STATS 310B: Theory of Probability II **Winter 2016/17**

Lecture 20: March 16

Lecturer: Sourav Chatterjee *Scribes: Kenneth Tay*

20.1 Martingale Inequalities

Theorem 20.1 (Azuma-Hoeffding Inequality) *Let $\{M_k\}_{0\leq k\leq n}$ be a martingale adapted to some filtration. Let $X_k = M_k - M_{k-1}$. Suppose that $|X_k| \leq c_k$ a.s. for each k, where $c_1, \dots, c_n$ are constants. Then for all $t > 0$,*

$$P(M_n - M_0 \geq t) \leq \exp\left(-\frac{t^2}{2\sum c_k^2}\right),$$
$$P(M_n - M_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum c_k^2}\right).$$

Proof: (Sketch.)

$$\begin{aligned} P(M_n - M_0 \geq t) &\leq e^{-\theta t}\mathbb{E}\left[e^{\theta(M_n - M_0)}\right] \\ &= e^{-\theta t}\mathbb{E}\left[e^{\theta\sum_{k=1}^n X_k}\right]. \end{aligned}$$

Note that if $\mathbb{E}X = 0$, $|X| \leq c$ a.s., then

$$\begin{aligned} \mathbb{E}\left[e^{\theta X}\right] &= \int_{-c}^{c} e^{\theta x}\mu(dx) && (\mu \text{ the law of } X) \\ &= \int_{-c}^{c} \exp\left[t(-c\theta) + (1-t)(c\theta)\right]\mu(dx) && (\text{where } t = \frac{c-x}{2c}) \\ &\leq \int_{-c}^{c} te^{-\theta c} + (1-t)e^{\theta c}\mu(dx) && (\text{Jensen}) \\ &= e^{-\theta c}\mathbb{E}\left(\frac{c-X}{2c}\right) + e^{\theta c}\mathbb{E}\left(1 - \frac{c-X}{2c}\right) \\ &= \cosh(\theta c) \\ &\leq \exp\left(\frac{\theta^2 c^2}{2}\right). \end{aligned}$$

Hence, using the conditional version of the above,

$$\begin{aligned} \mathbb{E}\left[e^{\theta\sum_{k=1}^n X_k}\right] &= \mathbb{E}\left[e^{\theta\sum_{k=1}^{n-1} X_k}\mathbb{E}\left[e^{\theta X_n}\middle|\mathcal{F}_{n-1}\right]\right] \\ &\leq \mathbb{E}\left[e^{\theta\sum_{k=1}^{n-1} X_k}e^{\theta^2 c^2/2}\right] \\ &\vdots \\ &\leq \exp\left(\frac{\theta^2\sum c_k^2}{2}\right), \end{aligned}$$

so

$$\begin{aligned} P(M_n - M_0 \geq t) &\leq e^{-\theta t}\mathbb{E}\left[e^{\theta \sum_{k=1}^n X_k}\right] \\ &\leq e^{-\theta t} \exp\left(\frac{\theta^2 \sum c_k^2}{2}\right). \end{aligned}$$

Optimizing over θ, we get the desired inequality. ■

Theorem 20.2 (Bounded Difference Inequality) *Let $X_1, \ldots, X_n$ be independent random variables. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function such that there exist $c_1, \ldots, c_n$ with the property that for all $x_1, \ldots, x_n$, $x'_1, \ldots, x'_n$,*

$$|f(x_1, \ldots, x_n) - f(x_1, \ldots, x'_i, \ldots, x_n)| \leq c_i.$$

Let $W = f(X_1, \ldots, X_n)$. Then for all $t > 0$,

$$\begin{aligned} P(W - \mathbb{E}W \geq t) &\leq \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right), \\ P(W - \mathbb{E}W \leq -t) &\leq \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right). \end{aligned}$$