

## 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  $f: \mathbb{R} \rightarrow \mathbb{R}$  then  $W \approx Z \sim N(0,1)$  in distribution"

↑ measured in Wasserstein metric

Precisely:

$$W_1(W, Z) = \sup_{\|h\|_{\text{Lip}} \leq 1} |\mathbb{E}h(W) - \mathbb{E}h(Z)|$$

by approximation, equivalent to  $\|h'\|_{\infty} \leq 1$

"Stein's Lemma" let  $Z \sim N(0,1)$  and let  $W$  be  $\forall$  r.v. Then  $\exists$  abs. const  $C$ :

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

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

Proof Fix  $\forall$   $h$  with  $\|h'\|_{\infty} \leq 1$ .

Consider Stein's differential equation

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

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

Example:  $h(x) = x \Rightarrow f(x) = -\frac{1}{2}$  works  $\left[ \text{SDE becomes } w = w \right]$

Proof of (\*):

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

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

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

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

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

$$\Rightarrow f(w) = e^{w^2/2} \int_{-\infty}^w e^{-x^2/2} (h(x) - \mathbb{E}h(Z)) dx \quad (*) \quad N(0,1)$$

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

$$\text{also } f(w) = -e^{w^2/2} \int_w^{\infty} e^{-x^2/2} (h(x) - \mathbb{E}h(Z)) dx =: \text{Stein}(h)(w) \quad (**)$$

(\*) is useful for  $w < 0$ , (\*\*) is useful for  $w > 0$ .

PROPERTIES OF THE SOLUTION ↑

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

hides an absolute constant factor

$$\textcircled{2} \quad \|\text{Stein}(h)\|_{\infty} \lesssim \|h'\|_{\infty}$$

$$\left[ \begin{array}{l} \text{wlog } h(0) = 0, \|h'\|_{\infty} = 1 \Rightarrow |h(x)| \leq |x|, |\mathbb{E}h(Z)| \leq \mathbb{E}|Z| \leq 1 \\ \Rightarrow \forall w > 0: |f(w)| \stackrel{(**)}{\leq} e^{w^2/2} \int_w^{\infty} e^{-x^2/2} (x+1) dx \lesssim 1. \end{array} \right.$$

$$\int_w^{\infty} e^{-x^2/2} x dx + \int_w^{\infty} e^{-x^2/2} dx$$

$\stackrel{\|}{\leq} e^{-w^2/2} \quad \stackrel{\|}{\leq} e^{-w^2/2}$  by Gaussian tail bound ("Mills ratio" p. 28)

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

$$\textcircled{3} \quad \|\text{Stein}(h')\|_{\infty} \lesssim \|h\|_{\infty}$$

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

$$\Rightarrow \forall w > 0: f'(w) \stackrel{(**)}{=} -we^{w^2/2} \int_w^{\infty} e^{-x^2/2} (h(x) - \mathbb{E}h(Z)) dx + e^{w^2/2} \cdot e^{-w^2/2} (h(w) - \mathbb{E}h(Z))$$

$1 \cdot 1 \leq 2 \quad \quad \quad 1 \cdot 1 \leq 2$

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

$$\int_w^{\infty} e^{-x^2/2} dx \leq \frac{1}{w} e^{-w^2/2} \quad \forall w > 0 \quad (o)$$

$\stackrel{\|}{\leq} e^{-w^2/2}$  since LHS  $\leq 1 \quad \forall w$

SKIP

$$\textcircled{4} \quad \| \text{Stein}(h)' \|_\infty \leq \| h' \|_\infty$$

↓  
wlog

Differentiate Stein's DE (here  $f = \text{solution}(\#)$ ):

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

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

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

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

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

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

$|H(x)| \leq |f(w)| + |h'(w)| \leq 1$  by  $\textcircled{2}$  & assumption

$$\Rightarrow |f'(w)| \leq \underbrace{e^{w^2/2} \int_w^\infty e^{-x^2/2} dx}_{\substack{\text{As} \\ \downarrow \text{by Mills ratio (p.2)}}} + \underbrace{Ce^{w^2/2}}_{\rightarrow 0}$$

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

$$\textcircled{5} \quad \| \text{Stein}(h)'' \|_\infty \leq \| h' \|_\infty$$

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

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

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

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

$$\Rightarrow \| f'' \|_\infty = \| \text{Stein}(H)' \|_\infty \stackrel{\textcircled{3}}{\leq} \| H \|_\infty \stackrel{\text{Stein}(h)}{\leq} \| f \|_\infty + \| h' \|_\infty \stackrel{\textcircled{2}}{\leq} \| h' \|_\infty$$

Stein's lemma is now completely proved.

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

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

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

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

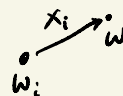
↑ absolute const.

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

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

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

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



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

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

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

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

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

$$\Rightarrow |\mathbb{E}Wf(W) - \mathbb{E}f'(W)| \leq \sum_i \underbrace{\mathbb{E}[X_i^2]}_{(\mathbb{E}|X_i|^3)^{2/3}} \mathbb{E}|X_i| + \frac{1}{2} \sum_i \mathbb{E}|X_i|^3 \leq \frac{3}{2} \sum_i \mathbb{E}|X_i|^3$$

□

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

Cor If  $Y_1, Y_2, \dots$  are iid r.v.'s with mean  $\mu$ , variance  $\sigma$ , and  $\mathbb{E}|Y_1|^3 =: \beta^3 < \infty$ , then

$$S_n = Y_1 + \dots + Y_n \text{ satisfies } W_1\left(\frac{S_n - \mu n}{\sigma \sqrt{n}}, Z\right) \leq \frac{C(\beta/\sigma)^3}{\sqrt{n}} \quad \forall n \in \mathbb{N}$$

↳ Berry-Esseen CLT

Remark A modification of this argument yields the same bound in Kolmogorov

$$\left| \text{metric } d_K(W, Z) = \sup_{x \in \mathbb{R}} |P\{W \leq x\} - P\{Z \leq x\}| \right. \quad \text{(use } h := \text{smoothing of } 1_{(-\infty, x]})$$

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