SOLUTIONS TO PROBLEM SET-4

EE6340

① Shuffles Increases Entropy :-

$$H(TX) \geq H(TX/T) \quad \left[\text{Conditioning cannot Increase entropy}\right]$$

$$= H(T^{-1}TX/T) \quad \left[\text{given 'T' we can reverse the shuffle}\right]$$

$$= H(X/T) \quad \left[T^{-1} \text{ and } T \text{ will result in same point}\right]$$

$$= H(X) \quad \left[X \,\&\, T \text{ are Independent}\right]$$

② Monotonic convergence of Emperical distribution :-

ⓐ $$\hat{P}_{2n}(x) = \frac{1}{2n}\sum_{i=1}^{2n} I(X_i = x)$$

$$= \frac{1}{2}\cdot\frac{1}{n}\sum_{i=1}^{n} I(X_i = x) + \frac{1}{2}\frac{1}{n}\sum_{i=n+1}^{2n} I(X_i = x)$$

$$= \frac{1}{2}\hat{P}_n(x) + \frac{1}{2}\hat{P}'_n(x).$$

P.T.O.

using convexity of $D(p||q)$ we have,

$$D(\hat{P}_{2n}||P) = D\left(\tfrac{1}{2}\hat{P}_n + \tfrac{1}{2}\hat{P}_n' \,\Big|\Big|\, \tfrac{1}{2}P + \tfrac{1}{2}P\right)$$

$$\leq \tfrac{1}{2}D(\hat{P}_n||P) + \tfrac{1}{2}D(\hat{P}_n'||P)$$

Taking expectations and using the fact that X_i's are Identically distributed we get,

$$ED(\hat{P}_{2n}||P) \leq ED(\hat{P}_n||P)$$

(b) The trick to this part is similar to part (a) and involves rewriting $\hat{P}_n$ in terms of $\hat{P}_{n-1}$.

$$\hat{P}_n = \frac{1}{n}\sum_{i=0}^{n-1} I(x_i = x) + \frac{I(x_n = x)}{n}$$

or, In general

$$\hat{P}_n = \frac{1}{n}\sum_{i \neq j} I(x_i = x) + \frac{I(x_j = x)}{n}$$

where j ranges from 1 to n.

Summing over 'j' we get,

$$n\hat{P}_n = \frac{n-1}{n}\sum_{j=1}^{n} \hat{P}_{n-1}^j + \hat{P}_n,$$

or,
$$\hat{P}_n = \frac{1}{n}\sum_{j=1}^{n} \hat{P}^j_{n-1}$$

where
$$\sum_{j=1}^{n} \hat{P}^j_{n-1} = \frac{1}{n-1}\sum_{i\neq j} I(x_i = x).$$

Again using convexity of $D(P\|Q)$ and the fact that the $D(\hat{P}^j_{n-1}\|P)$ are identically distributed for all j and hence have the same expected value, we obtain the final result.

(3) 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,
$$\log_n V_n^{1/n} = \frac{1}{n}\log_e V_n = \frac{1}{n}\sum \log_e x_i$$
$$\to E(\log_e(x_i)) \quad \left[\begin{array}{l}\text{By strong} \\ \text{laws of Number}\end{array}\right.$$

Since x_i and $\log_e(x_i)$ are i.i.d and $E(\log_e(x)) < \infty$.

Now,
$$E[\log_e(x_i)] = \int_0^1 \log_e(x)\,dx = -1.$$

P.T.O.

Hence, since 'e^x' is a Continuous function,

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

Thus the "effective" edge length of this solid is e^{-1}.

Note that since the X_i's are Independent, $E[V_n] = \prod E[X_i] = \left(\frac{1}{2}\right)^n$.

Also '$\frac{1}{2}$' is the arithmetic mean of the random variable and '$\frac{1}{e}$' is the Geometric mean.

