36710 - 36752

ADVANCED PROBABILITY OVERVIEW

FALL 2020

LECTURE 20: MON, NOV 9, 2020

■ DELTA METHOD

<u>Theorem</u> LET $f: \mathbb{R}^d \to \mathbb{R}^k$ BE DIFFERENTIABLE AT θ WITH TOTAL DERIVATIVE f'_θ. IF $\{X_n\}$ IS A SEQUENCE OF RV'S TAKING VALUES OVER THE DOMAIN OF f S.T. $r_n(X_n - \theta) \xrightarrow{D} X$ AS $r_n \to \infty$

r_n ↳ SEQUENCE OF NUMBERS

THEN

i) $$r_n\left(f(X_n) - f(\theta)\right) \xrightarrow{D} f'_\theta(X)$$

AND

ii) $$r_n\left(f(X_n) - f(\theta)\right) - f'_\theta\left(r_n(X_n - \theta)\right) = o_p(1)$$

Pf/ FIRST NOTICE THAT $r_n(X_n - \theta)$ IS TIGHT $(O_p(1))$ SO, SINCE $r_n \to \infty$, $X_n - \theta = o_p(1)$. LET $R(h) = f(\theta + h) - f(\theta) - f'_\theta(h)$

THEN $R(h) = o(\|h\|)$ AS $h \to 0$. SO, USING A RESULT PROVED LAST TIME, $R(X_n - \theta) = o_p(\|X_n - \theta\|)$. BY TAYLOR SERIES EXPANSION,

$$f(X_n) - f(\theta) - f'_\theta(X_n - \theta) = R(X_n - \theta) = o_p(\|X_n - \theta\|)$$

MULTIPLY BOTH TERMS BY r_n AND NOTICE THAT

$$o_p(r_n \|X_n - \theta\|) = o_p(1) \qquad \left[\text{BECAUSE } o_p(O_p(1)) = o_p(1)\right]$$

THIS PROVES ii). CLAIM i) FOLLOWS FROM SLUTSKY THEOREM SINCE THE TOTAL DERIVATIVE IS A LINEAR MAPPING. ∎

<u>REMARK</u> IF $d = k = 1$ AND $\sqrt{n}\,(X_n - \theta) \xrightarrow{D} X$ THEN

$$\sqrt{n}\left(f(X_n) - f(\theta)\right) \xrightarrow{D} \sqrt{n}\; f'(\theta)\,(X_n - \theta)$$

↳ VALUE OF DERIVATIVE OF f AT θ

IF $k = 1$

$$\sqrt{n}\left(f(X_n) - f(\theta)\right) \xrightarrow{D} \sqrt{n}\; \nabla f(\theta)^T (X_n - \theta)$$

↳ GRADIENT OF f AT θ

FOR ARBITRARY d AND k

$$\sqrt{n}\left(f(X_n) - f(\theta)\right) \xrightarrow{D} \sqrt{n}\; \overbrace{\underbrace{f'_\theta}_{k \times d} \cdot (X_n - \theta)}^{\text{MATRIX PRODUCT}}$$

MATRIX OF PARTIAL DERIVATIVES OF f AT θ

THE MOST COMMON APPLICATION IS WHEN $\sqrt{n}\,(X_n - \theta) \xrightarrow{D} N(0, \sigma^2(\theta))$

THEN $$\sqrt{n}\left(f(X_n) - f(\theta)\right) \xrightarrow{D} N\left(0, \sigma^2(\theta)\,(f'(\theta))^2\right)$$

<u>PROBLEM</u>: LIMITING VARIANCE DEPENDS ON θ !

<u>EXAMPLE</u> $$(X_1, Y_1), \dots, (X_n, Y_n) \overset{iid}{\sim} \left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

WE ARE INTERESTED IN CONFIDENCE INTERVAL FOR ρ.

