Mathematical Statistical Notes **Yajie Bao**

Chapter 1: Probability Theory

1.1 Probability Inequalities

Theorem 1.1 (The Gaussian Tail Inequality) *Let $X \sim N(0,1)$. Then*

$$\mathbb{P}(|X| > \epsilon) \leq \frac{2e^{-\epsilon^2/2}}{\epsilon}.$$

If $X_1, ..., X_n i.i.d \sim N(0,1)$, then

$$\mathbb{P}\left(\left|\bar{X}_n\right| > \epsilon\right) \leq \frac{2}{\sqrt{n}\epsilon} e^{-n\epsilon^2/2} \overset{large\ n}{\leq} e^{-n\epsilon^2/2}. \tag{1.1}$$

Lemma 1.2 *Supposed that $a \leq X \leq b$. Then*

$$\mathbb{E}\left(e^{tX}\right) \leq e^{t\mu} e^{\frac{t^2(b-a)^2}{8}}. \tag{1.2}$$

Theorem 1.3 (Hoeffding's Inequality) *Let $Y_1, ..., Y_n$ be i.i.d observations such that $\mathbb{E}(Y_i) = \mu$ and $a \leq Y_i \leq b$. Then for any $\epsilon > 0$,*

$$\mathbb{P}\left(\left|\bar{Y}_n - \mu\right| \geq \epsilon\right) \leq 2e^{-2n\epsilon^2/(b-a)^2}. \tag{1.3}$$

Theorem 1.4 (Bernstein Inequality) *If $X_1, ..., X_n$ are independent and $\mathbb{E}(X_i) = 0$. if $|X_i| \leq b$,then we have*

$$P\left(|\sum_{i=1}^n X_i| \geq \epsilon\right) \leq \exp\left[-\frac{\epsilon^2}{2\sum_{i=1}^n \mathbb{E}\left(X_i\right)^2 + 2\epsilon/3b}\right]. \tag{1.4}$$

Remark: If $\sum_{i=1}^n E(X_i^2) = O(n)$, then let $\epsilon = C\frac{\sqrt{\log n}}{\sqrt{n}}$ we have

$$P\left(|\frac{1}{n}\sum_{i=1}^n X_i| \geq C\frac{\sqrt{\log n}}{\sqrt{n}}\right) \leq n^{-C'}. \tag{1.5}$$

Note that for i.i.d. random variables $X_i, i = 1, 2, ..., n$ (satisfy CLT conditions), we have $\frac{1}{n}\sum_{i=1}^n (X_i - E(X_i)) = O_p(1/\sqrt{n})$, but it's not a uniform result.

For example, we want to obtain the convergence rate of $\max_{k\leq N} \frac{1}{n}\sum_{i\in H_k}(X_i - E(X_i))$,

$$P\left(\max_{k\leq N} \frac{1}{n}\sum_{i\in H_k}(X_i - E(X_i)) \geq ca_n\right) \leq \sum_{k=1}^N P\left(\frac{1}{n}\sum_{i\in H_k}(X_i - E(X_i))\right),$$

then we need Bernstein inequality (1.4) rather than CLT. Combine (1.4) and (1.5), we have

$$\max_{k\leq N} \frac{1}{n}\sum_{i\in H_k}(X_i - E(X_i)) = O_p(\frac{\sqrt{\log n}}{\sqrt{n}})$$

.

Theorem 1.5 (Mcdiarmid) *Let $X_1, ..., X_n$ be independent r.v. Supposed that*

$$\sup_{x_1,\ldots,x_n,x_i'} \left|g\left(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n\right)-g\left(x_1,\ldots,x_{i-1},x_i',x_{i+1},\ldots,x_n\right)\right| \le c_i$$

for $i = 1, 2, ..., n$. Then

$$\mathbb{P}\left(g\left(X_1,\ldots,X_n\right)-\mathbb{E}\left(g\left(X_1,\ldots,X_n\right)\right)\ge\epsilon\right)\le\exp\left\{-\frac{2\epsilon^2}{\sum_{i=1}^n c_i^2}\right\}. \tag{1.6}$$

1.2 Moment Inequalities

1. H'older's inequality: $E|XY| \le \left(E|X|^p\right)^{1/p}\left(E|Y|^q\right)^{1/q}$

2. Liapounov's inequality: $\left(E|X|^r\right)^{1/r} \le \left(E|X|^s\right)^{1/s}$, where r and s are constants satisfying $1 \le r \le s$

3. Minkowski's inequality: $\left(E|X+Y|^p\right)^{1/p} \le \left(E|X|^p\right)^{1/p} + \left(E|Y|^p\right)^{1/p}$, where X and Y are random variables and $p \ge 1$ is a constant.

4. Esseen and von Bahr (1965): Let $X_1, ..., X_n$ be independent random variables with mean 0 and $E\left|X_i\right|^p < \infty, i = 1, 2, ..., n$, where p is a constant in $[1, 2]$, then

$$E\left|\sum_{i=1}^n X_i\right|^p \le C_p \sum_{i=1}^n E\left|X_i\right|^p \tag{1.7}$$

   where C_p is a constant depending only on p.

