# Math 261B

Notes Lecture 13 - 21 Feb 2012

## Martingales

A *Martingale* is a sequence of random variables $X_0, X_1, ..., X_n$ such that $E[X_i \mid X_0, ..., X_{i-1}] = X_{i-1}$. The sequence $Y_0, ..., Y_n$ is a martingale with respect to $X_0, ..., X_n$ if $E[Y_i \mid X_0, ..., X_{i-1}] = Y_{i-1}$. These satisfy the *bounded difference constraint* if there exists $a_i \le b_i$ so that $a_i \le Y_i - Y_{i-1} \le b_i$.

**Theorem.** *(Azuma-Hoeffding) Let $Y_0, ..., Y_n$ be a martingale with respect to $X_0, ..., X_n$ satisfying the bounded difference constraint. Then*

$$P(Y_n > Y_0 + t), P(Y_n < Y_0 - t) \le e^{-2t^2/\sum(b_i - a_i)^2}.$$

*Proof.* Let $D_i = Y_i - Y_{i-1}$. Assume $Y_0 = 0$. For $\lambda > 0$,

$$P(Y_n > t) = P(e^{\lambda Y_n} > e^{\lambda t}) \le \frac{E[e^{\lambda Y_n}]}{e^{\lambda t}}$$

by Markov's Inequality.

Aside: Since $e^{\lambda x}$ is convex, then for all $x \in [a, b]$, $e^{\lambda x} \le \frac{x-a}{b-a} e^{\lambda b} + \frac{b-x}{b-a} e^{\lambda a} = \frac{b}{b-a} e^{\lambda a} = \frac{a}{b-a} e^{\lambda b} + \frac{x}{b-a}(e^{\lambda b} - e^{\lambda a})$. So if $Z$ is a random variable with $a \le Z \le b$ and $E[Z] = 0$, then $E[e^{\lambda Z}] \le \frac{b}{b-a} e^{\lambda a} - \frac{a}{b-a} e^{\lambda b} + \frac{E[Z]}{b-a}(e^{\lambda b} - e^{\lambda a}) = \frac{b}{b-a} e^{\lambda a} - \frac{a}{b-a} e^{\lambda b} \le e^{\frac{1}{8}\lambda^2 (b-a)^2}$. The last inequality here follows after some calculus.

Using this then,

$$\begin{aligned} E[e^{\lambda Y_n}] &= E[e^{\lambda Y_{n-1} + \lambda D_n}] \\ &= E[E[e^{\lambda Y_{n-1} + \lambda D_n}] \mid X_0, ..., X_{n-1}] \\ &= E[e^{\lambda Y_{n-1}} E[e^{\lambda D_n} \mid X_0, ..., X_{n-1}]] \\ &\le E[e^{\lambda Y_{n-1}}] e^{\lambda^2 (b_i - a_i)^2/8} \le e^{\lambda^2 \sum(b_i - a_i)^2/8}. \end{aligned}$$

Then form above,

$$P(Y_n > t) \le \frac{E[e^{\lambda Y_n}]}{e^{\lambda t}} \le e^{\lambda^2 \sum(b_i - a_i)^2/8 - \lambda t}.$$

Optimizing over $\lambda \ge 0$ gives the upper bound $e^{-2t^2/\sum(b_i - a_i)^2}$. □

The bounded difference condition can be replaced with other conditions:

The Averaged Lipschitz Condition: $|E[f \mid X_0, ..., X_{i-1}, X_i = a] - E[f \mid X_0, ..., X_{i-1}, X_i = a']| \le c_i$. In this case, in the result, we replace $\sum(b_i - a_i)^2$ with $\sum c_i^2$.

Averaged Bounded Differences Condition: $|E[f \mid X_0, ..., X_i] - E[f \mid X_0, ..., X_{i-1}]| \le c_i$.

Lipschitz Condition: $|f(a) - f(a')| \le d_i$ and $a$ and $a'$ differ only in the $i$-th coordinate. (This requires also that $X_0, ..., X_n$ be independent.)

**Example:** Throw $m$ balls into $n$ bins uniformly at random. Let $Z_i$ be the event that bin $i$ is empty, $Z = \sum Z_i$. Then $E[Z] = nE[Z_i] = n(1 - \frac{1}{n})^m \sim ne^{-m/n}$. (So for $m = 2n \log n$, this goes to 0, so we expect no empty bins.) We view $Z$ as a function of $X_1, ..., X_m$. Then Azuma-Hoeffding gives $P(|Z - E[Z]| > t) \le 2e^{-2t^2}$. But is this really a martingale?

We introduce the *Doob sequence* (or *Doob martingale*) of $f$ over $X_0, ..., X_n$. This is defined by $Y_i = E[f \mid X_0, ..., X_i]$ so $Y_0 = E[f]$ and $Y_n = f(X_0, ..., X_n)$, and $E[Y_i \mid X_0, ..., X_{i-1}] = E[E[f \mid X_0, ..., X_i] \mid X_0, ..., X_{i-1}] = \sum_a P(X_i = a) E[f \mid X_0, ..., X_{i-1}, X_i = a] = E[f \mid X_0, ..., X_{i-1}] = Y_{i-1}$. So the sequence $Y_i$ is a martingale.

