Quiz#3, Mathematical Statistics I, 2013 Fall

Time 9:00-10:10 102275011

Name: 潘奇鴻 PAN, chi-hung 9/15

1) **Clearly prove and derive. Reduce scores for missing explanation.**
2) **If you use some notations (e.g., sample mean, sample variance), you should define their formulas.**

+1/3.

1. [+3] Let X_n and Y_n be sequences of random variables with $X_n \xrightarrow{P} x$ and $Y_n \xrightarrow{P} y$. Let $a \in R$ and $b \in R$ be constants (not necessary positive). Show that $aX_n + bY_n \xrightarrow{P} \theta$ for some $\theta \in R$.

[Every line of your proof, please explain which inequality is used. Score is reduced for missing lines, inequalities, and explanations].

$X_n \xrightarrow{P} x,\ Y_n \xrightarrow{P} y$

$\forall \varepsilon > 0$

$\Rightarrow 0 \le P(|aX_n + bY_n - (ax+by)| \ge \varepsilon) \le P(|aX_n - ax| + |bY_n - by| \ge \varepsilon)$ $(\because A \subseteq B \Rightarrow P(A) \le P(B))$

$\le P(|aX_n - ax| \ge \frac{\varepsilon}{2}$ or $|bY_n - by| \ge \frac{\varepsilon}{2})$ ✓ ✗ More clearly. explain.

$\le P(|aX_n - ax| \le \frac{\varepsilon}{2}) + P(|bY_n - by| \le \frac{\varepsilon}{2})$ $(\because P(A\cup B) = P(A)+P(B)-P(A\cap B) \Rightarrow P(A)+P(B) \ge P(A\cup B))$

$= P(|X_n - x| \le \frac{\varepsilon}{2|a|}) + P(|Y_n - y| \le \frac{\varepsilon}{2|b|})$

$\Rightarrow 0 \le \lim_{n\to\infty} P(|aX_n + bY_n - (ax+by)| \ge \varepsilon)$

$\le \lim_{n\to\infty} P(|X_n - x| \le \frac{\varepsilon}{2|a|}) + \lim_{n\to\infty} P(|Y_n - y| \le \frac{\varepsilon}{2|b|})$ $(\because X_n \xrightarrow{P} X,\ Y_n \xrightarrow{P} Y)$

$= 0 \quad \forall \varepsilon \quad \therefore aX_n + bY_n \xrightarrow{P} \theta$ with $\theta = ax + by$, $a, b \in R$ #

+2/2

2. Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean. Let $(\bar{X})^2$ be an estimator of μ^2. Prove or disprove the following.

1) [+1] $(\bar{X})^2$ is unbiased for μ^2.
2) [+1] $(\bar{X})^2$ is consistent for μ^2.

1) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ with $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$\Rightarrow E(\bar{X}^2) = Var(\bar{X}) + [E(\bar{X})]^2 = \frac{\sigma^2}{n} + \mu^2 \ne \mu^2$

$\therefore \bar{X}^2$ is not unbiased for μ^2

2) $0 \le P(|\bar{X} - \mu| \ge \varepsilon) = P(|\bar{X} - \mu|^2 \ge \varepsilon^2) \le \frac{E(\bar{X}-\mu)^2}{\varepsilon^2} = \frac{Var(\bar{X})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$

$0 \le \lim_{n\to\infty} P(|\bar{X} - \mu| \ge \varepsilon) \le 0 \Rightarrow \lim_{n\to\infty} P(|\bar{X} - \mu| \ge \varepsilon) = 0$

$\therefore \bar{X} \xrightarrow{P} \mu$ let $h(x) = x^2$ by continuous mapping theorem

$h(\bar{X}) \xrightarrow{P} h(\mu) \quad \therefore \bar{X}^2 \xrightarrow{P} \mu^2$ #

+2/2

3. [+2] Let $X_1,\dots,X_n \overset{iid}{\sim} \text{uniform}(0,\theta)$ and $X_{(n)} = \max\{X_1,\dots,X_n\}$. Then, as $n \to \infty$, the random variable $n(\theta - X_{(n)})$ converges *in distribution* to what distribution?

