MSc Exam Asymptotic Statistics

Date: Friday November 4, 2016 (pnt+8/7.8

Time: 9.00-12.00 hours

Place: LB 5173.0055

Progress code: WBMA14003

Rules of the exam:

- It is allowed to use the enumeration of theorems as well as a simple calculator e.g. CASIO FX-82 or TI30.
- Provide each page with your name and student number.
- The number of points per question are indicated by a box.
- We wish you a lot of success with the completion of the exam!

1. Suppose that $H_1, H_2, \cdots$ is a sequence of independent Bernoulli random variables with $P(H_n = 1) = 1/n$. Let $X_n = (-1)^n \, n \, H_n$, for $n = 1, 2, \cdots$.

   (a) $\boxed{7}$ Show that X_n tends to zero in probability.

   (b) $\boxed{8}$ Show that X_n does not tend to zero almost surely.
   Hint: Consider $\limsup_{n\to\infty} X_n$ and $\liminf_{n\to\infty} X_n$.

2. Let $X, X_1, X_2, \cdots,$ be i.i.d. with probability mass function,

$$P(X = j) = P(X = -j) = \frac{c}{j^2 \log j}, \text{ for, } j = 3, 4, \cdots; \quad c = 2\sum_{j=3}^{\infty} \frac{1}{j^2 \log j},$$

   where c is the normalizing constant. In the question you will show that $\bar{X}_n \xrightarrow{P} 0$. This gives an example of a distribution that obeys the weak law of large numbers even though E(X) does not exist. Note that Theorem 4(c) implies that $\bar{X}_n$ does not converge a.s. to 0. Show that $\bar{X}_n \xrightarrow{P} 0$ in the following three steps.

   (a) $\boxed{5}$ Let $Y_{n,k} = X_k I(|X_k| \leq n)$ for $k = 1, \cdots, n$. Show that $\bar{Y}_n \xrightarrow{q.m.} 0$ so that $\bar{Y}_n \xrightarrow{P} 0$.

   (b) $\boxed{5}$ Show that $P(\bar{X}_n \neq \bar{Y}_n) \leq \sum_{k=1}^{n} P(X_k \neq Y_{n,k}) \to 0$, as $n \to \infty$.

   (c) $\boxed{5}$ Use (a) and (b) to conclude $\bar{X}_n \xrightarrow{P} 0$.

3. Assume for a triangular array of independent variables that the X_{nj} are uniformly bounded, say $|X_{n,j}| < A$ for all n and j and some fixed constant A. Let $S_n = \sum_{j=1}^{n} X_{nj}$.

   (a) $\boxed{10}$ Suppose that $\mathrm{Var}(S_n) \to \infty$, as $n \to \infty$. Show that

$$\frac{S_n - E(S_n)}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{D} N(0,1).$$

   (b) $\boxed{10}$ Apply this to the binomial random variable, $Y_n \in \mathcal{B}(n, p_n)$, (which is a sum of independent Bernoulli random variables) in the case $p_n = 1/\sqrt{n}$ to show that

$$\sqrt[4]{n}\left(\frac{Y_n}{\sqrt{n}} - 1\right) \xrightarrow{D} N(0,1).$$

4. (Sequential Information Inequality.) Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables whose density, $f(x|\theta)$, exists and satisfies the conditions of Theorem 19. Let N be a stopping time (i.e. N is a random variable taking positive integer values such that for all n, the event $\{N = n\}$ depends only on the variables $X_1, \cdots, X_n$). Let $\{\hat{\theta}_n(X_1, \cdots, X_n)\}_{n=1}^{\infty}$ be a sequence of estimates of θ with finite expectations and consider the estimate $\sum_{n=1}^{\infty} \hat{\theta}_n(X_1, \cdots, X_n) I\{N = n\} = \hat{\theta}_N$. Let $g(\theta) = E_\theta \hat{\theta}_N$.

