Lecture 18

Recap

- **Modes of Convergence:**

⇒ **Convergence in Probability**

$X_n \xrightarrow{P} X$ if $\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1 \quad \forall \varepsilon > 0$

- Chebyshev's Inequality is useful for solving

⇒ **Convergence in distribution**

$X_n \xrightarrow{d} X$ if $\lim_{n\to\infty} F_n(t) = F(t) \quad \forall t$

or

$\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$

$\forall t \in (-\delta, \delta)$ for some $\delta > 0$

- WLLN

Let $X_1, \dots, X_n$ be iid RV with finite mean and variance

Let $\bar{X}_n = \frac{\sum X_i}{n}$

$$\Rightarrow \forall \varepsilon > 0 \quad \lim_{n \to \infty} P(|\bar{X}_n - E(X)| < \varepsilon) = 1$$

$$\bar{X}_n \xrightarrow{P} E(X)$$

- Central Limit Thm

Let $(X_n)_{n=1}^{\infty}$ be a sequence of iid RV with finite $E(X) = \mu$

$Var(X) = \sigma^2$

Then the RV $Z_n = \dfrac{\bar{X}_n - \mu}{(\sigma/\sqrt{n})} \xrightarrow{d} N(0,1)$

i.e.

$\forall \; -\infty \leq a \leq b \leq \infty$:

$$\lim_{n \to \infty} \left(a \leq \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq b \right) = \underbrace{\frac{1}{\sqrt{2\pi}} \int_a^b e^{-z^2/2} \, dz}_{P(a \leq z \leq b) \text{ for } z \sim N(0,1)}$$

• (Practical) CLT

Let $X_1, \dots, X_n$ be iid RV with finite mean and variance.

Let $\bar{X}_n = \frac{\sum X_i}{n}$, $\hat{\sigma}_n^2 = \frac{\sum (X_i - \bar{X}_n)^2}{n-1}$

Then $$Z_n = \frac{\bar{X}_n - \mu}{\hat{\sigma}_n / \sqrt{n}} \xrightarrow{d} N(0,1)$$

$$\hookrightarrow \frac{\bar{X}_n - \mu}{\hat{\sigma}_n / \sqrt{n}}$$

$$= \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \xrightarrow{d} U \sim N(0,1)$$

$$\Rightarrow \underbrace{\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} V \sim N(0, \sigma^2)}_{\text{Useful form}}$$

9.4

Confidence Intervals

Def:

Let X be a RV with unknown parameter θ.

Let $\hat{\theta}_n$ be an estimator for θ from n iid observations $X_1, \dots, X_n$ of X.
The α-confidence interval (or $1-\alpha$ conf. int) of $\hat{\theta}_n$ is defined by $\varepsilon > 0$ s.t.:

$$P(|\theta - \hat{\theta}_n| \leq \varepsilon) \geq 1-\alpha$$

$$\Rightarrow P(-\varepsilon \leq \theta - \hat{\theta}_n \leq \varepsilon) \geq 1-\alpha$$

Goal

- We want to use the CLT to create confidence intervals using $\bar{X}_n$

Confidence Intervals with Normal Approx

- Let $X_1, \dots, X_n$ be iid obs. from a distribution with finite mean + Variance (can apply CLT)

We want:

- $$P(-\varepsilon \le \overline{X}_n - \mu \le \varepsilon) = \underbrace{P\left(-\frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n} \le \frac{\sqrt{n}(\overline{X}_n - \mu)}{\hat{\sigma}_n} \le \frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n} \right)}$$

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\hat{\sigma}_n} \xrightarrow{d} N(0,1)$$

$$\approx \Phi\left(\frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n}\right) - \Phi\left(\frac{-\varepsilon\sqrt{n}}{\hat{\sigma}_n}\right)$$

$$= 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n}\right) - 1$$

If we want $2\Phi\left(\frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n}\right) - 1 \ge 1 - \alpha$

we must find:

$$2\Phi\left(\frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n}\right) - 1 = 1 - \alpha$$

$$\Rightarrow \Phi(z_\alpha) = 1 - \frac{\alpha}{2}$$

$$\Rightarrow z_\alpha = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

- A common choice for $\alpha = 5\%$

So $z_\alpha = \phi^{-1}\left(1 - \frac{.05}{2}\right) = \phi^{-1}(.975) = 1.96$

Where $z_\alpha = \frac{\varepsilon\sqrt{n}}{\hat{\sigma}_n}$

$$\Rightarrow \varepsilon = \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}$$

$\rightarrow$ use a normal table or calculator

So if $\varepsilon_\alpha = \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}$, $z_\alpha = \phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ then:

$$P(-\varepsilon_\alpha \leq \overline{X}_n - \mu \leq \varepsilon) \geq 1 - \alpha$$

$$\Rightarrow P\left(\frac{-z_\alpha \hat{\sigma}_n}{\sqrt{n}} \leq \overline{X}_n - \mu \leq \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}\right)$$

$$= P\left(\frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}} \geq \mu - \overline{X}_n \geq -\frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}\right)$$

$$= P\left(\overline{X} + \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}} \geq \mu \geq \overline{X}_n - \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}\right)$$

$$\geq 1 - \alpha$$

So $P\left(\mu \in \left[\overline{X}_n - \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}, \overline{X}_n - \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}}\right]\right) \approx 1 - \alpha$

$\Rightarrow$ As $n \to \infty$ the bound tightens around $\overline{x}$

Ex)

Let $\overline{X}_n = \dfrac{X_1 + \dots + X_n}{n}$, $\hat{\sigma}_n^2 = \dfrac{\sum (X_i - \overline{X}_n)^2}{n-1}$

After $n = 10^6$ observations we have

$\overline{X}_n = 3$, $\hat{\sigma}_n^2 = \frac{1}{4}$

Find a $1-\alpha$ % CI for $\mu = E(X)$.

$$P\left(\frac{-z_\alpha \hat{\sigma}_n}{\sqrt{n}} \le \mu - \overline{X} \le \frac{z_\alpha \hat{\sigma}_n}{\sqrt{n}} \right)$$

$$\Rightarrow \varepsilon_\alpha = \frac{z_\alpha \cdot \frac{1}{2}}{\sqrt{10^6}} = \frac{z_\alpha}{2 \cdot 10^3} = \frac{z_\alpha}{2000}$$

So a $1-\alpha$ % CI for μ is $3 \pm \dfrac{z_\alpha}{2000}$