Exercises **14.7.2019**

The exercises start on page 2. This initial discussion may be helpful.

One thing we will see on Monday is that it is sometimes necessary to go beyond the usual scaling of time that occurs in the differential equations method. Indeed, this was the key to the advance achieved by Bohman.

In the first exercise we shall do a similar analysis for a slightly contrived problem. There then follow some variants of this problem. Finally, there is an exercise which is very similar in nature to the triangle-free process and will allow us to practise applying the approach shown in the course to a similar problem.

We will generally use Freedman's inequality rather than Hoeffding-Azuma. The sum which occurs in the denominator in the Hoeffding-Azuma inequality is $\sum_{i=1}^m c_i^2$ where the martingale increments X_i are such that $|X_i| \leq c_i$ almost surely. This sum is an upper bound on the variance of the martingale. In Freedman's inequality we effectively get to replace this sum by a quantity related to variance.

Freedman's inequality Let $(S_m)_{m=0}^M$ be a supermartingale with increments $(X_i)_{i=1}^M$ with respect to a filtration $(\mathcal{F}_m)_{m=0}^M$, let $R \in \mathbb{R}$ be such that $\max_i |X_i| \leq R$ almost surely, and let

$$V(m) := \sum_{i=1}^{m} \mathbb{E}\left(|X_i|^2 \,\middle|\, \mathcal{F}_{i-1} \right) .$$

Then, for every $\alpha, \beta > 0$, we have

$$\mathbb{P}\big(S_m - S_0 \geq \alpha \quad \text{and} \quad V(m) \leq \beta \quad \text{for some } m\big) \leq \exp\left(\frac{-\alpha^2}{2(\beta + R\alpha)} \right) .$$

1. Consider a random Battle Royale which starts with $N_0 = N$ participants numbered $1, \ldots, N$. At each time step a number $k \in [N] = \{1, \ldots, N\}$ is selected uniformly at random, and player k (if they are still alive) shoots one other player. If player k has already died then nothing happens. Let us write N_i for the number of participants alive after i steps.

   (a) We shall write $\bar{N}_i$ for $\mathbb{E}(N_i)$ show that $\bar{N}_i = N(1 - 1/N)^i \approx Ne^{-i/N}$.

   (b) We now aim to show that N_i stays close to $\bar{N}_i$ throughout the process. Specifically, we aim to show that (with high probability)

   $$\bar{N}_i - f_i \le N_i \le \bar{N}_i + f_i \qquad \text{for all } i \ge 0$$

   for some sequence f_i. We define the event E_s^b to be the event that N_i becomes larger than $\bar{N}_i + f_i$ at step $b + s$ after having increased from $\bar{N}_i + \frac{2}{3} f_i$ at step b_i (and never falling below). More formally, let E_s^b be the event that

   - $N_{b-1} < \bar{N}_{b-1} + \frac{2}{3} f_{b-1}$
   - $N_{b+i} \ge \bar{N}_{b+i} + \frac{2}{3} f_{b+i}$ for all $i = 0, \ldots s$, and
   - s is minimal such that $N_{b+s} > \bar{N}_{b+s} + f_{b+s}$

   Show that the event that $N_i > \bar{N}_i + f_i$ for some $i \le N \log N$ is contained in the union $\bigcup_{b,s: b+s \le N \log N} E_s^b$.

   (c) We now define a sequence of random variables associated with these deviation events. We wish to study

   $$N_{b+i} - \bar{N}_{b+i} - f_{b+i}$$

   as this sequence becomes positive exactly when we break the inequality $N_i \le \bar{N}_i + f_i$. However, it is useful to stop the process if we ever fall too far. Let τ be the stopping time defined to be the minimum i such that

   $$N_{b+i} < \bar{N}_{b+i} + \frac{2}{3} f_{b+i} \, .$$

   We write $i \wedge \tau$ for the minimum of i and τ (so that effectively the process is halted at the stopping time). Let

   $$Z_i^b := N_{b+i\wedge\tau} - \bar{N}_{b+i\wedge\tau} - f_{b+i\wedge\tau} \, .$$

   Show that E_s^b is contained in the event that

   $$Z_s^b > Z_0^b + \frac{1}{3} f_b - 1 \, .$$

   (When using this in the rest of the question feel free to ignore the -1, it really doesn't matter).

