STAT 620: Asymptotic Statistics **Spring 2022**

Lecture: Mar 24

Lecturer: Xianyang Zhang

1 Martingale and martingale difference sequences

1.1 Filtration

A filtration on a probability space $(\Omega, \mathcal{B}, P)$ is a sequence of sub-sigma fields $\{\mathcal{F}_n : n = 0, 1, 2, \dots\}$ such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$.

1.2 Martingale

A process $\{(Y_n, \mathcal{F}_n) : n = 0, 1, 2, \dots\}$ is a martingale if:

1. $\{\mathcal{F}_n\}$ is a filtration and $Y_n \in \mathcal{F}_n$.
2. Y_n is integrable.
3. For each n, $\mathbb{E}[Y_n|\mathcal{F}_{n-1}] = Y_{n-1}$.

An example. Let $Y_n = \sum_{i=1}^n X_i$, where $X_i \overset{i.i.d}{\sim} P$ with $E[X_i] = 0$ and $\mathcal{F}_i = \sigma(X_1, X_2, \cdots, X_i)$. Then $\{(Y_n, \mathcal{F}_n) : n = 0, 1, 2, \dots\}$ is a martingale.

1.3 Martingale difference sequence

A sequence $\{(X_n, \mathcal{F}_n) : n = 0, 1, 2, \dots\}$ is a martingale difference sequence if:

1. $\{\mathcal{F}_n\}$ is a filtration and $X_n \in \mathcal{F}_n$.
2. X_n is integrable.
3. $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = 0 \ \forall \ n \in \mathbb{N}$.

Let Y_n be a martingale and $X_n = Y_n - Y_{n-1}$. We then have

$$\begin{aligned}\mathbb{E}[X_n|\mathcal{F}_{n-1}] =& \mathbb{E}[Y_n|\mathcal{F}_{n-1}] - \mathbb{E}[Y_{n-1}|\mathcal{F}_{n-1}] \\ =& Y_{n-1} - Y_{n-1} = 0.\end{aligned}$$

On the other hand, if $Y_n = \sum_{i=1}^n X_i$ and X_i is a martingale difference sequence (MGD), then Y_n is a martingale.

2 Sub-Gaussianity

Let X_i be a MGD. Then, it is σ_i^2-sub-Gaussian if:

$$\mathbb{E}[\exp(\lambda X_i)|\mathcal{F}_{i-1}] \leq \exp(\lambda^2 \sigma_i^2/2) \ \forall \ i \in \mathbb{N}.$$

An example. If X_i is a MGD and $L_i \leq X_i \leq U_i$, where $L_i, U_i \in \mathcal{F}_{i-1}$, and $U_i - L_i \leq c_i$ (c_i is a constant). Then, X_i is $\frac{c_i^2}{4}$-sub-Gaussian MGD.

2.1 Theorem

If $\{X_i\}$ is σ_i^2-sub-Gaussian MGD, then for any $t \geq 0$,

$$P\left(\sum_{i=1}^{n}(X_i - \mathbb{E}X_i) \geq t\right) \vee P\left(\sum_{i=1}^{n}(X_i - \mathbb{E}X_i) \leq -t\right) \leq \exp\left(-\frac{t^2}{2\sum_{i=1}^{n}\sigma_i^2}\right).$$

Proof.

$$\begin{aligned}
\mathbb{E}[\exp(\lambda \sum_{i=1}^{n} X_i)] =& \mathbb{E}[\mathbb{E}[\exp(\lambda \sum_{i=1}^{n} X_i)|\mathcal{F}_{n-1}]] \\
=& \mathbb{E}[\exp(\lambda \sum_{i=1}^{n-1} X_i)\mathbb{E}[\exp(\lambda X_n|\mathcal{F}_{n-1})]] \\
\leq& \mathbb{E}[\exp(\lambda \sum_{i=1}^{n-1} X_i)] \exp\left(\frac{\lambda^2\sigma_n^2}{2}\right) \\
\leq& \exp\left(\frac{\lambda^2 \sum_{i=1}^{n}\sigma_i^2}{2}\right).
\end{aligned}$$

Therefore, $\sum_{i=1}^{n} X_i$ is $\sum_{i=1}^{n}\sigma_i^2$-sub-Gaussian. Now, apply the Chernoff bounds to complete the proof.

