Lecture 13 -- CLT

Tuesday, October 22, 2024 11:01 AM

Theorem: Central Limit Theorem

If $X_n \overset{iid}{\sim} f$ where $\mu = EX_n$ and $\sigma^2 = \operatorname{Var} X_n < \infty$

then

$$\sqrt{N}\left(\frac{\bar{X}_N - \mu}{\sigma}\right) \xrightarrow{d} N(0,1).$$

Intuition: $\bar{X} \approx N(\mu, \sigma^2/N)$.

Might like to say $\boxed{\bar{X} \xrightarrow{d} N(\mu, \sigma^2/N)}$

can't have N in limit

proper ways to write CLT

(1) $\sqrt{N}\left(\frac{\bar{X} - \mu}{\sigma}\right) \xrightarrow{d} N(0,1)$

(2) $\sqrt{N}(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$

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

(3) $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{N}} \xrightarrow{d} N(0,1)$

defn

(4) $\bar{X} \sim AN(\mu, \sigma^2/N)$

asymptotically normal

__Ex.__ $X_n \overset{iid}{\sim} \text{Bern}(p)$, $0 \le p \le 1$

$$\mu = E[X_n] = p$$

$$\sigma^2 = \text{Var}\,X_n = p(1-p)$$

$$\sigma = \sqrt{p(1-p)}$$

CLT says something about

$\bar{X} = \hat{p}$ = pct. of 1s in sample

CLT: $\sqrt{N}(\hat{p}-p) \xrightarrow{d} N(0, p(1-p))$

CLT: $\sqrt{N}(\hat{p}-p) \to N(0, p(1-p))$

So for large N

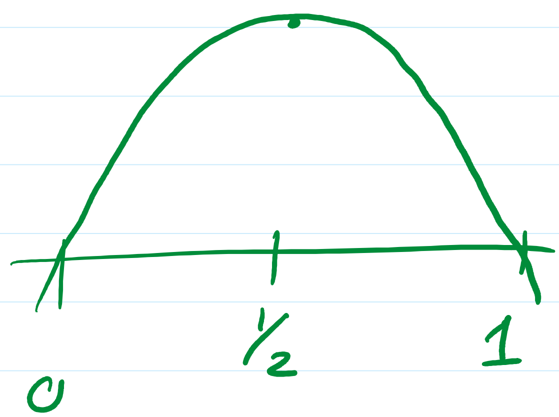
$$\hat{p} \sim AN\left(p, \frac{p(1-p)}{N}\right)$$

For a normal dist ~95% of vals fall w/in ± 2 s.d. of mean:

$$\hat{p} \pm 2\sqrt{\frac{p(1-p)}{N}}$$

replace p w/ $\hat{p}$

$$\hat{p} \pm \underbrace{2\sqrt{\frac{\hat{p}(1-\hat{p})}{N}}}_{\text{MOE}}$$



Ex. $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$

$$\mu = EX_n = \lambda$$

$$\mu = E X_n = \lambda$$

$$\sigma^2 = \operatorname{Var} X_n = \lambda \implies \sigma = \sqrt{\lambda}$$

CLT: $$\sqrt{N}\left(\frac{\bar{X} - \lambda}{\sqrt{\lambda}}\right) \xrightarrow{d} N(0,1)$$

i.e.

$$\bar{X} \approx N(\lambda, \lambda/N) \text{ for large } N.$$

<u>Theorem: MGFs and Convergence</u>

Let (X_n) be a seq of RVs w/ MGFs M_n.

Let X be a RV w/ MGF M.

If $M_n(t) \to M(t)$ as $N \to \infty$ $\forall t$ in some neighborhood of zero

then $X_n \xrightarrow{d} X$.

Taylor Series:

If g is k-times diff'able then the k^{th}-order Taylor poly about a is

$$T_k(x) = \sum_{r=0}^{k} \frac{g^{(r)}(a)}{r!}(x-a)^k$$

e.g. second order about zero

$$T_2(x) = g(0) + g'(0)\,x + \frac{g''(0)}{2}x^2$$

Under some conditions

$$g(x) \approx T_k(x) \text{ when } x \simeq a$$

---

pf. of CLT

$$\boxed{Y_N = \sqrt{N}\left(\frac{\bar{X}-\mu}{\sigma}\right)}$$ <u>want</u>: $Y_N \xrightarrow{d} N(0,1)$

$Z_n = \dfrac{X_n - \mu}{\sigma}$ then $EZ_n = 0$ and $\text{Var}Z_n = 1$

$z_n = \frac{X_n - \mu}{\sigma}$ then $E z_n = 0$ and $\text{var}\, z_n = 1$

$$Y_N = \sqrt{N}\left(\frac{\bar{X} - \mu}{\sigma}\right)$$

$$= \sqrt{N}\left(\frac{\frac{1}{N}\sum_n X_n - \frac{1}{N}\sum_n \mu}{\sigma}\right)$$

$$= \frac{\sqrt{N}}{N}\left(\frac{\sum_n X_n - \sum_n \mu}{\sigma}\right)$$

$$= \frac{\sqrt{N}}{N}\sum_n \left(\frac{X_n - \mu}{\sigma}\right)$$

show

$$Y_N = \frac{1}{\sqrt{N}}\sum_n z_n \qquad \left(\xrightarrow{d} N(0,1)\right)$$