**Example:** The *chromatic index*, $\chi'(G)$, of a graph $G$, is the least number of colors needed to color the edges properly. If the maximum degree is $\Delta$, there is a polynomial time algorithm to color the edeges of $G$

using $\Delta + 1$ colors. However, it is NP-complete to tell if $\chi'(G) = \Delta$. Suppose $G$ is a $\Delta$-regular bipartite graph. We describe an algorithm to color the edges of $G$:

- Each edge randomly chooses a color from $\{1, 2, ..., \Delta\}$.
- If needed, fix the color of the edge.

How many edges at a vertex are fixed in one step? Let $Z_e$ be the event that edge $e$ needs to be fixed. $E[Z_e] = (1 - \frac{1}{\Delta})^{2\Delta-2} \sim e^{-2}$. Consider $Z_1, ..., Z_\Delta$, $Z_1^1, ..., Z_1^{\Delta-1}$,...,$Z_\Delta^1, ..., Z_\Delta^{\Delta-1}$. The first ones are Lipschitz with value 2, and the rest with value 1. $P(|Z - E[Z]| > t) \leq 2e^{-2t^2/(4\Delta+\Delta(\Delta-1))}$ So $t$ must be greater than $\Delta$ to get concentration here, which is not very useful. However, if we define $Y_1, ..., Y_\Delta$ by $Y_i = Z_i^1 + \cdots + \mathbb{Z}_i^{\Delta-1}$, then each $Y_i$ is 1-Lischitz as well, so the bound becomes $2e^{-2t^2/5\Delta}$ which is much better.

As an example of edge exposure and vertex exposure martingales, consider the chromatic number of $G(n,p)$. We have that $P(|\chi(G(n,p)) - E[\chi(G(n,p))]| > t) \leq 2e^{-2t^2/n-1}$. So we get concentration around the mean, but we don't really know what the mean is. We will investigate $E[\chi(G(n, \frac{1}{2}))]$. Define $f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$. Define $k_0$ so that $f(k_0 - 1) > 1 > f(k_0)$. Let $k = (1 + o(1)) 2 \log_2 n$ ($k = k_0 - 4$) so $f(k) \geq n^{3+o(1)}$. We want to show that $P(\omega(G) < k) = e^{-n^{2+o(1)}}$. Let $Y(H)$ be the maximum number of edge-disjoint $k$-cliques in $H$.

**Lemma.** $E[Y] \geq \frac{n^2}{2k^4}(1 + o(1))$.

*Proof.* Let $\mathcal{K}$ be the family of $k$-cliques in $G$. $f(u) = \mu = E[|\mathcal{K}|]$. Let $W$ be the number of unordered pairs $\{A, B\}$ where $2 \leq |A \cap B| \leq k$ and $A, B$ are $k$-cliques. $E[W] = \dfrac{1}{2}\dbinom{n}{k} \displaystyle\sum_{i=2}^{k-1} \dbinom{k}{i}\dbinom{n-k}{k-i} 2^{-\binom{k}{2}} 2^{-\binom{k}{2}+\binom{i}{2}} = \dfrac{\Delta}{2} \sim \mu^2 k^4 n^{-2}$. Let $\mathcal{C}$ be a random subfamily of $\mathcal{K}$. $P[A \in \mathcal{C}] = q$. Let $W'$ be the number of unordered pairs $\{A, B\}$ with $2 \leq |A \cap B| \leq k - 1$ and $A, B \in \mathcal{C}$. Then $E[W'] = q^2 E[W]$. Now delete from $\mathcal{C}$ one set from each pair $\mathcal{C}^*$. Set $q = \mu/\Delta$, so $E[Y] \geq E[|\mathcal{C}^*|] = E[|\mathcal{C}|] - E[W'] = \mu q - \frac{\Delta}{2} q^2 = \frac{\mu^2}{\Delta} \sim \frac{n^2}{2k^4}$ $\square$

Then $P(\omega(G) < k) < e^{-(c+o(1))\frac{n^2}{\ln^8 n}}$ where $c > 0$. Let $Y_0, ..., Y_m$, $m = \binom{n}{2}$, be the edge exposure martingale. $P(\omega(G) < k) = P(Y = 0) = P(Y - E[Y] < E[Y]) \leq e^{-2E(Y)^2/\binom{n}{2}} \leq \frac{-n^4}{2k^8\binom{n}{2}}(1 + o(1)) = e^{-(c+o(1))\frac{n^2}{\log_2^8 n}}$. $\chi(G) \geq \frac{n}{\alpha(G)} = \frac{n}{\omega(\overline{G})} \geq \frac{n}{2\log_2 n}(1 + o(1))$. Let $m = \left\lfloor \frac{n}{log_2 n} \right\rfloor$ and $G|_S$ the restriction of $G(n, \frac{1}{2})$ to a set $S$ of size $m$. Define $k_0(m)$ so that $\dbinom{m}{k_0(m)} 2^{-\binom{k(m_0)}{2}} > 1 > \dbinom{m}{k_0(m)} 2^{-\binom{k(m_0)-1}{2}}$ where $k(m) = k_0(m) - 4$, $k \sim 2\log_2 m \sim 2\log_2 n$. $P(\alpha(G|_S) < k) < e^{-m^{2+o(1)}}$. There are $\binom{n}{m} < 2^n = 2^{m(1+o(1))}$ such sets. Therefore $P(\alpha(G|_S) < k \text{ for some } mset) < 2^{m^{1+o(1)}} e^{-m^{2+o(1)}} = o(1)$. Therefore $\chi(G) \leq \lceil \frac{n-m}{k} \rceil + m = \frac{n}{2\log_2 n} + o\left(\frac{n}{\log n}\right)$.