   (a) $\boxed{15}$ Show that

$$\mathrm{Var}_\theta(\hat{\theta}_N) \geq \frac{g'(\theta)^2}{\mathcal{F}(\theta) E_\theta N}.$$

   Note: You may use the Wald equations: If $Y_1, Y_2, \cdots$ are i.i.d. with finite mean, μ, and if N is a stopping time depending on the Ys such that $EN < \infty$, then

$$E(Y_1 + \cdots + Y_N) = \mu E(N).$$

   If in addition $\mu = 0$ and the variance is finite, then

$$E(Y_1 + \cdots + Y_N)^2 = \mathrm{Var}(Y_1) E(N).$$

MSc

1(a) $P(|X_n - 0| > \varepsilon) = P(|(-1)^n n H_n| > \varepsilon) = P(H_n = 1)$ 5

$= \frac{1}{n} \to 0, \ n \to \infty$ 1

$X_n \xrightarrow{P} 0$ 1

(b) $H_1, H_2, \ldots$ indep. $\Rightarrow X_1, X_2, \ldots$ indep. 2

$\sum_{n=1}^{\infty} P(X_n = n) = \sum_{n=1}^{\infty} \frac{1}{2} \frac{1}{n} = \infty$ (n even) 2

$\overset{BC}{\Longrightarrow} P(\{X_n = n\} \text{ i.o.}) = 1$ 2

$\Longrightarrow P(\limsup_{n\to\infty} X_n = \infty) = 1$

$\Longrightarrow \limsup_{n\to\infty} X_n \xrightarrow{a.s.} \infty$ 2

similarly $\liminf_{n\to\infty} X_n = -\infty$

MSc

2 (a) $X_1, X_2, \dots$ iid $\Rightarrow Y_{n,1}, Y_{n,2}, \dots$ iid

$$E[X_k I(|X_k| \leq n)] = \sum_{\substack{j=-n \\ |j| \geq 3}}^{n} j\, P(X_k = j) = 0 \quad (*) \text{ by } P(X_k = j) = P(X_k = -j)$$ 2

$$E[Y_{n,k}^2] = E[X_k^2\, I(X_k \leq n)]$$

$$= \sum_{j=3}^{n} 2 \cdot j^2\, P(X_k = j) = 2c \sum_{j=3}^{n} \frac{1}{\log j} < \infty \quad \forall_{k=3}^{n<\infty}$$ 2

$$\overset{T4(b)}{\Longrightarrow} \frac{1}{n} \sum_{k=1}^{n} Y_{n,k} = \bar{Y}_n \xrightarrow{qm} E[Y_{n,k}] = 0 \quad \text{by } (*)$$ 1

(b) $$P(\bar{X}_n \neq \bar{Y}_n) = P\left(\sum_{k=1}^{n} X_k \neq \sum_{k=1}^{n} Y_{n,k}\right)$$ 1

$$\leq P\left(\bigcup_{k=1}^{n} X_k \neq Y_{n,k}\right) = \sum_{k=1}^{n} P(X_k \neq Y_{n,k})$$

$$= n P(X_k \neq Y_{n,k}) = n\, P(|X_k| > n)$$ 1

$$= n \sum_{j=n+1}^{\infty} P(|X_k| = j)$$ 1

$$= n \cdot 2 \sum_{j=n+1}^{\infty} \frac{c}{j^2 \log j}$$ tail of distribution. 1

$$< \varepsilon \quad \text{if} \quad \sum_{j=n+1}^{\infty} \frac{2c}{j^2 \log j} < \frac{\varepsilon}{n}$$ 1

(c) $$\left.\begin{array}{l} \bar{Y}_n \xrightarrow{P} 0 \\ \bar{Y}_n - \bar{X}_n \xrightarrow{P} 0 \end{array}\right\} \overset{T6b}{\Longrightarrow} \bar{X}_n \xrightarrow{P} 0$$ 5

MSc

3 $B_n^2 = Var[S_n] = \sum_{j=1}^{n} Var[X_{nj}] \nearrow \infty$ by supposition.

(a) Fix $\varepsilon > 0$.

