36-788, Fall 2015
Homework 3

Due Oct 15

1. (Configuration functions). Let Π be a property defined over the union of finite products of a set $\mathcal{X}$, that is, a sequence of sets
$$\Pi_1 \subset \mathcal{X}, \Pi_2 \subset \mathcal{X} \times \mathcal{X}, \ldots, \Pi_n \subset \mathcal{X}^n.$$
The point $(x_1, \ldots, x_m) \in \mathcal{X}^m$ satisfies Π if $(x_1, \ldots, x_m) \in \Pi_m$. We also assume that Π is *hereditary* in the sense that if $(x_1, \ldots, x_m)$ satisfies Π, then so does any sub-sequence $x_{i_1}, \ldots, x_{i_k})$ of $(x_1, \ldots, x_m)$. The function f that maps any vector $(x_1, \ldots, x_n)$ to the size of its largest sub-sequence satisfying Π is the *configuration functions* of Π. (Notice that, with some abuse of notation, f can take as input any n-tuple $(x_1, \ldots, x_n)$, for any n.) Show that the configuration function has the self-bounding property. This is Corollary 3.18 in BLM.

2. Let $\mathcal{X}^n = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$, where, for $i = 1, \ldots, n$, $(\mathcal{X}_i, d_i)$ is a metric space with diameter D_i. Let $f \colon \mathcal{X}^n \to \mathbb{R}$ satisfy the Lipschitz condition
$$|f(x) - f(y)| \leq \sum_{i=1}^n d_i(x_i, y_i),$$
for all $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathcal{X}^n$. Let P be a product measure on $\mathcal{X}^n$ and $X = (X_1, \ldots, X_n) \sim P$.

   (a) Show that, for all $x > 0$,
$$\mathbb{P}\left(f(X) - \mathbb{E}[f(X)] > x\right) \leq \exp\left\{ - \frac{2x^2}{\sum_{i=1}^n D_i^2} \right\}. \tag{1}$$

   (b) Show that
$$\mathbb{P}\left(f(X) - \mathbb{E}[f(X)] \geq \sum_{i=1}^n D_i \right) = 0.$$
This means (1) is off when x is large. For a sharpening of the bounded difference inequality for large x in metric spaces of bounded diameter, see: E. Rio, On McDiarmid's concentration inequality Emmanuel Rio, Electronic Communications in Probability, 18(44), 1-11.

   (c) Let m_f be a median for $f(x)$. Show that
$$\left| \mathbb{E}[f(X)] - m_f \right| \leq \sqrt{\frac{\log 2}{2} \sum_{i=1}^n D_i^2}.$$

   (d) Define on $\mathcal{X}^n$ the distance $d : (x, y) \to \sum_{i=1}^n d_i(x_i, y_i)$ and let $A \subset \mathcal{X}^n$ be measurable and such that $P(A) > 0$. For $x \in \mathcal{X}^n$, let
$$d(x, A) = \inf \left\{ d(x, y), y \in A \right\}$$

be the distance function from x to A. Show that, for $X \sim P$,

$$\mathbb{E}\left[d(X,A)\right] \leq \sqrt{\frac{\log(1\ P(A))}{2}\sum_{i=1}^{n} D_i^2},$$

and that

$$\mathbb{P}\left(d(X,A) \geq x + \sqrt{\frac{\log(1/P(A))}{2}\sum_{i=1}^{n} D_i^2}\right) \leq \exp\left\{-\frac{2x^2}{\sum_{i=1}^{n} D_i^2}\right\}, \quad \forall x > 0.$$

Now consider the distance d given by the normalized Hamming metric:

$$d(x,y) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{x_i \neq y_i\}}.$$

(In this case $\mathcal{X}^n$ needs not be of bounded diameter). What happens to the previous bound? This is a classic result in the (older) literature on concentration of measure.