Feynman Liang
STAT 205A
Lecture #23
11/10/2016

1 Patterns in coin-tossing

Say in words (exercise X's general pattern):

Fix pattern $HTTHT$. Toss fair coin until see this pattern — requires W tosses where W is random, $t \leq W < \infty$ a.s.

What is $\mathbb{E}W$?

Consider the strategy

- bet \$1 that toss i is H
- if win, bet \$2 that toss $i+1$ is T
- if win, bet \$4 that toss $i+2$ is T
- if win, bet \$8 that toss $i+3$ is H
- if win, bet \$16 that toss $i+4$ is T

Do strategy for each $1 \leq i \leq W$, stop after toss W.

W is a stopping time, so the optional sampling theorem implies $\mathbb{E}[\text{profit}] = 0$.

$$\text{cost} = W \tag{1.1}$$
$$\text{return} = 32 + 4 = 36 \tag{1.2}$$
$$\text{profit} = \text{return} - \text{cost} \tag{1.3}$$
$$0 = \mathbb{E}[\text{profit}] = \mathbb{E}[36 - W] = 36 - \mathbb{E}W \tag{1.4}$$
$$\mathbb{E}W = 36 \tag{1.5}$$

2 MG proof of Radon-Nikodym

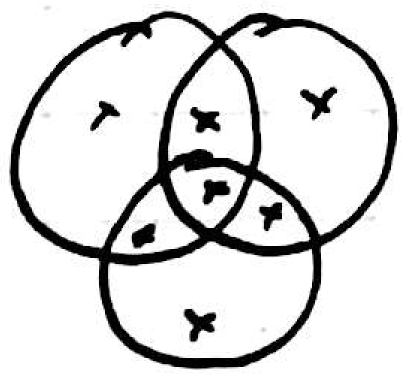
Theorem 2.1 (Radon-Nikodym)**.** *Let $(S, \mathcal{S}, \mu)$ be a probability space, $\mathcal{S} = \sigma(A_i, i \geq 0)$ countable events.*

If $\nu \ll \mu$, $\nu(S) < \infty$, then $\exists$measurable $h : S \to [0, \infty]$ such that $\nu(A) = \int_A h d\mu$ for all $A \in \mathcal{S}$.

$h = \frac{d\nu}{d\mu}$ *is the* Radon-Nikodym density of $\nu wrt \mu$.

Heuristics: $h(s) = \frac{d\nu}{d\mu}(s) = \lim_{A \downarrow \{s\}} \frac{\nu(A)}{\mu(A)}$.

Proof. Define $\mathcal{F}_n = \sigma(A_i, 1 \leq i \leq n)$ finite field with 2^n atoms.



Define $X_n(s) = \frac{\nu(F)}{\mu(F)}$ for atom $F \ni s$.

= $\nu(F)$ for each $F \in \mathcal{F}_n$

Claim: $(X_n, \mathcal{F}_n)$ is a MG.
Justification: Take $G \in \mathcal{F}_{n-1}$.

$$G = \underbrace{(G \cap A_n)}_{G_1} \cup \underbrace{(G \cap A_n^c)}_{G_2} \tag{2.1}$$

$$\mathbb{E}X_n 1_G = \mathbb{E}X_n 1_{G_1} + \mathbb{E}X_n 1_{G_2} \tag{2.2}$$

$$\text{Equation (2.3)} \implies = \nu(G_1) + \nu(G_2) = \nu(G) = \mathbb{E}X_{n-1}1_G \tag{2.3}$$

$$X_{n-1} = \mathbb{E}[X_n \mid \mathcal{F}_{n-1}] \tag{2.4}$$

By MG convergence theorem, $X_n \to X_\infty$ for some $X_\infty \geq 0$ a.s. If we prove $(X_n, n \geq 1)$ is UI, then Theorem from Lecture 20 implies $X_n = \mathbb{E}[X_\infty \mid \mathcal{F}_n]$ hence

$$\mathbb{E}X_\infty 1_F = \mathbb{E}X_n 1_F = \nu(F) \qquad \text{eq. (2.3)} \tag{2.5}$$

$$\mathbb{E}X_\infty 1_F = \nu(F) \qquad \forall F \in \cup_n \mathcal{F}_n \tag{2.6}$$

$$\mathbb{E}X_\infty 1_F = \nu(F) \qquad \forall F \in \sigma\left(\cup_n \mathcal{F}_n\right) = \mathcal{S} \tag{2.7}$$

$$\nu(F) = \mathbb{E}_\mu X_\infty 1_F = \int_F X_\infty d\mu \tag{2.8}$$

$$\tag{2.9}$$

So X_∞ is R-N density $\frac{d\nu}{d\mu}$ and we are done.

