Probability and Computing
Hilary Term 2018
Exercise Sheet 4

1. 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 $|X_i - X_{i-1}| \leq c$.

2. 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.

   Hint: Use the statement of the previous problem.

3. A particle takes a random walk on the line starting at position $i \geq 0$. What is the probability that it reaches 0 before reaching n?

4. A particle takes a random walk on the line starting at position 0. Let X_t be its position at time t and let $Y_n = \max_{t=1,\ldots,n} |X_t|$ be its maximum distance from 0 during the first n steps. We have seen that with high probability $Y_n = O(\sqrt{n \ln n})$.

   Show the opposite. More precisely show that $\mathbb{E}[Y_n] = \Omega(\sqrt{n})$.

   [**Hint:** Use the analysis of the 2SAT algorithm.]

5. Let n equidistant points be marked on a circle. Without loss of generality, we think of the points as being labelled clockwise from 0 to $n-1$. Initially, a wolf begins at 0, and there is one sheep at each of the remaining $n-1$ points. The wolf takes a random walk on the circle. At each step it moves with probability $1/2$ to its clockwise neighbour and with probability $1/2$ to its anticlockwise neighbour. At each visit to a point the wolf eats a sheep if it is still there. Which sheep is most likely to be the last eaten?

   [**Hint:** Model the situation as a one-dimensional random walk and use Problem 3.]