Probabilistic Method: Problem Set IV

Youngduck Choi
CIMS
New York University
yc1104@nyu.edu

Abstract

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

Question 1.

1. Let P, Q, R, S be uniformly and independently selected from the unit square. Let $f(\epsilon)$ be the probability that triangles PQR and QRS *both* have area less than ϵ. Find the asymptotics of $f(\epsilon)$ (neglecting constant factors) as ϵ approaches zero. [Idea: Integrate over $r = |QR|$.] [Warning: For all r the probability that PQR, QRS have area less than ϵ is at most one.]

Solution. From the analysis of the combinatorial geometry section 3.3, it follows that

$$Pr(b \leq |QR| \leq b + db) \leq 2\pi b db,$$

where b denotes the distance between P and R. Given the distance b, we must have $b < \frac{2\epsilon}{b}$ to ensure that the area is less than ϵ. An upper bound to the area of such region is $4\frac{2\epsilon}{b}\sqrt{2} = \frac{4\sqrt{2}\epsilon}{b}$, which can be seen from using the $\sqrt{2}$ middle strip. Now, the probability that S lies in such region is thus, $\max(\frac{4\sqrt{2}\epsilon}{b}, 1)$. As we need to compute the probability of both PQR and QRS having an area smaller than ϵ, the total probability is bounded by

$$f(\epsilon) \quad \leq \quad \int_{b=0}^{\sqrt{2}} 2\pi b[\max(\frac{4\sqrt{2}\epsilon}{b}, 1)]^2 db.$$

When $b \leq \frac{\epsilon}{4\sqrt{2}}$, we have the max term is simply 1. Hence, using the additivity of integral, we have

$$f(\epsilon) \quad \leq \quad \int_{b=0}^{4\sqrt{2}\epsilon} 2\pi b[\max(\frac{4\sqrt{2}\epsilon}{b}, 1)]^2 db. + c\int_{b=4\sqrt{2}\epsilon}^{1} \frac{\epsilon^2}{b} db,$$

where c is the constant associated from the original integral. Now, observe that the first integral is $O(\epsilon^2)$ and the second integral is $O(\epsilon^2 \ln(\epsilon))$. Therefore, we have shown that $f(\epsilon) = O(\epsilon^2 \ln(\epsilon))$. $\square$

Question 2.

2. Let X be the number of triangles in $G(n,p)$ with $p = c/n$. Find both the precise and the asymptotic [c fixed, $n \to \infty$, in terms of c] values for the expectation and variance of X.

Solution. For every 3-set S of vertices in $G(n,p)$, let A_S be the event that S is a triangle. In particular, we have $X = \sum_S X_S$. By Linearity of Expectation, obtain

$$E[X] = \sum_S E[X_S] = \binom{n}{3} p^3 = \binom{n}{3} (\frac{c}{n})^3 \sim \frac{1}{6} c^3.$$

Now, by definition of variance, we have

$$Var[X] = \sum_S Var[X_S] + \sum_{S \neq T} Cov[X_S, X_T].$$

Using the variance formula for discrete random variable, we have that

$$\sum_S Var[X_S] = \binom{n}{3} p^3 (1 - p^3).$$

As $p = o(1)$, we have that $(1 - p^3) = o(1)$. Therefore, we can further deduce

$$Var[X_S] = p^3(1-p^3) \sim p^3 \text{ and } Var[X_S] \sim E[X_S] \sim \frac{1}{6}c^3.$$

Now, observe that covariance is 0 for S, T pair, where $|S \cap T| \neq 2$. Now, for S, T pair, where $|S \cap T| = 2$, we have, by definition of covariance,

$$Cov(X_S, X_T) = E[X_S X_T] - E[X_S]E[X_T] = p^5 - p^6.$$

Since there are $\binom{n}{3} 3(n-3)$ choices (fix the first triangle, pick the one that will not be shared, and choose the remaining one from the rest of the graph), we finally have

$$\sum_{S \neq T} Cov(X_S, X_T) = \binom{n}{3} 3(n-3)(p^5 - p^6) = o(1).$$

Therefore, we can conclude that $Var[X] \sim \frac{c^3}{6}$ as well. Reminds me of Poisson, but not gonna think too hard about it for now. $\square$

Question 3.

5. Let X_i, $1 \le i \le n$, be i.i.d. uniform on $\{1, \ldots, 6\}$ (throws of a fair die), $Y_i = X_i - \frac{7}{2}$ (to move to zero mean) and $Y = \sum_{i=1}^{n} Y_i$. Use Chernoff Bounds to give $A = A(n)$ as small (asymptotically) as possible (include the constant factor!) so that

   (a) $\Pr[Y > A] < n^{-1}$

   (b) $\Pr[Y > A] < n^{-10}$

   (c) $\Pr[Y > A] < e^{-\sqrt{n}}$

Solution. With simple computation, we can see that

$$\sigma_i^2 = \frac{35}{12} \text{ and } \sigma_2 = \frac{35}{12}.$$

As Y_i are uniformly bounded, by the use of Chernoff bound, it follows that

$$P(Y > a\sigma) < e^{-\frac{a^2}{2}(1+o(1))}.$$

Therefore, we must have $\frac{a^2}{2} = \ln(n), 10\ln(n), \sqrt{n}$. Solving these respectively, we obtain that $a = (2\ln(n))^{\frac{1}{2}}, (20\ln(n))^{\frac{1}{2}}, 2^{\frac{1}{2}}n^{\frac{1}{4}}$. $\square$