Where independence was spent

Lecture 4 ended with a claim it did not support: that Stein's method survives dependence. The claim is credible because the method has one moving part — every theorem in that lecture is a bound on the single quantity $\mathbb{E}[f'(W) - Wf(W)]$ — and because independence entered the proof in exactly one place. This assignment locates that place and then pays for it twice: with a *coupling*, and with *locality*. It ends on a sum of dependent indicators that nothing proved so far can touch.

recalled from lecture For h bounded and Lipschitz, *Stein's equation* $f'(w) - wf(w) = h(w) - \mathbb{E}h(g)$ has the explicit solution f_h, which satisfies $\|f_h'\|_\infty \le 4\|h - \mathbb{E}h(g)\|_\infty$ and $\|f_h''\|_\infty \le 2\|h'\|_\infty$. Evaluating at W and taking expectations gives the *master equation*

$$\mathbb{E}h(W) - \mathbb{E}h(g) \;=\; \mathbb{E}\big[f_h'(W) - Wf_h(W)\big].$$

Here $g \sim \mathcal{N}(0,1)$. In Theorem 2.4 we bounded the right-hand side for $W = n^{-1/2}\sum_i X_i$ with X_i i.i.d., mean 0, variance 1.

1. One quantity to bound

Problem 1. core Suppose $\big|\mathbb{E}[f'(W) - Wf(W)]\big| \le \varepsilon(\|f'\|_\infty + \|f''\|_\infty)$ for every bounded f with two bounded derivatives. Deduce $|\mathbb{E}h(W) - \mathbb{E}h(g)| \le C\varepsilon(\|h\|_\infty + \|h'\|_\infty)$ for every bounded Lipschitz h, C absolute. *Everything below is therefore an estimate of one expectation.*

Problem 2. core Reread the proof of Theorem 2.4. Write $W^{(i)} = W - X_i/\sqrt{n}$ and identify the **single** step at which independence was used. Then check that $\mathbb{E}[Wf(W)] = n^{-1/2}\sum_i \mathbb{E}[X_i f(W)]$ requires no independence at all, so that what is lost is one factorization and not the strategy.

2. Paying with a coupling: exchangeable pairs

A pair (W, W') is **exchangeable** if (W, W') and (W', W) have the same law, and is a **Stein pair** if in addition $\mathbb{E}[W' - W \mid W] = -aW$ for some $a > 0$. The following one-line fact is the whole mechanism.

Problem 3. core Let (W, W') be exchangeable and let F be antisymmetric, $F(x,y) = -F(y,x)$, with $\mathbb{E}|F(W,W')| < \infty$. Show $\mathbb{E}F(W,W') = 0$.

Problem 4. Apply Problem 3 to $F(x,y) = (y-x)\big(f(x) + f(y)\big)$, expand $f(W')$ about W to first order, and use the Stein-pair property to obtain

$$\mathbb{E}\big[f'(W) - Wf(W)\big] \;=\; \mathbb{E}\Big[f'(W)\Big(1 - \tfrac{1}{2a}\,\mathbb{E}\big[(W'-W)^2 \mid W\big]\Big)\Big] \;+\; \text{err}, \qquad |\text{err}| \le \frac{\|f''\|_\infty}{4a}\,\mathbb{E}|W' - W|^3.$$

So a Gaussian approximation theorem now asks for two things: that $\frac{1}{2a}\mathbb{E}[(W'-W)^2 \mid W]$ be close to 1, and that $W' - W$ be small.

Problem 5. core *The canonical pair.* Let $W = n^{-1/2}\sum_{i\le n} X_i$ with X_i i.i.d., mean 0, variance 1. Let I be uniform on $\{1,\dots,n\}$ and X_I' an independent copy of X_I, both independent of everything else, and put $W' = W - (X_I - X_I')/\sqrt{n}$. Show (W,W') is exchangeable and $\mathbb{E}[W' - W \mid X_1,\dots,X_n] = -W/n$, so by the tower property it is a Stein pair with $a = 1/n$.

Problem 6. core For that pair, show $\mathbb{E}\big[(W'-W)^2 \mid X_1,\dots,X_n\big] = n^{-2}\sum_i (X_i^2 + 1)$, so that $\frac{1}{2a}\mathbb{E}[(W'-W)^2 \mid X_1,\dots,X_n] = \frac{1}{2}\big(1 + \frac{1}{n}\sum_i X_i^2\big)$. Bound its distance from 1 by $\frac{1}{2}\sqrt{\mathrm{Var}(X_1^2)/n}$ and the remainder in Problem 4 by $\|f''\|_\infty \mathbb{E}|X_1 - X_1'|^3/(4\sqrt{n})$, and conclude the central limit theorem with rate $n^{-1/2}$. *Conditional Jensen replaces conditioning on W by conditioning on $X_1,\dots,X_n$; check the inequality goes the right way.*

Independence was used twice above — say where. Neither use was the product formula $\mathbb{E}[X_i f(W^{(i)})] =$

$\mathbb{E}[X_i]\mathbb{E}[f(W^{(i)})]$ of Problem 2, which never appeared.

