Assignment 1 covering material from Lectures 1-15.
Due by 11:59pm on 3 May.

Instructions: Please write detailed solutions to the following problems. Quality of exposition will comprise a significant number of marks. Clarity, neatness, and style count.

1. Fix $k \geq 1$ and let $M_k = M_k(n)$ denote the number of vertices in $G(n) \sim \mathrm{ER}(n, c_n/n)$ having degree k.

   (a) Show that if $c_n = a\log(n)$ for $a > 1$, then $M_k \xrightarrow[n\to\infty]{prob} 0$.

   (b) Show that if $c_n = a\log(n)$ for $a < 1$, then $M_k \xrightarrow[n\to\infty]{prob} \infty$.

   (c) For given $\mu > 0$, find a sequence b_n such that if $c_n = \log(n) + b_n$, then

   $$\mathbb{E}[M_k] \xrightarrow[n\to\infty]{} \mu.$$

   For this choice of c_n, to what distribution do you suppose that $\mathscr{L}(M_k)$ converges? (No need to prove the convergence.)

2. This exercise is related to Lecture 9.

   (a) Let $D, D_1, D_2, \ldots$ be i.i.d. integer-valued random variables. For $k \in \mathbb{Z}$, set $p_k := \mathbb{P}(D = k)$ and

   $$P_n^{(k)} := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}[D_i = k].$$

   Show that as $n \to \infty$, if $n\,\varepsilon_n^2 \to \infty$, then

   $$\mathbb{P}\Big(\max_{k\geq 0}\big|P_n^{(k)} - p_k\big| \geq \varepsilon_n\Big) \to 0.$$

   Show that this is optimal in the sense that for any $c > 0$ and $k \in \mathbb{Z}$ with $p_k > 0$,

   $$\limsup_n \mathbb{P}\Big(\big|P_n^{(k)} - p_k\big| \geq c/\sqrt{n}\Big) > 0.$$

   (b) Let $Y_1, Y_2, \ldots$ be i.i.d. and $\varphi : \mathbb{R}^2 \to \mathbb{Z}$. For $i = 1, 2, \ldots$, now set $D_i = \varphi(Y_i, Y_{i+1})$, $p_k := \mathbb{P}(D_i = k)$, and $P_n^{(k)}$ defined as above. Note that the D_i have the same distribution, but may be dependent. Show that as $n \to \infty$, if $n\,\varepsilon_n^2 \to \infty$, then

   $$\mathbb{P}\Big(\max_{k\geq 0}\big|P_n^{(k)} - p_k\big| \geq \varepsilon_n\Big) \to 0.$$

   [Such D_i arise in statistics as, for example, moving averages $D_i = (Y_i + Y_{i+1})/2$, or as measurements of "clumping" in data. More generally, we could set $D_i = \varphi(Y_i, Y_{i+1}, \ldots, Y_{i+m-1})$ for some $\varphi : \mathbb{R}^m \to \mathbb{Z}$ and $m \geq 2$, and then the same conclusion of the exercise holds.]

3. (a) For $0 < p < q < 1$, let T_p be the number of triangles in $G_p \sim \mathrm{ER}(n, p)$ and T_q be the number of triangles in $G_q \sim \mathrm{ER}(n, q)$. Show that $T_q \geq_{\mathrm{st}} T_p$.

   (b) Let X and Y be non-negative random variables with finite means. Show that if $X \leq_{\mathrm{st}} Y$, then $\mathbb{E}[X] \leq \mathbb{E}[Y]$. Construct a (simple) counterexample where $\mathbb{E}[X] \leq \mathbb{E}[Y]$, but $X \not\leq_{\mathrm{st}} Y$.