Lecture 26: Poisson approximation for subgraph counts

Last time we showed Theorem 25.3 giving a general bound on the total variation distance between the distribution of a sum of indicators and a Poisson distribution.

Theorem 25.3. *Let $(X_i)_{i=1}^n$ be a collection of Bernoulli random variables with $\mathbb{P}(X_i = 1) = p_i$, and for each $i = 1, \dots, n$, let $(X_j^{(i)})_{j \neq i}$ be defined on the same space as $(X_j)_{j=1}^n$ such that*

$$\mathscr{L}\big((X_j^{(i)})_{j\neq i}\big) = \mathscr{L}\big((X_j)_{j\neq i} | X_i = 1\big).$$

Then for $W = \sum_{i=1}^n X_i$ and $\mu := \mathbb{E}[W] = \sum_{i=1}^n p_i$, we have

$$d_{\mathrm{TV}}\big(\mathscr{L}(W), \mathrm{Po}(\mu)\big) \leq \min\big\{1, \mu^{-1}\big\} \sum_{i=1}^n p_i \mathbb{E}\big| X_i + \sum_{j\neq i} \big(X_j - X_j^{(i)}\big)\big|.$$

We then applied the theorem to derive a general bound on the total variation distance between the number of isolated vertices in an Erdős-Rényi random graph and a Poisson distribution. The computations in that case were particularly easy because in the coupling we used, $X_j^{(i)} \geq X_j$. This is a general phenomenon which we capture in the following result.

Theorem 26.2. *Let $(X_i)_{i=1}^n$ be a collection of Bernoulli random variables with $\mathbb{P}(X_i = 1) = p_i$, and for each $i = 1, \dots, n$, let $(X_j^{(i)})_{j \neq i}$ be defined on the same space as $(X_i)_{i=1}^n$ such that*

$$\mathscr{L}\big((X_j^{(i)})_{j\neq i}\big) = \mathscr{L}\big((X_j)_{j\neq i} | X_i = 1\big).$$

If $X_j^{(i)} \geq X_j$ for all $i \neq j$, then for $W = \sum_{i=1}^n X_i$ and $\mu := \mathbb{E}[W] = \sum_{i=1}^n p_i$, we have

$$d_{\mathrm{TV}}(W, Z) \leq \min\{1, \mu^{-1}\} \bigg(\mathrm{Var}(W) - \mu + 2 \sum_{i=1}^n p_i^2 \bigg).$$

Proof. Using the bound of Theorem 25.3, the triangle inequality, and that $X_j^{(i)} \geq X_j$ in the first line below, and then that $\mathscr{L}(X_j^{(i)}) = \mathscr{L}(X_j | X_i = 1)$ in the first equality, we have

$$\begin{aligned}
d_{\mathrm{TV}}\big(\mathscr{L}(W), \mathrm{Po}(\mu)\big) &\leq \min\big\{1, \mu^{-1}\big\} \sum_{i=1}^n p_i \mathbb{E}\Big[X_i + \sum_{j\neq i} \big(X_j^{(i)} - X_j\big)\Big] \\
&= \min\big\{1, \mu^{-1}\big\} \sum_{i=1}^n \Big(p_i^2 + \sum_{j\neq i} \big(\mathbb{P}(X_j^{(i)} = 1) p_i - p_i p_j \big) \Big) \\
&= \min\big\{1, \mu^{-1}\big\} \sum_{i=1}^n \Big(p_i^2 + \sum_{j\neq i} \big(\mathbb{P}(X_j = 1, X_i = 1) - p_i p_j \big) \Big) \\
&= \min\big\{1, \mu^{-1}\big\} \bigg(2 \sum_{i=1}^n p_i^2 + \sum_{i=1}^n \sum_{j\neq i} \mathbb{P}(X_j = 1, X_i = 1) - \sum_{i,j=1}^n p_i p_j \bigg) \\
&= \min\big\{1, \mu^{-1}\big\} \Big(2 \sum_{i=1}^n p_i^2 + \mathbb{E}[W^2] - \mathbb{E}[W] - \mathbb{E}[W]^2 \Big),
\end{aligned}$$

and rearranging gives the theorem. □

As a check, if W is the number of isolated vertices in $G \sim \mathrm{ER}(n,p)$, then $\mathbb{E}[W] = n(1-p)^{n-1}$,

$$\begin{aligned}\mathrm{Var}(W) - \mathbb{E}[W] &= n(n-1)(1-p)^{2n-3} - \mathbb{E}[W]^2 \\ &= n^2(1-p)^{2n-3}p - n(1-p)^{2n-3},\end{aligned}$$

and

$$2\sum_{i=1}^{n} p_i^2 = 2n(1-p)^{2n-2}.$$

Some straightforward manipulations show that the bound of Theorem 26.2 equals the intermediate bound (25.4) of Proposition 25.4 from Lecture 25 previously derived.

