Probability and Statistics

Tsinghua Math Camp Prof. Paul Horn

Homework 5: Due Wednesday 8/2/2017

Note: You should get to know your coaches, Gavin St. John (`gavin.stjohn@du.edu`) and Zhe Liu (刘喆) (`z-liu16@mails.tsinghua.edu.cn`). If you need to contact me, my email address is `paul.horn@du.edu` or you may ask via WeChat.

I want very much to learn all of your names, but please be patient with me as I am not so good at pronouncing Chinese and I am very bad at remembering names (even non-Chinese names). So please help me by telling me your name very slowly and being patient with me as I try to learn and pronounce it.

Homework problems: To be turned in.

1. Suppose that G is a graph on $2n$ vertices where each edge is present independently with probability $\frac{1}{2}$. Note that if X is a subset of n vertices of G, the expected number of edges between X and $\bar{X}$ is n^2. Prove that, with probability tending to one as $n \to \infty$, *every* subset of n vertices has at least $n^2 - f(n)$ edges leaving it for $f(n)$ as small as possible.

   Notes/hints: Reworded: show that for every X of size n, if $e(X, \bar{X})$ is the number of edges between X and $\bar{X}$, then $|e(X, \bar{X}) - n^2| \leq f(n)$ with probability $1 - o(1)$. You may find it useful to know that $\binom{2n}{n} \sim \sqrt{\frac{2}{\pi}} \frac{2^{2n}}{\sqrt{2n+1}}$, although it may be enough to know just that $\binom{2n}{n} \leq 4^n$. Use the Chernoff bound proven in class. (You need not prove these estimates although the second is easy. (Why?))

2. Suppose X is $N(0, 1)$.

   (a) Find the pdf and expectation container of $Y = X^2$. (Hint/note: Recall that $\mathbb{P}(X^2 \leq x) = \mathbb{P}(-\sqrt{x} \leq X \leq \sqrt{x})$.

   (b) Let $Y_k = X_1^2 + X_2^2 + \cdots + X_k^2$, where X_i are independent $N(0, 1)$ random variables. Find the expectation container $\varphi_{Y_k}(t)$.

   (c) Find $\mathbb{E}[Y_k]$ (where Y_k is as in the previous part.) and $\mathrm{Var}(Y_k)$. (Use your expectation container).

3. Let $X = X_1 + \cdots + X_n$ where X_i are independent with $\mathbb{P}(X_i = 1) = p_i$ and $\mathbb{P}(X_i = 0) = 1 - p_i$. Prove the Chernoff upper estimate

$$\mathbb{P}(X - \mathbb{E}[X] \geq \lambda) \leq e^{-\lambda^2/2(\mathbb{E}[X]+\lambda/3)}.$$

   To do this, try the following:

   - Follow the proof of the Chernoff bounds, applying e^{tx} instead of e^{-tx} to both sides.

- At some point you need to bound

$$\mathbb{E}[e^{t(X_i-p_i)}] = (p_i e^{t(1-p_i)} + (1-p_i)t^{-tp_i}) = e^{tp_i}(p_i e^t + 1 - p_i).$$

Use the fact that $e^t = 1 + t + t^2/2 + t^3/3! + \cdots \leq 1 + t + \frac{t^2}{2}(\frac{1}{1-t/3})$ if $t < 3$, and then continue as the proof previously and choose a new t to minimize the result.

4. Suppose X and a random variable with $\mathbb{E}[X] = 0$ and $\mathrm{Var}(X) = \sigma^2$. Prove that for all $\lambda \geq 0$,

$$\mathbb{P}(X \geq \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

5. A random variable has the Beta(α, β) if

$$f_X(x) = Cx^{\alpha-1}(1-x)^{\beta-1}$$

for $0 \leq x \leq 1$.

(a) Find the value of C that makes this a PDF for integral α, β.

(b) Find $\mathbb{E}[X]$.