

# Quantitative CLT via Stein's Method

We've seen that, for any random variable  $W$ ,  
if  $Z \stackrel{d}{=} \mathcal{N}(0,1)$ , then

$$d_{W,1}(W, Z) \leq \sup_{\mathcal{F}} |\mathbb{E}[f'(W) - Wf(W)]|$$

where  $\mathcal{F} = \{f: \mathbb{R} \rightarrow \mathbb{R} : \sup|f| \leq 2, \sup|f'| \leq \sqrt{\frac{2}{\pi}}, \sup|f''| \leq 2\}$

We will apply this to  $W = \frac{S_n}{\sqrt{n}}$  where

$S_n = X_1 + \dots + X_n$ ,  $\{X_n\}_{n=1}^{\infty}$  iid  $\mathbb{E}[X_j] = 0$ ,  $\mathbb{E}[X_j^2] = 1$ ,  $\mathbb{E}[|X_j|^3] < \infty$ .

$$\mathbb{E}[Wf(W)] = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}[X_j f(W)]$$

Key idea:  $W_j :=$

$$\therefore \mathbb{E}[X_j f(W_j)]$$

$$\mathbb{E}[X_j f(w)] = \mathbb{E}[X_j (f(w) - f(w_j))] ]$$

Now, putting on Calculus hats:  $f(w) - f(x) \approx f'(x)(w-x)$

More precisely: if  $f$  is twice differentiable @  $x$ , then:

$$\parallel \\ (w-x)f'(x)$$

$$\begin{aligned} &= \mathbb{E} \left[ X_j (f(w) - f(w_j) - (w-w_j)f'(w_j) + (w-w_j)f'(w_j)) \right] \\ &= \mathbb{E} \left[ X_j \cdot \right] + \mathbb{E} \left[ X_j (w-w_j)f'(w_j) \right] \end{aligned}$$

$$(1) \quad \mathbb{E}[Wf(W)] = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}[X_j f(W)] = \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{1}{2n} \mathbb{E}[X_j^3 f''(\xi_j)] + \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{1}{\sqrt{n}} \mathbb{E}[f'(w_j)]$$

for some rv's  $\xi_j$ .

Now,  $f'(w_j) - f'(w) =$

$$\begin{aligned} \therefore & \left( \frac{1}{n} \sum_{j=1}^n \mathbb{E}[f'(w_j)] \right) - \mathbb{E}[f'(w)] \\ &= \frac{1}{n} \sum_{j=1}^n (\mathbb{E}[f'(w_j)] - f'(w)) = \end{aligned}$$

$$(2) \quad \mathbb{E}[f'(w)] = \left( \frac{1}{n} \sum_{j=1}^n \mathbb{E}[f'(w_j)] \right) +$$

Combining (1) & (2):

$$\mathbb{E}[Wf(W) - f'(w)] = \frac{1}{2n^{3/2}} \sum_{j=1}^n \mathbb{E}[X_j^3 f''(\xi_j)] - \frac{1}{n^{3/2}} \sum_{j=1}^n \mathbb{E}[f''(\eta_j) X_j]$$

$$\therefore |\mathbb{E}[Wf(W) - f'(w)]| \leq \frac{1}{2n^{3/2}} + \frac{1}{n^{3/2}}$$

Now (and only now) using the fact that the  $X_j$  are identically distributed, we have

$$\frac{1}{n} \sum_{j=1}^n \mathbb{E}[|X_j|^3] = \mathbb{E}[|X_1|^3] \quad \text{and} \quad \frac{1}{n} \sum_{j=1}^n \mathbb{E}[|X_j|] = \mathbb{E}[|X_1|]$$

$$\therefore |\mathbb{E}[w f(w) - f'(w)]|$$

$$\leq \frac{1}{2\sqrt{n}} \sup |f''| \mathbb{E}[|X_1|^3] + \frac{1}{\sqrt{n}} \sup |f''| \mathbb{E}[|X_1|]$$

**Theorem:** If  $\{X_n\}_{n=1}^{\infty}$  are iid  $L^3$  random variables with  $\mathbb{E}[X_j] = 0$ ,  $\mathbb{E}[X_j^2] = 1$ , and  $S_n = X_1 + \dots + X_n$ , then if  $Z \stackrel{d}{=} \mathcal{N}(0,1)$ ,

$$d_{w_1}\left(\frac{S_n}{\sqrt{n}}, Z\right) \leq \frac{\mathbb{E}[|X_1|^3] + 2\mathbb{E}[|X_1|]}{\sqrt{n}}$$

Pf.

**Cor:**  $d_{Kol}\left(\frac{S_n}{\sqrt{n}}, Z\right) \leq \frac{\text{Const.}}{n^{1/4}}$