Probability and Computing
Hilary Term 2019
Exercise Sheet 4

1. A gambler with an initial amount of k units plays a sequence of unfair games: at each time step, the gambler bets a unit and wins it back with probability $p < 1/2$. The gambler quits when he either becomes bankrupt or he reaches a fixed amount n.

   (a) What is the probability that he becomes bankrupt?

   (b) What is the expected number of steps of the game? You may assume as given that the expected number of steps is bounded.

2. A fair coin is tossed repeatedly. Show that the expected number of tosses until the pattern HTH appears for the first time is 10. What about the pattern consisting of n heads? Describe a way to compute the expected number of tosses before any given pattern B appears for the first time.

3. The conclusion of the statement of the Azuma-Hoeffding inequality is that

$$P(X_n - X_0 \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{2\sum_{i=1}^{n} c_i^2}\right),$$

   for $\lambda > 0$.

   In which part(s) of the proof did we use the constraint $\lambda > 0$?

4. Let G be a regular graph of degree d. We color its nodes red or blue independently and uniformly at random.

   (a) What is the expected number of monochromatic edges?

   (b) Show that the actual number of monochromatic edges is concentrated around its expectation.

5. We say that a function $f(x_1, \ldots, x_n)$ is c-Lipschitz if when we change the value of any coordinate causes f to change by at most c.

   Show that if f is c-Lipschitz and Z_1 is independent of $Z_2, \ldots, Z_n$

$$|E[f(Z_1, \ldots, Z_n)|Z_1] - E[f(Z_1, \ldots, Z_n)]| \leq c.$$

   Using this, show that if f is c-Lipschitz and Z_i is independent of $Z_{i+1}, \ldots, Z_n$ conditioned on $Z_1, \ldots, Z_{i-1}$, then the Doob martingale $X_i = E[f(Z_1, \ldots, Z_n) \,|\, Z_1, \ldots, Z_i]$ satisfies $|X_i - X_{i-1}| \leq c$.

6. Let $a = a_1 \ldots a_n$ and $b = b_1 \ldots b_n$ be two binary sequences of length n. A longest common subsequence (lcs) of a and b is a subsequence of maximum length common to both; a subsequence is any sequence resulting from keeping only a part (not necessarily contiguous) of a sequence. For example, 001101 is an lcs of 101011010 and 000110111. Suppose now that the symbols of a and b are all chosen independently and uniformly at random from $\{0, 1\}$. Let the random variable X_n denote the length of a lcs of a and b.

(a) Use the Azuma-Hoeffding inequality to show that

$$P(|X_n - E[X_n]| \geq \lambda) \leq 2\exp(-\lambda^2/8n)$$

for any $\lambda > 0$.

(b) How, if at all, does each of the following changes to the problem affect your bound of the first part? Justify your answers rigorously.

   i. The symbols of a, b are not binary, but are chosen uniformly at random from an alphabet of size $k > 1$.

   ii. The symbols of a, b are not independent.

   iii. There are three strings a, b, c instead of just two.