Lemma 2.2. *Suppose $\nu \ll \mu$. $\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0$ such that*

$$\mu(A) \leq \delta(\varepsilon) \implies \nu(A) \leq \varepsilon \tag{2.10}$$

Proof. If false for ε, $\exists A_n$ such that $\mu(A_n) \leq 2^{-n}$ and $\nu(A_n) > \varepsilon$. Consider $\Lambda = \{A_n \text{ i.o.}\}$. By Borel-Cantelli, $\mu(\Lambda) = 0$ but $\nu(\Lambda) \geq \varepsilon$, contradicting $\nu \ll \mu$. □

Claim: (X_n) is UI.

Justification: By eq. (2.3), $\mathbb{E}X_n 1_{X_n \geq b} = \nu(X_n \geq b)$. Given $\varepsilon > 0$, take b such that $\frac{\nu(S)}{b} \leq \delta(\varepsilon)$. Then

$$\nu(X_n \geq b) \underset{\text{Markov}}{\leq} \frac{\mathbb{E}X_n}{b} = \frac{\nu(S)}{b} \leq \delta(\varepsilon) \tag{2.11}$$

$$\text{lemma 2.2}\nu(X_N \geq b) \leq \varepsilon \tag{2.12}$$

$$\implies \sup_n \mathbb{E}X_n 1_{X_n \geq b} \leq \varepsilon \tag{2.13}$$

So UI. □

The above proof relies on martingale convergence theorem for existence of R-N density X_∞. It also only holds for countable events.

3 Azuma's inequality

Theorem 3.1 (Azuma's inequality). *Let $S_n = \sum_{i=1}^n X_i$ be a MG with $|X_i| \leq 1$ a.s.*

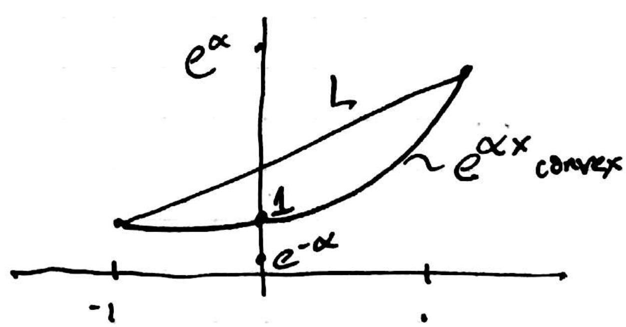
Then for $\lambda > 0$,

$$P(S_n \geq \lambda\sqrt{n}) \leq e^{-\lambda^2/2} \tag{3.1}$$

so

$$P(|S_n| \geq \lambda\sqrt{n}) \leq 2e^{-\lambda^2/2} \tag{3.2}$$

Lemma 3.2. *If $\mathbb{E}Y = 0$ and $|Y| \leq 1$, then $\mathbb{E}e^{\alpha Y} \leq e^{\alpha^2/2}$ for all α.*



Proof of Lemma 3.2.

$$\mathbb{E}e^{\alpha Y} \underset{\text{convexity}}{\leq} \mathbb{E}L(Y) \underset{\text{linear}}{=} L(\mathbb{E}Y) = L(0) = (e^\alpha + e^{-\alpha})/2 \underset{\text{calculus}}{\leq} e^{\alpha^2/2} \tag{3.3}$$

Calculus: coefficient of α^{2n} in series expansion

$$\frac{1}{(2n)!} \leq \frac{1}{2^n n!} \tag{3.4}$$

□

Proof of Azuma (Theorem 3.1). Apply lemma 3.2 to conditional distribution of X_i given $\mathcal{F}_{i-1}$

$$\mathbb{E}[e^{\alpha X_i} \mid \mathcal{F}_{i-1}] \leq e^{\alpha^2/2} \tag{3.5}$$

$$\mathbb{E}[e^{\alpha S_n} \mid \mathcal{F}_{n-1}] = e^{\alpha S_{n-1}} \mathbb{E}[e^{\alpha X_n} \mid \mathcal{F}_{n-1}) \leq e^{\alpha^2/2} e^{\alpha S_{n-1}}] \tag{3.6}$$

$$\tag{3.7}$$

Take $\mathbb{E}$ and apply tower property

$$\mathbb{E}[e^{\alpha S_n}] \leq e^{\alpha^2/2} \mathbb{E} e^{\alpha S_{n-1}} \tag{3.8}$$

$$\mathbb{E}[e^{\alpha S_n}] \leq \left(e^{\alpha^2/2}\right)^n = e^{\alpha^2 n/2} \tag{3.9}$$