$\exists n_\varepsilon : \varepsilon B_n > A > |X_{nj}| \quad \forall j, \forall n > n_\varepsilon$

$\Rightarrow I_{\{|X_{nj}| \geq \varepsilon B_n\}}(x) = 0 \quad \forall x$ 3

$\Rightarrow E\{X_{nj}^2 I_{(|X_{nj}| \geq \varepsilon B_n)}\} = 0 \quad \forall j$ Lindeberg condition holds 5

$$\overset{(T5)}{\Rightarrow} \frac{Z_n}{B_n} = \frac{\sum X_{nj} - E[\sum X_{nj}]}{\sqrt{\sum_{j=1}^{n} Var[X_{nj}]}} = \frac{S_n - ES_n}{\sqrt{Var[S_n]}} \xrightarrow{D} n(0,1)$$ 2

(b) $\sum_{j=1}^{n} X_{nj} = Y_n = S_n \in \text{Binomial}(n, n^{-\frac{1}{2}})$

$\Rightarrow EY_n = n \cdot n^{-\frac{1}{2}} = n^{\frac{1}{2}}$ 2

$Var[Y_n] = Var[S_n] = n \cdot n^{-\frac{1}{2}}(1 - n^{-\frac{1}{2}}) = n^{\frac{1}{2}} - 1 \sim n^{\frac{1}{2}}$ 2

$$\overset{L\text{-}FT}{\Rightarrow} \frac{Y_n - EY_n}{\sqrt{Var[Y_n]}} = \frac{\sqrt{n^{1/2}}}{\sqrt{n^{1/2}}} \cdot \frac{Y_n - n^{\frac{1}{2}}}{\sqrt{n^{1/2}}}$$ 2

$$= \sqrt[4]{n}\left(\frac{Y_n}{\sqrt{n}} - 1\right)$$ 2

$$\xrightarrow{D} n(0,1)$$ 2

MSc

4(a) $\frac{\partial}{\partial\theta} \log f(x_1,\dots,x_n|\theta) = \frac{\frac{\partial}{\partial\theta} f(x_1,\dots,x_n|\theta)}{f(x_1,\dots,x_n|\theta)} \Longrightarrow$

$$\frac{\partial}{\partial\theta} f(x_1,\dots,x_n|\theta) = \left(\frac{\partial}{\partial\theta} \log f(x_1,\dots,x_n)\right) \cdot f(x_1,\dots,x_n|\theta)$$

$$= \Psi(x_1,\dots,x_n|\theta) \cdot f(x_1,\dots,x_n|\theta)$$ 5

depends only on $x_1,\dots,x_n$ not on θ

$$g'(\theta) = \frac{\partial}{\partial\theta} E\hat{\theta}_n = \frac{\partial}{\partial\theta} \int \hat{\theta}_n \cdot f(x_1,\dots,x_n|\theta)\, d\nu(x_1,\dots,x_n)$$

assoTig

$$= \int \hat{\theta}_n \frac{\partial}{\partial\theta} f(x_1,\dots,x_n|\theta)\, d\nu(x_1,\dots,x_n)$$

$$= \int \hat{\theta}_n\ \Psi(x_1,\dots,x_n) \cdot f(x_1,\dots,x_n|\theta)\, d\nu(x_1,\dots,x_n)$$

$$= E\left[\hat{\theta}_n(X_1,\dots,X_n) \cdot \Psi(X_1,\dots,X_n|\theta)\right] = Cov[\hat{\theta}_n, \Psi]$$ 5

$$; E[\Psi(X_1,\dots,X_n|\theta] = 0$$

$$(g'(\theta))^2 = \left(Cov[\hat{\theta}_n \cdot \Psi(X_1,\dots,X_n)]\right)^2$$

$$\leq Var[\hat{\theta}_n] \cdot Var[\Psi(X_1,\dots,X_n)]$$

$$= Var[\hat{\theta}_n] \cdot Var[\Psi(X_1)] \cdot E[n]$$

$$\frac{[g'(\theta)]^2}{\mathcal{I}(\theta) E[n]} \leq Var[\hat{\theta}_n]$$ 5