Convergence

Stat 510

WHAT HAPPENS WITH MORE DATA?

MAIN IDEAS

- LAW OF LARGE NUMBERS

  HIGH PROBABILITY THAT $\overline{X}_n$ CLOSE TO μ

- CENTRAL LIMIT THEOREM

  $\overline{X}_n$ IS APPROX NORMAL

WHEN n "LARGE"

DEFN

GIVEN A SEQUENCE OF RVS $X_1, X_2, \ldots$

AND ANOTHER RV X

WE SAY X_n CONVERGES IN PROBABILITY TO X,

$$X_n \xrightarrow{P} X$$

IF FOR ANY $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \longrightarrow 0 \quad \text{AS} \quad n \longrightarrow \infty$$

DEFN

GIVEN A SEQUENCE OF RVS $X_1, X_2, \ldots$

AND ANOTHER RV X

LET F_n BE THE CDF OF X_n AND F BE THE CDF OF X

WE SAY X_n CONVERGES IN DISTRIBUTION TO X,

$$X_n \xrightarrow{d} X$$

IF

$$\lim_{n \to \infty} F_n(t) = F(t)$$

FOR ALL t WHERE F IS CONTINUOUS

Point Mass Distribution

$$P(X = c) = 1$$

$$X_n \xrightarrow{P} X \qquad X_n \xrightarrow{P} c$$

$$X_n \xrightarrow{d} X \qquad X_n \xrightarrow{d} c$$

Example

$X_n \sim N(0, 1/n)$ $\qquad X_n \xrightarrow{?} 0$

in dist? Let F be CDF for point mass @ 0.

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \qquad Z = \sqrt{n}\, X_n \sim N(0,1)$$

$t < 0$ $\quad F_n(t) = P[X_n \leq t] = P[\sqrt{n}\, X_n \leq \sqrt{n}\, t] = P[Z \leq \sqrt{n}\, t] \to 0$

$t > 0$ $\quad F_n(t) = P[Z \leq \sqrt{n}\, t] \to 1$

$F_n(t) \to F(t)$ $\quad \underline{t \neq 0}$

$t = 0$ $\quad F_n(0) = 1/2 \neq F(0) = 1$

$X_n \xrightarrow{d} 0$

not needed.

EXAMPLE

$X_n \sim N(0, 1/n)$ $\quad X_n \xrightarrow{?} 0$

IN PROB? LET $\varepsilon > 0$

$$P(|X_n| > \varepsilon) = P(|X_n|^2 > \varepsilon^2)$$

$$\le \frac{\mathbb{E}[X_n^2]}{\varepsilon^2} = \frac{\mathbb{V}[X_n]}{\varepsilon^2} = \frac{1/n}{\varepsilon^2} \longrightarrow 0$$

$X_n - 0$

$$X_n \xrightarrow{P} 0$$

$$X_n \xrightarrow{d} X \quad \not\Rightarrow \quad X_n \xrightarrow{P} X$$

$$E[X_n] = -1 \cdot 0 = 0$$

$$V[X_n] = (-1)^2 \cdot 1 = 1 \quad \checkmark$$

$$X \sim N(0,1) \qquad X_n = -X \quad \text{FOR } n = 1, 2, 3, \dots$$

$$N(0,1)$$

SAME DISTRIBUTION

$$\lim_{n \to \infty} F_n(x) = F(x) \qquad X_n \xrightarrow{d} X$$

$$P(|X_n - X| > \varepsilon) = P(|2X| > \varepsilon) = P[|X| > \varepsilon/2] \neq 0$$

$$\cancel{\xrightarrow{P}}$$

DEFN

X_n CONVERGES TO X IN QUADRATIC MEAN

$$X_n \xrightarrow{qm} X$$

IF $\mathbb{E}\left[(X_n - X)^2\right] \longrightarrow 0$ AS $n \longrightarrow \infty$

