Lecture 16

2024 Spring

July 9, 2024

Last Few Lectures

- Convergence in distribution & Convergence in probability

$\lim_{n\to\infty} F_n(x) = F(x)$. $\lim_{n\to\infty} Pr(|X_n - X| \geq \varepsilon) = 0$

for all x where $F(x)$ is continuous. or

$\lim_{n\to\infty} Pr(|X_n - X| < \varepsilon) = 1$

This Lecture

- Markov Inequality
- Weak Law of Large Number
- Central Limit Theorem

Markov Inequality

1. Suppose. $X \geq 0$, $Pr(X \geq a) \leq \frac{E(X)}{a}$, $a > 0$

proof: $E(X) = \int_0^\infty x f(x) dx \geq \int_a^\infty x f(x) dx \geq \int_a^\infty a f(x) dx$

$= a \int_a^\infty f(x) dx = a Pr(X \geq a)$

$\Rightarrow Pr(X \geq a) \leq \frac{E(X)}{a}$

2. Suppose $X \geq 0$, $Pr(X \geq a) \leq \frac{E(X^k)}{a^k}$, $k = 1, 2, \cdots$, $a > 0$

proof: $E(X^k) = \int_0^\infty x^k f(x) dx \geq \int_a^\infty x^k f(x) dx \geq \int_a^\infty a^k f(x) dx$

$= a^k \int_a^\infty f(x) dx = a^k Pr(X \geq a)$

$\Rightarrow Pr(X \geq a) \leq \frac{E(X^k)}{a^k}$.

3. For $X \in \mathbb{R}$. , $Pr(|X| \geq a) \leq \frac{E(|X|^k)}{a^k}$, $a > 0$. $k = 1, 2, \cdots$

useful in proving convergence in prob.

as long as you can show the bound will go to zero.

$\Rightarrow \lim_{n \to \infty} Pr(|\cdot| \geq \varepsilon) = 0$

4. For $X \in \mathbb{R}$. $E(X) = \mu$. take $k = 2$

$Pr(|X - \mu| \geq a) \leq \frac{E\{(X - \mu)^2\}}{a^2} = \frac{var(X)}{a^2}$

bounded by its variance.

Weak Law of Large Numbers

1. Suppose $X_1 \cdots X_n$ independent. $E(X_i)=\mu$, $Var(X_i)=\sigma^2$
   Then. $\bar{X}_n = \frac{X_1+\cdots+X_n}{n} \xrightarrow{P} \mu$
   proof: using the Markov Inequality (version 4)
   $$Pr(|\bar{X}_n-\mu| \geq \varepsilon) \leq \frac{Var(\bar{X}_n-\mu)}{\varepsilon^2} = \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{Var\{\frac{1}{n}(X_1+\cdots+X_n)\}}{\varepsilon^2}$$
   $$= \frac{1}{n^2} \cdot \frac{n\sigma^2}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0, \text{ as } n\to\infty$$
   $\Rightarrow \bar{X}_n \xrightarrow{P} \mu$.

2. Suppose. $X_1 \cdots X_n$ independent. $E(X_i)=\mu$ $Var(X_i)=\sigma_i^2 \leq c$, for $i=1,2,\cdots$
   $$Pr(|\bar{X}_n-\mu| \geq \varepsilon) = \frac{\frac{1}{n^2} \cdot Var(X_1+\cdots+X_n)}{\varepsilon^2} \leq \frac{1}{n^2} \cdot \frac{nc}{\varepsilon^2} = \frac{c}{n\varepsilon^2} \to 0, \text{ as } n\to\infty$$
   $\Rightarrow \bar{X}_n \xrightarrow{P} \mu$.