(d) We shall define the sequence f_i as follows:

$$f_i := N^{1/2}(\log N)\left(1 - \frac{1}{2N}\right)^i \approx N^{1/2}(\log N)\, e^{-i/2N}.$$

Show that (with this choice of f_i) the sequence $Z_i^b : i \geq 0$ is a supermartingale.

(e) Using Freedman's inequality show that $\mathbb{P}\left(E_s^b\right) \leq \exp(-c(\log N)^2)$ for some constant $c > 0$ and for all $b + s \leq N \log N$.

(f) Show that with high probability

$$\bar{N}_i - f_i \leq N_i \leq \bar{N}_i + f_i \qquad \text{for all } i \geq 0.$$

(g) Would it be possible to replace f_i by some sequence that is significantly smaller?

2. This question is a variant of first question. Suppose now there are two opposing armies (not necessarily the British and the French ;). This time one number is selected on each side at each step. When a soldier is activated he shoots someone in the opposing army. Let M_i and N_i denote the number of soldiers left in each army after i steps, and suppose $M_0 = N_0 = N$. We wish to show that both M_i and N_i remain somewhat close to $\bar{N}_i$ for some time. We will use an error function of the form

$$f_i = N^{1/2}(\log N)\, e^{\eta i/N}$$

for some $\eta \in \mathbb{R}$. For which values of η is it possible to prove that with high probability

$$\bar{N}_i - f_i \leq M_i, N_i \leq \bar{N}_i + f_i \qquad \text{for all } i \geq 0\,?$$

3. Now consider yet another variant. Again we have two opposing armies. However, now an activated soldier shoots someone on the other side with probability α and his own side with probability $1 - \alpha$ for some $\alpha \in (0, 1)$. Let $\eta(\alpha)$ be the smallest value of η we could take in question 2 for this variant of the game. Find $\eta(\alpha)$ for all $\alpha \in (0, 1)$.

(*) 4. How about if there are k equal sized armies and the shooting probabilities are given by some doubly stochastic $k \times k$ matrix A?

5. This final question will be based on the approach described on Monday. Let us consider the bipartite C_4-free process. In other words we start with two sets A, B of n vertices and we shall only ever add edges between the two sides. The process runs by adding at each step a uniformly random edge that can be added without creating a C_4. Let us define $p = p(i) = i/n^2$ to be the density after i steps, it may also be useful to use the scaled time $t = in^{-4/3}$ and $q = q(i) = e^{-i^3/n^4} = e^{-t^3}$ which will be approximately the probability that a pair is open ater i steps.

   (a) Let $Q(i)$ be the number of open edges in G_i. Explain why we should expect $Q(i) \approx n^2 q = n^2 e^{-t^3}$.

   (b) Given a pair $u \in A, v \in B$ let $X_{u,v}$ be the number of copies of C_4 containing u and v such that G_i contains one of the other three edges and the remaining two are open. Explain why we should expect $X_{u,v}(i) \approx 3pn^2q^2 = 3tn^{4/3}e^{-2t^3}$.

   (c) Given a pair $u \in A, v \in B$ let $Y_{u,v}$ be the number of copies of C_4 containing u and v such that G_i contains two of the other three edges and the remaining edge is open. Explain why we should expect $Y_{u,v}(i) \approx 3p^2n^2q = 3t^2n^{2/3}e^{-t^3}$.

   (d) Assume[1] (unrealistically) that the approximate values in (a) and (b) hold precisely, prove that with high probability the approximation in (c) holds for all time, up to an error of the form $f_i = n^{3/5}e^{Ct^3}$, for some constant C.

   (e) Now assume (d), show deterministically that the process runs for at least $cn^{4/3}(\log n)^{1/3}$ steps for some constant $c > 0$.

   (f) Conjecture the correct (optimal) value of the constant c.

[1] You may also assume all common neighbourhoods have size at most $(\log n)^2$