$$P(n(\theta - X_{(n)}) \le t) = P(\theta - X_{(n)} \le \tfrac{t}{n})$$
$$= P(X_{(n)} > \theta - \tfrac{t}{n}) = 1 - P(X_{(n)} \le \theta - \tfrac{t}{n})$$
$$= 1 - \prod_{i=1}^{n} P(X_i \le \theta - \tfrac{t}{n}) = 1 - \left(\int_0^{\theta - \frac{t}{n}} \frac{1}{\theta}\, dx\right)^n = 1 - \left[\tfrac{1}{\theta}(\theta - \tfrac{t}{n})\right]^n$$
$$\Rightarrow \lim_{n\to\infty} P(n(\theta - X_{(n)}) \le t) = \lim_{n\to\infty} 1 - \left(1 - \frac{t}{n\theta}\right)^n = 1 - e^{-\frac{t}{\theta}}\ \forall t, \text{ and } X \text{ is continuous}$$

$n(\theta - X_{(n)}) \xrightarrow{d} Y$ with $Y \sim \exp(\frac{t}{\theta})$

$\text{Exp}(\theta)$

+3/5

4. Let $X_1,\dots,X_n \overset{iid}{\sim} N(\mu_X, \sigma_X^2)$ and $Y_1,\dots,Y_n \overset{iid}{\sim} N(\mu_Y, \sigma_Y^2)$. Assume that $X_1 \perp Y_1$, $\forall i = 1,\dots,n$.

1) [+1]Derive $\theta = P(X_1 \le Y_1)$. +1
2) [+1]Find a consistent estimator $\hat{\theta}$ of θ. +1
3) [+3]Prove the consistency of $\hat{\theta}$. +1

1) $X_1 - Y_1 \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$

$$\theta = P(X_1 \le Y_1) = P(X_1 - Y_1 \le 0) = P\left(\frac{X_1 - Y_1 - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \le \frac{-(\mu_X - \mu_Y)}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right)$$
$$= \Phi\left(\frac{-(\mu_X - \mu_Y)}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) \text{ with } \Phi: \text{c.d.f of } N(0,1).$$

2) by previous. $\bar{X} \xrightarrow{P} \mu_X$, $\bar{Y} \xrightarrow{P} \mu_Y$

Define $\bar{X} :=$, $\bar{Y} =$

by chebyshev inequality

$$P(|S_n^2 - \sigma^2| \ge \varepsilon) = \frac{Var(S_n^2)}{\varepsilon^2} = \frac{2\sigma^4}{\varepsilon^2(n-1)} \quad \text{with } \frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{(n-1)}$$
$$\Rightarrow \lim_{n\to\infty} P(|S_n^2 - \sigma^2| \ge \varepsilon) = 0 \quad \therefore S_n^2 \xrightarrow{P} \sigma^2$$
$$\Rightarrow \hat{\theta} = \Phi\left(\frac{-(\bar{X}_n - \bar{Y}_n)}{\sqrt{S_{n1}^2 + S_{n2}^2}}\right) \text{ with } S_{n1}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})}{n-1},\ S_{n2}^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}$$

3) $\because \bar{X}_n \xrightarrow{P} \mu_X$, $\bar{Y}_n \xrightarrow{P} \mu_Y$, $S_{n1}^2 \xrightarrow{P} \sigma_X^2$, $S_{n2}^2 \xrightarrow{P} \sigma_Y^2$

4 dimensional continuous fuction. → let $h(x,y,z,w) = \Phi\left(\frac{-(x-y)}{\sqrt{z^2 + w^2}}\right)$ with Φ: c.d.f of $N(0,1)$ is continuous function.

by continuous mapping theorem

$$\Rightarrow h(\bar{X}_n, \bar{Y}_n, S_{n1}^2, S_{n2}^2) = \Phi\left(\frac{-(\bar{X}_n - \bar{Y}_n)}{\sqrt{S_{n1}^2 + S_{n2}^2}}\right) \xrightarrow{P} \Phi\left(\frac{-(\mu_X - \mu_Y)}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right)$$

not correct.

$$\therefore \hat{\theta} = \Phi\left(\frac{-(\bar{X}_n - \bar{Y}_n)}{\sqrt{S_{n1}^2 + S_{n2}^2}}\right) \text{ is a consistent estimator of } \theta$$

More details are necessary

◎ Use 1 dimensional continuous fuction rather than 4 dimensional fuction.