Examples:

1. ① $X_1 \cdots X_n \overset{iid}{\sim} \chi_1^2$, then $\bar{X}_n \xrightarrow{P} 1$
   mean of χ_1^2
   ② $Y_n \sim \chi_n^2$, then $\frac{Y_n}{n} \xrightarrow{P} 1$
   $Y_n \overset{d}{=} X_1+\cdots+X_n$, $X_i \overset{iid}{\sim} \chi_1^2$
   $\frac{Y_n}{n} \overset{d}{=} \bar{X}_n$ $\quad \bar{X}_n \xrightarrow{P} 1 \Rightarrow \bar{X}_n \xrightarrow{d} 1 \Rightarrow \frac{Y_n}{n} \xrightarrow{d} 1 \Rightarrow \frac{Y_n}{n} \xrightarrow{P} 1$.
   ③ $X_1 \cdots X_n \overset{iid}{\sim} Poi(\mu)$. $\bar{X}_n \xrightarrow{P} \mu$
   $$\frac{X_1^2+\cdots+X_n^2}{n} \xrightarrow{P} E(X^2) = Var(X)+\{E(X)\}^2 = \mu+\mu^2$$

2. $X_1 \cdots X_n$ iid with cdf. $F(x) = \begin{cases} 1 - e^{-\mu x^2}, & x \geq 0, \ \mu > 0 \\ 0, & x < 0 \end{cases}$

Find the value of a s.t. $\lim_{n\to\infty} \Pr\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i^2 - a\right| \geq \varepsilon\right\} = 0$. for any $\varepsilon > 0$.

$\frac{\sum_{i=1}^n X_i^2}{n} \xrightarrow{P} a = E(X^2)$. (fourth moment exists. $f(x) = 2\mu x \exp(-\mu x^2)$, $x \geq 0$.

$E(X^2) = \int_0^\infty 2\mu x^3 \exp(-\mu x^2)\,dx = \frac{1}{\mu}$

using the "change of variable" technique. to get rid of integration with Gamma function.

3.

Suppose. $X_1 \cdots X_n \sim \text{Cauchy}(0,1)$ $f(x) = \frac{1}{\pi(1+x^2)}$

will $\bar{X}_n \xrightarrow{P}$ constant ?

→ Var(X) does not exist. cannot use the WLLN.

→ In fact. $\bar{X}_n \stackrel{d}{=} \text{Cauchy}(0,1)$

– variance needs to exist to use the WLLN.

4. (Chernoff Bounds) $X \geq 0$, $a > 0$

$\Pr(X \geq a) = \Pr(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}} = \frac{M_X(t)}{e^{ta}}$, for $t > 0$

$\Pr(X \geq a) \leq \inf_{t>0} \frac{M_X(t)}{e^{ta}}$.

$\Pr(X \leq a) = \Pr(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}} = \frac{M_X(t)}{e^{ta}}$, $t < 0$

$\Pr(X \leq a) \leq \inf_{t<0} \frac{M_X(t)}{e^{ta}}$

For a sequence of iid. Bernoulli trials with $\Pr(X_i = 1) = p$, letting $Y = \sum_{i=1}^n X_i$. Letting $\mu = np$, prove the following bound for all $\delta > 0$.

$$\Pr\{Y \geq (1+\delta)\mu\} \leq \left\{\frac{e^\delta}{(1+\delta)^{1+\delta}}\right\}^\mu$$

proof: $\Pr(Y \geq (1+\delta)\mu) \leq \frac{M_Y(t)}{e^{t(1+\delta)\mu}}$ (By the Chernoff bound)

For $M_Y(t) = \{M_X(t)\}^n = \{pe^t + (1-p)\}^n = \{1 + p(e^t - 1)\}^n$ $\boxed{e^x \geq 1+x}$

$\leq \{e^{p(e^t-1)}\}^n = e^{np(e^t-1)} = e^{\mu(e^t-1)}$

$\Rightarrow \Pr(Y \geq (1+\delta)\mu) \leq \frac{e^{\mu(e^t-1)}}{e^{t(1+\delta)\mu}}$, $t > 0$.