Applying Markov inequality with $\phi = \exp$

$$P(S_n \geq \lambda\sqrt{n}) \leq \frac{\mathbb{E} e^{\alpha S_n}}{e^{\alpha\lambda\sqrt{n}}} \leq e^{\alpha^2 n/2 - \alpha\lambda\sqrt{n}} \tag{3.10}$$

Minimize over α by taking $\alpha = \lambda/\sqrt{n}$

$$P(S_n \geq \lambda\sqrt{n}) \leq e^{\alpha^2 n/2 - \alpha\lambda\sqrt{n}} = e^{-\lambda^2/2} \tag{3.11}$$

□

4 Method of bounded differences

Corollary 4.1. *Take $(\xi_i, 1 \leq i \leq n)$ independent, arbitrary state spaces.*

Take $\mathbb{R}$-valued $Z = f(\xi_1, \xi_2, \ldots, \xi_n)$ such that if $\tilde{x} = (x_1, \ldots, x_n)$ and $\tilde{y} = (y_1, \ldots, y_n)$ differ in one coordinate only (i.e. $|\{i : y_i \neq x_i\}| = 1$), then $|f(\tilde{x}) - f(\tilde{y})| \leq 1$.

Then $P(|Z - \mathbb{E}Z| \geq \lambda\sqrt{n}) \leq 2e^{-\lambda^2/2}$ for $\lambda > 0$.

This is useful for analysis of random algorithms: consider randomized traveling salesman where the tour $\tilde{x}$ is changed at a single location $y_i \neq x_i$.

Proof. WLOG assume $\mathbb{E}Z = 0$. Write $S_m = \mathbb{E}[Z \mid \mathcal{F}_m]$ where $\mathcal{F}_m = \sigma(\xi_i, 1 \leq i \leq m)$, so $(S_m, 1 \leq m \leq n)$ is a MG.

If we can show "S_m has bounded differences"

$$|S_m - S_{m-1}| \leq 1 \tag{4.1}$$

then Azuma's inequality (theorem 3.1) yields the desired conclusion.

Lemma 4.2. *If Y is such that any 2 possible values within 1, then $|Y - \mathbb{E}Y| \le 1$.*

Proof. $\min \operatorname{supp} Y \le Y \le \max \operatorname{supp} Y$ and $\min \operatorname{supp} Y \le \mathbb{E}Y \le \max \operatorname{supp} Y$ so

$$|Y - \mathbb{E}Y| \le \max \operatorname{supp} Y - \min \operatorname{supp} Y \le 1 \tag{4.2}$$

□

If we know all $(\xi_i, i \ne m)$ then apply lemma 4.2 conditionally

$$|Z - \underbrace{\mathbb{E}[Z \mid \xi_i, i \ne m]}_{Z^*}| \le 1 \tag{4.3}$$

$$\tag{4.4}$$

Lemma 4.3. *If W is independent of $(Y, \mathcal{G})$, then $\mathbb{E}[Y \mid \mathcal{G}, W] = \mathbb{E}[Y \mid \mathcal{G}]$.*

Proof. Exercise. □

By lemma 4.3

$$\mathbb{E}[Z^* \mid \mathcal{F}_m] = \mathbb{E}[Z^* \mid \mathcal{F}_{m-1}, \xi_m] = \mathbb{E}[Z^* \mid \mathcal{F}_{m-1}] \tag{4.5}$$

$$\text{tower property} \implies = \mathbb{E}[Z \mid \mathcal{F}_{m-1}] \tag{4.6}$$

$$\tag{4.7}$$

This is nice: Z and Z^* are both conditioned on the same σ-field $\mathcal{F}_m$ so

$$\begin{aligned} |S_m - S_{m-1}| &= |\mathbb{E}[Z \mid \mathcal{F}_m] - \mathbb{E}[Z \mid \mathcal{F}_{m-1}]| && (4.8)\\ &= |\mathbb{E}[Z \mid \mathcal{F}_m] - \mathbb{E}[Z^* \mid \mathcal{F}_m]| && (4.9)\\ &\le \mathbb{E}[|Z - Z^*| \mid \mathcal{F}_m] && (4.10)\\ \text{Equation (4.3)} \implies &\le 1 && (4.11) \end{aligned}$$

□