Thm

- $X_n \xrightarrow{qm} X \implies X_n \xrightarrow{P} X$
- $X_n \xrightarrow{P} X \implies X_n \xrightarrow{D} X$
- IF $P[X = c] = 1$ AND $X_n \xrightarrow{D} X$

  THEN $X_n \xrightarrow{P} X$

REVERSES NOT TRUE IN GENERAL

THM X_n, X, Y_n, Y RVs, g continuous function

- $X_n \xrightarrow{p} X,\ Y_n \xrightarrow{p} Y \implies X_n + Y_n \xrightarrow{p} X + Y$
- $X_n \xrightarrow{d} X,\ Y_n \xrightarrow{d} c \implies \underline{X_n + Y_n} \xrightarrow{d} \underline{X_n + c}$

SLUTSKY'S THM

- $X_n \xrightarrow{p} X,\ Y_n \xrightarrow{p} Y \implies X_n Y_n \xrightarrow{p} XY$
- $X_n \xrightarrow{d} X,\ Y_n \xrightarrow{d} c \implies \underline{cX_n} \xrightarrow{d} \underline{cX}$
- $X_n \xrightarrow{p} X \implies g(X_n) \xrightarrow{p} g(X)$

CONTINUOUS MAPPING THM

- $X_n \xrightarrow{d} X \implies g(X_n) \xrightarrow{d} g(X)$

THM WEAK LAW OF LARGE NUMBERS (WLLN)

GIVEN $X_1, X_2, \ldots X_n$ IID $E[X_i] = \mu$

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

PROOF ASSUME $\sigma^2 < \infty$.

$$P\left(|\overline{X}_n - \mu| > \varepsilon\right) \leq \frac{\mathbb{V}[\overline{X}_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0 \checkmark$$

Thm CENTRAL LIMIT THEOREM (CLT)

"CLASSICAL"

$X_1, X_2, \ldots X_n$ IID $\mathbb{E}[X_i] = \mu$, $\mathbb{V}[X_i] = \sigma^2 < \infty$

DEFINE $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

THEN

$$Z_n = \frac{(\bar{X}_n - \mu)}{\sqrt{\mathbb{V}[\bar{X}_n]}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z$$

$Z \sim N(0,1)$

APPROXIMATING PROBABILITIES

STATEMENTS LIKE $P(\overline{X}_n > c) \approx$ NORMAL

→ "FOR ANY DISTRIBUTION OF X_i"

→ "WHEN n LARGE"

Alternative Forms

$$Z_n \approx N(0,1)$$

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

$$\bar{X}_n - \mu \approx N(0, \sigma^2/n)$$

$$\sqrt{n}(\bar{X}_n - \mu) \approx N(0, \sigma^2)$$

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \approx N(0,1)$$

$$\bar{X}_n \xrightarrow{d} N(\mu, \sigma^2/n)$$

because of this

can't say

CLT WITH S_n

$$\frac{\sqrt{n}\,(\bar{X}_n - \mu)}{S_n} \xrightarrow{d} N(0,1)$$

PROOF BY EXERCISE!

THM BERRY - ESSÈEN INEQUALITY

GIVEN $\mathbb{E}\left[|X_1|^3\right] < \infty$ THEN

$$\sup_z \left| P(Z_n \leq z) - \Phi(z) \right| \leq \frac{33}{4} \frac{\mathbb{E}\left[|X_1 - \mu|^3\right]}{\sqrt{n}\,\sigma^3}$$

CDF OF $N(0,1)$

MAYBE TRY SIMULATIONS....

THM DELTA METHOD

GIVEN $\dfrac{\sqrt{n}\,(Y_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$

AND g IS DIFFERENTIABLE WITH $g'(\mu) \neq 0$

THEN $\dfrac{\sqrt{n}\,\big(g(Y_n) - g(\mu)\big)}{|g'(\mu)|\,\sigma} \xrightarrow{d} N(0,1)$

IF $Y_n \simeq N(\mu, \sigma^2/n)$

THEN $g(Y_n) \simeq N\left(g(\mu), (g'(\mu))^2 \sigma^2/n\right)$