Problem Sheet 3 - Part C Probabilistic Combinatorics - Oxford HT 2008

---

1. Let X be a non-negative random variable. Prove that

$$\mathbb{P}(X = 0) \le \frac{\operatorname{Var}(X)}{\mathbb{E}\,(X^2)}.$$

   [Hint: Cauchy-Schwarz.]

   This is a slight strengthening of the bound proved in lectures. Deduce that if $X^{(n)}$ is a sequence of non-negative random variables indexed by n, such that $\mathbb{E}\,X^{(n)} \to \mu$ and $\operatorname{Var} X^{(n)} \to \sigma^2$ for some positive constants μ and σ^2, then $\mathbb{P}(X^{(n)} > 0)$ is bounded away from 0 as $n \to \infty$.

2. Let $k \ge 3$ be fixed. Let $S_{n,p}$ be a subset of $\{1, 2, \dots, n\}$ chosen by including each element independently with probability p. Show that as $n \to \infty$, $p = n^{-2/k}$ is a threshold function for the property "$S_{n,p}$ contains an arithmetic progression of length k".

   [Hint: first find the asymptotics as $n \to \infty$ of:

   (i) the number of APs of length k made up of elements of $\{1, 2, \dots, n\}$;

   (ii) the number of *pairs* of such APs which share exactly one element;

   (iii) the number of pairs which share two or more elements.]

3. Show that there exists a positive constant c with the following property. If $n \in \mathbb{N}$ and $a_1, \dots, a_n$ satisfy $\sum_{i=1}^n a_i^2 = 1$, and if $\epsilon_1, \dots, \epsilon_n$ are i.i.d. and each equal to $+1$ with probability $1/2$ and -1 with probability $1/2$, then

$$\mathbb{P}\left(\left| \sum_{i=1}^n \epsilon_i a_i \right| \le 1 \right) \ge c.$$

   [Hint: a direct application of Chebyshev's inequality to the sum doesn't quite work. Try looking at two cases: (i) $a_i > 1/2$ for some i; (ii) $a_i \le 1/2$ for all i.]

4. (i) Let $V = \{1, 2, \dots, n\}$. Suppose $G_1 = (V, E_1), G_2 = (V, E_2), \dots, G_m = (V, E_m)$ are i.i.d., each with distribution $G(n, p)$. What is the distribution of $G = (V, E)$ where $E = \bigcup_{i=1}^m E_i$?

   (ii) Let Q be any increasing property (i.e. if a graph $G = (V, E)$ has property Q then so does the graph $\tilde{G} = (V, \tilde{E})$ whenever $E \subset \tilde{E}$).

   Prove that there exists a threshold function for Q in the model $G(n, p)$.

   [Hint: choose $p = p_\epsilon(n)$ so that $\mathbb{P}(G(n, p)$ has property $Q) = \epsilon$ and consider part (i). If each G_i has distribution $G(n, p)$, what can you say about the probability that the graph G has property Q?]

5. Use the Symmetric Local Lemma to show that any k-SAT formula in which no variable lies in more than $2^{k-2}/k$ clauses is satisfiable.

   [Reminder of definitions: an example of a 3-SAT formula with 4 clauses is

$$(x_1 \text{ OR } x_4 \text{ OR } \overline{x_6}) \text{ AND } (\overline{x_1} \text{ OR } x_2 \text{ OR } x_5) \text{ AND } (\overline{x_2} \text{ OR } \overline{x_3} \text{ OR } \overline{x_4}) \text{ AND } (\overline{x_4} \text{ OR } \overline{x_5} \text{ OR } x_6).$$

   The variables x_i take values TRUE or FALSE and $\overline{x_i}$ means NOT x_i.]

Course webpage: `http://www.stats.ox.ac.uk/~martin/PC.html`