Large Deviations
Homework 1

Deadline for handing in solutions: March 8.

1. **An analogue for Gaussian random variables of Theorem I.3 in den Hollander's book.** It is well-known (see e.g. page 7 of the 1996 edition of Durrett's book *Probability: Theory and Examples*) that

$$(x^{-1} - x^{-3})\exp(-x^2/2) \le \int_x^\infty \exp(-y^2/2)dy \le x^{-1}\exp(-x^2/2)\,. \tag{1}$$

   Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution $N(0,1)$, and let, as usual, $M_n = \frac{1}{n}\sum_{i=1}^n X_i$. Use (1) to prove that

$$\mathbf{P}(M_n > a) \simeq \exp(-a^2 n/2) \tag{2}$$

   for any $a > 0$.

2. **Analysis of a branching random walk.** Consider the following branching random walk model on the real line. At time 0, there is only one individual, denoted $A_\emptyset$, situated at the origin. At time 1, $A_\emptyset$ dies and is replaced by two individuals A_0 and A_1, and their locations are independent random variables with distributions $N(0,1)$. Think of A_0 and A_1 as children of $A_\emptyset$. The process then continues inductively as follows. Each individual A_b alive at time n is dead at time $n+1$, and replaced by two children A_{b0} and A_{b1}, whose positions relative to where A_b was are independent $N(0,1)$ random variables.

   Let $X_{b_1 b_2 \cdots b_n}$ denote the displacement of the individual $A_{b_1 b_2 \cdots b_n}$ from its parent $A_{b_1 b_2 \cdots b_{n-1}}$. All the displacements $X_0, X_1, X_{00}, X_{01}, X_{10}, X_{11}, X_{000}, \ldots$ are assumed to be indenpendent of each other.

   Clearly, at any time n, there will be 2^n individuals in the population. The most interesting thing to analyze in this model is how fast the population spreads on the unit line. Let R_n denote the position of the rightmost individual at time n. The point of this problem is to use large deviations to obtain an upper bound on how fast R_n grows with n.

   Use the result of Problem 1 to find some constant C such that

$$\limsup_{n\to\infty} \frac{R_n}{n} \le C$$

   with probability 1. This is tantamount to showing that the position of the rightmost individual grows at most linearly in time[1].

[1] In fact, the growth is precisely linear. To show this, we need also to show that $\liminf_{n\to\infty} \frac{R_n}{n} \ge c$ with probability 1, for some $c > 0$. This is done as follows. Define R'_1 to be the position of the rightmost individual at time 1. For each $n \ge 2$, define R'_n to be the position of the rightmost of the two children of the individual standing at position R'_{n-1} at time $n-1$. Clearly $R'_n \le R_n$. Also, for each n, $R'_n - R'_{n-1}$ has the distribution of the maximum of two independent $N(0,1)$ random variables. That maximum has a strictly positive expected value μ. The strong law of large numbers gives $\lim_{n\to\infty} \frac{R'_n}{n} = \mu$, which, since $R'_n \le R_n$, implies $\liminf_{n\to\infty} \frac{R_n}{n} \ge \mu$.

Some hints: Let $Y_{b_1 b_2 \cdots b_n}$ denote the position of the individual $A_{b_1 b_2 \cdots b_n}$ at time n, and note that

$$Y_{b_1 b_2 \cdots b_n} = X_{b_1} + X_{b_1 b_2} + \ldots + X_{b_1 b_2 \cdots b_n} .$$

Use (2) to bound the probability that a given individual's position $Y_{b_1 b_2 \cdots b_n}$ is greater than Cn, and then bound the probability that *at least one* of the 2^n individuals at time n is at a position greater than Cn. For which values of C does this give a useful bound for large n? Finally, sum over all n and use Borel–Cantelli.

3. This problem is a more advanced variant of the large deviations-based proof of the SLLN[2] for fair coin tosses that I did in the first lecture.

