Solutions for CS174 Homework 2

Solution 1. As in the lecture notes, let X_k be the number of random edges added while there are k connected components, until there are $k-1$ connected components. Let X be the number of edges added to the graph in total until the graph has $\sqrt{n}$ or fewer connected components. So $\mathrm{E}[X] = \sum_{\sqrt{n}\le k\le n} \mathrm{E}[X_k] \le \sum_{\sqrt{n}\le k\le n} \frac{n-1}{k-1} = (n-1)(H_{n-1} - H_{\sqrt{n}-1}) \sim (n\ln n)/2$.

For the variance, $\mathrm{Var}[X] = \sum_{\sqrt{n}\le k\le n} \mathrm{Var}[X_k] \le \sum_{1\le k\le n} \mathrm{Var}[X_k]$. From lecture 9, the latter sum is approximately $n^2\pi^2/6$. Therefore σ_X is at most $n\pi/\sqrt{6}$.

To Apply Chebyshev, we set the probability of exceeding the mean at 0.01, then $t = 10$ in the Chebyshev formula: $\Pr[|X - \overline{X}| \ge t\sigma_x] \le \frac{1}{t^2}$, so $X \ge (n\ln n)/2 + 10n\pi/\sqrt{6}$

Solution 2. The definition of covariance is $\mathrm{Cov}[X,Y] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$.

(a) It's easy to see that $\mathrm{E}[X] = \mathrm{E}[Y] = 1/2, \mathrm{Var}[X] = \mathrm{Var}[Y] = 1/4$. And $\Pr[XY = 1] = 1/3, \Pr[XY = 0] = 2/3$. So $\mathrm{E}[XY] = 1/3$, and $\mathrm{Cov}[X,Y] = 1/12$.

(b) $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + \mathrm{Cov}[X,Y] = 7/12$.

Solution 3. As in the lecture notes, define X_s for each subset $S \subset V$ of 5 vertices as:$X_s = 1$ if S are the vertices of a 5-clique, and $X_s = 0$ otherwise. So $X = \sum X_s$ is the total number of 5-cliques in the graph, and $\mathrm{E}[X] = \binom{n}{5}p^{10} \approx n^5 p^{10}/120$.

At the threshold probability, $\mathrm{E}[X]$ is close to 1, so the threshold value is $p_0 = 120^{0.1}n^{-0.5} \approx 1.6n^{-0.5}$.

Solution 4. Let X to denote the sum of the 1000 tosses. $\mathrm{E}[X] = 3500$. $\mathrm{Var}[X] = 1000 \times ((1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2)/6 = 2900$. So using Markov bound, $\Pr[X > 5000] < 3500/5000 = 0.7$;

Using Chebyshev bound, $\Pr[X > 5000] < \Pr[|X - 3500| > 1500] < 2900/(1500^2) = 1.3 \times 10^{-3}$.

Using Chernoff bound, let Y_i be a dice toss that takes value $1/6, 2/6, 3/6, 4/6, 5/6, 1$. Thus $\mathrm{E}[Y_i] = 3.5/6$. Let Z_i be a bernolli variable with $\Pr[Z_i = 1] = 3.5/6$. Note that function $f(x) = e^{tx}$ for $t > 0$ is convex. For a convex function f, we can see that $E[f(Y_i)] \le E[f(Z_i)]$. So we can extend the Chernoff bound to apply to random variables take values in $[0, 1]$. This bound is called Hoeffding Bound. So $Y = \sum Y_i = X/6$. $\Pr[X > 5000] = \Pr[Y > 5000/6] = \Pr[Y > (1 + 1500/3500)3500/6] < e^{-(3500/6)(1500/3500)^2/4} \doteq e^{-27}$.