LET
$$\rho_n = \frac{\sum_i (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_i (X_i - \bar{X}_n)^2}\sqrt{\sum_i (Y_i - \bar{Y}_n)^2}}$$
EMPIRICAL CORRELATION

THEN
$$\sqrt{n}\,(\rho_n - \rho) \xrightarrow{D} N\left(0, (1 - \rho^2)^2\right)$$

<u>IDEA</u>: CHOOSE f S.T. $f(\theta) = \int \frac{1}{\sigma(\theta)}$

IN THIS CASE

$$f(e) = \int \frac{1}{1-e^2} = \frac{1}{2} \log \frac{1+e}{1-e} = \operatorname{arctanh} e$$

SO, BY THE DELTA METHOD

$$\sqrt{n}\,(\operatorname{arctanh} e_n - \operatorname{arctanh} e) \xrightarrow{D} N(0,1)$$

AND

$$\left[\tanh\left(\operatorname{arctanh}(e_n) - z_\alpha/\sqrt{n}\right), \tanh\left(\operatorname{arctanh}(e) + z_\alpha/\sqrt{n}\right) \right]$$

z_α UPPER α QUANTILE OF $N(0,1)$ IS $1-\alpha$ ASYMPTOTIC CI FOR e.

<u>VARIANCE STABILIZING TRANSFORMATION</u>

- SECOND ORDER DELTA METHOD

ASSUME $f: \mathbb{R}^d \to \mathbb{R}$. WHAT IF $\nabla f(\theta) = 0$?

<u>EXAMPLE:</u> $X_1, \dots, X_n \overset{iid}{\sim} (\theta, \Sigma)$ AND $\theta \mapsto f(\theta) = \frac{\|\theta\|^2}{2}$

$\nabla f(\theta) = \theta$

SO $\nabla f(0) = 0$

BY CLT $\leftarrow$ SINCE $\sqrt{n}\,(\bar{X}_n - \theta) \xrightarrow{D} N(0, \Sigma)$

THEN

$$\sqrt{n}\,\left(f(\bar{X}_n) - f(\theta)\right) \xrightarrow{D} N\left(0, \nabla f(\theta)^T \Sigma \nabla f(\theta)\right)$$

$$\overset{d}{=} 0 \qquad \text{FOR } \theta = 0$$

TAKE A 2^{nd} ORDER TAYLOR SERIES EXPANSION.

IF $r_n (X_n - \theta) \xrightarrow{D} X$ AND $\nabla f(\theta) = 0$ THEN

$$r_n^2 \left(f(X_n) - f(\theta)\right) \xrightarrow{D} \frac{1}{2} X^T H_\theta X$$

HESSIAN OF f AT θ

$d \times d$ WITH i,j ENTRY $\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_\theta$

THEN IN THE ABOVE EXAMPLE

$$n\left(\frac{\|\bar{X}_n\|^2}{2} - \frac{\|A\|^2}{2}\right) \xrightarrow{D} \frac{1}{2} X^T X$$

WHERE $X \sim N(0, \Sigma)$

EXAMPLE ASSUME $\sqrt{n}\,\bar{X}_n \xrightarrow{D} N(0,1)$

CONSIDER THE FUNCTION $x \mapsto \cos(x)$ [$=1$ AT $x=0$]

THEN

$$\sqrt{n}\left[\cos(\bar{X}_n) - \overset{1}{\overset{\shortparallel}{\cos(0)}}\right] \xrightarrow{P} 0$$

SO $\sqrt{n}$ IS NOT THE RIGHT SCALING! BUT

$$\cos(\bar{X}_n) - \cos(0) = (\bar{X}_n - 0)\cdot 0 + \frac{1}{2}(\bar{X}_n - 0)^2 \left.\frac{\partial^2 \cos}{\partial^2 x}\right|_{x=0} + \dots$$

MULTIPLY BY n AND NOTICE THAT $\left.\frac{\partial^2 \cos}{\partial^2 x}\right|_{x=0} = -1$ TO GET

$$-2n\left(\cos(\bar{X}_n) - 1\right) \xrightarrow{D} \chi^2_1$$

BECAUSE $n\bar{X}_n^2 \xrightarrow{D} \chi^2_1$ BY CMT

CENTRAL LIMIT THEOREM AND CHARACTERISTIC FUNCTIONS

WE ARE STILL CONSIDERING CONVERGENCE IN DISTRIBUTION, WHICH RECALL IS DEFINED THROUGH BOUNDED CONTINUOUS FUNCTIONS:

$$X_n \xrightarrow{D} X \quad \text{MEANS} \quad \mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$$

