Walk through Combinatorics: Exposure martingales*

Exposure martingales Suppose $\Omega_1, \dots, \Omega_n$ are probability spaces on discrete sets $A_1, \dots, A_n$ respectively, and we have a function $f \colon A_1 \times \cdots \times A_n \to \mathbb{R}$. We can then sample independently b_1 from A_1, b_2 from A_2, etc, and consider the random variable

$$f(b_1, \dots, b_n).$$

To this random variable we can associate a sequence of random variables $X_0, \dots, X_n$ as follows:

$$\begin{aligned}
X_0(\vec{a}) &= \mathbb{E}_{\vec{b}}[f(\vec{b})],\\
X_1(\vec{a}) &= \mathbb{E}_{\vec{b}}[f(\vec{b}) \mid a_1 = b_1],\\
X_2(\vec{a}) &= \mathbb{E}_{\vec{b}}[f(\vec{b}) \mid a_2 = b_2 \wedge a_1 = b_1],\\
&\ \ \vdots\\
X_i(\vec{a}) &= \mathbb{E}_{\vec{b}}\big[f(\vec{b}) \mid \bigwedge_{j \le i} (a_j = b_j)\big],
\end{aligned}$$

where the subscript in $\mathbb{E}_{\vec{b}}$ indicates that the expectation is taken over the random choice of $\vec{b}$. Note that $X_n = f(a_1, \dots, a_n)$.

Lemma 1. *The sequence of random variables $X_0, \dots, X_n$ is a martingale.*

Proof. We need to verify that, for arbitrary values of $x_0, \dots, x_i$, we have

$$\mathbb{E}\big[X_{i+1} | (X_i = x_i) \wedge \cdots \wedge (X_0 = x_0)\big] = x_i.$$

Let E be the event $(X_i = x_i) \wedge \cdots \wedge (X_0 = x_0)$. Since

$$\mathbb{E}_{\vec{b}}[X_{i+1}(\vec{b}) \mid E] = \sum_{\vec{a}} \mathbb{E}\big[X_{i+1} \mid E \wedge \bigwedge_{j \le i} (a_j = b_j)\big] \Pr\big[\bigwedge_{j \le i} (a_j = b_j) \mid E\big]$$

the summands for which $\vec{a} \notin E$ vanish, and so the above is equal to

$$= \sum_{\vec{a} \in E} \mathbb{E}_{\vec{b}}\big[X_{i+1} \mid \bigwedge_{j \le i} (a_j = b_j)\big] \Pr\big[\bigwedge_{j \le i} (a_j = b_j) \mid E\big]$$

*These notes are available from the course webpage, and directly from `http://www.borisbukh.org/DiscreteMath14/notes_exposure_martingales.pdf` it suffices to verify that, whenever $X_i(\vec{a}) = x_i$, we have

$$\mathbb{E}_{\vec{b}}\big[X_{i+1}(\vec{b}) \mid \bigwedge_{j\le i}(a_j = b_j)\big] = x_i.$$

Let $\Pr_i$ denote the probability on the Ω_i. Since $X_{i+1}(\vec{b})$ depends only on the $(b_1, \dots, b_{i+1})$ we have

$$\begin{aligned}
\mathbb{E}_{\vec{b}}\big[X_{i+1}(\vec{b}) \mid \bigwedge_{j\le i}(a_j = b_j)\big] &= \sum_{b_{i+1}\in A_{i+1}} \Pr{}_{i+1}[b_{i+1}] X_{i+1}(a_1, \dots, a_i, b_{i+1}) \\
&= \sum_{b_{i+1}\in A_{i+1}} \Pr{}_{i+1}[b_{i+1}] \mathbb{E}_{\vec{c}}\big[f(\vec{c}) \mid c_{i+1} = b_{i+1} \wedge \bigwedge_{j\le i} c_j = a_j\big] \\
&= \mathbb{E}_{\vec{c}}\big[f(\vec{c}) \mid \bigwedge_{j\le i} c_j = a_j\big] \\
&= X_i(\vec{a}). \qquad \square
\end{aligned}$$

The martinagle $X_0, \dots, X_n$ is called *exposure martingale* for the function f and the sequence of $A_1, \dots, A_n$. The reason for the name is that one can think of the sequence X_i as the expectations of f when the first i variables $b_1, \dots, b_i$ have been exposed, but the rest remain unknown.

Two examples of particular interest to us are associated to exposure of edges and vertices in a random graph $G(n, p)$:

Edge-exposure martingale Let $f\colon 2^{\binom{[n]}{2}} \to \mathbb{R}$ be a function on n-vertex graphs, and suppose we are interested in the random variable $f(G)$ where G is sampled from $G(n, p)$.

We can write

$$2^{\binom{[n]}{2}} = \prod_{e\in\binom{[n]}{2}} A_e$$

where the set $A_e = [2]$ records whether the edge e is in a graph or not. Let $b_e \in A_e$ be equal to 2 with probability p and 1 with probability $1 - p$. Then the vector $\vec{b}$ is naturally identified with the graph G. If we order the edges of $\binom{[n]}{2}$ as $e_1, \dots, e_{\binom{n}{2}}$, then we obtain an exposure martingale

$$X_i(H) = \mathbb{E}_G\Big[f(G) \mid \bigwedge_{j\le i}\big((e_j \in G) \iff (e_j \in H)\big)\Big].$$

This martingale is known as an *edge-exposure martingale* for the function f. It has length $\binom{n}{2}$.

