Random chromatic numbers, some statistical folklore, and some puzzling inequalities†

David Pollard

Yale University

Applied Probability Day‡ *in honor of Chris C. Heyde*
28 June 2008

† Hartigan, Pollard, and Tatikonda (2008)
http://www.stat.yale.edu/˜pollard/Papers/chromatic.*

‡ http://www.cap.columbia.edu/

Achlioptis and Naor (2005)

- random (multi)graph $\approx$ Erdös-Rényi

  (i) vertex set $[n] := \{1, 2, \ldots, n\}$, with $n \to \infty$

  (ii) $m = cn$ edges $\{v_1, v_1'\}, \ldots, \{v_m, v_m'\}$
  with $v_1, v_1', \cdots \sim$ iid uniform on [n]

- with prob $\to 1$, chromatic number concentrates on $\{k_c, k_c + 1\}$ where k_c is smallest integer such that

$$c < k_c \log k_c$$

- hard part: if $c < (k-1)\log(k-1)$ show

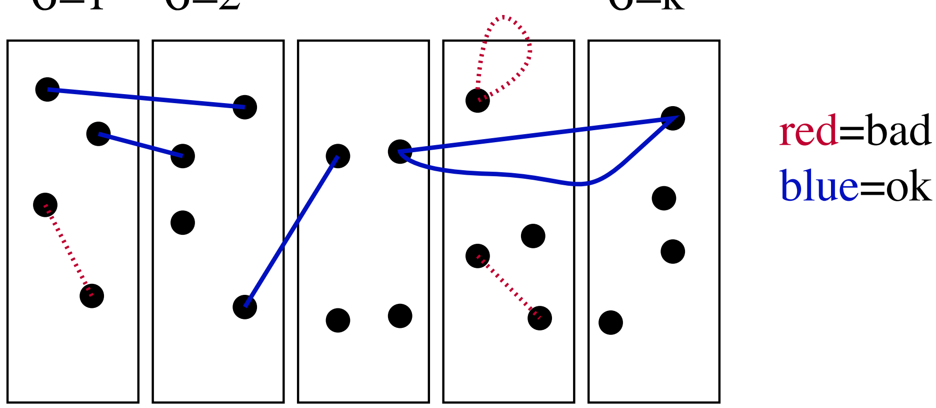
$$\liminf_{n\to\infty} \mathbb{P}\{\text{graph is } k\text{-colorable}\} > 0$$

- A&N used "second moment method" and threshold property of distribution of chromatic number

- $\mathcal{S}$ = all (balanced) 'k-partitions' σ of [n]; that is, $\sigma : [n] \to [k]$ and

  $\#\{v \in [n] : \sigma(v) = i\} = B := n/k$ for each $i \in [k]$.

- $S_\sigma = \{$no edges between vertices with same $\sigma\}$



- $Z := \sum_{\sigma \in \mathcal{S}} S_\sigma$ = number of balanced k-colorings

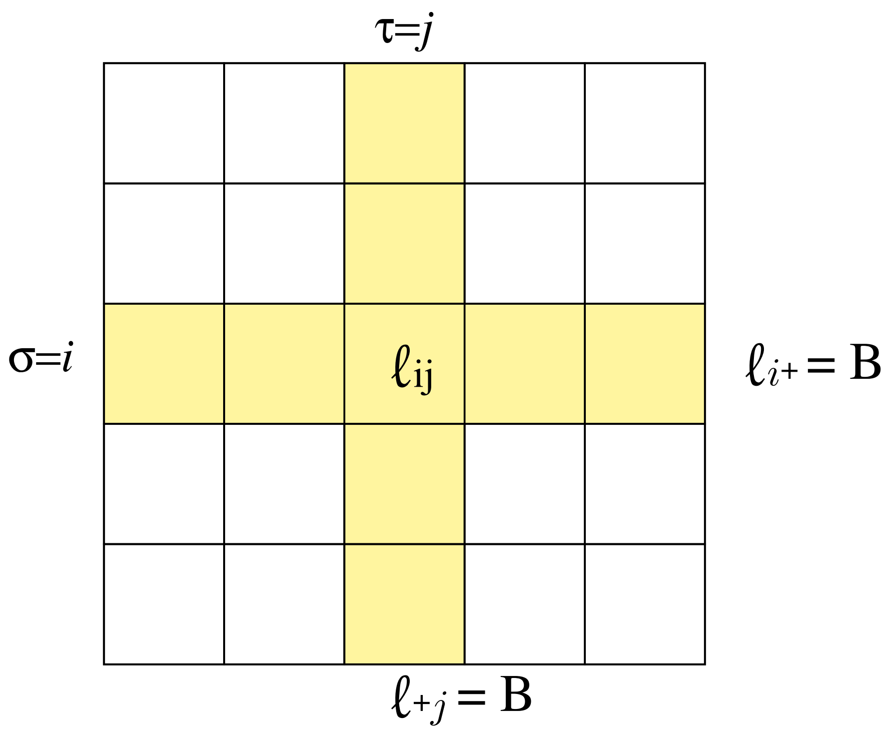
  $$\mathbb{P}\{\text{graph is } k\text{-colorable}\} \geq \mathbb{P}\{Z \geq 1\} \geq \frac{(\mathbb{P}Z)^2}{\mathbb{P}Z^2}$$

