An introduction to Stein's method

Nathan Ross (University of Melbourne)

Three lectures

1. Basics and normal approximation
2. Poisson approximation
3. Multivariate and process approximation

References:

- Ross (2011). Fundamentals of Stein's method. *Probability Surveys*.
- Barbour, Holst, Janson (1992). Poisson approximation.
- Chen, Goldstein, Shao (2011). Normal approximation by Stein's method.

A collection of exercises from Greg Terlov and Partha Dey:

- `https://sites.google.com/a/illinois.edu/gterlov`

An introduction to Stein's method

2. Poisson approximation

https://tinyurl.com/mrpd826w

Nathan Ross (University of Melbourne)

Classical Law of Small Numbers

Assume

- $X_1, X_2, \ldots$ independent with $X_i \sim \mathrm{Bernoulli}(p_i)$,
- $W = W_n = \sum_{i=1}^n X_i$.

Then, for $Z = Z_n \sim \mathrm{Poisson}\big(\sum_{i=1}^n p_i\big)$ and $A \subseteq \{0, 1, \ldots, \}$,

$$|\mathbb{P}(W \in A) - \mathbb{P}(Z \in A)| \leq \sum_{i=1}^n p_i^2.$$

Erdős-Rényi random Graph

Erdős-Rényi random graph $G_n \sim \mathrm{ER}(n, p)$:

- n vertices,
- Each of the $\binom{n}{2}$ possible edges is present with probability p, independent between edges.

Structural statistics:

- Number of vertices of degree k, $k = 0, 1, \ldots$
- Number of small subgraphs such as triangles and two-stars (related to clustering coefficient).

We can write these statistics as

$$W = \sum_\alpha X_\alpha,$$

where X_α is the indicator that the structure occurs at "position" α.

Q: When is $\sum_\alpha X_\alpha$ close in distribution to a Poisson distribution?

Total variation distance

For random variables W and Z, define the total variation distance between their distributions by

$$d_{\mathrm{TV}}(W, Z) = \sup_{\text{event } A} |\mathbb{P}(W \in A) - \mathbb{P}(Z \in A)|$$
$$= \sup_{\text{event } A} |\mathbb{E} h_A(W) - \mathbb{E} h_A(Z)|,$$

where $h_A(x) = \mathbb{I}[x \in A]$.

The second expression has the right form for Stein's method.

Stein's Method Framework

Three steps to Stein's method for a given target distribution of Z.

1. Characterising operator $\mathcal{A} = \mathcal{A}_Z$ on real valued functions:

$$\mathbb{E}\mathcal{A}f(X) = 0 \text{ wide class of functions } f \iff X \stackrel{d}{=} Z.$$

2. For given h, find Stein solution f_h:

$$\mathcal{A}f_h(x) = h(x) - \mathbb{E}h(Z) =: \tilde{h}(x).$$

3. Use structure of W and properties of f_h to bound

$$|\mathbb{E}\mathcal{A}f_h(W)| = |\mathbb{E}h(W) - \mathbb{E}h(Z)|.$$

For bound on d_{TV}, take $h(x) = \mathbb{I}[x \in A]$ with $A \subseteq \mathbb{Z}$.

Stein's Method for Poisson Approximation

For integer-valued random variable W and $Z \sim \mathrm{Poisson}(\lambda)$,

$$d_{\mathrm{TV}}(W, Z) \leq \sup_{f \in \mathcal{F}_\lambda} |\mathbb{E}[\lambda f(W+1) - Wf(W)]|,$$

where

$$\mathcal{F}_\lambda := \left\{ f : \|f\|_\infty < \lambda^{-1/2}; \|\Delta f\|_\infty \leq (1 - e^{-\lambda})/\lambda \right\},$$

Size-bias Bound

For $W \geq 0$ an integer-valued random variable with $\mathbb{E}W = \lambda$, we say the random variable W^s has the size-biased distribution of W if

$$\mathbb{P}(W^s = k) = \frac{k\mathbb{P}(W = k)}{\lambda}.$$

If (W, W^s) are defined on the same space and $Z \sim \mathrm{Poisson}(\lambda)$, then

$$d_{\mathrm{TV}}(W, Z) \leq \min\{1, \lambda\}\mathbb{E}|W + 1 - W^s|.$$

Size-bias Construction

Assume

- $W = \sum_\alpha X_\alpha$, with $X_\alpha \sim \mathrm{Bernoulli}(p_\alpha)$ (any dependence),
- $\lambda = \sum_\alpha p_\alpha < \infty$.

If for each α,

$$\mathscr{L}\big((X_\beta^{(\alpha)})_{\beta\neq\alpha}\big) = \mathscr{L}\big((X_\beta)_{\beta\neq\alpha} | X_\alpha = 1\big)$$

and I is independent random variable with $\mathbb{P}(I = \alpha) = p_\alpha/\lambda$, then

$$W^s := 1 + \sum_{\beta\neq I} X_\beta^{(I)}$$

has the size-biased distribution of W.

If variables above are on the same space and $Z \sim \mathrm{Poisson}(\lambda)$,

$$d_{\mathrm{TV}}(W, Z) \leq \min\{\lambda^{-1}, 1\} \textstyle\sum_\alpha p_\alpha \, \mathbb{E}\big|X_\alpha - \sum_{\beta\neq\alpha}(X_\beta^{(\alpha)} - X_\beta)\big|.$$

Erdős-Rényi application

$$d_{\mathrm{TV}}(W,Z) \leq \min\{\lambda^{-1},1\}\sum_\alpha p_\alpha\,\mathbb{E}\Big|X_\alpha - \sum_{\beta\neq\alpha}(X_\beta^{(\alpha)} - X_\beta)\Big|.$$

If the construction is such that $X_\beta^{(\alpha)} \geq X_\beta$, then

$$d_{\mathrm{TV}}(W,Z) \leq \frac{\mathrm{Var}(W)}{\lambda} - 1 + \frac{2}{\lambda}\sum_\alpha p_\alpha^2.$$

This result applies to W equal to the count in an Erdős-Rényi random graph of any of

- vertices having degree no greater than k,
- vertices having degree no less than k,
- copies of a fixed graph H.

Still work to analyze the mean and variance to determine when this is small in terms of parameter of the Erdős-Rényi graph.

Some notes

- Poisson bounds under local dependence (Ross 2011, Sections 4.1 and 4.2).
- Can also use size-biasing in Stein's method for normal approximation (Ross 2011, Section 3.4) and (Chen, Goldstein, Shao 2011, Section 2.3.4, and applications throughout).
- More general construction for size-biasing a sum of random variables, where size-bias a random summand chosen proportional to its mean, and adjust the rest conditional on that summand having the size-biased value.

Exercises

- (Easy) Derive a general bound in terms of mean and variance in the case that the size-bias coupling satisfies $X_\beta^{(\alpha)} \leq X_\beta$.
- (Easy) Use the previous result to bound the total variation distance between the number of empty bins when distributing k balls into n bins uniformly at random, and a Poisson distribution with the same mean.
- (Hard) Derive a bound on the total variation distance between the number of vertices having degree exactly k in an Erdős-Rényi random graph, and a Poisson distribution with the same mean.

References and Further Reading

Basic introduction:

- Ross (2011). Fundamentals of Stein's method. *Probability Surveys*. **Sections 4.0 and 4.3**.

Research monograph:

- Barbour, Holst, Janson (1992). Poisson approximation. **Chapter 1, Sections 2.1, 5.1, and 5.2**.