Vertex-exposure martingale As in the previous example we consider a graph function $f\colon 2^{\binom{[n]}{2}} \to \mathbb{R}$, and we are interested in $G(n, p)$. This time, for each $i = 1, \dots, n$ we define

$$A_i = \prod_{j<i} A_{\{j,i\}},$$

where $A_{\{j,k\}}$ are as in the previous example. The resulting martingale

$$X_i(H) = \mathbb{E}_G\Big[f(G) \mid \bigwedge_{j,k\le i} \big(\{j,k\}\in G \iff \{j,k\}\in H\big)\Big]$$

is known as the *vertex-exposure martingale*. It corresponds to exposing in i'th step all edges going back from the vertex i. This martingale has length n.

Lipschitz functions Azuma's inequality requires martingales satisfying $|X_{i+1} - X_i| \le 1$. The following condition on f is the most common way to satisfy this requirement.

A function $f\colon A_1 \times \cdots \times A_n \to \mathbb{R}$ is called 1-*Lipschitz* if $|f(\vec{a}) - f(\vec{a}')| \le 1$ whenever $\vec{a}$ and $\vec{a}'$ differ in at most one coordinate.

Lemma 2. *If $f\colon A_1\times\cdots\times A_n \to \mathbb{R}$ is 1-Lipschitz, then the associated exposure martingale satisfies $|X_{i+1} - X_i| \le 1$.*

Proof. Fix $\vec{a}$ arbitrarily. We will show that $|X_{i+1}(\vec{a}) - X_i(\vec{a})| \le 1$. We have

$$\begin{aligned} X_i(\vec{a}) &= \mathbb{E}_{\vec{b}}\big[f(\vec{b}) \mid \bigwedge_{j\le i}(a_j = b_j)\big], & (1)\\ X_{i+1}(\vec{a}) &= \mathbb{E}_{\vec{b}}\big[f(\vec{b}) \mid (a_{i+1} = b_{i+1}) \wedge \bigwedge_{j\le i}(a_j = b_j)\big]. \end{aligned}$$

Define function $g\colon A_1 \times \cdots \times A_n \to \mathbb{R}$ by

$$g(\vec{b}) = f(b_1, \ldots, b_i, a_{i+1}, b_{i+2}, \ldots, b_n).$$

Since f is 1-Lipschitz, we have

$$|g(\vec{b}) - f(\vec{b})| \le 1 \quad \text{for all } \vec{b}. \tag{2}$$

We can replace f by g in the definition of X_{i+1}, and then get rid of conditioning on b_{i+1} to obtain

$$\begin{aligned} X_{i+1}(\vec{a}) &= \mathbb{E}_{\vec{b}}\big[g(\vec{b}) \mid (a_{i+1} = b_{i+1}) \wedge \bigwedge_{j\le i}(a_j = b_j)\big]\\ &= \mathbb{E}_{\vec{b}}\big[g(\vec{b}) \mid \bigwedge_{j\le i}(a_j = b_j)\big]. & (3) \end{aligned}$$

The lemma follows by subtracting (1) from (3), and then applying (2). $\square$

In the cases of edge- and vertex-exposure martingales, the functions f satisfying the assumption of the lemma are called *edge-* and *vertex-Lipschitz* respectively.

Application to the chromatic number of a graph The combination of the Azuma's lemma, and the preceding inequality is incredibly powerful. Here is a simple example. Let $\chi(G)$ be the chromatic number of a graph G. Note that G is vertex-Lipschitz. Indeed, if G and G' differ only in the edges emanating from some vertex v, then $\chi(G) \leq \chi(G')+1$ because we can take a coloring of G' in $\chi(G')$ colors, and re-color v in a totally new color to obtain a proper coloring of G. Similarly, $\chi(G') \leq \chi(G)+1$, and so $|\chi(G)-\chi(G')| \leq 1$. We then deduce from Azuma's inequality that

$$\Pr\bigl[|\chi(G) - \mathbb{E}[\chi(G)]| > \lambda] < 2\exp(-\lambda^2/2n).$$

So, the chromatic number is concentrated in an interval of length $\Theta(\sqrt{n})$ around its mean. However, this does not tell us what the mean is!

Bonus material on martingales: If $X_0, X_1, \ldots, X_n$ is a sequence of random variables satisfying the condition $\Pr[X_{i+1}|X_i = x_i] = x_i$ for all i and x_i, then it *does not* follow that $X_0, \ldots, X_n$ is a martingale. Here is a counterexample. Consider a random variable Y sampled from $\{-1,+1\}^3$ according to the following weird rule

$$\begin{aligned}
\Pr[(-1,-1,-1)] &= 1/8,\\
\Pr[(-1,-1,+1)] &= 1/8,\\
\Pr[(-1,+1,-1)] &= 1/4,\\
\Pr[(+1,-1,+1)] &= 1/4,\\
\Pr[(+1,+1,-1)] &= 1/8,\\
\Pr[(+1,+1,+1)] &= 1/8.
\end{aligned}$$

Let $f = \sum_{j=1}^{3} Y_j$, and let $X_i = \sum_{j\leq i} Y_j$. We have $\mathbb{E}[X_{i+1}|X_i = x_i] = x_i$ for all i, but $\mathbb{E}[X_3|X_2 = 0, X_1 = x_1] = x_1$.