Applications of Stochastics — Exercise sheet 1: Erdős-Rényi, Borel-Cantelli, Chung-Fuchs

Gábor Pete
http://www.math.bme.hu/~gabor

October 2, 2020

Notation. The probability measure for the Erdős-Rényi random graph $G(n,p)$ is denoted by $\mathbf{P}_{n,p}$ or $\mathbf{P}_p$.

Subsets of a base set S are sometimes denoted by $\omega \in \{0,1\}^S$, thinking that $\omega(s) = 1$ iff $s \in \omega$.

The comparisons $\sim$, $\asymp$, $\ll$, $\gg$ are used as agreed in class.

Bonus exercises are marked with $*$, and can be handed in for extra points any time before the exam period.

▷ **Exercise 1.** An event for the Erdős-Rényi random graph, $A \subset \{0,1\}^{\binom{n}{2}}$, is called *upward closed* or *increasing* if, whenever $\omega \in A$ and $\omega' \supseteq \omega$, then also $\omega' \in A$. Show that, for any such event A, other than the empty or the complete set, the function $p \mapsto \mathbf{P}_p[\,A\,]$ is a strictly increasing polynomial of degree at most $\binom{n}{2}$, with $\mathbf{P}_p[\,A\,] = p$ for $p \in \{0,1\}$. In particular, there exists a unique p such that $\mathbf{P}_p[\,A\,] = 1/2$; this value is usually called the *critical* (or *threshold*) *density*, and will be denoted by $p_c(n) = p_c^A(n)$.

▷ **Exercise 2.** Prove carefully that choosing M edges one-by-one between n vertices, always uniformly at random, independently of previous choices, but resampling the edge if a multiple edge was created, we get the model $G(n,M)$.

▷ **Exercise 3.** Find the order of magnitude of the critical density $p_c(n)$ for the random graph $G(n,p)$ containing a copy of the cycle C_4. Same with K_4. (Hint: as in class, use the 1st and 2nd Moment Methods.)

▷ **Exercise 4.** Let H be the following graph with 5 vertices and 7 edges: a complete graph K_4 with an extra edge from one of the four vertices to a fifth vertex. Show that if $5/7 > \alpha > 4/6$, and $p = n^{-\alpha}$, then the expected number of copies of H in $G(n,p)$ goes to infinity, but nevertheless the probability that there is at least one copy goes to 0. What goes wrong with the 2nd Moment Method?

▷ **Exercise 5.** Let $X_k(n)$ be the number of degree k vertices in the Erdős-Rényi random graph $G(n,\lambda/n)$, with any $\lambda \in \mathbb{R}_+$ fixed. Show that $X_k(n)/n$ converges in probability, as $n \to \infty$, to $\mathbf{P}[\,\mathsf{Poisson}(\lambda) = k\,]$. (Hint: the 1st moment of $X_k(n)$ should be clear; then use the 2nd moment method.)

▷ **Exercise 6.** Assume that Y_n, $n = 1,2,\dots$, are non-negative integer valued random variables, with $\mathbf{E}[\,Y_n\,] \le K < \infty$, independently of n. Show that $\mathsf{Binom}\left(n - Y_n, \frac{\lambda}{n}\right) \xrightarrow{d} \mathsf{Poi}(\lambda)$, as $n \to \infty$.

Recall that we used this exercise to show that the limit of the first any fixed number of steps in our exploration random walk for $G(n,\lambda/n)$ converges in distribution to the first steps in the random walk with iid jumps $\mathsf{Poi}(\lambda) - 1$.

▷ **Exercise 7.** Prove that $1 - x \le e^{-x}$ for all $x \in \mathbb{R}$, and $e^{-x} \le 1 - x/2$ for all $0 < x < \epsilon$, if $\epsilon > 0$ is small enough. Conclude that, for any sequence $\epsilon_n \in (0,1)$, we have: $\sum_n \epsilon_n = \infty \iff \prod_n (1 - \epsilon_n) = 0$.

▷ **Exercise 8.** Show by example that, in the Second Borel-Cantelli lemma, the assumption of independence cannot be omitted: construct some events A_n on some probability space with $\sum_{n=1}^{\infty} \mathbf{P}[\,A_n\,] = \infty$ but $\mathbf{P}[\,A_n \text{ infinitely often}\,] = 0$.

The following is called "Monte Carlo integration."

