
Probabilistic Method: Problem Set I

Youngduck Choi
CIMS
New York University
yc1104@nyu.edu

Abstract

This work contains solutions to the problem set I of Probabilistic Method 2016 at Courant Institute of Mathematical Sciences.

Question 1.

1. (-) Let X_1, \dots, X_n be independent random variables with $\Pr[X_i = +1] = \Pr[X_i = -1] = \frac{1}{2}$. Set $X = X_1 + \dots + X_n$. Find $E[X^2]$ precisely. Find $E[X^4]$ precisely. [Idea: Expand and use linearity of expectation.]

Solution. By expanding the terms and using linearity of expectation, we have

$$\begin{aligned} E[X^2] &= E\left[\left(\sum_{i=1}^n X_i\right)^2\right] \\ &= E\left[\sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j\right] \\ &= \sum_{i=1}^n E[X_i^2] + 2 \sum_{1 \leq i < j \leq n} E[X_i X_j]. \end{aligned}$$

For

Question 2.

2. Find an asymptotic formula for

$$\sum_{k=n^{1/2}}^{2n^{1/2}} (n)_k n^{-k}$$

by parametrizing $k = cn^{1/2}$ and turning it into an integral which can be evaluated numerically. (You can leave it in the form of a definite integral if you wish.) (See Asymptopia, Chapter 4)

Solution.

Question 3.

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

(a) Show

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} n^{-1/2} \sum_{k=1}^{\epsilon n^{1/2}} (n)_k n^{-k} = 0$$

by using an appropriate upper bound for the addends.

(b) (*) Show

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} n^{-1/2} \sum_{k=Kn^{1/2}}^n (n)_k n^{-k} = 0$$

by using an appropriate upper bound for the addends.

(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. (You may assume the previous parts.)

Solution.

Question 4.

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 3×3 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. Consider a random $n \times n$ matrix, M_n , obtained by assigning each entry independently either 1 or 0, where the probability of assigning 1 is p . Let X be a random variable, which counts number of 1s in the matrix, and Y be a random variable, which counts number of 3×3 submatrices of all 1s. For any 3×3 submatrix S , let Y_S be the indicator random variable for the event for which the submatrix S has entries of all 1s, so that $Y = \sum Y_S$. By Linearity of Expectation, we have

$$E[Y] = \sum E[Y_S] = \binom{n}{3}^2 p^9.$$

Clearly, $E[X] = n^2 p$. Therefore, again by Linearity of Expectation, it follows that

$$E[X - Y] = n^2 p - \binom{n}{3}^2 p^9 = f(p).$$

Hence, there exists a random assignment, for which the number of 1s minus the number of 3×3 submatrices of 1 is at least $f(p)$. Fix such a coloring. Select one entry from each submatrix and change to 0. This leaves the matrix with at least $f(p)$ entries with 1.

We now optimize this result by maximizing $f(p)$ with respect to p . Observe that f is concave with respect to $p \in [0, 1]$. Solving for the local maxima by setting the first-order derivative equals to 0, we get that f is maximized at $p^* = (\frac{n^2}{9\binom{n}{3}^2})^{\frac{1}{8}} = (\frac{2}{(n-1)(n-2)})^{\frac{1}{4}}$. Substituting p^* back into $f(p)$, we obtain

$$\begin{aligned} f(p^*) &= n^2 \left(\frac{2}{(n-1)(n-2)} \right)^{\frac{1}{4}} - \binom{n}{3}^2 \frac{n^2}{9\binom{n}{3}^2} \left(\frac{2}{(n-1)(n-2)} \right)^{\frac{1}{4}} \\ &= \frac{8}{9} n^2 \left(\frac{2}{(n-1)(n-2)} \right)^{\frac{1}{4}} \\ &\sim \frac{8}{9} 2^{\frac{1}{4}} n^{\frac{3}{2}}. \end{aligned}$$

Recall that $m(n)$ be the minimum number of 1 in $n \times n$ matrix, such that there must exist a 3×3 submatrix of all 1s. We have shown that $m(n) = \Omega(n^{\frac{3}{2}})$. \square

Question 5.

5. We are given $m = 2^{n-1}k$ sets, each of size n , in a universe Ω . Consider the following randomized algorithm for a 2-coloring: First color each point $v \in \Omega$ randomly. Now, for each monochromatic set e , select a random vertex $v \in e$. Each such selected v (regardless of how often it was selected) has its color (definitely, no probability here) flipped. 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. Note that we use the language of n -uniform hypergraph, which is an equivalent problem as the problem under consideration. Before preceeding with the main part of the proof, we define an object, called a conflicting pair, in a two-coloring scheme. We say that an ordered pair of edges (e, f) a conflicting pair, if e is monochromatic with some color k , and for some $v \in e \cap f$, $f \setminus \{v\}$ is monochromatic with the color not k . We call the vertex v the peacebreaker of (e, f) .

We now consider the randomized algorithm. Fix (e, f) an ordered pair of edges from the hypergraph. Notice that the algorithm will fail, if, after the initial coloring, (e, f) is a conflicting pair with a peacebreaker v , and at the second stage chooses the peacebreaker v to flip its color, exactly when it reviews e . The event just described, which we denote as $T_{(e,f)}$, is hence a sufficient condition for a failure to occur. Therefore, the probability of $T_{(e,f)}$ will serve as an upper bound to the failure probability. Note that the probability of $T_{(e,f)}$ is given by

$$\begin{aligned} P(T) &= P(\{(e, f) \text{ is a conflicting pair after the initial coloring}\}) \\ &\cdot P(\{v \text{ is chosen at the review of } e\} | \{(e, f) \text{ is a conflicting pair after the initial coloring}\}) \\ &\leq 2^{2-2n} n^{-1}. \end{aligned}$$

One should note that the last quantity is an inequality, as it might be that there is a no intersection between e and f to begin with, which gives 0 probability for (e, f) being a conflicting pair after the initial coloring. The equality is achieved when e and f share exactly one vertex. As there are m^2 ordered pairs of edges in the hypergraph, by the subadditivity of probability, we obtain

$$\begin{aligned} P(\{\text{failure probability}\}) &\leq P\left(\bigcup_{(e,f) \in E \times E} T_{(e,f)}\right) \\ &\leq \sum_{(e,f) \in E \times E} P(T_{(e,f)}) \\ &\leq m^2 2^{2-2n} n^{-1}, \end{aligned}$$

where E denotes the edge set. If $m^2 2^{2-2n} n^{-1} < 1$, we can ensure that the failure probability is less than 1. Substituting $2^{n-1}k$ for m , we have

$$(2^{n-1}k)^2 2^{2-2n} n^{-1} < 1,$$

which is equivalent to $k < \sqrt{n}$. Hence, by taking $k = \sqrt{n} - 1$, we can have a failure probability less than 1. We have shown that $k = \Omega(\sqrt{n})$. \square

Question 6.

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 $E[e^{\lambda X_i}]$.
- (b) Find a closed form for $E[e^{\lambda X}]$.
- (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 λ .
- (d) Compare this with the lower bound

$$\Pr[X \geq \frac{n}{2}] \geq \Pr[X = \frac{n}{2}] = 2^{-n} \binom{n}{\frac{3n}{4}}$$

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

Solution.