  (by Cauchy-Schwarz)

- $$\mathbb{P}Z = \frac{n!}{(B!)^k}\left(1 - \frac{1}{k}\right)^m$$

- $\mathbb{P}Z^2 = \sum_{\sigma,\tau\in\mathcal{S}} \mathbb{P}S_\sigma S_\tau$
- For $\sigma, \tau \in \mathcal{S}$, with $\ell_{ij} = \#\{v : \sigma(v) = i, \tau(v) = j\}$,

$$\mathbb{P}S_\sigma S_\tau = \left(1 - \frac{2}{k} + \sum_{i,j} \left(\frac{\ell_{ij}}{n}\right)^2\right)^m$$



- cf. $\mathcal{H} = k \times k$ matrices with nonnegative integer entries (the ℓ_{ij}), each row and column sum $= B = n/k$.

- Need to show $A_n(c) := \mathbb{P}Z^2/(\mathbb{P}Z)^2$ stays bounded as $n \to \infty$ if $c < (k-1)\log(k-1)$

$$(\mathbb{P}Z)^2 \sim \frac{k^{2n}}{n^{k-1}}\left(1-\frac{1}{k}\right)^{2nc} \qquad \text{by Stirling}$$

- Collect σ, τ pairs with same ℓ_{ij}'s:

$$\mathbb{P}Z^2 = \sum_{\ell \in \mathcal{H}} \frac{n!}{\prod_{i,j} \ell_{ij}!}\left(1-\frac{2}{k}+\sum_{i,j}\left(\frac{\ell_{ij}}{n}\right)^2\right)^{nc}$$

$\mathcal{H} = k \times k$ matrices with nonnegative integer entries (the ℓ_{ij}), each row and column sum $= B = n/k$.

- The hard work begins.

☺ FOLKLORE cf. Haberman (1974)

- Compare with $k \times k$ table of independent Poisson(n/k^2) counts, $Y = \{Y_{ij} : 1 \le i, j, \le k\}$. For $\ell \in \mathcal{H}$,

$$p(\ell) := \mathbb{P}\{Y = \ell\} = \frac{e^{-n}(n/k^2)^n}{\prod_{i,j} \ell_{ij}!} = \frac{n!}{\prod_{i,j} \ell_{ij}!} \frac{n^n e^{-n}}{n!} k^{-2n}$$

☺ $\mathbb{P}\{Y \in \mathcal{H}_k\} \sim n^{-(2k-1)/2}$. Thus

$$p_2(\ell) := \mathbb{P}\{Y = \ell \mid Y \in \mathcal{H}_k\} \sim n^{k-1} k^{-2n} \frac{n!}{\prod_{i,j} \ell_{ij}!}$$

- goodness-of-fit: $X_n^2 = \sum_{ij}(Y_{ij} - n/k^2)^2/(n/k^2)$

☺ Under $\mathbb{P}_2$ conditional distribution, X_n^2 approximately $\chi^2_{k^2-(2k-1)}$ distributed

- Several substitutions later: $A_n(c)$ is bounded by a constant times

$$\left(1 - \frac{1}{k}\right)^{-2nc} \mathbb{P}_2 \left(\left(1 - \frac{1}{k}\right)^2 + \frac{X_n^2}{nk^2} \right)^{nc}$$

$$\le \mathbb{P}_2 \exp\left(\frac{cX_n^2}{(k-1)^2}\right)$$

NOTE THE PATTERN:

$$(\mathbb{P}Z)^2 \sim \frac{k^{2n}}{n^{k-1}}\left(1-\frac{1}{k}\right)^{2nc}$$

$$\mathbb{P}Z^2 = \sum_{\ell\in\mathcal{H}} \frac{n!}{\prod_{i,j}\ell_{ij}!}\left(1-\frac{2}{k}+\sum\nolimits_{i,j}\left(\frac{\ell_{ij}}{n}\right)^2\right)^{nc}$$

$$p_2(\ell) \sim n^{k-1}k^{-2n}\frac{n!}{\prod_{i,j}\ell_{ij}!} \qquad \text{[defines } \mathbb{P}_2\text{]}$$

$$\frac{\mathbb{P}Z^2}{(\mathbb{P}Z)^2} \leq \text{const } \mathbb{P}_2 \exp\left(\frac{cX_n^2}{(k-1)^2}\right)$$
$$\stackrel{?}{=} O(1) \qquad \text{if } c < (k-1)\log(k-1)?$$

