STAT C206A / MATH C223A : Stein's method and applications 1

Lecture 3

Lecture date: Aug 31, 2007 *Scribe: David Rosenberg*

1 First step in Stein's method

Suppose you have a r.v. W and a standard gaussian r.v. Z and you want to bound

$$\sup_{g\in\mathcal{D}} |\mathbf{E}\, g(W) - \mathbf{E}\, g(Z)|,$$

where $\mathcal{D}$ is some given class of functions. The first step is to find another class of functions $\mathcal{D}'$ such that

$$\sup_{g\in\mathcal{D}} |\mathbf{E}\, g(W) - \mathbf{E}\, g(Z)| \le \sup_{f\in\mathcal{D}'} \left|\mathbf{E}\left(f'(W) - Wf(W)\right)\right|. \tag{1}$$

Stein's idea: If $\mathcal{D}'$ is a class of functions such that for every $g \in \mathcal{D}$, $\exists f \in \mathcal{D}'$ s.t.

$$f'(x) - xf(x) = g(x) - \mathbf{E}\, g(Z) \tag{2}$$

for $Z \sim N(0,1)$, then (1) holds. (This o.d.e. is sometimes called the 'Stein equation'.) Indeed, take any $g \in \mathcal{D}$, and find $f \in \mathcal{D}'$ that solves the above equation. Then

$$\begin{aligned}\mathbf{E}\, g(W) - \mathbf{E}\, g(Z) &= \mathbf{E}\left[g(W) - \mathbf{E}\, g(Z)\right] \\ &= \mathbf{E}\left(f'(W) - Wf(W)\right).\end{aligned}$$

Clearly, given $\mathcal{D}$ it is in our interest to have $\mathcal{D}'$ as small as possible.

Lemma 1 (Stein) *Given a function $g : \mathbb{R} \to \mathbb{R}$ that is bounded, $\exists$ absolutely continuous f solving $f'(x) - xf(x) = g(x) - \mathbf{E}\, g(Z)$ for all x, satisfying*

$$|f|_\infty \le \sqrt{\frac{\pi}{2}}\, |g - Ng|_\infty \quad \textit{and} \quad \left|f'\right|_\infty \le 2\, |g - Ng|_\infty$$

(where $|f|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$, $Ng := \mathbf{E}\, g(Z), Z \sim N(0,1)$).

If g is Lipschitz, but not necessarily bounded, then

$$|f|_\infty \le \left|g'\right|_\infty, \quad \left|f'\right|_\infty \le \sqrt{\frac{2}{\pi}}\left|g'\right|_\infty, \quad \textit{and} \quad \left|f''\right|_\infty \le 2\left|g'\right|_\infty.$$

Actually, the third and fourth inequalities are not due to Stein, but were obtained later.

Exercise 1: Show that all five constants are the best possible.

So we can now take $\mathcal{D}'$ to be these f's (which in particular have the given bounds).

The above lemma tells us, for instance, that

$$\mathrm{Wass}(W,Z) \le \sup\left\{\left|\mathbf{E}\left(f'(W) - Wf(W)\right)\right| : |f|_\infty \le 1, |f'|_\infty \le \sqrt{2/\pi}, |f''|_\infty \le 2\right\}.$$

2 Example: Ordinary CLT in the Wasserstein metric

Suppose $X_1, X_2, \ldots, X_n$ are independent, mean 0, variance 1, $\mathbf{E}|X_i|^3 < \infty$. Let $S_n = \sum_1^n X_i$. Take any $f \in C^1$ with f' absolutely continuous, and satisfying $|f| \le 1$,$|f'| \le \sqrt{2/\pi}$, and $|f''| \le 2$. First, note that

$$\mathbf{E}\,Wf(W) = \frac{1}{\sqrt{n}} \sum \mathbf{E}\left(X_i f(W)\right). \tag{3}$$

Now let

$$W_i = W - \frac{X_i}{\sqrt{n}} = \frac{\sum_{j\ne i} X_j}{\sqrt{n}}$$

Then X_i, W_i are independent. Thus

$$\mathbf{E}\,X_i f(W_i) = \underbrace{\mathbf{E}(X_i)}_{=0}\,\mathbf{E}\,f(W_i) = 0$$

and so

$$\begin{aligned}\mathbf{E}\left(X_i f(W)\right) &= \mathbf{E}\left(X_i\left(f(W) - f(W_i)\right)\right)\\ &= \mathbf{E}\left(X_i\left(f(W) - f(W_i) - (W-W_i)f'(W_i)\right)\right)\\ &\quad + \mathbf{E}\left[X_i(W-W_i)f'(W_i)\right].\end{aligned}$$

