Probabilistic method in combinatorics and algorithmics
WS 2017/18

Exercise sheet 5

Exercises for the exercise session on 6/12/2017

Problem 5.1. Let $X \sim \mathrm{Bin}(n,p)$ for some fixed $p \in (0,1)$. Recall that in the proof of Theorem 4.1 in the lecture, we used that the function $\varphi\colon (-1,\infty) \to \mathbb{R}$ defined by $\varphi(y) := (1+y)\ln(1+y) - y$ satisfies

$$\varphi(y) \geq \frac{y^2}{2\left(1+\frac{y}{3}\right)} \geq 0 \qquad \forall y > -1.$$

(a) Complete the proof of Theorem 4.1 by proving the 'lower tail bound'

$$\mathbb{P}[X \leq \mu - t] < \exp\left(-\frac{t^2}{2\mu}\right) \tag{1}$$

for $\mu = \mathbb{E}[X]$ and all $t > 0$.

To that end, prove first that this bound is trivial for $t > \mu$ and verify directly that it is also true for $t = \mu$. (*Note:* The corresponding argument for $t = n - \mu$ in the proof of the 'upper tail bound' was missing in the lecture.) For $t < \mu$, show that

$$\mathbb{P}[X \leq \mu - t] \leq \exp\big(u(\mu - t)\big)(pe^{-u} + 1 - p)^n$$

for every $u \geq 0$. Choose u such that

$$e^{-u} = (\mu - t)\frac{1-p}{p(n-\mu+t)}$$

(check that this value of u is indeed non-negative!) and deduce that

$$\mathbb{P}[X \leq \mu - t] \leq \exp\left[-\mu\varphi\left(-\frac{t}{\mu}\right) - (n-\mu)\varphi\left(\frac{t}{n-\mu}\right)\right].$$

(The corresponding calculations in the lecture were rather long; you do not have to repeat them in full length, just point out the differences.) Use this to prove (1).

(b) Prove that

$$\mathbb{P}[X \geq \mathbb{E}[X] + t] < \exp\left(-\frac{t^2}{2\left(\mathrm{Var}[X] + \frac{t}{3}\right)}\right)$$

for all $t > 0$ (i.e. the 'upper tail' case of the Chernoff bound 2 in the special case that the X_i are i.i.d.).

Hint. As in the proof for the Chernoff bound 1, you might assume that $t < n - \mathbb{E}[X]$ (the case $t = n - \mathbb{E}[X]$ again being a non-trivial calculation, which you are allowed to skip). Start from

$$\mathbb{P}[X \leq \mu - t] \leq \exp\left[-\mu\varphi\left(\frac{t}{\mu}\right) - (n-\mu)\varphi\left(-\frac{t}{n-\mu}\right)\right]$$

to prove the desired bound (up to that point, the proofs are identical).

Problem 5.2. Prove the following variant of Chernoff bounds.
Suppose that $X_1, \dots, X_n$ are independent random variables, where each X_i has only a finite number of possible values $x_{i,1}, \dots, x_{i,m_i} \in [-1, 1]$ and satisfies $\mathbb{E}[X_i] = 0$. Let $X = X_1 + \dots + X_n$. Then for $0 \le t \le 2\mathrm{Var}[X]$,

$$\mathbb{P}[X \ge t] \le \exp\left(-\frac{t^2}{4\mathrm{Var}[X]}\right).$$

Hint. Start as in the proof of the Chernoff bound 1 from the lecture. Prove that $\mathbb{E}[\exp(uX_i)]$ is bounded by $1 + u^2\mathrm{Var}[X_i]$ if $u \le 1$.

Problem 5.3. Let $p = p(n) \in (0, 1)$ be given. For $t > 0$ and a fixed vertex v of $G(n, p)$, compare the bounds on $\mathbb{P}[|d(v) - \mathbb{E}[d(v)]| \ge t]$ provided by Chebyshev's inequality and by the Chernoff bounds 1 and 2. In each of the three cases, how large does t have to be in order to deduce that

$$\mathbb{P}[|d(v) - \mathbb{E}[d(v)]| \ge t] = o(1)?$$

How large does t have to be if we want to prove that

$$\mathbb{P}[\exists v\colon\ |d(v) - \mathbb{E}[d(v)]| \ge t] = o(1)? \tag{2}$$

Are there functions $p(n)$ for which the minimum requirements for t in (2) from the two Chernoff bounds coincide?

Problem 5.4. Denote by T the number of triangles in $G(n, \frac{1}{\sqrt{n}})$. Prove that for every $\varepsilon > 0$, we have

$$T \in \left[\frac{1-\varepsilon}{6} n^{\frac{3}{2}}, \frac{1+\varepsilon}{6} n^{\frac{3}{2}}\right]$$

with probability $1 - o(1)$. If we let $\varepsilon = \varepsilon(n) \to 0$, how fast can ε tend to 0 if we still want the same result?
Hint. Prove first that the number of triangles containing a fixed vertex is concentrated around its expectation. To this end, you can use that for $p(n) = \frac{1}{\sqrt{n}}$, (2) holds for some value $t = \Theta(n^{1/4}\sqrt{\ln n})$.

Problem 5.5. Suppose we place n balls in n bins, where each ball chooses its bin uniformly at random and independently from the other balls. Prove that for each $\varepsilon > 0$,

$$\mathbb{P}\left[\exists \text{ a bin with at least } \left(\frac{3}{2} + \varepsilon\right) \ln n \text{ balls}\right] = o(1).$$

By how much can we improve the lower bound on the number of balls in a bin if we want to use Chernoff bounds?
If we have n^2 balls in total, for what k can we prove that

$$\mathbb{P}\left[\exists \text{ a bin with at least } k \text{ balls}\right] = o(1)?$$