(4) Monotonicity of entropy per element :-

(a) By the chain rule for entropy,

$$\frac{H(X_1, X_2 \ldots X_n)}{n} = \frac{\sum_{i=1}^{n} H(X_i / X^{i-1})}{n}$$

$$= \frac{H(X_n / X^{n-1}) + \sum_{i=1}^{n-1} H(X_i / X^{i-1})}{n}$$

$$= \frac{H(X_n / X^{n-1}) + H(X_1, X_2, \ldots X_{n-1})}{n} \quad -(1)$$

From stationarity, It follows that for all $1 \leq i \leq n$,

$$H\left(X_n / X^{n-1}\right) \leq H\left(X_i / X^{i-1}\right)$$

which further implies, by averaging both sides,

$$H\left(X_n / X^{n-1}\right) \leq \sum_{i=1}^{n} \frac{H\left(X_i / X^{i-1}\right)}{n-1}$$

$$= \frac{H\left(X_1, X_2, \dots X_{n-1}\right)}{(n-1)} \quad -②$$

From ① and ②,

$$\frac{H\left(X_1, X_2 \dots X_n\right)}{n} \leq \frac{1}{n}\left[\frac{H\left(X_1, X_2 \dots X_{n-1}\right)}{n-1} + H\left(X_1, X_2 \dots X_{n-1}\right)\right]$$

$$= \frac{H\left(X_1, X_2 \dots X_{n-1}\right)}{n-1}.$$

(b) By stationarity $H\left(X_n / X^{n-1}\right) \leq H\left(X_i / X^{i-1}\right)$

$$\Rightarrow H\left(X_n / X^{n-1}\right) = \sum_{i=1}^{n} \frac{H\left(X_n / X^{n-1}\right)}{n}$$

$$\leq \sum_{i=1}^{n} \frac{H\left(X_i / X^{i-1}\right)}{n} = \frac{H\left(X_1, X_2, \dots X_n\right)}{n}$$

5.) Doubly Stochastic Matrices.

(a) Note that $b_j = \sum_i a_i P_{ij} \geq 0$.

$$\sum_j b_j = \sum_j \sum_i a_i P_{ij} = \sum_i a_i \sum_j P_{ij} = 1.$$

Hence 'b' is a probability vector

$$H(b) - H(a) = -\sum_j b_j \log b_j + \sum_i a_i \log a_i$$

$$= \sum_j \sum_i a_i P_{ij} \log\left(\sum_k a_k P_{kj}\right) + \sum_i a_i \log a_i$$

$$= \sum_i \sum_j a_i P_{ij} \log \frac{a_i}{\sum_k a_k P_{kj}} \geq \left(\sum_{i,j} a_i P_{ij}\right) \log \frac{\sum_{ij} a_i}{\sum_{ij} b_j}$$

$$= 1 \log \frac{m}{m}$$

$$= 0$$

follows from log-sum inequality.

(b) If the matrix is doubly stochastic, the substituting $\mu_i = \frac{1}{m}$, we can check that $\mu = \mu P$

(c) If the uniform is a stationary distribution, then

$$\frac{1}{m} = \mu_i = \sum_j \mu_j P_{ji} = \frac{1}{m} \sum_j P_{ji}$$

or $\sum_j P_{ji} = 1$ so that the matrix is doubly stochastic.

6.) The entropy of dog looking for bone:

(a) Chain rule:

$$H(X_0 X_1 \cdots X_n) = \sum_{i=0}^{n} H(X_i / X^{i-1})$$

$$= H(X_0) + H(X_1/X_0) + \sum_{i=2}^{n} H(X_i / X_{i-1}, X_{i-2})$$

since for $i>1$, the next position depends only on the previous 2 (i.e. dog's walk is 2nd order Markov, if the dog's position is the state). Since $X_0 = 0$ deterministically $H(X_0) = 0$ and since the first step is equally likely to be positive or negative,

$H(X_1/X_0) = 1$. Further for $i>1$,

$$H(X_i / X_{i-1}, X_{i-2}) = H(0.1, 0.9)$$

$\Rightarrow$

$$H(X_0 X_1 \cdots X_n) = 1 + (n-1) H(0.1, 0.9)$$

(b)

$$\frac{H(X_0 X_1 \cdots X_n)}{n+1} = \frac{1 + (n-1) H(0.1, 0.9)}{n+1} \rightarrow H(0.1, 0.9)$$

(c) The dog must take at least one step to establish the direction of travel from which it ultimately reverses. Letting S be the no. of steps taken between reversals, we have

$$E(S) = \sum_{s=1}^{\infty} s(0\cdot9)^{s-1}(0\cdot1) = 10$$

Starting at time 0, the expected no. of steps to the first reversal is 11. If we don't count the last step in which the dog reverses his direction, the answer is 10.