PCMI GSS Asymptotic Enumeration: Problem Set 1

Lecturer: Jinyoung Park; TA: Bob Krueger

July 7, 2025

The $\star$'s roughly indicate difficulty. Hints are given as footnotes. Please find a group to work with and don't be afraid to ask questions!

1. $\star$ Recall this theorem from the lecture: if G is an n-vertex d-regular bipartite graph, then $\log_2 i(G) \le (1 + O(1/d))\, n/2$. Give an example which shows that the dependence on d is best possible. ($\star\star$ If we've prove this theorem already, try to follow the proof on your example and notice how the inequalities become sharp.)

2. $\star - \star\star$ You might skim through this exercise if you are already familiar with Shannon entropy. In what follows $\mathbf{X}, \mathbf{Y}, \ldots$ are finitely-supported random variables. Recall that the (binary) entropy of $\mathbf{X}$ is

$$H(\mathbf{X}) = \sum_x \mathbb{P}(\mathbf{X} = x) \log \frac{1}{\mathbb{P}(\mathbf{X} = x)},$$

where the log is taken base 2. Entropy is a measure of the 'information' stored in a random variable, here measured in bits; another interpretation is that entropy is the average 'surprise' upon learning the realization of a random variable. Verify the following starting with only the above definition of entropy.

   (a) *(Uniform maximizes entropy)* Suppose $\mathbf{X}$ is supported on a finite set S. Show that $H(\mathbf{X}) \le \log |S|$, with equality if and only if $\mathbf{X}$ is uniform on S. (This is why entropy is helpful for counting problems: if we want to count $|S|$, we may equivalently calculate the entropy of the uniform random variable on S.)[1]

   (b) *(Chain rule)* Recall the definition of conditional entropy:

$$H(\mathbf{X}|\mathbf{Y}) = \sum_y \mathbb{P}(\mathbf{Y} = y) \sum_x \mathbb{P}(\mathbf{X} = x|\mathbf{Y} = y) \log \frac{1}{P(\mathbf{X} = x|\mathbf{Y} = y)}.$$

   Show that

$$H(\mathbf{X}, \mathbf{Y}) - H(\mathbf{X}) = H(\mathbf{Y}|\mathbf{X}).$$

[1] There is an elementary proof using the inequality $\log(x) \le x - 1$. Alternatively, one can use Jensen's inequality.

(c) *(Additivity for independent variables)* We say that $\mathbf{X}$ and $\mathbf{Y}$ are independent if $\mathbb{P}(\mathbf{X} = x, \mathbf{Y} = y) = \mathbb{P}(\mathbf{X} = x)\mathbb{P}(\mathbf{Y} = y)$ for all x and y. Show that if $\mathbf{X}$ and $\mathbf{Y}$ are independent, then $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{Y})$ and $H(\mathbf{X}, \mathbf{Y}) = H(\mathbf{X}) + H(\mathbf{Y})$.

(d) *(Dropping conditioning)* Show that $H(\mathbf{Y}|\mathbf{X}) \leq H(\mathbf{Y})$, and characterize equality.[2]

(e) *(Subadditivity)* Show that $H(\mathbf{X}, \mathbf{Y}) \leq H(\mathbf{X}) + H(\mathbf{Y})$. Furthermore $H(\mathbf{X}) \leq \sum_i H(\mathbf{X}_i)$.

(f) *(Dropping conditioning, part 2)* Show that $H(\mathbf{Y}|\mathbf{X}, \mathbf{Z}) \leq H(\mathbf{Y}|\mathbf{X})$.

(g) *(Inserting conditioning)* Show that $H(\mathbf{X}|\mathbf{Z}) \leq H(\mathbf{X}|\mathbf{Y}) + H(\mathbf{Y}|\mathbf{Z})$.

(h) *(Refinement)* If $\mathbf{Y}$ determines $\mathbf{X}$, then $H(\mathbf{X}|\mathbf{Y}) = 0$ and $H(\mathbf{Z}|\mathbf{Y}) \leq H(\mathbf{Z}|\mathbf{X})$.

(i) *(Shearer's Inequality)* Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_k)$ be a random vector, and let $\alpha : 2^{[k]} \to \mathbb{R}_{\geq 0}$ be such that $\sum_{A \ni i} \alpha(A) \geq 1$ for all $i \in [k]$. Show that

$$H(\mathbf{X}) \leq \sum_{A \subseteq [k]} \alpha(A) H(\mathbf{X}_A),$$

where $\mathbf{X}_A = (\mathbf{X}_i)_{i \in A}$.[3]

3. $\star\star$ Let $0 \leq \alpha \leq 1/2$. Show that the number of subsets of $[n]$ of size at most αn is at most $2^{h(\alpha)n}$, where $h(\alpha) = \alpha \log_2 \left(\frac{1}{\alpha}\right) + (1 - \alpha) \log_2 \left(\frac{1}{1-\alpha}\right)$.[4] Derive that $\binom{n}{\leq k} \leq \left(\frac{en}{k}\right)^k$.[5]

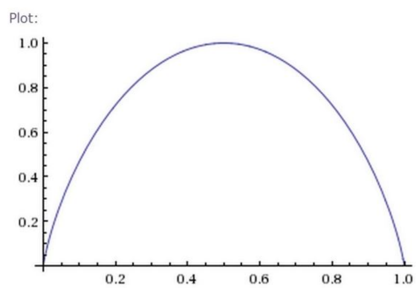

Figure 1: a plot of $h(\alpha)$

4. $\star\star$ Find a bijection between the following two collections.

- (proper) 3-colorings of Q_d with the color of a prefixed vertex v_0 is fixed;
- the graph homomorphisms from Q_d to $\mathbb{Z}$ (in which two vertices a, b are adjacent iff $|a - b| = 1$) with $f(v_0) = 0$.

[2] Use $\log(x) \leq x - 1$ or Jensen's Inequality *judiciously*, in addition to Bayes' Rule.

[3] For each $A \subseteq [k]$, we have $H(\mathbf{X}_A) = \sum_{i \in A} H(\mathbf{X}_i | \mathbf{X}_{A \cap [i-1]}) \geq \sum_{i \in A} H(X_i | X_{[i-1]})$.

[4] Image from D. Galvin, *Three tutorial lectures on entropy and counting*, arXiv 1406.7872

[5] Where have you seen $h(\alpha)$ before? Express the counting problem as an entropy problem and use subadditivity.