5. C_r inequality: $E|X+Y|^r \le C_r\left(E|X|^r + E|Y|^r\right)$, where $C_r = 0$ if $0 < r \le 1$ and $C_r = 2^{r-1}$ if $r > 1$.

6. Marcinkiewicz and Zygmund: For $p \ge 2$,

$$E\left|\sum_{i=1}^n X_i\right|^p \le \frac{C_p}{n^{1-p/2}} \sum_{i=1}^n E\left|X_i\right|^p. \tag{1.8}$$

   where C_p is a constant depending only on p.

7. H ájek and Rènyi: Let $Y_1, ..., Y_n$ be independent random variables having finite variances, then we have

$$P\left(\max_{1\le l\le n} c_l\left|\sum_{i=1}^l\left(Y_i-EY_i\right)\right|>t\right)\le\frac{1}{t^2}\sum_{i=1}^n c_i^2\,\mathrm{Var}\left(Y_i\right), \quad t>0 \tag{1.9}$$

   where c_i's are positive constants satisfying $c_1 \le c_2 \le ... \le c_n$. If $c_i = 1$, then inequality reduces to famous Kolmogorov's inequality.

Corollary 1.6 *For independent random variables $X_1, ..., X_n$ with $\sup_n \mathbb{E}(X_n) < \infty$, we have*

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n\left(X_i-EX_i\right)\right|>t\right)\le\begin{cases} O\left(n^{1-p}\right) & \text{if } 1<p<2 \\ O\left(n^{-p/2}\right) & \text{if } p\ge 2\end{cases}$$

1.3 Asymptotic theory

1.3.1 Convergence Mode

Theorem 1.7 *For random k-vectors $X, X_1, X_2, ...$ on a probability space, $X_n \to a.s.X$ if and only if for every $\epsilon > 0$*

$$\lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty}\{\|X_m - X\| > \epsilon\}\right) = 0.$$

Theorem 1.8 *Let $\{X_n\}$ be a sequence of* ***independent*** *random variables, then $X_n \to a.s.0$ if and only if for every $\epsilon > 0$*

$$\sum_{i=1}^{n} P(|X_n| \geq \epsilon) < \infty.$$

Remark:

1. $X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow Y$ does not imply that $X_n + Y_n \rightsquigarrow X + Y$.
2. $X_n \xrightarrow{P} b$ does not imply that $\mathbb{E}(X_n) \to b$.
3. If $X_n \xrightarrow{P} X$ and $|X_n|^r$ is uniformly integrable, then $X_n \xrightarrow{r} X$

Theorem 1.9 *Let $X_1, X_2, ...$ be a sequence of identically distributed random variables with a finite $E|X_1|$ and let $Y_n = n^{-1}\max_{i\leq n}|X_i|$, then (1) $Y_n \to_{L_1} 0$ (2) $Y_n \to_{a.s.} 0$*

1.3.2 Stochastic convergence rate

Theorem 1.10 *Let $X, X_1, X_2, ..., Y_1, Y_2, ..., Z_1, Z_2, ...$ be random variables, then the following statements hold:*

(a) If $X_n \to_d X$, then $X_n = O_p(1)$.

(b) If $X_n = O_p(Z_n)$ and $P(Y_n = 0) = 0$, then $X_nY_n = O_p(Y_nZ_n)$.

(c) If $X_n = O_p(Z_n)$ and $Y_n = O_p(Z_n)$, then $X_n + Y_n = O_p(Z_n)$.

(d) If $E|X_n| = O(a_n)$, then $X_n = O_p(a_n)$, where $a_n \in (0, \infty)$.

(e) If $X_n \to_{a.s.} X$, then $sup_n|X_n| = O_p(1)$.

1.4 The law of large numbers

Theorem 1.11 (Jun Shao p62) *Let $X_1, X_2, ...$ be a sequence of independent identically distributed random variables,*

(i) (The WLLN) A necessary and sufficient condition for the existence of a sequence of real numbers $\{a_n\}$ for which

$$\frac{1}{n}\sum_{i=1}^{n} X_i - a_n \to_p 0$$

is that $nP\left(|X_1| > n\right) \to 0$*, in which case take* $a_n = E\left(X_1 I_{\{|X_1| \le n\}}\right)$.

(ii) (The SLLN) A necessary and sufficient condition for the existence of a constant c for which

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to_{a.s.} c$$

is that $\mathbb{E}(|X_1|) < \infty$*, in which case take* $c = E(X_1)$ *and*

$$\frac{1}{n}\sum_{i=1}^{n} c_i\left(X_i - EX_1\right) \to_{a.s.} 0$$

for any bounded sequence of real numbers $\{c_i\}$.

Theorem 1.12 (Jun Shao p65) *Let* $X_1, X_2, \ldots$ *be a sequence of independent random variables with finite expectations.*

(i) (The SLLN) If there is a constant $p \in [1, 2]$ *such that*

$$\sum_{i=1}^{\infty} \frac{E\left|X_i\right|^p}{i^p} < \infty$$

then

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - EX_i\right) \to_{a.s.} .$$

(ii) (The WLLN) If there is a constant $p \in [1, 2]$ *such that*

$$\lim_{n\to\infty} \frac{1}{n^p}\sum_{i=1}^{n} E\left|X_i\right|^p = 0$$

then

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - EX_i\right) \to_p 0.$$