Lecture 10 Notes

Agenda: more nonparametric estimation

- recap from end of last time
- convergence rate
- localizing functionals
- implementation

Recap:

graphon function: ω
node locations $U_i \underset{i.i.d.}{\sim}$Uniform[0,1]
Ball $B(x, y, h) := \{(a, b) \in [0, 1]^2 : |x - a| < h \;\&\; |y - b| < h\}$

Method for approximating ω
1. Pretend for now we observe all U's
2. Take all $(u_i, u_j) \in B(x, y, h)$
3. Average A_{ij} for those dyads
4. Return as $\hat{\omega}(x, y)$

recall:

$$\begin{aligned} \mathbb{P}(A_{ij} = 1 | u_i = x, u_j = y) &= \omega(x, y) \\ &= \mathbb{E}[A_{ij} | u_i = x, u_j = y] \end{aligned} \tag{1}$$

If w is smooth, then $\omega(x + \epsilon, y + \nu)$ is close to (x, y).
So averaging A_{ij} from points near (x, y) should approximate $\omega(x, y)$

$$\langle A; B(x, y, z)\rangle = \frac{1}{n^2|B|} \sum_{(i,j):(u_i,u_j)\in B(x,y,h)} A_{ij} \tag{2}$$

$$\mathbb{E}[\langle A; B\rangle] = \frac{1}{n^2|B|} \sum_{(i,j)\in B} \mathbb{E}[A_{ij}] = \frac{1}{n^2|B|} \sum_{(i,j)\in B} \omega(u_i, u_j) \tag{3}$$

Assume that ω is a smooth function - specifically that

$$\frac{1}{|B|} \int_{B(x,y,h)} |\omega(u, v) - \omega(x, y)|\, du\, dv \leq kh^{\gamma} \tag{4}$$

for some $k, \gamma \in \mathbb{R}^+$

$$\frac{1}{n^2|B|} \sum_{(i,j)\in B} \omega(u_i, u_j) \xrightarrow{\substack{L^2 \\ almost\ surely}} \frac{1}{|B|} \int_{B(x,y,h)} \omega(u, v)\, du\, dv = \overline{\omega}(x, y, h) \tag{5}$$

As a consequence of this assumption

$$|\overline{\omega}(x,y,h) - \omega(x,y)| \leq O(h^{\gamma}) \tag{6}$$

What about variance?

$$\mathbb{V}ar[\langle A; B\rangle] = \mathbb{V}ar[\langle \omega; B\rangle + \langle \epsilon, B\rangle] \tag{7}$$

where:

$$\epsilon_{ij} = A_{ij} - \omega(u_i, u_j) \tag{8}$$

ϵ_{ij} is dependent on $\omega(u_i, u_j)$ but also has conditional mean 0
$A_{ij} \in \{0, 1\}$
$\omega(u_i, u_j) \in (0, 1)$

Therefore ϵ_{ij} is uncorrelated with $\omega(u_i, u_j)$

Recall: for random variables X and Y...

$$\begin{aligned} \mathbb{C}[X,Y] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[X\mathbb{E}[Y|X]] - \mathbb{E}[X]\mathbb{E}\mathbb{E}[Y|X]] \end{aligned} \tag{9}$$

Also:

$$\mathbb{V}ar[X+Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y] + 2\mathbb{C}ov[X,Y] \tag{10}$$

So:

$$\mathbb{V}ar[\langle A; B\rangle] = \mathbb{V}ar[\langle \omega; B\rangle] + \mathbb{V}ar[\langle \epsilon; B\rangle] \tag{11}$$

Now we take a closer look at (11).
Since ϵ_{ij} is uncorrelated with ϵ_{kl}, it follows that

$$\mathbb{V}ar[\langle \epsilon_{ij}; B\rangle] = O\left(\frac{1}{n^2 h^2}\right) \tag{12}$$

And (since $\mathbb{V}ar[Bern] = p(1-p)$) we have

$$\mathbb{V}ar[\epsilon_{ij}] \leq \frac{1}{4} \tag{13}$$

$\mathbb{V}ar[\langle \omega; B\rangle]$ is much more *annoying*
$\mathbb{V}ar[\frac{1}{n^2|B|}\sum_{(i,j)\in B}\omega(u_i, u_j)]$
$\omega(u_i, u_j)$ is correlated with $\omega(u_i, u_k)$

There are two approaches for us to work with
Approach 1: bound $\mathbb{C}ov[\omega(u_i, u_j), \omega(u_i, u_k)]$ using smoothness and finite size h
Approach 2: steal result on generalized U-statistics

