STEIN'S METHOD

- Recall "Stein's identity" [HW 1]:

$$\mathbb{E}f'(Z) = \mathbb{E}Zf(Z) \quad \text{if } Z \sim N(0,1)$$

- Stein's identity characterizes $N(0,1)$ distribution.

Moreover, a stability result holds:

"if $\mathbb{E}f'(W) \approx \mathbb{E}Wf(W)$ $\forall$ bounded, smooth function $f: \mathbb{R} \to \mathbb{R}$
then $W \approx Z \sim N(0,1)$ in distribution"

measured in Wasserstein metric

$$W_1(W, Z) = \sup_{\|h'\|_\infty \le 1} |\mathbb{E}h(W) - \mathbb{E}h(Z)|$$

Precisely:

"<u>Stein's Lemma</u>" Let $Z \sim N(0,1)$ and let W be $\forall$ r.v. Then

$$W_1(W,Z) \le \sup_{f \in \mathcal{F}} |\mathbb{E}f'(W) - \mathbb{E}Wf(W)|$$

where $\mathcal{F} = \{ f: \mathbb{R} \to \mathbb{R}: \|f\|_\infty \le C, \|f'\|_\infty \le C, \|f''\|_\infty \le C \}$

abs. constant

<u>Proof</u> Fix $\forall$ h with $\|h'\|_\infty \le 1$.

Consider Stein's differential equation

$$f'(w) - wf(w) = h(w) - \mathbb{E}h(Z), \qquad w \in \mathbb{R}$$

If $\exists$ solution $f \in \mathcal{F}$ $(*)$

then we can set $w := W$ and take expectation on both sides. $\Rightarrow$

$$\mathbb{E}f'(W) - \mathbb{E}Wf(W) = \mathbb{E}h(W) - \mathbb{E}h(Z)$$

Take $|\cdot|$ on both sides, we complete the proof, modulo $(*)$. $\square$

Proof of $(*)$:

① SOLVING STEIN'S D.E. by the method of integrating factors:

$e^{\int(-w)dw} = e^{-w^2/2}$ in our case

- Multiply both sides by $e^{-w^2/2}$ $\Rightarrow$

$$\left(e^{-w^2/2} f(w)\right)' = e^{-w^2/2}\left(h(w) - \mathbb{E}h(Z)\right)$$

$$\Rightarrow e^{-w^2/2} f(w) = \int_{-\infty}^{w} e^{-x^2/2}\left(h(x) - \mathbb{E}h(Z)\right) dx + \underset{0}{\underset{\shortparallel}{C}}$$

$$\Rightarrow f(w) = e^{w^2/2}\int_{-\infty}^{w} e^{-x^2/2}\left(h(x) - \mathbb{E}h(Z)\right) dx \qquad (*)$$

- Note that the total integral $\int_{-\infty}^{\infty} e^{-x^2/2}(h(x) - \mathbb{E}h(Z)) = \mathbb{E}h(\overbrace{X}^{N(0,1)}) - \mathbb{E}h(Z) = 0$. Thus,

$$\boxed{f(w) = -e^{w^2/2}\int_{w}^{\infty} e^{-x^2/2}\left(h(x) - \mathbb{E}h(Z)\right) dx =: \mathrm{Stein}(h)(w)} \qquad (**)$$

$(*)$ is useful for $w \le 0$, $(**)$ is useful for $w \ge 0$.

Unique up to $ce^{w^2/2}$
$\Rightarrow$ unique if we require $f(w) = o(e^{w^2/2})$ as $w \to \pm\infty$

PROPERTIES OF THE SOLUTION:

hides an absolute constant factor

② $\boxed{\|\mathrm{Stein}(h)\|_\infty \lesssim \|h'\|_\infty}$

WLOG $h(0) = 0$, $\|h'\|_\infty = 1$ $\Rightarrow |h(x)| \le |x|$, $|\mathbb{E}h(Z)| \le \mathbb{E}|Z| \le 1$

$$\Rightarrow \forall w > 0: |f(w)| \overset{(**)}{\le} e^{w^2/2} \underbrace{\int_{w}^{\infty} e^{-x^2/2}(x+1)\,dx}_{\underbrace{\int_w^\infty e^{-x^2/2} x\,dx}_{= e^{-w^2/2}} + \underbrace{\int_w^\infty e^{-x^2/2}\,dx}_{\lesssim e^{-w^2/2}}} \lesssim 1.$$

by Gaussian tail bound ("Mills ratio" p. 28)