▷ **Exercise 9.** Let $f : [0,1] \longrightarrow \mathbb{R}$ be a measurable function with $\int_0^1 |f(x)|^4 dx < \infty$, and let $U_1, U_2, \dots$ be i.i.d. $\mathsf{Unif}[0,1]$ variables. Prove that $(f(U_1) + \cdots + f(U_n))/n$ converges almost surely to $\int_0^1 f(x)dx$.

I argued intuitively in class that we cannot have almost sure convergence in the Central Limit Theorem. To make this precise, here is a bonus exercise:

▷ **Exercise 10.** * Let $X_1, X_2, \dots$ be iid variables with $\mathbf{E}X_i = \mu$ and $\mathrm{Var}X_i = \sigma^2 < \infty$, and let $Z_n := \frac{X_1+\cdots+X_n-n\mu}{\sqrt{n}\sigma}$. Show that $\liminf_n Z_n = -\infty$ and $\limsup_n Z_n = \infty$ almost surely.

The following exercise gives an example where the converse of Chung-Fuchs is not true:

▷ **Exercise 11.** Recall the standard Cauchy density: $\frac{1}{\pi(1+x^2)}$ for $x \in \mathbb{R}$. Accepting the fact that if $X_1, \dots, X_n$ are i.i.d. Cauchy variables, then the sum $S_n = X_1 + \cdots + X_n$ has the distribution of nX_1, show the following:

(a) $S_n/n \xrightarrow{\mathbf{P}} 0$ does not hold.

(b) For any $\epsilon > 0$, the expected number of returns to the interval $(-\epsilon, \epsilon)$ by the Cauchy walk S_n is infinite.

The following exercise originates from a question from the audience during the proof of Chung-Fuchs. I know how to do (a), it needs a trick, but is doable — it is for bonus. Part (b) is easy and compulsory. I have not really thought about (c) and (d). My guess about (c) is that the answer is "no", and then it should not be hard to find an example. For (d) I do not have a guess, and the solution might actually be hard.

▷ **Exercise 12.** Let $X_1, X_2, \dots$ be iid integer-valued variables with positive variance, and let $S_n = X_1 + \cdots + X_n$ be the resulting random walk. Let M_n be the largest (rightmost) modus of the distribution of S_n.

(a) * Assume that the X_i's are symmetric; i.e., $X_i \overset{d}{=} -X_i$. Show that $M_{2n} = 0$ for any $n \geq 0$.

(b) Give an example with $\mathbf{E}X_i = 0$, but $M_{2k} < 0$ for some k.

(c) * Assume that $\mathbf{E}X_i = 0$ but $M_{2k} < 0$ for some k. Does it follow that $M_{2n} \to -\infty$?

(d) * Whenever $M_{2n} \to -\infty$, does $\limsup_n M_{2n}/n < 0$ also hold?

The goal of the last bonus exercise is to present one way to pass from $G(n,p)$ to the $G(n,M)$ model.

▷ **Exercise 13.** * Fix $\delta > 0$ arbitrary, and let $p_n \in (0,1)$ and $M_n \in \left\{0, 1, \dots, \binom{n}{2}\right\}$ be two sequences satisfying $\binom{n}{2}p_n \to \infty$ and $(1+\delta)\binom{n}{2}p_n < M_n$ for all n. Let $A_n \subset \{0,1\}^{\binom{n}{2}}$ be a sequence of upward closed events such that $\mathbf{P}_{p_n}[\,A_n\,] \to 1$. Prove that $\mathbf{P}\big[\,G(n, M_n) \text{ satisfies } A_n\,\big] \to 1$, as $n \to \infty$.

In more detail:

(a) Show that $\mathbf{P}\big[\,\mathsf{Binom}\big(\binom{n}{2}, p_n\big) < M_n\,\big] \to 1$.

(b) Let $\mathcal{E}_n$ denote the number of edges in $G(n,p)$. Deduce from part (a) that $\mathbf{P}_{p_n}[\,A_n \mid \mathcal{E}_n < M_n\,] \to 1$.

(c) Show that, for any $M \in \left\{0, 1, \dots, \binom{n}{2}\right\}$, we have $\mathbf{P}_{p_n}\big[\,A_n \bigm| \mathcal{E}_n = M\,\big] = \mathbf{P}\big[\,G(n,M) \text{ satisfies } A_n\,\big]$.

(d) Deduce from part (c) that $\mathbf{P}_{p_n}\big[\,A_n \bigm| \mathcal{E}_n < M_n\,\big] \leq \mathbf{P}\big[\,G(n, M_n) \text{ satisfies } A_n\,\big]$.

Combining parts (b) and (d) concludes the exercise.