   The setup considered in the lecture is the usual one with a sequence of i.i.d. random variables $X_1, X_2, \ldots$ taking values 0 or 1 with probability $\frac{1}{2}$ each. Think of the coins as being lined up on the real line, with the index n in X_n representing the position of coin n. In joint work[3] with Benjamini, Peres and Steif, we introduced a kind of time-dynamics on top of this. We let the sequence of coins evolve over time, and $(X_1(t), X_2(t), \ldots)$ represents the coin sequence at time t. The evolution is as follows. First $(X_1(0), X_2(0), \ldots)$ is chosen randomly like $(X_1, X_2, \ldots)$ above. Then each coin independently evolves as follows. First it waits for an exponential amount of time with rate 1 until it changes its value, then it waits another (independent) exponential amount of time with rate 1 before switching again, and so on. Hence, each coin independently behaves like a stationary continuous time Markov chain with intensity matrix

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} .$$

   This implies that for any fixed t the distribution of $(X_1(t), X_2(t), \ldots)$ is the same as that of $(X_1, X_2, \ldots)$.

   Let M_n denote $\frac{1}{n}\sum_{i=1}^n X_i$, and, similarly, let $M_n(t)$ denote $\frac{1}{n}\sum_{i=1}^n X_i(t)$.

   Now let $A \in \{0,1\}^{\mathbf{N}}$ be some event which has probability 1 for the sequence $(X_1, X_2, \ldots)$ of i.i.d. coin tosses. We call such an A an *almost sure property* of the sequence of coin tosses. For instance, A could be the event that infinitely many of the coins take value 1, or it could be the event that $\lim_{n\to\infty} M_n = \frac{1}{2}$. For such an event A, we have for each fixed t that

$$\mathbf{P}((X_1(t), X_2(t), \ldots) \in A) = 1 .$$

   An important observation now is that, since the set of times t is *uncountable*, it does *not* immediately follow that

$$\mathbf{P}((X_1(t), X_2(t), \ldots) \in A \text{ for all } t \in [0,\infty)) = 1 . \tag{3}$$

   Somewhat interestingly, there do exist almost sure properties A such that (3) *fails*, i.e. such that $(X_1(t), X_2(t), \ldots) \in \neg A$ for some (random) set of exceptional times t. Examples of this are given in the aforementioned paper by Benjamini et al.

[2] the strong law of large numbers

[3] I. Benjamini, O. Häggström, Y. Peres and J. Steif: Which Properties of a Random Sequence are Dynamically Sensitive?, to appear in the next issue of the *Annals of Probability*, preprint available at `http://www.math.chalmers.se/~olleh/papers.html`

On the other hand, it turns out that (3) holds for the event $A = \{\lim_{n\to\infty} M_n = \frac{1}{2}\}$, and the purpose of this problem is to prove this.

Consider the first n coins during the time interval $[0,1]$. Define $(Y_1^1, \ldots, Y_n^1) = (X_1(0), \ldots, X_n(0))$, and then define $(Y_1^2, \ldots, Y_n^2)$ to be the sequence seen after the first time one of the n coin flips, let $(Y_1^3, \ldots, Y_n^3)$ be the sequence seen after the second time this happens, and so on.

(a) Prove by induction that $(Y_1^k, \ldots Y_n^k)$ has the same distribution as $(X_1, \ldots, X_n)$ for every k.

(b) Note that there will only be approximately n switches among the n first coins during the time interval $[0,1]$. Make a more precise statement along those lines, and combine it with den Hollander's Theorem I.3 to get a good bound for

$$\mathbf{P}\left(M_n(t) > \frac{1}{2} + \varepsilon \text{ for some } t \in [0,1] \right) . \tag{4}$$

(c) Sum your estimate of (4) over all n, and use this together with Borel–Cantelli to show that with probability 1, there are at most finitely many n such that, for some $t \in [0,1]$, $M_n(t)$ exceeds $\frac{1}{2} + \varepsilon$.

(d) Use the result in (c) to conclude that

$$\mathbf{P}\left(\lim_{n\to\infty} M_n(t) = \frac{1}{2} \text{ for all } t \in [0,1] \right) = 1 .$$

(e) Finally, use the result in (d) to conclude that

$$\mathbf{P}\left(\lim_{n\to\infty} M_n(t) = \frac{1}{2} \text{ for all } t \geq 0 \right) = 1 .$$