Get MGF of Y_N:

all same

$$M_{Y_N}(t) = \prod_n M_{z_n}(t/\sqrt{N})$$

$$= M(t/\sqrt{N})^N$$

MGF of any z_n

Second order taylor approx. of $M(t)$ about $a=0$

$$M(t) \approx M(0) + M'(0)t + M''(0)\frac{t^2}{2}$$

$Ee^{0x} = E1 = 1$

$E[z_n] = 0$

$E[z_n^2] = Var(z_n) = 1$

$$= 1 + \frac{t^2}{2}$$

So

$$M_{Y_N}(t) = M(t/\sqrt{N})^N \approx \Big(1 + \underbrace{\frac{t^2}{2}}_{c}\frac{1}{N}\Big)^N$$

$$\boxed{\left(1+\frac{c}{n}\right)^n \to e^c \text{ as } n \to \infty}$$

as $N \to \infty$

$$M_{Y_N}(t) \to e^{t^2/2}$$

$M_{Y_N}(t) \to e^{t/2}$

as $n \to \infty$

MGF of $N(0,1)$

So $Y_N \xrightarrow{d} N(0,1)$.

Delta Method

Theorem: First Order Delta Method

If Y_N is a seq of RVs where

$$\sqrt{N}\left(Y_N - \theta\right) \xrightarrow{d} N(0, \psi^2(\theta))$$

Think $Y_N = \bar{X}$, $\theta = \mu$, $\psi^2(\theta) = \sigma^2$.

Then this is CLT

then if g is a diff'able function and

$g'(\theta) \neq 0$ then

$$\sqrt{N}\left(g(Y_N) - g(\theta)\right) \xrightarrow{d} N\left(0, [g'(\theta)]^2 \psi^2(\theta)\right)$$

Intuition:

$$Y_N \sim AN(\theta, \psi^2/N)$$

then $g(Y_N) \sim AN\left(g(\theta), [g'(\theta)]^2 \psi^2/N\right)$

Ex. $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$, $\lambda > 0$

CLT: $\sqrt{N}\left(\dfrac{\bar{X} - \lambda}{\sqrt{\lambda}}\right) \xrightarrow{d} N(0,1)$

alt. $\boxed{\sqrt{N}(\bar{X} - \lambda) \xrightarrow{d} N(0, \lambda)}$

Consider $g(x) = \log(x)$

$g'(x) = 1/x \longrightarrow [g'(x)]^2 = 1/x^2$

$g'(x) = 1/x \longrightarrow [g'(\lambda)]^2 = 1/\lambda^2$

So by FO Δ-method

$$\sqrt{N}\left(g(\bar{x}) - g(\lambda)\right) \xrightarrow{d} N\left(0, [g'(\lambda)]^2 \lambda\right)$$

i.e.

$$\sqrt{N}\left(\log(\bar{x}) - \log(\lambda)\right) \xrightarrow{d} N\left(0, \underbrace{1/\lambda^2 \, \lambda}_{1/\lambda}\right)$$

i.e.

$$\boxed{\begin{array}{l} \text{CLT: } \bar{X} \sim AN(\lambda, \lambda/N) \\ \Delta\text{-meth: } \log(\bar{x}) \sim AN(\log(\lambda), 1/N\lambda) \end{array}}$$

<u>Pf.</u> <u>of Δ-method</u>

Consider FO taylor approx of g about θ

$$g(x) \approx g(\theta) + g'(\theta)(x - \theta)$$

$\underline{\text{Know}}$: $\underbrace{\sqrt{N}(Y_N - \theta) \xrightarrow{d} N(0, \psi^2)}$

plug Y_N into taylor approx:

$$g(Y_N) \approx g(\theta) + g'(\theta)(Y_N - \theta)$$

$$\Rightarrow g(Y_N) - g(\theta) \approx g'(\theta)(Y_N - \theta)$$

$$\Rightarrow \sqrt{N}(g(Y_N) - g(\theta)) \approx g'(\theta)\underbrace{\sqrt{N}(Y_N - \theta)}_{\xrightarrow{d} N(0, \psi^2)}$$

$$\boxed{\sqrt{N}(g(Y_N) - g(\theta)) \xrightarrow{d} N(0, [g'(\theta)]^2 \psi^2)}$$

$\underline{\text{ex.}}$ $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, $\mu \neq 0$

What is asymptotic dist of $1/\bar{X}$?

CLT: $\sqrt{N}(\bar{X}-\mu) \xrightarrow{d} N(0, \sigma^2)$

Δ-method: $g(x) = 1/x$

$g'(x) = -1/x^2$, $g'(\mu) = -1/\mu^2 \neq 0$

So $\sqrt{N}(g(\bar{X}) - g(\mu)) \xrightarrow{d} N(0, [g'(\mu)]^2 \sigma^2)$

$g'(\mu)^2 = 1/\mu^4$

i.e.

$$\sqrt{N}\left(\frac{1}{\bar{X}} - \frac{1}{\mu}\right) \xrightarrow{d} N(0, \sigma^2/\mu^4)$$

or equiv.

$$\frac{1}{\bar{X}} \sim AN(1/\mu, \sigma^2/N\mu^4)$$

Theorem: Second-Order Δ-method

Assume $\sqrt{N}(Y_N - \theta) \xrightarrow{d} N(0, \psi^2)$

Assume $\sqrt{N}(Y_N - \theta) \xrightarrow{d} N(0, \psi^2)$

let g be twice diff'able and $g'(\theta) = 0$

then

$$N(g(Y_N) - g(\theta)) \xrightarrow{d} \frac{\psi^2 g''(\theta)}{2} \chi^2(1)$$

so long as $g''(\theta) \neq 0$.