3. Paying with locality: dependency graphs

Let $Y_1, \ldots, Y_n$ have mean 0, and $W = \sum_i Y_i$ with $\operatorname{Var} W = 1$. A graph G on $\{1, \ldots, n\}$ is a **dependency graph** for the Y_i if $(Y_i)_{i \in A}$ and $(Y_j)_{j \in B}$ are independent whenever A, B are disjoint with no edge of G between them. Write N_i for i together with its neighbours, $D = \max_i |N_i|$, and $W_i = \sum_{j \notin N_i} Y_j$.

> **Problem 7.** CORE Show $Y_i \perp\!\!\!\perp W_i$, hence $\mathbb{E}[Y_i f(W_i)] = 0$ for every bounded f. Show also $\mathbb{E}[Y_i Y_j] = 0$ whenever $j \notin N_i$, and deduce
> $$\sum_i \sum_{j \in N_i} \mathbb{E}[Y_i Y_j] \;=\; \operatorname{Var} W \;=\; 1.$$
> *This pair of facts is what replaces independence: W_i plays the role $W^{(i)}$ played in Theorem 2.4.*

> **Problem 8.** Rerun the proof of Theorem 2.4 with W_i in place of $W^{(i)}$: write $\mathbb{E}[W f(W)] = \sum_i \mathbb{E}[Y_i f(W)]$, expand $f(W)$ about W_i to first order, and use both facts from Problem 7. Show that if $|Y_i| \le M$ for every i, the result is
> $$\big|\mathbb{E}[f'(W) - W f(W)]\big| \;\le\; C\,\|f''\|_\infty\, n\, D^2 M^3.$$
> Identify the step that produced the factor D^2, and the step that would have produced D alone had f'' not been needed.

> **Problem 9.** *The simplest dependent case.* Let $X_1, \ldots, X_{n+1}$ be i.i.d. bounded with mean 0 and variance 1, and set $Y_i = X_i X_{i+1}$ for $i \le n$. Show that Y_i and Y_j are independent whenever $|i-j| \ge 2$, so $D = 3$; that the Y_i are *uncorrelated*, so $\operatorname{Var}\big(\sum_i Y_i\big) = n$; and that Problem 8 gives $n^{-1/2} \sum_i Y_i \Rightarrow \mathcal{N}(0,1)$ with rate $n^{-1/2}$. *Compare Homework 3, Problem 11: pairwise uncorrelatedness alone did not save the characteristic-function proof, and does not save this one either. Locality does.*

4. Where independence fails outright

Let π be a uniformly random permutation of $\{1, \ldots, n\}$, let $I_i = \mathbf{1}\{\pi(i) = i\}$, and let $W = \sum_i I_i$ be the number of fixed points.

> **Problem 10.** CORE Show $\mathbb{E}W = 1$ and $\operatorname{Var} W = 1$ for $n \ge 2$. Then show $\mathbb{P}(I_1 = I_2 = 1) \ne \mathbb{P}(I_1 = 1)\mathbb{P}(I_2 = 1)$, so the I_i are not independent and Theorem 2.7 (Le Cam) does not apply. *Nor does anything in Homework 3: every Poisson statement there assumed independent summands.*

IMPORTED, WITHOUT PROOF **Chen–Stein, coupling form.** Let $W = \sum_{i \le n} I_i$ with $I_i \sim \operatorname{Bernoulli}(p_i)$, not necessarily independent, and $\lambda = \sum_i p_i$. Suppose that for each i one can construct, on the same probability space as W, a variable V_i whose law is that of $W - 1$ conditioned on $I_i = 1$. Then

$$d_{\mathrm{TV}}\big(W, \operatorname{Poisson}(\lambda)\big) \;\le\; \min(1, \lambda^{-1}) \sum_i p_i\, \mathbb{E}\big|W - V_i\big|.$$

> **Problem 11.** Fix i. If $\pi(i) = i$ put $\sigma = \pi$; otherwise put $\sigma = \pi \circ (i\;\; \pi^{-1}(i))$. Show $\sigma(i) = i$ always, that σ is uniform among permutations fixing i, and that $W(\sigma) - W(\pi) \in \{1, 2\}$ when $\pi(i) \ne i$, the value 2 occurring exactly when π has a 2-cycle at i. Deduce $\mathbb{E}|W - V_i| \le 2/n$ with $V_i = W(\sigma) - 1$, and hence
> $$d_{\mathrm{TV}}\big(W, \operatorname{Poisson}(1)\big) \;\le\; \frac{2}{n}.$$

Finally, by inclusion–exclusion $\mathbb{P}(W = 0) = \sum_{k=0}^{n} (-1)^k / k!$, whose distance from the Poisson value e^{-1} is 7.1×10^{-3}, 1.8×10^{-4} and 2.3×10^{-8} at $n = 4, 6, 10$. The bound just proved gives 0.2 at $n = 10$. *By how many orders of magnitude is the method wrong here — and what does it have that inclusion–exclusion does not?*

CORE problems are short and self-contained. The quiz on Wednesday Sep 16 draws on the core problems of this set and of Homework 3. AI assistance is permitted and encouraged throughout.