5. [+3] Let X_n and Y_n be sequences of random variables and X be a <u>continuous</u> random variable. Let $\varepsilon > 0$. Show

1) $P(X_n + Y_n \le x)$ is bounded below by $P(X_n \le x - \varepsilon) + P(|Y_n| > \varepsilon)$.

   $P(X_n + Y_n \le x)$ is bounded above by $P(X_n \le x + \varepsilon) + P(|Y_n| > \varepsilon)$.

2) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} 0$, then derive $\lim_{n\to\infty} P(X_n + Y_n \le x)$.

3) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} a$, then derive $\lim_{n\to\infty} P(X_n + Y_n \le x)$. Also, $X_n + Y_n \xrightarrow{d} ?$

+1/3.

Note:
$Y_n \xrightarrow{d} k$, k: constant

$\lim_{n\to\infty} P(|Y_n - k| \le t) = 0$

$= P(0 \le t) = 0$

$\Rightarrow Y_n \xrightarrow{d} k$

$\Leftrightarrow Y_n \xrightarrow{P} k$

1) $\forall \varepsilon$ $P(X_n \le x - \varepsilon) = P(X_n \le x - \varepsilon, |Y_n| \le \varepsilon) + P(X_n \le x - \varepsilon, |Y_n| > \varepsilon)$

Explain $\le P(X_n + Y_n \le x) + P(|Y_n| > \varepsilon)$

$\Rightarrow P(X_n + Y_n \le x) \ge P(X_n \le x - \varepsilon) - P(|Y_n| > \varepsilon)$ — ① (LB)

$P(X_n + Y_n \le x) = P(X_n + Y_n \le x, |Y_n| \le \varepsilon) + P(X_n + Y_n \le x, |Y_n| > \varepsilon)$

Explain

$\le P(X_n \le x + \varepsilon) + P(|Y_n| > \varepsilon)$ — ② (UB)

2) If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} 0$. by Note, ①, ②

1) is true for $\varepsilon > 0$. (not $\varepsilon \ne 0$).

$P(X_n \le x - \varepsilon) - P(|Y_n| > \varepsilon) \le P(X_n + Y_n \le x) \le P(X_n \le x + \varepsilon) + P(|Y_n| > \varepsilon)$

Let $\varepsilon = 0$ ✗

$P(X < x) \le \lim_{n\to\infty} P(X_n + Y_n \le x) \le P(X < x)$

and X: continuous $\Rightarrow P(X < x) = P(X \le x)$

$\therefore P(X \le x) \le \lim_{n\to\infty} P(X_n + Y_n \le x) \le P(X \le x)$ $\Rightarrow \lim_{n\to\infty} P(X_n + Y_n \le x) = P(X \le x)$

3) If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} a$. by Note,

✗

$P(X_n \le x - a - \varepsilon) = P(X_n \le x - a - \varepsilon, |Y_n - a| \le \varepsilon) + P(X_n \le x - a - \varepsilon, |Y_n - a| > \varepsilon)$

$\le P(X_n + Y_n \le x) + P(|Y_n - a| > \varepsilon)$

Similar

$P(X_n + Y_n \le x) = P(X_n + Y_n \le x, |Y_n - a| \le \varepsilon) + P(X_n + Y_n \le x, |Y_n - a| > \varepsilon)$

$\le P(X_n \le x - a + \varepsilon) + P(|Y_n - a| > \varepsilon)$

take $\varepsilon = 0$.

$\Rightarrow P(X + a \le x) \le \lim_{n\to\infty} P(X_n + Y_n \le x) \le P(X + a \le x)$. $\forall x$, X: continuous

$\Rightarrow X_n + Y_n \xrightarrow{d} X + a$.