log_e(X)) \text{ a.e.}$$

by the Strong Law of Large Numbers, since X_i and $\log_e(X_i)$ are i.i.d. and $E(\log_e(X)) < \infty$. Now

$$E(\log_e(X_i)) = \int_0^1 \log_e(x)\, dx = -1$$

Hence, since e^x is a continuous function,

$$\lim_{n\to\infty} V_n^{\frac{1}{n}} = e^{\lim_{n\to\infty} \frac{1}{n} \log_e V_n} = \frac{1}{e} < \frac{1}{2}.$$

Thus the "effective" edge length of this solid is e^{-1}. Note that since the X_i's are independent, $E(V_n) = \prod E(X_i) = (\frac{1}{2})^n$. Also $\frac{1}{2}$ is the arithmetic mean of the random variable, and $\frac{1}{e}$ is the geometric mean.

8) **Products.** Let

$$X = \begin{cases} 1, & \frac{1}{2} \\ 2, & \frac{1}{4} \\ 3, & \frac{1}{4} \end{cases}$$

Let $X_1, X_2, \ldots$ be drawn i.i.d. according to this distribution. Find the limiting behavior of the product

$$(X_1 X_2 \cdots X_n)^{\frac{1}{n}} \quad .$$

8) *Products*. Let

$$P_n = (X_1 X_2 \dots X_n)^{\frac{1}{n}} . \tag{156}$$

Then

$$\log P_n = \frac{1}{n} \sum_{i=1}^{n} \log X_i \to E \log X, \tag{157}$$

with probability 1, by the strong law of large numbers. Thus $P_n \to 2^{E \log X}$ with prob. 1. We can easily calculate $E \log X = \frac{1}{2} \log 1 + \frac{1}{4} \log 2 + \frac{1}{4} \log 3 = \frac{1}{4} \log 6$, and therefore $P_n \to 2^{\frac{1}{4} \log 6} = 1.565$.