Homework 14

Exercise 1.* A function $f : \mathcal{X}^n \to \mathbb{R}$ has c-bounded differences (where $c := (c_1, c_2, \cdots, c_n)$) on $\mathcal{X}^n$ iff

$$\left| f(x_1, x_2, \cdots, x_i, \cdots, x_n) - f(x_1, x_2, \cdots, x_i', \cdots, x_n) \right| \leq c_i$$

$\forall\ (x_1, x_2, \cdots, x_i, \cdots, x_n), (x_1, x_2, \cdots, x_i', \cdots, x_n) \in (\mathcal{X}^n)^2$. i.e., if the $i-th$ entry of any input to the function is modified, the function changes by at most c_i. From now on, with slight abuse of notation, we denote a vector $(X_1, \cdots, X_n)$ as $X \in \mathcal{X}^n$.

Recall from the class that the McDiarmid's Inequality tells us that if a function f has c-bounded differences, then $f(X)$ concentrates around its expected value $\mathbb{E}[f(X)]$.

In this question, we will explore the concentration of the function $f(X)$, when it has c-bounded differences, but only on subset $\mathcal{Y}^n \subset \mathcal{X}^n$. As a takeaway, we will see that for a function $f(X)$ with c-bounded differences on a high probability set, the function concentrates around its conditional expectation i.e., $\mathbb{E}[f(X)|X \in \mathcal{Y}^n]$.

Theorem: Consider a function $f : \mathcal{X}^n \to \mathbb{R}$ with c-bounded differences on $\mathcal{Y}^n \subset \mathcal{X}^n$, then for every $\epsilon \geq 0$

$$\mathbb{P}[\big(f(X) - \mathbb{E}[f(X)|X \in \mathcal{Y}^n]\big) \geq \epsilon + p\bar{c}] \leq p + \exp\left(-\frac{\epsilon^2}{\sum_{i=1}^n c_i^2}\right)$$

where $\bar{c} = \sum_{i=1}^n c_i$ and $p = \mathbb{P}[X \notin \mathcal{Y}^n]$.

Let's try to prove it now!

a). Show that the function $f : \mathcal{X}^n \to \mathbb{R}$ has c-bounded differences over the subset $\mathcal{Y}^n$ iff it is 1-Lipschitz over $\mathcal{Y}^n$ with respect to the weighted hamming distance d_c i.e.,

$$|f(x) - f(y)| \leq d_c(x, y)\ \forall\ (x, y) \in (\mathcal{Y}^n)^2$$

where

$$d_c(x, y) = \sum_{i=1}^n c_i \mathbf{1}(x_i \neq y_i)$$

Next, we define a Lipschitz extension of f i.e., function g (such that it is 1-Lipschitz over $\mathcal{X}$ with respect to d_c) as follows

$$g(x) = \inf_{y \in \mathcal{Y}^n} \{f(y) + d_c(x, y)\}$$

for all $x \in \mathcal{X}^n$.

b). Show that the following hold for the function $g(X) : \mathcal{X}^n \to \mathbb{R}$:

$$g(X) \leq \begin{cases} f(X) & \text{if } X \in \mathcal{Y}^n \\ \mathbb{E}[f(X)|X \in \mathcal{Y}^n] + \bar{c} & \text{if } X \notin \mathcal{Y}^n \end{cases}$$

c). How would you compute an upper bound on $\mathbb{E}[g(X)]$ from here?

d). Prove the above theorem, by using the results from the previous parts. Try to upper bound $\mathbb{P}[f(X) - \mathbb{E}[g(X)] \geq \epsilon]$. *Hint: You might want to use the law of total probability.*

Exercise 2. Let $0 < p < 1$ and $M = (M_n,\, n \in \mathbb{N})$ be the process defined recursively as

$$M_0 = x \in]0,1[, \quad M_{n+1} = \begin{cases} p\, M_n, & \text{with probability } 1 - M_n \\ \\ (1-p) + pM_n, & \text{with probability } M_n \end{cases}$$

and $(\mathcal{F}_n,\, n \in \mathbb{N})$ be the filtration defined as $\mathcal{F}_n = \sigma(M_0, \ldots, M_n)$, $n \in \mathbb{N}$.

a) For what value(s) of $0 < p < 1$ is the process M is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? Justify your answer.

b) In the case(s) M is a martingale, compute $\mathbb{E}(M_{n+1}\,(1 - M_{n+1})\,|\,\mathcal{F}_n)$ for $n \in \mathbb{N}$.

c) Deduce the value of $\mathbb{E}(M_n\,(1 - M_n))$ for $n \in \mathbb{N}$.

d) Does there exist a random variable M_∞ such that

(i) $M_n \underset{n\to\infty}{\to} M_\infty$ a.s. ? (ii) $M_n \overset{L^2}{\underset{n\to\infty}{\to}} M_\infty$? (iii) $\mathbb{E}(M_\infty|\mathcal{F}_n) = M_n$, $\forall n \in \mathbb{N}$?

e) What can you say more about M_∞? (No formal justification required here; an intuitive argument will do.)

Exercise 3. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables such that $\mathbb{P}(\{X_n = +1\}) = p$ and $\mathbb{P}(\{X_n = -1\}) = 1 - p$ for some fixed $0 < p < 1/2$.

Let $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$, $n \geq 1$. Let also $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \geq 1$.

Preliminary question. Deduce from Hoeffding's inequality that for any $0 < p < 1/2$,

$$\mathbb{P}(\{|S_n - n\,(2p-1)| \geq nt\}) \leq 2\exp\left(-\frac{nt^2}{2}\right) \quad \forall t > 0,\, n \geq 1.$$

This inequality will be useful at some point in this exercise. Let now $(Y_n,\, n \in \mathbb{N})$ be the process defined as $Y_n = \lambda^{S_n}$ for some $\lambda > 0$ and $n \in \mathbb{N}$.

a) *Using Jensen's inequality only*, for what values of λ can you conclude that the process Y is a submartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$?

b) Identify now the values of $\lambda > 0$ for which it holds that the process $(Y_n = \lambda^{S_n},\, n \in \mathbb{N})$ is a martingale / submartingale / supermartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

c) Compute $\mathbb{E}(|Y_n|)$ and $\mathbb{E}(Y_n^2)$ for every $n \in \mathbb{N}$ (and every $\lambda > 0$).

d) For what values of $\lambda > 0$ does it hold that $\sup_{n\in\mathbb{N}} \mathbb{E}(|Y_n|) < +\infty$? $\sup_{n\in\mathbb{N}} \mathbb{E}(Y_n^2) < +\infty$?