Random Graphs
Assignment 3. Solutions

1. Let $X_1, \ldots, X_n$ be independent random variables with $\Pr[X_i = +1] = \Pr[X_i = -1] = \frac{1}{2}$. Set $X = X_1 + \ldots + X_n$. Find $E[X^2]$ precisely. Find $E[X^4]$ precisely. [Idea: Expand and use linearity of expectation.]
Solution: $E[X^2] = \sum_{i,j} E[X_i X_j]$. But $E[X_i X_j]$ is zero unless $i = j$ in which case it is one so $E[X^2] = \sum_1^n 1 = n$. Similarly $E[X^4] = \sum_{i,j,k,l} E[X_i X_j X_k X_l]$. This expectation is zero except when it is a perfect square. This occurs if $i = j = k = l$ ($n$ times) or $i = j \neq k = l$ ($n(n-1)$ times) or $i = l \neq j = k$ or $i = k \neq j = l$ so a total of $n + 3n(n-1)$ times. Each time the expectation is then one so $E[X^4] = n + 3n(n-1)$.

2. Find an asymptotic formula for 
$$\frac{2n^{1/2}}{\sum_{k=n^{1/2}}^{2n^{1/2}} (n)_k n^{-k}}$$
by parametrizing $k = cn^{1/2}$ and turning it into an integral which can be evaluated numerically.
Solution: Set $f(k) = (n)_k n^{-k}$ for convenience. Parametrizing $k = xn^{1/2}$ we have $f(k) \sim e^{-x^2/2}$. We transform the sum into an integral (you don’t have to be this formal) as follows: Let $x_0 = 1, \ldots, x_u = 2$ be the values of $kn^{-1/2}$ for $k$ integral, $n^{1/2} \leq k \leq 2n^{1/2}$, set $\Delta x = x_{i+1} - x_i = n^{-1/2}$, so the sum is $n^{1/2} \sum_{i=0}^u g(x_i)(\Delta x)$ where $g(x) := f(xn^{1/2}) \sim e^{-x^2/2}$. The value $i = u$ is asymptotically negligible (compared to the whole sum) and so as $n \to \infty$ the sums approach the integral (you may recall from First Year Calculus that this is really the definition of the definite integral) and so the sum approaches $\int_{1}^{2} e^{-x^2/2} dx$.

3. Now we go to the complete sum by showing the edge effects are negligible.

(a) Show 
$$\lim_{\epsilon \to 0} \lim_{n \to \infty} n^{1/2} \sum_{k=1}^{\epsilon n^{1/2}} (n)_k n^{-k} = 0$$
by using an appropriate upper bound for the addends.

**Solution:** Bound the addends by one! The sum is then at most $\epsilon n^{1/2}$.

(b) (*) Show
\[\lim_{K \to \infty} \lim_{n \to \infty} n^{-1/2} \sum_{k=K^{n^{1/2}}}^{n} (n)_{k}n^{-k} = 0\]

by using an appropriate upper bound for the addends.

**Solution:** Bounding this upper sum is harder. Even though $f(k)$ is decreasing there are $\Theta(n)$ terms so bounding by the term at $K\sqrt{n}$ doesn’t work. Fortunately, the bounding of $\ln(1 - \frac{i}{n})$ by $-\frac{i}{n}$ is an overestimate (all the terms in the Taylor Series for $\ln(1 - \epsilon)$ are negative) and so we have

\[f(k) \leq \exp\left(\sum_{i=1}^{k-1} \frac{i}{n}\right) = e^{-k(k-1)/2n}\]

for all $k$ (not just $k = \Theta(\sqrt{n})$. That $k = 1$ is annoying so let us give ground (part of the art is that when we want to show something is negligible we can give ground to make it simpler and if it is still negligible we are done – we are not looking for an asymptotic formula for the negligible parts, only that they are negligible!) and say

\[f(k) \leq h(k) := e^{-k^2/3n}\]

which will hold for all $k \geq K\sqrt{n}$, indeed, all $k \geq 3$. Now $h(k)$ is a decreasing function for its sum from $K\sqrt{n}$ to $n$ is at most the sum from $K\sqrt{n}$ to $\infty$ which is bounded (to be really formal) by the integral from $K\sqrt{n} - 1$ to $\infty$ and that approaches $\sqrt{n} \int_{K\sqrt{n}}^{\infty} e^{-t^2/3}dt$ which can be less than $\epsilon_1 \sqrt{n}$ for arbitrarily small $\epsilon_1$ by making $K$ large.

**A More General Approach:** This argument was serendipitous in that the error term in the $\ln(1 - \epsilon)$ approximation was going in the right direction. Here is another approach that works quite widely. We know that

\[f(k) \sim e^{-k^2/2n}\]

as long as $k = o(n^{2/3})$. Let us cut off the sum at

\[L = 10n^{1/2} \sqrt{\ln n}\]
At \( k = L \) the term \( f(L) \sim n^{-50} \). As \( f \) is decreasing we have
\[
\sum_{k=L}^{n} f(k) \leq nf(L) = o(1)
\]
So now we can just look at
\[
\sum_{k=K\sqrt{n}}^{L} f(k)
\]
But in that range (as \( L \ll n^{2/3} \)) we do have \( f(k) \sim e^{-k^2/2n} \) and so we can bound it, indeed, we can find its asymptotic value. which is \( \sim \sqrt{n}\sqrt{\pi/2} \) as desired.

