Feynman Liang
STAT 210A
Lecture #19
10/27/2016

1 Asymptotics

So far: *finite-sample* results.
Pros: Exactly correct/optimal
Cons: Computations can be

- Complicated
- Intractable
- Specialized
- Reliant on particular assumptoins

Example 1.1 (Normal approximation to binomial). $X \sim \text{Binom}(n, \theta)$, $n = 2000$
Computing CI for θ

$$\begin{aligned} X &\approx N(n\theta, n\theta(1-\theta)) && (1.1) \\ &\approx N(n\theta, n\frac{x}{n}(1-\frac{x}{n}) && (1.2) \end{aligned}$$

$$\begin{aligned} &\frac{X}{n} \pm z_{\alpha/2}\sqrt{\frac{\theta(1-\theta)}{n}} && (1.3) \\ &\frac{X}{n} \pm z_{\alpha/2}\sqrt{\frac{\theta(1-\theta)}{n}} && (1.4) \\ & && (1.5) \end{aligned}$$

Warning: Can go wrong if $\theta \approx 0$ or $\theta \approx 1$.

Example 1.2. $X_i \overset{\text{iid}}{\sim} p_\theta(X) \quad i = 1, \dots, n$

2 Convergence in probability

Definition 2.1. A sequence $X_1, X_2, \dots$ of r.v.s *converges in probability* to a r.v. X if

$$\forall \varepsilon > 0 : \mathbb{P}(|X_n - X| > \varepsilon) \to 0 \tag{2.1}$$

Usually, X is constant c. Write as $X_n \xrightarrow{p} X$.

Proposition 2.2 (Chebyshev). *For any r.v. X, const $a > 0$,*

$$\mathbb{P}(|X| > a) \leq \frac{\mathbb{E}X^2}{a^2} \tag{2.2}$$

Proof.

$$1_{|X|>a} \leq \frac{X^2}{a^2} \tag{2.3}$$

$$\mathbb{P}(|X| > a) \leq \frac{\mathbb{E}X^2}{a^2} \tag{2.4}$$

□

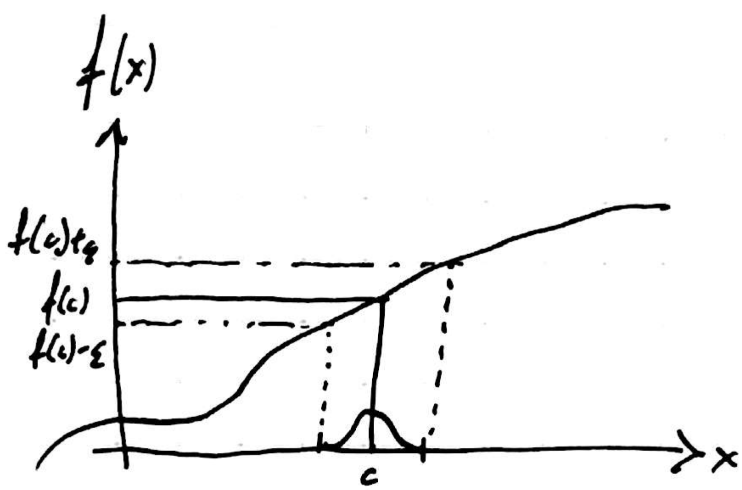
Corollary 2.3. $\mathbb{P}(|X - \mathbb{E}X| > a) \leq \frac{Var(X)}{a^2}$

Example 2.4. $X_1, X_2, \ldots \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$

Then $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\text{p}} \mu$.

$\mathbb{E}\bar{X}_n = \mu$, $\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n} \to 0$ (WLLN)

Proposition 2.5 (Continuous mapping theorem). *If f is continuous at c, and $X_n \xrightarrow{p} c$, then $f(X_n) \xrightarrow{p} f(c)$.*



Proof. Fix $\varepsilon > 0$. $\exists \delta_\varepsilon > 0$ with $|X - c| \leq \delta_\varepsilon \implies |f(x) - f(c)| \leq \varepsilon$.

$$\mathbb{P}(|f(X_n) - f(c)| > \varepsilon) \leq \mathbb{P}(|X_n - c > \delta_\varepsilon) \to 0 \tag{2.5}$$

□

Definition 2.6. A sequence $\delta_1, \delta_2, \ldots$ is *consistent for $g(\theta)$* if

$$\delta_n \xrightarrow{\text{p}} g(\theta), \quad \forall\theta \tag{2.6}$$

Note: $\text{MSE}(\theta, \delta_n) = \text{Bias}_\theta(\delta_n)^2 + \text{Var}_\theta(\delta_n)$ so if both Bias $\to 0$ and Var $\to 0$, then

$$\text{MSE} = \mathbb{E}[(\delta_n - g(\theta))^2] = \text{RHS of Chebyshev} \to 0 \tag{2.7}$$

so δ_n is consistent.

3 Convergence in Distribution

a.k.a. *weak convergence*.

