B52 Nov 26 Lec 2 Notes

Theorem: Central limit theorem (CLT)

The standardized average of independent RVs with finite mean & variance converge in distribution to (standard) Normal (0,1).

$$Z_n = \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{D} N(0,1)$$

Result can be used to find approximate probabilities of $\bar{X}_n$ for "large" n, based on Normal distribution.

$$\Leftrightarrow \quad \bar{X}_n \sim N(\mu, \sigma^2/n)$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

Types of Convergence

Consider sequence of continuous RVs $X_1, X_2, \dots$ and RV Y.

X_n converges in probability to Y, as $n \to \infty$, if

$$\lim_{n\to\infty} P(|X_n - Y| \geq \varepsilon) = 0, \quad \forall \varepsilon > 0 \qquad (\text{denoted } X_n \xrightarrow{P} Y)$$

X_n diverges in probability to Y, as $n \to \infty$, if

$$\lim_{n\to\infty} P(X_n \leq x) = P(Y \leq x), \quad \forall x \in \mathbb{R} \qquad (\text{denoted } X_n \xrightarrow{P} Y)$$

Theorem: Levy's Theorem

If MGF $m_{X_n}(t) \to m_Y(t)$ as $n \to \infty$, then $X_n \xrightarrow{D} Y$

Proof: CLT

WTS $m_{Z_n}(t) \to m_Z(t)$, where $Z \sim N(0,1)$

MGF of $N(0,1)$: $m_Z(t) = E[e^{tZ}] = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2 - 2tx)}\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2 - 2tx + t^2 - t^2)}\, dx$$

$$= e^{t^2/2} \cancel{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2}\, dx} \quad 1, \text{ since } N(t,1)$$

$$= e^{t^2/2}(1) = e^{t^2/2}, \quad \forall t \in \mathbb{R}$$

Proof (continued...): CLT

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n} \mu}{\sigma/\sqrt{n}}$$

$$= \sum_{i=1}^{n} \frac{\frac{1}{n}(x_i - \mu)}{\sigma/\sqrt{n}}$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \left(\frac{x_i - \mu}{\sigma}\right)$$

$$m_{Z_n}(t) = \left[m_{\left(\frac{x_i-\mu}{\sigma\sqrt{n}}\right)}(t) \right]^n = \left[\sum_{k=0}^{\infty} \frac{t^k}{k!} \mu_k \right]^n$$

$$= \Big[1 + \frac{t^1}{1!} \cdot \underbrace{\mu}_{E\left[\frac{x_i-\mu}{\sigma\sqrt{n}}\right] = 0} + \frac{t^2}{2!} \underbrace{\mu_2}_{E\left[\left(\frac{x_i-\mu}{\sigma\sqrt{n}}\right)^2\right] = V\left(\frac{x_i-\mu}{\sigma\sqrt{n}}\right) = \frac{1}{n}} + \underbrace{\frac{t^3}{3!}\mu_3 + \ldots}_{=R} \Big]^n$$

$$\Rightarrow m_{Z_n}(t) = \left(1 + \frac{t^2}{2} \cdot \frac{1}{n} + R\right)^n$$

$$\lim_{n\to\infty} m_{Z_n}(t) = \lim_{n\to\infty} \Big(1 + \frac{t^2/2}{n} + \underbrace{R}\Big)^n \qquad \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$$

All terms in R are divided by higher orders of n. Thus they approach 0 much faster.

$$= e^{t^2/2}$$

$$= m_Z(t)$$

Ex 1:

Assume you flip a coin 25 times. Find the approximate probability that the proportion of Heads is greater than 0.6

$$\text{Let } I_i = \begin{cases} 1, & \text{heads} \\ 0, & \text{tails} \end{cases}$$

$$\text{Define } \bar{X}_{25} = \frac{1}{25} \sum_{i=1}^{25} I_i$$

$$= \frac{(\#\text{ heads in 25 flips})}{25}$$

$$E[I_i] = P(I_i = 1) = \tfrac{1}{2}$$

$$V(I_i) = (p \cdot q) = \left(\tfrac{1}{2}\right)^2 = \tfrac{1}{4}$$

By CLT, $\bar{X}_{25} \sim N\left(\tfrac{1}{2}, \tfrac{1/4}{25} = \tfrac{1}{100}\right)$

$$P(\bar{X}_{25} > 0.6) \Rightarrow P\left(\frac{\bar{X}_{25} - \frac{1}{2}}{\sqrt{\frac{1}{100}}} > \frac{.6 - \frac{1}{2}}{\sqrt{\frac{1}{100}}}\right)$$

$$= P(z > 1)$$

$$= 1 - P(z \le 1) = .1586 \approx 16\%$$