(c) Find an asymptotic formula for
\[
\sum_{k=1}^{n} (n)k n^{-k}
\]
by splitting it into the ranges \( k < \epsilon n^{1/2}, \epsilon n^{1/2} \leq k \leq Kn^{1/2} \) and \( Kn^{1/2} < k \leq n \) and then taking appropriate limits.
Solution: The side ranges become negligible as \( \epsilon \to 0 \) and \( K \to \infty \) so we get the sum is asymptotic to \( S\sqrt{n} \) with
\[
S = \int_{0}^{\infty} e^{-t^2/2} dt = \sqrt{\pi/2}
\]

4. Prove, for \( m = m(n) \) as large as you can, the existence of an \( n \times n \) matrix \( A \) of zeroes and ones with \( m \) ones which does not contain a \( 3 \times 3 \) submatrix of all ones. Use the alteration method: make each entry one with probability \( p \) and then for each such submatrix change a one to zero. When you optimize [using Calculus!] your final answer should be of the form \( m \sim an^b \) for some reasonable \( a, b \).
Solution: Let \( X \) be the number of ones and \( Y \) the number of \( 3 \times 3 \) submatrices of all ones. Then \( E[X] = n^2p \) and \( E[Y] = \binom{n}{3}^2 p^9 \sim n^6p^9/36 \). Thus
\[
E[X - Y] \sim n^2p - n^6p^9/36
\]
We use calculus to optimally pick \( p = 2^{1/4} n^{-1/2} \) giving \( E[X - Y] \sim (8/9)2^{1/4}n^{3/2} \) so there is an \( A \) with at least that many ones having no \( 3 \times 3 \) submatrix of all ones.
5. We are given \(m = 2^{n-1}k\) sets, each of size \(n\), in a universe \(\Omega\). Consider the following randomized algorithm for coloring: First color each point \(v \in \Omega\) randomly. Now, for each monochromatic set \(e\), select a random vertex \(v \in e\) and switch its color. Call the algorithm a failure if some set \(e\) originally had all or all but one vertex the same color and ended with all vertices that color. Find \(k\) as large as you can (as an asymptotic function of \(n\)) so that the failure probability is less than one. (Note that this, unfortunately, does not give us any result on \(m(n)\) since there are other ways that a set \(e\) could end up monochromatic.)

**Solution:** It is tautologically impossible for \(e\) to be originally all Red and stay Red. Let \(B_e\) be that \(e\) was originally all but one vertex Red and became Red. Say \(e\) blames \(f\) (event \(BL_{e,f}\)) if \(B_e, f\) was Blue, \(e, f\) overlap in a single vertex \(v\) and \(v\) was selected by \(f\) to change color. If \(B_e\) then some \(BL_{e,f}\). Now \(\Pr[BL_{e,f}] \leq 2^{2-2n}n^{-1}\) as the coloring of \(e \cup f\) is one of two possibilities (as Red could be Blue) and then \(f\) must select \(v\). Thus

\[
\Pr[\lor B_e] = \Pr[\lor BL_{e,f}] \leq \sum \Pr[BL_{e,f}] \leq m^2 2^{2-2n}n^{-1} = k^2 n^{-1} < 1
\]

when \(k < \sqrt{n}\).

6. Set \(X = \sum_{i=1}^{n} X_i\) where \(X_i = \pm 1\) uniformly and independently. Bound \(\Pr[X > \frac{n}{2}]\) as follows.

(a) Find a closed form for \(E[e^{\lambda X_i}]\).

**Solution:** \(\cosh \lambda\).

(b) Find a closed form for \(E[e^{\lambda X}]\).

**Solution:** \(\cosh^n \lambda\)

(c) Use the Chernoff Bound \(\Pr[X > a] < E[e^{\lambda X}] e^{-\lambda a}\) with \(a = \frac{n}{2}\). Use Calculus (this gets a little messy to put in closed form, full points for numerical answers) to select the optimal \(\lambda\).

**Solution:** \(\Pr[X > a] < (\cosh \lambda e^{-\lambda/2})^n\). We minimize by minimizing the logarithm. The \(n\) goes out and we need to minimize \(g(\lambda) = \ln(\cosh \lambda) - \frac{1}{2} \lambda\). The minimum occurs (we actually need some further argument to show this is a minimum) when

\[
0 = g'(\lambda) = \tanh \lambda - \frac{1}{2}
\]

\((\tanh \lambda := (e^\lambda - e^{-\lambda})/(e^\lambda + e^{-\lambda})\) is an increasing function and this has a unique solution.)
(d) Compare this with the lower bound

$$Pr[X \geq \frac{n}{2}] \geq Pr[X = \frac{n}{2}] = 2^{-n} \left( \frac{n}{3n^4} \right)$$

showing that the upper and lower bounds have the same main terms.

**Solution:** Stirling gives

$$2^{-n} \left( \frac{n}{3n^4} \right) \sim cn^{-1/2}2^{nH(1/4)}$$

where

$$H(p) := -p\log p - (1 - p)\log(1 - p)$$

is the entropy function. The Chernoff bound was $e^{ng(\lambda)}$ where $\lambda$ is the inverse hyperbolic tangent of one half and some (omitted!) calculations show that $g(\lambda) = H(1/4)\ln(2)$. Actually, there is nothing special about $1/2$, the Chernoff bounds give the right exponential term for $Pr[X > bn]$ for any fixed $b \in (0, 1)$. 