For $w < 0$, proceed similarly but use $(*)$.

$$\int_{w}^{\infty} e^{-x^2/2}dx \le \frac{1}{w} e^{-w^2/2} \quad \forall w > 0 \quad (\circ)$$

$\lesssim e^{-w^2/2}$ since LHS ≤ 1 $\forall w$

③ $\boxed{\|\mathrm{Stein}(h)'\|_\infty \lesssim \|h\|_\infty}$

WLOG $\|h\|_\infty = 1$ $\Rightarrow |\mathbb{E}h(Z)| \le 1$

$$\Rightarrow \forall w > 0: f'(w) \overset{(**)}{=} -we^{w^2/2}\cdot\int_{w}^{\infty} e^{-x^2/2}\underbrace{\left(h(x) - \mathbb{E}h(Z)\right)}_{|\cdot| \le 2}dx + e^{w^2/2}\cdot e^{-w^2/2}\underbrace{\left(h(w) - \mathbb{E}h(Z)\right)}_{|\cdot| \le 2}$$

$$\Rightarrow |f'(w)| \le 2we^{w^2/2}\int_{w}^{\infty} e^{-x^2/2}dx + 2 \lesssim 1 \text{ by Gaussian tail bd } (\circ).$$

④ $\|\mathrm{Stein}(h)'\|_\infty \lesssim \underset{\substack{\| \\ 1 \text{ WLOG}}}{\|h'\|_\infty}$

Differentiate Stein's DE (here f = solution (**)):

$$f''(w) - f(w) - wf'(w) = h'(w)$$

$$\Rightarrow \quad f''(w) - wf'(w) = f(w) + h'(w) =: H(w)$$

Use again the method of integrating factors. Multiply by $e^{-w^2/2}$ $\Rightarrow$

$$\left(e^{-w^2/2} f'(w)\right)' = e^{-w^2/2} H(w)$$

$$\Rightarrow \quad e^{-w^2/2} f'(w) = -\int_w^\infty e^{-x^2/2} H(x)\,dx + C$$

$$\Rightarrow \quad f'(w) = -e^{w^2/2} \int_w^\infty e^{-x^2/2} \underbrace{H(x)}_{|H(x)| \le |f(w)| + |h'(w)| \lesssim 1 \text{ by ② \& assumption}}\,dx + Ce^{w^2/2}$$

$$\Rightarrow \quad |f'(w)| \le \underbrace{e^{w^2/2} \int_w^\infty e^{-x^2/2}\,dx}_{\lesssim 1 \text{ by Mills ratio (p.2)}} + \underbrace{Ce^{w^2/2}}_{=0}$$

(If $C \neq 0$, $f'(w) \approx Ce^{w^2/2}$ as $w \to +\infty$ $\Rightarrow$ $f(w)$ is unbounded, contradicting ②)

⑤ $\|\mathrm{Stein}(h)''\|_\infty \le \|h'\|_\infty$

As in ④, $f''(w) - wf'(w) = f(w) + h'(w) =: H(w)$. Substitute $w = Z$, take $\mathbb{E}$ $\Rightarrow$

$\Rightarrow \mathbb{E}H(Z) = \mathbb{E}f''(Z) - \mathbb{E}Zf'(Z) = 0$ by Stein's identity for f'.

$$\Rightarrow \quad f''(w) - wf'(w) = H(w) - \mathbb{E}H(Z)$$

i.e. $f' = \mathrm{Stein}(H)$

$$\Rightarrow \|f''\|_\infty = \|\mathrm{Stein}(H)'\|_\infty \overset{③}{\lesssim} \|H\|_\infty \overset{\triangle}{\le} \|\underset{}{\overset{\mathrm{Stein}(h)}{\overset{\|}{f}}}\|_\infty + \|h'\|_\infty \overset{②}{\lesssim} \|h'\|_\infty.$$

Stein's lemma is now completely proved.