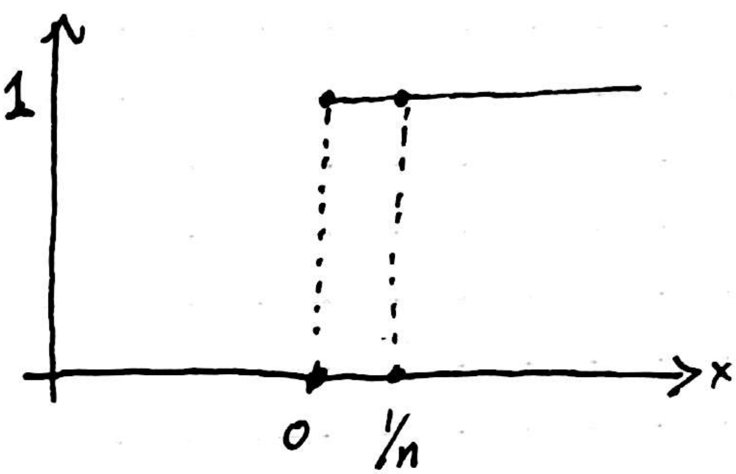
Definition 3.1. A sequence of r.v.s $X_1, X_2, \ldots$ converges in distribution to X if

$$F_{X_n}(X) \to F_X(x) \tag{3.1}$$

$\forall x$ such that $F_X(x)$ cts at $X = x$.

Usually written as $X_n \Rightarrow X$, $X_n \Rightarrow N(0,1)$, or $X_n \Rightarrow F$.

Example 3.2. $X_n = \frac{1}{n}$ w.p 1 $F_{X_n}(0) = 0$.
$X = 0$ w.p 1 $F_{X_n}(0) = 1$.



Theorem 3.3. $X_n \Rightarrow X$ *iff*

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \tag{3.2}$$

for any f bounded, continuous.

Corollary 3.4. *If g is continuous, and $X_n \Rightarrow X$, then $g(X_n) \Rightarrow g(X)$.*

Proof. f bounded and continuous
$\implies f \circ g$ bounded and continuous
$\implies \mathbb{E}[f(g(X_n))] \to \mathbb{E}[f(g(X))]$ □

Theorem 3.5 (Slutsky's Theorem)**.** *If $X_n \Rightarrow X$ and $Y_n \xrightarrow{p} c$, then:*

(a) $X_n + Y_n \Rightarrow X + c$

(b) $X_n \cdot Y_n \Rightarrow c \cdot X$

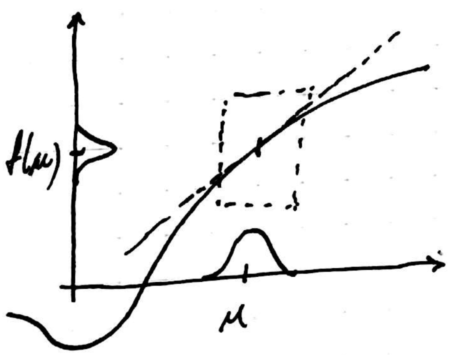
(c) $X_n / Y_n \Rightarrow X / c$ *if* $c \neq 0$

Proof. $(X_n, Y_n) \Rightarrow (X, c)$. All (a)-(c) are continuous functions (provided $c \neq 0$). Apply continuous mapping theroem. □

Theorem 3.6 (Central Limit Theorem (CLT)). *$X_i \overset{\text{iid}}{\sim} (\mu, \sigma^2)$, $i \geq 1$, $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.*
Then $\sqrt{n}(\bar{X}_n - \mu) \Rightarrow N(0, \sigma^2)$
($\bar{X}_n \approx N(\mu, \frac{\sigma^2}{n})$, need to subtract off mean to prevent exploding to ∞)

Example 3.7. $X_n \sim \text{Binom}(n, \theta)$. $X_n = n\bar{B}$ where $B_1, B_2, \ldots \overset{\text{iid}}{\sim} \text{Ber}(\theta)$.
Estimator $\hat{\theta} = \frac{X_n}{n} = \bar{B}_n$.
By WLLN, $\hat{\theta} \xrightarrow{\text{p}} \theta$.
By CLT, $\sqrt{n}(\hat{\theta} - \theta) \Rightarrow N(0, \theta(1-\theta))$.
$\frac{\sqrt{n}(\hat{\theta}-\theta)}{\sqrt{\hat{\theta}(1-\hat{\theta})}} \Rightarrow N(0,1)$

Theorem 3.8 (Delta Method). *If $\sqrt{n}(X_n - \mu) \Rightarrow N(0, \sigma^2)$ and $f(x)$ is diff. at μ, then $\sqrt{n}(f(X_n) - f(\mu)) \Rightarrow N(0, f'(u)^2\sigma^2)$.*



$$n(f(X_n) - f(X)) = \frac{1}{2} f''(\mu) X^2$$

Proof.

$$f(X_n) = f(\mu) + f'(\mu)(X_n - \mu) + o(X_n - \mu) \tag{3.3}$$
$$\sqrt{n}(f(X_n) - f(\mu)) = f'(\mu)\sqrt{n}(X_n - \mu) + \sqrt{n}o(X_n - \mu) \tag{3.4}$$

□