Homework 6 Solutions

Note: These solutions are not necessarily model answers. Rather, they are designed to be tutorial in nature, and sometimes contain a little more explanation than an ideal solution. Also, bear in mind that there may be more than one correct solution. The maximum total number of points available is 45.

1. (a) (i) Fix any subset S of 5 vertices. The probability S is a clique in G is p^{10} since S contains 10 edges. There are $\binom{n}{5}$ possible choices for S, so by linearity of expectation, the expected number of 5-cliques in G is $\binom{n}{5}p^{10}$. Solving for $\binom{n}{5}p^{10} = 1$ yields $p^{10} = \Theta(1/n^5)$, so $p = \Theta(1/\sqrt{n})$. *3pts*

(ii) Fix any unordered pair of disjoint subsets of 5 vertices; there are $\frac{1}{2}\binom{n}{5}\binom{n-5}{5}$ choices. The probability that this choice contains a $K_{5,5}$ is p^{25}. Hence, the expected number of $K_{5,5}$ is $\frac{1}{2}\binom{n}{5}\binom{n-5}{5}p^{25}$. Setting this quantity to 1 yields $p^{25} = \Theta(1/n^{10})$, which solves to $p = \Theta(1/n^{2/5})$. *3pts*

(iii) Fix an ordered sequence of n vertices; there are $\frac{1}{2}(n-1)!$ choices (because we may always assume that the first vertex is fixed, and each order is equivalent to its reversal). The probability this induces a Hamiltonian cycle is p^n. Hence, the expected number of Hamiltonian cycles is $\frac{1}{2}(n-1)! \cdot p^n$. Setting this quantity to 1 yields $p = \Theta(1/n)$, using the bounds $(n/e)^n \leq n! \leq n^n$. *3pts*

(b) The result of part (iii) implies that if $p = o(1/n)$ then the expected number of Hamilton cycles tends to zero, and if $p = \omega(n)$ then this expected number tends to ∞. The first of these facts implies, by the usual argument based on Markov's inequality, that $\Pr[G \text{ contains a Ham. cycle}] \to 0$ when $p = o(1/n)$. So we might expect $p = 1/n$ to be a threshold for this property. However, part (iii) doesn't tell us anything about the variance, so we have no way of concluding that $\Pr[G \text{ contains a Ham. cycle}] \to 1$ when $p = \omega(1/n)$; and in fact that is not true! In reality we need p to be larger than $\frac{\ln n}{n}$ to conclude this (by a more sophisticated argument that we won't give here). And again, by a more sophisticated argument, we can also conclude that $\Pr[G \text{ contains a Ham. cycle}] \to 0$ for p all the way up to $\frac{\ln n}{n}$ (which again doesn't contradict the weaker fact in part (iii)). *2pts*

2. (a) Let X_i be the r.v. indicating whether vertex i is isolated. Then $\mathrm{E}[X_i] = (1-p)^{n-1}$, and by linearity of expectation, $\mathrm{E}[X] = n(1-p)^{n-1}$. *2pts*

(b) Write $p = a \cdot \frac{\ln n}{n}$. Note that $\mathrm{E}[X] \sim ne^{-p(n-1)} \sim ne^{-a\ln n} = n^{1-a}$. The case $p = o(\frac{\ln n}{n})$ is equivalent to $a = o(1)$, and thus $\mathrm{E}[X] \sim n^{1-o(1)} \to \infty$. The second case, $p = \omega(\frac{\ln n}{n})$, is equivalent to $a = \omega(1)$ and thus $\mathrm{E}[X] \sim n^{-(\omega(1)-1)} \to 0$. *2pts*

(c) By Markov's inequality we have $\Pr[X \geq 1] \leq \mathrm{E}[X]$, which by part (b) goes to zero in the case $p = \omega(\frac{\ln n}{n})$. Hence $\Pr[X > 0] \to 0$, as required. *2pts*

(d) For any $i \neq j$, $\mathrm{E}[X_iX_j] = (1-p)^{2n-3}$ (there are $2n-3$ possible edges adjacent to either i or j). Hence, $\mathrm{E}[X^2] = n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3}$. Therefore, $\mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = n(1-p)^{n-1} + n(1-p)^{2n-3}(np-1)$. *3pts*

(e) By Chebyshev's inequality, $\Pr[X = 0] \leq \Pr[|X - \mathrm{E}[X]| \geq \mathrm{E}[X]] \leq \mathrm{Var}[X]/\mathrm{E}[X]^2 = \frac{1}{\mathrm{E}[X]} + \frac{np-1}{n(1-p)}$. Now for $p = o(\frac{\ln n}{n})$ we know from part (b) that $\mathrm{E}[X] \to \infty$, so the first term here goes to zero. And the second term is $\frac{np-1}{n(1-p)} \leq \frac{p}{1-p}$, which certainly goes to zero for $p = o(\frac{\ln n}{n})$. Hence we have $\Pr[X = 0] \to 0$, i.e., $\Pr[X > 0] \to 1$, as required. *2pts*

3. (a) Fix a tournament T, and pick a ranking R u.a.r., namely, a random ordering of the n vertices. For each edge e in T, the probability that R disagrees with e is $\frac{1}{2}$. By linearity of expectation, the expected number of edges (over random R) in which R and T disagree is $\frac{1}{2}\binom{n}{2}$. Hence, there must exist a ranking R such that R and T disagree on at most 50% of the edges. *3pts*