• As an application, let's strengthen quantitative CLT (lem p. 118 for 3 times diff'ble h):

Wasserstein CLT Let $X_1, \dots, X_n$ be independent mean zero random variables with $\mathbb{E}|X_i|^3 < \infty$. Then $W = X_1 + \cdots + X_n$ satisfies

$$W_1(W, Z) \le \underset{\text{absolute const.}}{C} \sum_{i=1}^n \mathbb{E}|X_i|^3 \quad \text{where} \quad Z \sim N(0, \mathrm{Var}(W))$$

Proof WLOG $\mathrm{Var}(W) = 1$.

By Stein's lemma (p. 1), it is enough to prove that

$$\left| \mathbb{E} W f(W) - \mathbb{E} f'(W) \right| \le C \sum_{i=1}^n \mathbb{E}|X_i|^3 \quad \text{whenever } \|f''\|_\infty \le 1.$$

$$W f(W) = \sum_i X_i f(W).$$

Taylor approximation around $W_i := \sum_{j: j \ne i} X_j = W - X_i$:

$W_i \xrightarrow{X_i} W$

$$f(W) = f(W_i) + (W - W_i) f'(W_i) + (W - W_i)^2 \frac{A_i^2}{2} \quad \text{where } |A_i| \le \|f''\|_\infty \le 1$$

Multiply both sides by X_i, sum over i, take $\mathbb{E}$

$$\mathbb{E} W f(W) = \underbrace{\sum_i \mathbb{E} X_i f(W_i)}_{\text{independence} \; = 0} + \sum_i \underbrace{\mathbb{E} X_i (\overbrace{W - W_i}^{= X_i}) f'(W_i)}_{\text{independ.} \; = \mathbb{E}[X_i^2]\, \mathbb{E} f'(W_i)} + \sum_i \mathbb{E} X_i (\overbrace{W - W_i}^{= X_i})^2 \frac{A_i^2}{2}$$

$$|f'(W_i) - f'(W)| \le |W_i - W| \cdot \|f''\|_\infty \le |X_i| \Rightarrow \mathbb{E} f'(W_i) = \mathbb{E} f'(W) + B_i \text{ where } |B_i| \le \mathbb{E}|X_i|$$

$$\Rightarrow \mathbb{E} W f(W) = \underbrace{\sum_i \mathbb{E}[X_i^2]}_{= \mathrm{Var}(W) = 1} \cdot \mathbb{E} f'(W) + \sum_i \mathbb{E}[X_i^2] B_i + \sum_i \mathbb{E} X_i^3 \frac{A_i^2}{2}.$$

$$\overset{\Delta}{\Rightarrow} \left| \mathbb{E} W f(W) - \mathbb{E} f'(W) \right| \le \sum_i \underbrace{\mathbb{E}[X_i^2]\, \mathbb{E}|X_i|}_{\le (\mathbb{E}|X_i|^3)^{2/3} (\mathbb{E}|X_i|^3)^{1/3} = \mathbb{E}|X_i|^3} + \frac{1}{2} \sum_i \mathbb{E}|X_i|^3 \le \frac{3}{2} \sum_i \mathbb{E}|X_i|^3$$

□

Use Thm for $X_i = \frac{Y_i - \mu}{\sigma}$ ⇓

Cor If $Y_1, Y_2, \dots$ are iid r.v.'s with mean μ, variance σ, and $\mathbb{E}|Y_1|^3 =: \rho^3 < \infty$, then $S_n = Y_1 + \cdots + Y_n$ satisfies

$$W_1\left(\frac{S_n - \mu n}{\sigma \sqrt{n}}, \underset{N(0,1)}{Z} \right) \le \frac{C (\rho/\sigma)^3}{\sqrt{n}} \quad \forall n \in \mathbb{N}.$$

Remark (Berry–Esseen CLT) A modification of this argument yields the same bound in Kolmogorov metric $d_K(W, Z) = \sup_{x \in \mathbb{R}} |P\{W \le x\} - P\{Z \le x\}|$ (use $h :=$ smoothing of $1_{(-\infty, x]}$)

[E. Bolthausen, An estimate of the remainder in a combinatorial CLT '1984]