What is useful and remarkable about the theorem is that it only requires that we compute the mean and the variance, rather than having to understand how the coupling works in detail. Also note again the advantage over the method of moments: less computation and bounds in a strong metric. We next provide a more interesting application of the theorem.

Subgraph counts. Let H be a fixed subgraph with no isolated vertices (e.g., a triangle) and define

$$r_H := \min\Big\{\frac{v_{H'}}{e_{H'}} : H' \text{ is a subgraph of } H\Big\},$$

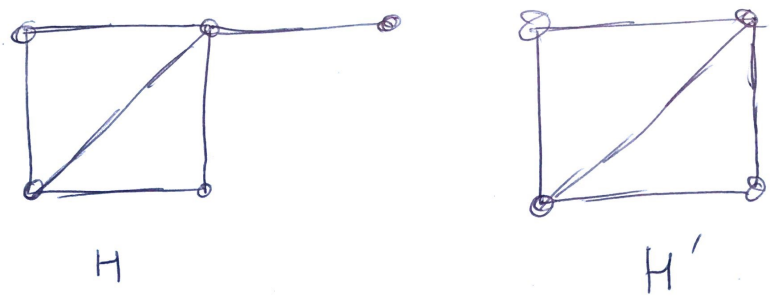
where generically $v_{H'}$ and $e_{H'}$ are the number of vertices and edges of H'. We showed in Lecture 3 that if W_n is the number of copies of H (meaning the number of subgraphs of the complete graph on n vertices which are isomorphic to H) that appear in G_n, then

$$W_n \xrightarrow[n\to\infty]{prob} \begin{cases}\infty, & p_n \gg n^{-r_H},\\ 0, & p_n \ll n^{-r_H}.\end{cases}$$

The proof used the first moment method and the Paley-Zygmund inequality. As usual, we may suspect that around the threshold of $p_n \approx n^{-r_H}$, that $\mathscr{L}(W_n)$ will converge to a Poisson distribution. However, we already saw that if $v_H/e_H > r_H$, then for $p_n = n^{-r_H}$,

$$\mathbb{E}[W_n] = \Theta(n^{v_H}p^{e_H}) = \Theta(n^{v_H - r_H e_H}) \to \infty,$$

because there is a strict subgraph $H' \subsetneq H$ such that $v_{H'}/e_{H'} = r_H$ and for each appearance of H' in G, there are many copies of H. So at the threshold there ought to be a finite but random number of copies of such an H', and then for each of these, a number of copies of H tending to infinity. See the picture and caption for an example.



For this H, $v_H/e_H = 5/6$, but $r_H = 4/5 < 5/6$, with the minimum uniquely achieved by H'. Thus when $p = n^{-r_H} = n^{-4/5}$, we expect a random but finite (actually Poisson as we will see) number of copies of H', and for each copy, we should see order $np = n^{1/5}$ copies of H.

As a result, we should at most expect Poisson convergence for H such that $r_H = v_H/e_H$. To further explore what is happening, we use Stein's method, and in particular Theorem 26.2 to obtain the following.

Theorem 26.3. *Fix n and p, let H be a graph with no isolated vertices and e_H edges, let W be the number of copies H in $G \sim \mathrm{ER}(n,p)$, and write $\mu = \mathbb{E}[W]$. Then*

$$d_{\mathrm{TV}}\big(\mathscr{L}(W), \mathrm{Po}(\mu)\big) \leq \min\{1, \mu^{-1}\}\left(\mathrm{Var}(W) - \mu + 2\mu p^{e_H}\right).$$

Proof. Let Γ be the set of all copies of H in K_n, the complete graph on n vertices and for $\alpha \in \Gamma$, let X_α be the indicator that there is a copy of H in G at α, and set $W = \sum_{\alpha\in\Gamma} X_\alpha$. We show that Theorem 26.2 applies to W by constructing a monotonic coupling. To this end, for a fixed $\alpha \in \Gamma$, let $X_\beta^{(\alpha)}$ be the indicator that there is a copy of H in $G \cup \{\alpha\}$ at β. Here, $G \cup \{\alpha\}$ means we add the minimum edges necessary to G to have a copy of H at α. The following three evident facts combined with Theorem 26.2 now imply the result:

1. $(X_\beta^{(\alpha)})_{\beta\neq\alpha}$ has the distribution of $(X_\beta)_{\beta\neq\alpha}$ given that $X_\alpha = 1$.
2. For all $\beta \in \Gamma/\{\alpha\}$, $X_\beta^{(\alpha)} \geq X_\beta$.
3. $\mathbb{E}[X_\alpha] = p^{e_H}$. □

Theorem 26.3 is a very general result, but it doesn't address the discussion above. What properties of a subgraph H make W approximately Poisson? We will answer this question in the next lecture by expressing the mean and variance in terms of properties of H.