(b) Fix n. Fix a ranking R on n vertices, and pick a tournament T u.a.r., i.e., assign a random direction to each edge independently. Let X be the r.v. for the number of edges in which R and T disagree (over random T). Again, $\mathrm{E}[X] = \frac{1}{2}\binom{n}{2}$, by writing X as the sum of independent 0-1 r.v.'s, one for each edge. Applying the Chernoff bound with $\delta = 0.02$ and $\mu = \frac{1}{2}\binom{n}{2}$ yields $\Pr[X < 0.49\binom{n}{2}] \leq \exp(-\frac{1}{2}(0.02)^2 \cdot \frac{1}{2}\binom{n}{2}) = e^{-\Omega(n^2)}$. Taking a union bound over all possible rankings R (there are $n!$ of them), we obtain, *4pts*

$$\Pr_T[\exists \text{ a ranking } R \text{ s.t. } R \text{ and } T \text{ disagree on} < 49\% \text{ of the edges in } T] < n! \cdot e^{-\Omega(n^2)} < 1$$

for sufficiently large n. This implies there exists a tournament T such that every ranking disagrees with at least 49% of the edges in T.

4. (a) The expected number of edges is $\frac{1}{2}pn(n-1) = 8(n-1)$. *1pt*

(b) Clearly, we may write the r.v. X for the number of edges in G as the sum of $\frac{1}{2}n(n-1)$ independent 0-1 r.v.'s (one for each edge). Applying the Chernoff bound with $\mu = 8(n-1)$ and $\delta = \frac{1}{4}$, we obtain $\Pr[X > 10(n-1)] < e^{-(1/4)^2 8(n-1)/3} = 2^{-\Omega(n)}$. Hence, $\Pr[X \leq 10(n-1)] \geq 1 - 2^{-\Omega(n)}$ and the required inequality follows. *2pts*

(c) Let r be the number of red vertices in the coloring. Then the number of pairs of vertices of the same color is $\binom{r}{2} + \binom{n-r}{2}$. Let Y be the r.v. for the number of violated edges in G. Then, $\mathrm{E}[Y] = p(\binom{r}{2} + \binom{n-r}{2}) \geq 2p\binom{n/2}{2} = 4(n-2)$ (assuming that n is even to avoid messy rounding details). *3pts*

Again, Y can be written as a sum of of independent 0-1 r.v.'s Applying the Chernoff bound with $\mu = 4(n-2)$ and $\delta = \frac{3}{4}$, we obtain $\Pr[Y \leq n-2] \leq e^{-(3/4)^2 \cdot 4(n-2) \cdot 1/2} = e^{-9(n-2)/8}$. The required result now follows.

(d) There are 2^n different ways to color the graph. By part (c) and a union bound over colorings, we have *2pts*

$$\Pr_G[\exists \text{ an assignment of colors to } G \text{ with} \leq n-2 \text{ violated edges}] \leq 2^n \cdot e^{-\frac{9}{8}(n-2)} \leq \frac{1}{4}$$

for $n \geq 9$. When an assignment has $n-1$ violated edges, then removing any $n-2$ edges still leaves at least one violated edge. Since with probability $\geq \frac{3}{4}$ every assignment has $\geq n-1$ violated edges, we see that if we remove any $n-2$ edges of G there is still no valid 2-coloring.

(e) Fix a sequence of k distinct vertices $v_1, \ldots, v_k$ in G. The probability (over G) that $(v_1, \ldots, v_k)$ is a cycle in G is p^k. There are at most n^k sequences of k distinct vertices, and for each such sequence, we may assign an indicator variable that tells us whether it corresponds to a k-cycle in G. By linearity of expectation, the expected number of k-cycles in G is at most $n^k \cdot p^k = 16^k$. *3pts*

The expected number of cycles of length at most $\frac{1}{8}\log n$ is bounded by $\sum_{k=3}^{\frac{1}{8}\log n} 16^k \leq 16^{\frac{1}{8}\log n + 1} = 16\sqrt{n}$.

(f) By Markov's inequality, with probability at least $3/4$ (over G), G has at most $64\sqrt{n}$ cycles of length at most $\frac{1}{8}\log n$. We may delete an edge from each of these cycles, thereby removing at most $64\sqrt{n}$ edges. *2pts*

(g) Taking a union bound over the *complements* of the events in parts (b), (d), (f), we know that for all sufficiently large n, with probability at least $1/2 - 2^{-\Omega(n)} > 0$ over G from $\mathcal{G}_{n,16/n}$, G satisfies the properties in parts (b), (d) and (f) simultaneously. This means that by deleting at most $64\sqrt{n}$ edges from G, we obtain a graph G' with the following properties: *3pts*

— the induced subgraph on any subset of $\frac{1}{8}\log n$ vertices of G' is 2-colorable (by part (f));

— G' has at most $10(n-1)$ edges (by part (b));

— G' is not 2-colorable even if we delete any $n - 2 - 64\sqrt{n}$ edges (by part (d); note that we may have already deleted $64\sqrt{n}$ edges from G itself in part (f)).

Finally, note that for sufficiently large n we have (very crudely) $n - 2 - 64\sqrt{n} \geq \frac{1}{2}(n-1) = \frac{1}{20} \times 10(n-1)$, so the number of edges we can delete is at least a 0.05 fraction of the edges of G (and thus also of G'). Thus G' is the required graph G_n, and we have verified the existence of G' by the probabilistic method.

Note: Many students did not use a union bound in this part, but instead multiplied the probabilities of good/bad events happening. This is invalid because none of these events are independent!!!