Statistics

MATH 414

Lecture 14

Vivek T

A Digression

① We say that a sequence (X_n) of random variables with cdfs $F_1, F_2, \ldots, F_k, \ldots$ converges in distribution to X if

$$\boxed{\lim_{n \to \infty} F_n(x) \longrightarrow \lim F_X(x)}$$

for all x where F_X is continuous.

We write

$$\boxed{X_n \xrightarrow{D} X}$$

<u>Theorem</u> Let $X_1, X_2, \dots, X_n$ be a sequence of iid random variables from a distribution with mean μ, standard deviation σ, and a moment generating function defined on an interval containing 0. Then

Theorem Let $X_1, X_2, \ldots, X_n$ be a sequence of iid random variables from a distribution with mean μ, standard deviation σ, and a moment generating function defined on an interval containing 0. Then

$$\boxed{\lim_{n \to \infty} P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z\right) = \Phi(z)}$$

where $\overline{X}_n = \frac{\sum X_i}{n}$, we write

$$(\overline{X}_n - \mu)\Big/ \sigma/\sqrt{n} \xrightarrow{D} \text{Norm}(0, 1)$$

2 We say that a sequence of random variables X_n converges in probability to X if for all $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0$$

and we write

$$\boxed{X_n \xrightarrow{P} X}$$

Theorem (weak law of large numbers)

Let X_n be a sequence of iid random variables with common mean μ and variance $\sigma^2 < \infty$. Let $\overline{X}_n = \dfrac{\sum X_i}{n}$.

Then

$$\boxed{\overline{X}_n \xrightarrow{P} \mu}$$

Theorem (weak law of large numbers)

Let X_n be a sequence of iid random variables with common mean μ and variance $\sigma^2 < \infty$. Let $\overline{X}_n = \frac{\sum X_i}{n}$.

Then

$$\overline{X}_n \xrightarrow{P} \mu$$

(the proof use Chebychev's inequality)

Some algebra of limits

1. $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ $\Rightarrow$ $X_n + Y_n \xrightarrow{P} X + Y$

2. $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

3. $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ $\Rightarrow$ $X_n Y_n \xrightarrow{P} XY$

Some algebra of limits

① $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X+Y$

② $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

③ $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_nY_n \xrightarrow{P} XY$

Proof Let $\varepsilon > 0$ be given. Then

$$|X_n - X| + |Y_n - Y| \geq |X_n + Y_n - (X+Y)|$$

Some algebra of limits

① $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X+Y$

② $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

③ $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_nY_n \xrightarrow{P} XY$

Proof Let $\varepsilon > 0$ be given. Then

$$\underbrace{|X_n - X| + |Y_n - Y|}_{\text{Event 2}} \geq \underbrace{|X_n + Y_n - (X+Y)| > \varepsilon}_{\text{Event 1}}$$

$> \varepsilon$

Then, Event 1 $\subseteq$ Event 2

Some algebra of limits

① $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ $\Rightarrow$ $X_n + Y_n \xrightarrow{P} X+Y$

② $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

③ $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ $\Rightarrow$ $X_nY_n \xrightarrow{P} XY$

Proof Let $\varepsilon > 0$ be given. Then

$$\underbrace{|X_n - X| + |Y_n - Y| > \varepsilon}_{\text{Event 2}} \geq \underbrace{|X_n + Y_n - (X+Y)| > \varepsilon}_{\text{Event 1}}$$

Then, $P(\text{Event 1}) \leq P(\text{Event 2})$

Some algebra of limits

① $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X+Y$

② $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

③ $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n Y_n \xrightarrow{P} XY$

Proof Let $\varepsilon > 0$ be given. Then

$$|X_n - X| + |Y_n - Y| \geq |X_n + Y_n - (X+Y)| > \varepsilon$$

$$P(|X_n + Y_n - (X-Y)| > \varepsilon) \leq P(|X_n - X| + |Y_n - Y| > \varepsilon)$$

Some algebra of limits

① $X_n \xrightarrow{P} X, \; Y_n \xrightarrow{P} Y \implies X_n + Y_n \xrightarrow{P} X+Y$

② $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

③ $X_n \xrightarrow{P} X, \; Y_n \xrightarrow{P} Y \implies X_nY_n \xrightarrow{P} XY$

Proof Let $\varepsilon > 0$ be given. Then

$$|X_n - X| + |Y_n - Y| \geq |X_n + Y_n - (X+Y)| > \varepsilon$$

$$P(|X_n + Y_n - (X-Y)| > \varepsilon) \leq P(|X_n - X| + |Y_n - Y| > \varepsilon)$$

$$\leq P(|X_n - X| > \varepsilon/2) + P(|Y_n - Y| > \varepsilon/2)$$

Some algebra of limits

① $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ $\Rightarrow$ $X_n + Y_n \xrightarrow{P} X+Y$

② $X_n \xrightarrow{P} X$ and a is a constant then

$$aX_n \xrightarrow{P} aX$$

③ $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ $\Rightarrow$ $X_nY_n \xrightarrow{P} XY$

Excercise!