U-Statistics

Given independent random variables $X_1, X_2, ..., X_n$ and a symmetric function ψ of two args, a **U-Statistic** is

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\psi(X_i, X_j) = U_\psi \tag{14}$$

All terms in here are dependent on eachother (for terms that share the same X_i)

general results on variance of U_ψ based on

$$\mathbb{V}ar[\psi(X_1, X_2)] \tag{15}$$

$$\mathbb{V}ar[\mathbb{E}[\psi(X_1, X_2)|X_1]] \tag{16}$$

(15)	Variance of individual summands
(16)	Covariance between summands

Generalized U-statistic: Given: $X_1, ..., X_n \underset{i.i.d.}{\sim}$ $Y_1, ..., Y_n \underset{i.i.d.}{\sim}$

$$U_\psi = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\psi(X_i, Y_j) \tag{17}$$

(note that Ys must have different distribution than Xs)

Generalized results on $\mathbb{V}ar[U_\psi]$ in terms of

$\mathbb{V}ar[\psi(X_1, X_2)]$	$Y \sim$ vertical coordinates
$\mathbb{V}ar[\mathbb{E}[\psi(X, Y)|X]]$	$X \sim$ U-coordinates with horizontal limit of box
$\mathbb{V}ar[\mathbb{E}[\psi(X, Y)|Y]]$	$\psi \sim \omega$

Use smoothness of ω function

$$\mathbb{V}ar[\langle\omega; B\rangle] = O(\frac{1}{nh}) \tag{18}$$

$$\begin{aligned}\mathbb{V}ar[; B\rangle] &= \mathbb{V}ar[\langle\epsilon; B\rangle] + \mathbb{V}ar[\langle\omega; B\rangle] \\ &= O\left(\frac{1}{n^2h^2}\right) + O\left(\frac{1}{nh}\right) \\ &= O\left(\frac{1}{nh}\right)\end{aligned} \tag{19}$$

$$\begin{aligned}\mathbb{E}[\langle A; B\rangle] &= \overline{\omega}(x, y, h) \\ &= \omega(x, y) + O(h^\gamma)\end{aligned} \tag{20}$$

So we have the *Mean-Squared Error (MSE)* as $[bias]^2 + [variance]$

$$\mathbb{E}[(\langle A; B\rangle - \omega(x,y))^2] = O(h^{2\gamma}) + O\left(\frac{1}{nh}\right) \tag{21}$$

Pick h that minimizes this.
note: there is a trade-off between bias & variance

$$\begin{aligned} \emptyset\left(h^{2\gamma-1}\right) + O\left(\frac{-1}{n^2h^2}\right) &= O \\ h^{2\gamma-1} &= \frac{1}{nh^2} \\ h^{2\gamma+1} &= \frac{1}{n} \\ h &= n^{\frac{-1}{2\gamma+1}} \end{aligned} \tag{22}$$

And thus:

$$\begin{aligned} MSE &= O\left(n^{\frac{-1\gamma}{2\gamma+1}}\right) + O\left(\frac{1}{n \cdot n^{-\frac{1}{2\gamma+1}}}\right) \\ &= O\left(n^{-\frac{2\gamma}{2\gamma+1}}\right) \end{aligned} \tag{23}$$

Recall that for parametric estimates we only get $\mathrm{O(n^{-1})}$

Topology Fact:
$dim(X) = dim(Y) \iff \exists \phi : X \to Y$ s.t. ϕ is continuous and $\exists \phi^{-1}$ s.t. ϕ^{-1} is also continuous.

Graphon Fact: Any CID model is equivalent to a ω-function $[0,1]^2 \to [0,1]$
i.e. $\mathbb{P}(x_i, x_j) = \omega(\phi(u_i), \phi(u_j))$ for some $\phi : \mathbb{X} \to [0,1]$

Suppose $\mathbb{X}$ is $\mathbb{R}^2$ or $\mathbb{R}^3$ as in Continuous Latent Space Models (CLSM). There cannot exist a homeomorphism between $\mathbb{R}^2$ and $[0,1]$.
$\therefore \phi$ in $\mathbb{P}(x_i, x_j) = \omega(\phi(u_i), \phi(u_j))$ must not be smooth.
$\therefore$ if $\mathbb{P}$ is smooth, ϕ is not smooth, ω must not be smooth.

Localizing Functionals and Implementation

These topics will be covered in Lecture 11.