Find out the minimal value of $\frac{e^{\mu(e^t-1)}}{e^{t(1+\delta)\mu}}$, with respect to t.

take the first derivative, and let it be 0. $t = \log(1+\delta)$

The minimal is $\left\{\frac{e^\delta}{(1+\delta)^{1+\delta}}\right\}^\mu$.

Central Limit Theorem Let $X_1, \ldots, X_n$ be a sequence of iid random variables having mean 0 and variance σ^2. Let $Y_n = \sum_{i=1}^{n} X_i$, then

$$\lim_{n\to\infty} \Pr\left(\frac{Y_n}{\sigma\sqrt{n}} \le x\right) = \Phi(x), \quad -\infty < x < \infty,$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. Norm(0,1)

1. notation $\frac{Y_n}{\sigma\sqrt{n}} \xrightarrow{d} \text{Norm}(0,1)$

   $Y_n = n\bar{X}_n$, $\frac{Y_n}{\sigma\sqrt{n}} = \frac{n\bar{X}_n}{\sigma\sqrt{n}} = \frac{\sqrt{n}\bar{X}_n}{\sigma} \xrightarrow{d} \text{Norm}(0,1)$

2. If $E(X_1) = \mu \neq 0$. $E(X_1 - \mu) = 0$

   $\Rightarrow \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} \text{Norm}(0,1)$

Comparison with WLLN. $1 \cdot \bar{X}_n \xrightarrow{P} \mu$, $n \to \infty$

$\sqrt{n}\bar{X}_n \xrightarrow{d} \text{Norm}(\cdot,\cdot)$, $n\to\infty$

$n\bar{X}_n \to \infty$, explode.

Proof: $\begin{cases} \text{mgf \& cdf one-to-one.} \\ \text{lim mgf \& lim cdf one-to-one.} \end{cases}$

Strategy: Show that the mgf will converge to the mgf of Norm(0,1) proof: $Z_n = \frac{Y_n}{\sigma\sqrt{n}} \xrightarrow{d} Norm(0,1)$, $Y_n = \sum_{i=1}^{n} X_i$. $\begin{cases} E(X_i) = 0 \\ Var(X_i) = \sigma^2 \end{cases}$

$E(X^2) = Var(X) = \sigma^2$

$M_X(t)$ is mgf of $X_1 \cdots X_n$.

$M_{Y_n}(t) = \{M_X(t)\}^n$

Because $Z_n = \frac{1}{\sigma\sqrt{n}} \cdot Y_n$, $M_{Z_n}(t) = \{M_X(\frac{t}{\sigma\sqrt{n}})\}^n$

$$M_X\left(\frac{t}{\sigma\sqrt{n}}\right) = M_X(0) + \left(\frac{t}{\sigma\sqrt{n}}\right) M'_X(0) + \frac{1}{2} \cdot \left(\frac{t}{\sigma\sqrt{n}}\right)^2 M''_X(0) + \varepsilon_n.$$

expand at 0

$$= 1 + 0 + \frac{1}{2} \cdot \frac{t^2}{n \cdot \cancel{\sigma^2}} \cdot \cancel{\sigma^2} + \varepsilon_n$$

$$= 1 + \frac{t^2/2 + n\varepsilon_n}{n}$$

$\boxed{n\varepsilon_n \to 0 \text{ as } n \to \infty}$

$$M_{Z_n}(t) = \left(1 + \frac{t^2/2 + n\varepsilon_n}{n}\right)^n$$

$\boxed{(1 + \frac{a}{n})^n \to \exp(a),}$

$$\lim_{n\to\infty} M_{Z_n}(t) = \exp(t^2/2) \sim Norm(0,1)$$

limit of mgf is the mgf of Norm(0,1)

$\Rightarrow$ limit of cdf is the cdf of Norm(0,1)

$\Rightarrow$ Convergence in distribution.

$$\Rightarrow Z_n \xrightarrow{d} Norm(0,1)$$