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.

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 $h < \frac{2\epsilon}{b}$ 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 $O(\epsilon^2 \ln(\epsilon^{-1}))$. Therefore, we have shown that $f(\epsilon) = O(\epsilon^2 \ln(\epsilon^{-1}))$. $\square$

1. Here is a problem from work done by Roberto Oliveira, who received his Ph.D. under my supervision some years ago and is now at IMPA in his home town (lucky guy!) of Rio de Janiero. The exponential distribution Exp(a) is the positive distribution with density function $f(t) = ae^{-at}$. We let X_a denote this distribution. Calculus exercise: $E[X_a] = a^{-1}$. Now set $X = \sum_{i=1}^{\infty} X_{i^2}$ where the X_{i^2} are assumed to be mutually independent. (This is finite as $E[X] = \sum i^{-2}$ converges.) Our object will be to use Chernoff bounds to get an asymptotic upper bound for $\Pr[X < \epsilon]$ as $\epsilon \to 0^+$. (This is the kind of problem where it is very easy to be "too precise" and then get totally lost. Please follow the bounds indicated parenthetically below. You should still get a rigorous asymptotic upper bound – it turns out that the bound achieved is quite good.)

   (a) Find $E[e^{-\lambda X_a}]$ in closed form. (Do this precisely.)
   (b) Find $E[e^{-\lambda X}]$ as an infinite product. (Do this precisely.)
   (c) Setting $\lambda = K^2$ with K a positive integer bound $E[e^{-\lambda X}]$ from above by using the first K terms of the infinite product. (You should have a numerator and denominator product. For the numerator use Stirling's Formula. For the denominator observe that all the terms are between K^2 and $2K^2$ and lower bound the denominator product by taking each term as K^2.)
   (d) Now you will apply the Chernoff bound ($\lambda > 0$) $\Pr[X \le \epsilon] \le E[e^{-\lambda X}]e^{\lambda\epsilon}$. Using the upper bound just found find a value of K (as a function of ϵ) that yields a "good" upper bound on $\Pr[X < \epsilon]$. (Ignore the requirement that K needs to be integral. Also, in finding a good value of K look at just the main terms from Stirling (i.e., not the square root term) and find that K that does best there. Your final answer should be something like $e^{-1/\epsilon}$. Good luck!)

Question 2.

2. Set $\Omega = [n] \times [n]$. Define a random set $C \subset \Omega$ by

$$\Pr[(x,y) \in C] = p = \frac{c}{n}$$

the events $(x,y) \in C$ mutually independent. A horozontal bond is a pair $(x,y),(x+1,y) \in C$ and a vertical bond is a pair $(x,y),(x,y+1) \in C$. Find the expected number of bonds. Use Janson's Inequality to bound the probability there are no bonds in both directions and find the limiting probability as $n \to \infty$.

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

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

Now, by definition of variance, we have

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

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

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

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

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

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,

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

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

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

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 $G \sim G(n,p)$ with $p = cn^{-2/3}$ and let v, w be two distinct fixed vertices of G. Use Janson's Inequality to find the limiting probability that v, w are *not* joined by a path of length three.

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 $a = (2\ln(n))^{\frac{1}{2}}, (20\ln(n))^{\frac{1}{2}}, 2^{\frac{1}{2}}n^{\frac{1}{4}}$. □