Chernoff bounds:

1. If (V_1, V_2) is a fixed partition of the vertices of $G(n, 1/2)$, what is the distribution of the number of edges of $G(n, 1/2)$ joining V_1 to V_2?

   Show that the probability that $G(n, 1/2)$ contains a bipartite subgraph with at least $n^2/8 + n^{3/2}$ edges is $o(1)$.

2. A *tournament* on a vertex set V is an orientation of the edges of the complete graph on V. Thus, for each pair $\{i, j\}$ of distinct elements of V, exactly one of the directed edges $i \to j$ and $j \to i$ is present. (Think of an all-play-all tournament whose players are the elements of V; the orientation of the edge between i and j indicates who wins the match between i and j.)

   Let σ be a permutation of $\{1, 2, \ldots, n\}$. The permutation can be seen as a ranking of the players. We say that an *upset* occurs if the edge $i \to j$ is present (i.e. j beats i) but $\sigma(i) < \sigma(j)$ (i.e. i is higher ranked than j).

   Show that there exists a tournament on $\{1, 2, \ldots, n\}$ such that, for all rankings σ, the difference between the number of upsets and the number of non-upsets is no greater than $2n^{3/2}\sqrt{\log n}$. (In other words, no ranking gives a correct prediction for significantly more than 50% of the matches.)

3. Let $H = (V, E)$ be a hypergraph. Let χ be a two-colouring (red/blue) of its vertices.

   The *discrepancy* of an edge $e \in E$ under the colouring χ is the absolute difference between the number of blue vertices in e and the number of red vertices in e. The *discrepancy* of H under χ, denoted $\mathrm{disc}(H, \chi)$, is the maximum over all edges e of the discrepancy of e under χ. Finally, the *discrepancy* of H, $\mathrm{disc}(H)$, is defined as $\min_\chi \mathrm{disc}(H, \chi)$.

   [For example, if H is k-uniform, $\mathrm{disc}(H) < k$ if and only if H is 2-colourable.]

   (i) Show that if H is k-uniform and has $m \geqslant 2$ edges, then $\mathrm{disc}(H) \leqslant 2\sqrt{k \log m}$.

   (ii) Show that if H is k-uniform and each edge intersects at most d other edges, then $\mathrm{disc}(H) \leqslant \sqrt{2k \log(6(d+1))}$.

Branching processes:

4. Using results from lectures, show that the survival probability $\rho(c) = 1 - \eta(c)$ of the Poisson branching process $\mathbf{X}_{\mathrm{Po}(c)}$ satisfies $\rho(1+\varepsilon) \sim 2\varepsilon$ as ε tends to zero from above.

   Can you obtain further terms in this expansion?

5. Let Y_k denote the number of k-vertex components of $G = G(n,p)$.

(i) Using Cayley's formula k^{k-2} for the number of trees on k (labelled) vertices, show directly that if k is fixed and $p = p(n)$ satisfies $np \to c$ with $c > 0$ constant, then

$$\mathbb{E}Y_k \sim \binom{n}{k} k^{k-2} p^{k-1} e^{-ck}.$$

(ii) Deduce that $\rho_k(c) = c^{k-1}k^{k-1}e^{-ck}/k!$. [You may like to give a direct proof of this formula.]

(iii) Deduce that

$$\sum_{k=1}^{\infty} c^{k-1}\frac{k^{k-1}}{k!}e^{-ck} = 1$$

if $0 \leqslant c \leqslant 1$, and that the sum is strictly less than 1 if $c > 1$. [You may not like to give a direct proof of this!]

6. (i) Show that for each $c \in (1,\infty)$ there is a unique $d \in (0,1)$ such that $ce^{-c} = de^{-d}$.

(ii) Let η be the extinction probability of $\mathbf{X}_{\mathrm{Po}(c)}$, the Galton–Watson branching process with offspring distribution Po(c). Show that $c\eta = d$ where d is related to c as in part (i).

(iii) Consider the first particle (the root) in the branching process $\mathbf{X}_{\mathrm{Po}(c)}$. What is the probability of extinction of the process conditional on the event that the root has k children (for $k \in \{0,1,2,\dots\}$)? Use this to find the conditional distribution of the number of children of the root, conditional on the event that the process dies out.

(iv) Hence or otherwise argue that the branching process $\mathbf{X}_{\mathrm{Po}(c)}$, conditioned on extinction, has the same distribution as the branching process $\mathbf{X}_{\mathrm{Po}(d)}$. What does this *suggest* about the random graphs $G(n,d/n)$ and $G(n,c/n)$?

Bonus question (compulsory for MFoCS students, optional for others):

7. (i) Let $X_1, X_2, \dots, X_n$ be independent random variables such that $0 \leqslant X_i \leqslant 1$. Let $S_n = \sum_{i=1}^n X_i$ and let $p = \sum \mathbb{E}X_i/n$, so that $\mathbb{E}S_n = np$. Show that

$$\mathbb{P}\left(S_n \geqslant xn\right) \leqslant e^{-uxn}\left(1 - p + pe^u\right)^n$$

for any $u > 0$, $x > p$, and deduce that the Chernoff bounds proved in lectures for the case $S_n \sim \mathrm{Bin}(n,p)$ also apply in this more general case.

(ii) Let $a_1, \dots, a_n$ be constants and let $c > 0$. Let $Y_1, \dots, Y_n$ be independent random variables such that $a_i \leqslant Y_i \leqslant a_i + c$, for all i. Give (with brief justification) a version of the Chernoff bound for $\mathbb{P}(S_n - \mathbb{E}S_n \geqslant t)$, where $S_n = \sum_{i=1}^n Y_i$.

If you find an error please check the website, and if it has not already been corrected, e-mail riordan@maths.ox.ac.uk