3 Martingale decomposition

Let $\{X_i\}_{i=1}^{n}$ be independent random variables (each X_i takes values in the space $\mathcal{X}$). Let $f : \mathcal{X}^n \to \mathbb{R}$. Our goal is to Control

$$f(X_1, \cdots, X_n) - \mathbb{E}[f(X_1, \cdots, X_n)].$$

To this end, Define $X_{1:n} = (X_1, \ldots, X_n)$ and let $\mathcal{F}_i = \sigma(X_1, \ldots, X_i)$. Also, define $D_i = \mathbb{E}[f(X_{1:n})|\mathcal{F}_i] - \mathbb{E}[f(X_{1:n})|\mathcal{F}_{i-1}]$.

We claim that D_i is MGD. Note

$$\mathbb{E}[D_i|\mathcal{F}_{i-1}] = \mathbb{E}[\mathbb{E}[f|\mathcal{F}_i]|\mathcal{F}_{i-1}] - \mathbb{E}[f|\mathcal{F}_{i-1}] = \mathbb{E}[f|\mathcal{F}_{i-1}] - \mathbb{E}[f|\mathcal{F}_{i-1}] = 0.$$

Moreover, we have

$$f - \mathbb{E}f = \sum_{i=1}^{n} D_i.$$

We aim to impose conditions on f such that $\sum_{i=1}^{n} D_i$ is sub-Gaussian.

3.1 Bounded difference

We say f satisfies the c_i bounded differences if:

$$|f(x_{1:i-1}, x_i, x_{i+1:n}) - f(x_{1:i-1}, x_i', x_{i+1:n})| \leq c_i \text{ for all } x_1, \ldots, x_n, x_i' \in \mathcal{X} \text{ and all } 1 \leq i \leq n.$$

In this case, $f(X_{1:n}) - \mathbb{E}f(X_{1:n})$ is $\frac{\sum_{i=1}^{n} c_i^2}{4}$-sub-Gaussian. To this end, we only need to show each D_i is $c_i^2/4$ sub-Gaussian. Recall that

$$D_i = \mathbb{E}[f(X_{1:n})|\mathcal{F}_i] - \mathbb{E}[f(X_{1:n})|\mathcal{F}_{i-1}].$$

Let

$$U_i = \sup_{\tilde{x}_i}\left[\int f(X_{1:i-1}, \tilde{x}_i, x_{i+1:n})dP(x_{i+1:n}) - \int f(X_{1:i-1}, x_{i:n})dP(x_{i:n})\right],$$

and

$$L_i = \inf_{\tilde{x}_i} \left[\int f(X_{1:i-1}, \tilde{x}_i, x_{i+1:n}) dP(x_{i+1:n}) - \int f(X_{1:i-1}, x_{i:n}) dP(x_{i:n}) \right].$$

We observe that

$$L_i \le D_i \le U_i$$

and

$$U_i - L_i \le c_i.$$

So D_i is $c_i^2/4$ sub-Gaussian.

3.2 Corollary

If $f : \mathcal{X}^n \to \mathbb{R}$ satisfies c_i-bounded difference, then for any $t \ge 0$,

$$P(f(X_{1:n}) - \mathbb{E}f(X_{1:n}) \ge t) \le \exp\left(-\frac{2t^2}{\sum_{i=1}^n c_i^2} \right).$$

Let $\mathcal{F}$ be a class of functions from $\mathcal{X}$ to $\mathbb{R}$. Assume that for any $f \in \mathcal{F}$ and any $x, x' \in \mathcal{X}$.

$$|f(x) - f(x')| \le B < \infty.$$

Under the above assumption,

$$\sup_{f \in \mathcal{F}} \left[\frac{1}{n} \sum_{i=1}^n (f(x_i) - \mathbb{E}f(X_i)) \right]$$

and

$$\sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n (f(x_i) - \mathbb{E}f(X_i)) \right|$$

as functions of $(x_1, \ldots, x_n)$ satisfy $\frac{B}{n}$ bounded difference.