FOR ALL BOUNDED CONTINUOUS FUNCTIONS

IN $\mathbb{R}^d$ INSTEAD OF CONSIDERING ALL BOUNDED CONTINUOUS FUNCTIONS

IT IS ENOUGH TO CONSIDER FUNCTIONS OF THE FORM

$$x \in \mathbb{R}^d \mapsto \exp\{i t^T x\} \quad \text{ALL } t \in \mathbb{R}^d$$

[RECALL $\exp(ix) = \cos(x) + i \sin(x)$]

THESE FUNCTIONS YIELD CHARACTERISTIC FUNCTIONS.

<u>Def</u> LET X BE A RV IN $(\mathbb{R}^1, \mathcal{B}^1)$. THEN

$$t \in \mathbb{R} \mapsto \mathbb{E}[\exp(itX)]$$

IS THE CHARACTERISTIC FUNCTION OF X

IF X IS A RANDOM VECTOR IN $\mathbb{R}^d$, ITS chf IS

$$t \in \mathbb{R}^d \to \mathbb{E}[\exp(i t^T X)]$$

<u>EXAMPLE</u>: IF $X \sim N(0,1)$ $\quad \phi_X(t) = \exp\{-t^2/2\}$

OF COURSE CHARACTERISTIC FUNCTIONS ARE EXAMPLES OF CONTINUOUS BOUNDE FUNCTIONS!

$$|e^{ix}| = \sqrt{\cos^2(x) + \sin^2(x)} = 1$$

$$|e^{ix} - e^{iy}| \leq 2$$ AND APPLY DOMINATED CONVERGENCE THEOREM

PROPERTIES OF CHARACTERISTIC FUNCTIONS:

1) $\phi_X(0) = 1$, $|\phi_X(t)| \leq 1$

2) $\phi_X(-t) = \overline{\phi}_X(t)$

3) $|\phi_X(t+h) - \phi_X(t)| \leq \mathbb{E}\,|e^{ihX} - 1|$

↳ UNIFORMLY CONTINUOUS FUNCTIONS

4) $\phi_{aX+b}(t) = e^{itb}\,\phi_X(at)$

5) IF $X \perp\!\!\!\perp Y$ THEN $\phi_{X+Y}(t) = \phi_X(t)\,\phi_Y(t)$

6) IF $\mathbb{E}\left[|X|^k\right] < \infty$ THEN $\phi_X^{(k)}(0) = i^k\,\mathbb{E}\left[X^k\right]$

k^{th} DERIVATIVE OF ϕ_X AT 0

INVERSION FORMULA AND UNIQUENESS

Thm LET ϕ_X BE THE chf OF THE RANDOM VECTOR $X \in \mathbb{R}^d$.

LET $A = \{(x_1, \dots, x_d) : a_j \leq x_j \leq b_j \text{ ALL } j\}$

WHERE $a_j < b_j$ ALL j. LET μ_X BE THE DISTRIBUTION OF X

AND ASSUME THAT $\mu_X(\partial A) = 0$. LET, FOR $T > 0$,

$$B_T = \{(y_1, \dots y_d) : |y_j| \leq T \text{ ALL } j\}$$

THEN

$$\mu_X(A) = \lim_{T \to \infty} \frac{1}{(2\pi)^d} \int_{B_T} \prod_{j=1}^{d} \left[\frac{\exp(-iy_ja_j) - \exp(-iy_jb_j)}{iy_j}\right] \phi_X(y)\, dy$$

$dy_1 \dots dy_d$

DISTINCT PROB. MEASURES HAVE DISTINCT chf's.

Corollary (CRAMÉR - WOLD) LET X AND Y BE TWO RANDOM VECTORS IN $\mathbb{R}^d$. THEN $X \overset{d}{=} Y$ IFF $t^T X \overset{d}{=} t^T Y$ FOR ALL $t \in \mathbb{R}^d$.