An introduction to Stein's method

Nathan Ross (University of Melbourne)

Four lectures

1. Basics and normal approximation
2. Poisson approximation
3. Multivariate and process approximation
4. Concentration inequalities

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

4. Concentration inequalities

https://tinyurl.com/bdds7dpe

Nathan Ross (University of Melbourne)

Hoeffding's inequality

Assume

- $X_1, X_2, \ldots, X_n$ independent random variables,
- $a \leq X_i \leq b$,
- $W = W_n = \sum_{i=1}^n X_i$.

Then for any $t > 0$,

$$\mathbb{P}(|W_n - \mathbb{E}W_n| \geq t) \leq 2e^{-\frac{2t^2}{n(b-a)}}.$$

Exchangeable pairs bound

Assume

- W is a random variable with $\mathbb{E}[W] = 0$ and $\mathrm{Var}(W) = \sigma^2$.
- (W, W') is an exchangeable pair, satisfying

$$\mathbb{E}[W' - W | W] = -\lambda W,$$

  some $\lambda > 0$.
- $\mathbb{E}\left[(W' - W)^2 | W\right] \leq 2\lambda(bW + c)$, some $b, c \geq 0$.
- $\mathbb{E}[e^{\theta W}] < \infty$, all $\theta \in \mathbb{R}$.
- $\mathbb{E}\left[e^{\theta W} |W' - W|\right] < \infty$, all $\theta \in \mathbb{R}$.

Then for all $t > 0$,

$$\mathbb{P}(W \geq t) \leq \exp\left\{\frac{-t^2}{2(c + bt)}\right\}, \quad \mathbb{P}(W \leq -t) \leq \exp\left\{\frac{-t^2}{2c}\right\}.$$

Application: Hoeffding's combinatorial statistic

- Fix $(a_{ij})_{1 \le i,j \le n}$ with $0 \le a_{ij} \le 1$.
- Let π be a uniform random permutation of $\{1, \ldots, n\}$.

Let

$$W = \sum_i a_{i\pi_i} - \frac{1}{n} \sum_{i,j} a_{i,j}.$$

Then for any $t > 0$,

$$\mathbb{P}(|W| \ge t) \le 2 \exp \left\{ \frac{-t^2}{\frac{4}{n} \sum_{i,j} a_{ij} + 2t} \right\}$$

Size bias bound

Assume

- $X \geq 0$ is a random variable with $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$.
- X^s is a size-biased coupling of X with $|X^s - X| \leq c$.

If $X^s \geq X$, then for all $t > 0$,

$$\mathbb{P}(X - \mu \leq -t) \leq \exp\left\{\frac{-t^2}{2c\mu}\right\}.$$

If $\mathbb{E}[e^{(2/c)X}] < \infty$, then for all $t > 0$,

$$\mathbb{P}(X - \mu \geq t) \leq \exp\left\{\frac{-t^2}{c(2\mu + t)}\right\}.$$

Size-bias Construction

Assume

- $X = \sum_\alpha Y_\alpha$, with $Y_\alpha \sim \mathrm{Bernoulli}(p_\alpha)$ (any dependence),
- $\mu = \sum_\alpha p_\alpha < \infty$.

If for each α,

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

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

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

has the size-biased distribution of X.

References and further reading

- Ross (2011). Fundamentals of Stein's method. *Probability Surveys*. **Section 7**