- Under $\mathbb{P}_2$ conditional distribution, X_n^2 approximately $\chi^2_{k^2-(2k-1)}$ distributed.
- Actually (Hartigan, Pollard, and Tatikonda 2008), a sharpened χ^2-approximation (cf. Haberman 1974) takes care of the contributions from $X_n^2 \leq \text{tiny} \cdot n$

CONTRIBUTIONS FROM $X_n^2 >$ tiny $\cdot n$

- for $-1 \le t$,

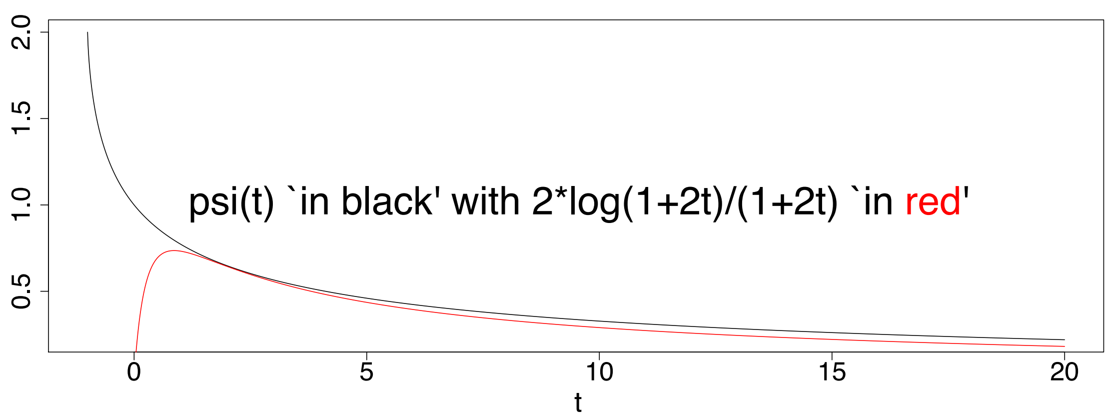
$$h(t) := (1+t)\log(1+t) - t = \frac{1}{2}t^2\psi(t)$$

- Control tail contribution by easy Poisson facts and: *If* $u = (u_1, \ldots, u_k) \in [-1, k-1]^k$ *and* $\sum_j u_j = 0$ *then*

$$\sum_j h(u_j) \ge \frac{\log(k-1)}{(k-1)} \sum_j u_j^2$$

- To reduce to a simpler case we need the inequality

$$\psi(t) \ge 2\log(1+2t)/(1+2t) \qquad \text{for all } t \ge 0$$



EASY POISSON FACTS

- for $-1 \le t$,

$$h(t) := (1+t)\log(1+t) - t = \frac{1}{2}t^2\psi(t)$$

- $\mathbb{N}_0 =$ the set of all nonnegative integers.

LEMMA: Suppose $W \sim \text{Poisson}(\lambda)$ with $\lambda \ge 1$.

(i) for all $\ell = \lambda + \lambda u \in \mathbb{N}_0$,

$$\begin{aligned}\sqrt{2\pi\lambda}\,&\mathbb{P}\{W = \ell\} \\ &= \exp\left(-\lambda h(u) - \tfrac{1}{2}\log(1+u) + O(1/\ell)\right) \\ &= \exp\left(-\tfrac{1}{2}\lambda u^2 + O\left(|u| + \lambda|u|^3\right)\right).\end{aligned}$$

(ii) for all $\ell = \lambda(1+u) \in \mathbb{N}_0$,

$$\mathbb{P}\{W = \ell\} \le \exp(-\lambda h(u))$$

(iii) For all $w \ge 0$,

$$\begin{aligned}\mathbb{P}\{|W - \lambda| \ge \lambda w\} &\le 2\exp(-\lambda h(w)) \\ &= 2\exp\left(-\tfrac{1}{2}\lambda w^2 + O\left(\lambda|w|^3\right)\right)\end{aligned}$$

References

Achlioptis, D. and A. Naor (2005). The two possible values of the chromatic number of a random graph. *Annals of Mathematics 162*, 1335–1351.

Haberman, S. J. (1974). *The Analysis of Frequency Data*. University of Chicago Press.

Hartigan, J., D. Pollard, and S. Tatikonda (2008). Conditioned Poisson distributions and the concentration of chromatic numbers. Technical report, Yale University. (Available at http://www.stat.yale.edu/˜pollard/Papers/).