Note that

$$\left|f(b) - f(a) - (b-a)f'(a)\right| \le \frac{1}{2}(b-a)^2\left|f''\right|_\infty$$

and that $W - W_i = X_i/\sqrt{n}$. Thus

$$\begin{aligned}&\left|\mathbf{E}\left[X_i\left(f(W) - f(W_i) - \frac{X_i}{\sqrt{n}}f'(W_i)\right)\right]\right|\\ &\le \frac{1}{2}\,\mathbf{E}\left|X_i\frac{X_i^2}{n}\right| \cdot \left|f''\right|_\infty \le \frac{1}{n}\,\mathbf{E}\left|X_i\right|^3.\end{aligned}$$

Again,

$$\mathbf{E}\left[X_i\,(W - W_i)\,f'(W_i)\right] = \frac{1}{\sqrt{n}}\,\mathbf{E}\,X_i^2 f'(W_i)$$
$$= \frac{1}{\sqrt{n}}\,\mathbf{E}\,f'(W_i)$$

since $\mathbf{E}\,X_i^2 = 1$ and X_i is independent of W_i.

From (3) and the above calculation we see that

$$\left|\mathbf{E}\,W f(W) - \frac{1}{n}\sum \mathbf{E}\,f'(W_i)\right| \le \frac{1}{n^{3/2}}\sum_{i=1}^{n}\mathbf{E}\,|X_i|^3 .$$

Finally, note that

$$\left|\frac{1}{n}\sum \mathbf{E}\,f'(W_i) - \mathbf{E}\,f'(W)\right| \le \frac{|f''|_\infty}{n}\sum \mathbf{E}\,|W - W_i|$$
$$= \frac{|f''|_\infty}{n^{3/2}}\sum \mathbf{E}\,|X_i| \le \frac{2}{n^{3/2}}\sum \mathbf{E}\,|X_i| .$$

Combining, we have

$$\left|\mathbf{E}\,f(W)W - \mathbf{E}\,f'(W)\right|$$
$$\le \frac{1}{n^{3/2}}\sum \mathbf{E}\,|X_i|^3 + \frac{2}{n^{3/2}}\sum \mathbf{E}\,|X_i| .$$

Since $\mathbf{E}\,X_i^2 = 1$ we can conclude that $\mathbf{E}\,|X_i|^3 \ge 1$ and hence $\mathbf{E}\,|X_i| \le \left(\mathbf{E}\,|X_i|^3\right)^{1/3} \le \mathbf{E}\,|X_i|^3$. We have now arrived at a 'Berry-Esséen bound' for the Wasserstein metric:

Theorem 2 *Suppose $X_1, \ldots, X_n$ are independent with mean 0, variance 1, and finite third moments. Then*

$$\mathrm{Wass}\left(\frac{\sum_1^n X_i}{\sqrt{n}},\, Z\right) \le \frac{3}{n^{3/2}}\sum_1^n \mathbf{E}\,|X_i|^3 ,$$

where $Z \sim N(0,1)$.

Unfortunately, this isn't a real Berry-Esséen bound, since it's a bound on the Wasserstein metric and not the Kolmogorov metric. From a lemma proved in Lecture 2, we can get

$$\mathrm{Kolm}(W,Z) \le 2\sqrt{\frac{1}{\sqrt{2\pi}}\mathrm{Wass}(W,Z)} = \frac{2}{(2\pi)^{1/4}}\sqrt{\mathrm{Wass}(W,Z)}.$$

But this is of order $n^{-1/4}$, which is suboptimal.

Exercise 2: Get the true Berry-Esséen bound using Stein's method. This involves analyzing the solution of the Stein equation (2) for $g(x) = 1_{\{x \leq t\}}$ for arbitrary $t \in \mathbb{R}$.

Exercise 3: Consider Erdős-Rényi graph $G(n,p)$. Has n vertices and $\binom{n}{2}$ possible edges, each edge being open or closed with prob p and $1-p$, independently of each other. Let T_n = number of triangles in this graph. Find a way to use Stein's method to prove the CLT for T_n when (a) p is fixed, and (b) p is allowed to go to zero with n.