# Applications of Stochastics — Exercise sheet 1

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

September 27, 2019

**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 a star. They can be handed in for extra points.

▷ **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 $X_k(n)$ is 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 $n \to \infty$.

▷ **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.** Accepting the fact that if $X_1, \dots, X_n$ are i.i.d. Cauchy variables, then the sum $S_n = X_1 + \dots + X_n$ has the distribution of $nX_1$, show the following:

**(a)** $S_n/n \xrightarrow{p} 0$ 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.

▷ **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 $(f(U_1) + \dots + f(U_n))/n$ converges almost surely to $\int_0^1 f(x)dx$.

▷ **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.

▷ **Exercise 11.** Flip a fair coin 60 times, and let $X \sim \mathsf{Binom}(60, 1/2)$ be the number of heads. Using Markov's inequality for $e^{tX}$ with the best possible $t$, which can be found by minimizing the convex function $f(t) = \log(1+e^t) - \frac{5}{6}t$, show that

$$\mathbf{P}\big[\,|X-30| \geq 20\,\big] \leq 2 \cdot 3^{60} \cdot 5^{-50} < 10^{-6}.$$

▷ **Exercise 12.** Prove that for any $\delta > 0$ there exist $c_\delta > 0$ and $C_\delta < \infty$ such that

$$\mathbf{P}\big[\,|\mathsf{Poisson}(\lambda) - \lambda| > \delta\lambda\,\big] < C_\delta\, e^{-c_\delta \lambda},$$

for any $\lambda > 0$. (Hint: use the moment generating function of $\mathsf{Poisson}(\lambda)$.)

▷ **Exercise 13.** Let $\xi_i \sim \mathsf{Expon}(\lambda)$ i.i.d. random variables, and let $S_n := \xi_1 + \cdots + \xi_n$. Prove that for any $\delta > 0$ there exist $c_\delta > 0$ and $C_\delta < \infty$ (also depending on $\lambda$, of course) such that

$$\mathbf{P}\big[\,|S_n - \mathbf{E}S_n| > \delta n\,\big] < C_\delta\, e^{-c_\delta n}.$$

Hint: use the moment generating function of $\mathsf{Expon}$ or the previous Poisson exercise!

▷ **Exercise 14.** Let $p, \alpha \in (0,1)$ arbitrary, and let $\alpha_n \to \alpha$ such that $\alpha_n n \in \mathbb{Z}$ for every $n$. Using Stirling's formula, show that

$$\lim_{n\to\infty} \frac{-\log \mathbf{P}\big[\,\mathsf{Binom}(n,p) = \alpha_n n\,\big]}{n} = \alpha \log\frac{\alpha}{p} + (1-\alpha)\log\frac{1-\alpha}{1-p}\,.$$

When $\alpha = p$, we are getting that $\mathbf{P}\big[\,\mathsf{Binom}(n,p) = \alpha_n n\,\big]$ is only subexponentially small. In particular, roughly how large is $\mathbf{P}\big[\,\mathsf{Binom}(n,p) = \lfloor pn \rfloor\,\big]$?

The next bonus exercise contains some analytic details regarding the moment generating function. The main tool will be the *Dominated Convergence Theorem (DCT)*: if $\{X_n\}_{n\geq 1}$ and $X$ and $Y$ are random variables on the same probability space, with the almost sure pointwise convergence $\mathbf{P}[\,X_n \to X\,] = 1$, plus $|X_n| \leq Y$ holds almost surely for all $n$, where $\mathbf{E}Y < \infty$, then $\mathbf{E}|X_n - X| \to 0$, and thus $\mathbf{E}X_n \to \mathbf{E}X < \infty$.

▷ **Exercise 15.*** Assume that $m_X(t) := \mathbf{E}[\,e^{tX}\,] < \infty$ for some $t = t_0 > 0$, and let $\kappa_X(t) := \log m_X(t)$.

**(a)** Show that $e^{tx} < 1 + e^{t_0 x}$ for all $0 \leq t \leq t_0$ and $x \in \mathbb{R}$. Deduce that $m_X(t) < \infty$ for all $0 \leq t \leq t_0$.

**(b)** Using part (a) and the DCT, show that if $t_n \to t$, all of them in $[0, t_0]$, then $m_X(t_n) \to m_X(t)$. Thus $m_X(t)$ and $\kappa_X(t)$ are continuous functions of $t \in [0, t_0]$.

**(c)** Show that $x < e^{tx}/t$ for any $t > 0$ and $x \in \mathbb{R}$. Deduce that $\mathbf{E}\big[\,Xe^{tX}\,\big] < \infty$ if $0 < t \leq t_0/2$.

**(d)** Using that $e^b - e^a = \int_a^b e^y\,dy$, show that $(e^{tx}-1)/t \leq xe^{tx}$ for any $t > 0$ and $x \in \mathbb{R}$. Using part (c) and the DCT, show that $m'_X(0) = \mathbf{E}X < \infty$.

**(e)** Deduce that $\kappa'_X(0) = \mathbf{E}X$. Deduce that if $\alpha > \mathbf{E}X$, then $\kappa_X(t) - \alpha t < 0$ for some $t \in (0, t_0)$.

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

▷ **Exercise 16.*** Fix $\delta > 0$ arbitrary, and let $p_n \in (0,1)$ and $M_n \in \big\{0, 1, \dots, \binom{n}{2}\big\}$ 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 $M \in \big\{0, 1, \dots, \binom{n}{2}\big\}$, we have $\mathbf{P}_{p_n}\big[\,A_n \,\big|\, \mathcal{E}_n = M\,\big] = \mathbf{P}\big[\,G(n,M) \text{ satisfies } A_n\,\big]$.

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

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