Proof Let $\varepsilon > 0$ be given. Then

$$|X_n - X| + |Y_n - Y| \geq |X_n + Y_n - (X+Y)| > \varepsilon$$

$$P(|X_n+Y_n-(X-Y)| > \varepsilon) \leq P(|X_n - X| + |Y_n - Y| > \varepsilon)$$

$$\leq P(|X_n - X| > \varepsilon/2) + P(|Y_n - Y| > \varepsilon/2)$$

Theorem. Let X_n, X, A_n be random variables, and let "a" be a constant

Then

if $X_n \xrightarrow{D} X$

$A_n \xrightarrow{P} a$

then $\dfrac{X_n}{A_n} \xrightarrow{D} \dfrac{X}{a}$ if $a \neq 0$.

Theorem Under suitable conditions on $f(x;\theta)$, any consistent sequence of solutions of mle equations satisfies

$$\boxed{\sqrt{n}\,(\hat{\theta}-\theta_0) \xrightarrow{D} N\left(0, \frac{1}{I(\theta_0)}\right)}$$

Sketch of Proof

Recall that

$$\ell(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta)$$

and

$$\ell'(\theta) = \sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta}.$$

Sketch of Proof

Recall that

$$l(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta)$$

and

$$l'(\theta) = \sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta}.$$

Expanding $l'(\theta)$ into a Taylor series of order 2 about θ_0 and evaluating at $\hat{\theta}_n$, we have

$$l'(\hat{\theta}_n) = l'(\theta_0) + (\hat{\theta}_n - \theta_0)\, l''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2\, l'''(\theta_n^*)$$

where $\theta_n^* \in (\theta_0, \hat{\theta}_n)$.

Rearranging

$$\ell'(\hat{\theta}_n) = \ell'(\theta_0) + (\hat{\theta}_n - \theta_0)\,\ell''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2\,\ell'''(\theta_n^*),$$

we get

Sketch of Proof

Recall that

$$l(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta)$$

and $$l'(\theta) = \sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta}.$$

Expanding $l'(\theta)$ into a Taylor series of order 2 about θ_0 and evaluating at $\hat{\theta}_n$, we have

$$l'(\hat{\theta}_n) = l'(\theta_0) + (\hat{\theta}_n - \theta_0)\, l''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2\, l'''(\theta_n^*)$$

where $\theta_n^* \in (\theta_0, \hat{\theta}_n)$.

This is zero since $\hat{\theta}_n$ solves the mle eqn.

Rearranging

$$l'(\hat{\theta}_n) = l'(\theta_0) + (\hat{\theta}_n - \theta_0)\, l''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2\, l'''(\theta_n^*),$$

we get

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) = \frac{n^{-1/2}\, l'(\theta_0)}{-n^{-1}\, l''(\theta_0) - (2n)^{-1}(\hat{\theta}_n - \theta_0)\, l'''(\theta_0)}$$

Rearranging

$$l'(\hat{\theta}_n) = l'(\theta_0) + (\hat{\theta}_n - \theta_0)\, l''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2\, l'''(\theta_n^*),$$

we get

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) = \frac{n^{-1/2}\, l'(\theta_0)}{-n^{-1}\, l''(\theta_0) - (2n)^{-1}(\hat{\theta}_n - \theta_0)\, l'''(\theta_0)}$$

By Central Limit Theorem,

$$\frac{1}{\sqrt{n}}\, l'(\theta_0) = \frac{\sqrt{n}}{n} \sum_{i=1}^{n} \frac{\partial \log f(x_i, \theta)}{\partial \theta}$$

$$\xrightarrow{D} N(0, I(\theta_0))$$

Rearranging

$$l'(\hat{\theta}_n) = l'(\theta_0) + (\hat{\theta}_n - \theta_0)\, l''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2 l'''(\theta_n^*),$$

we get

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) = \frac{n^{-1/2}\, l'(\theta_0)}{-n^{-1}\, l''(\theta_0) - (2n)^{-1}(\hat{\theta}_n - \theta_0)\, l'''(\theta_0)}$$

By Central limit Theorem,

$$\frac{1}{\sqrt{n}}\, l'(\theta_0) = \frac{\sqrt{n}}{n} \sum_{i=1}^{n} \frac{\partial \log f(x_i, \theta)}{\partial \theta}$$

$$\xrightarrow{D} N(0, I(\theta_0))$$

var of $\frac{\partial \log f}{\partial \theta}$

Also, by the law of large numbers,

$$-\frac{1}{n} \ell''(\theta_0) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f(x_i;\theta_0)}{\partial \theta^2}$$

$$\xrightarrow{P} I(\theta_0)$$

Also, by the law of large numbers,

$$-\frac{1}{n} l''(\theta_0) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f(x_i;\theta_0)}{\partial \theta^2}$$

$$\xrightarrow{P} I(\theta_0)$$

Finally for the third expression

$\hat{\theta}_n - \theta_0 \xrightarrow{P} 0$, and the expression

$$\frac{l'''(\theta_0)}{n} \le \frac{\sum M(x_i)}{n} \longrightarrow E_{\theta_0}(M(x))$$

(Therefore the expression $\frac{l'''(\theta_0)}{n} (\hat{\theta}_n - \theta_0) \xrightarrow{P} 0.$

as $n \to \infty$

Therefore,

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{D} \frac{N\left(0, I(\theta)\right)}{I(\theta)}$$

$$= N\left(0, \frac{1}{I(\theta)}\right)$$

$\blacksquare$