More proof Sketches Regarding the Closure of Limiting Probabilities in Sparse Random Hyper-graphs

Alberto Larrauri

February 4, 2020

1 Preliminaries

$G^d(N,p)$ denotes the binomial model of random d-uniform hyper-graphs. Consider $d \geq 2$ fixed for the rest of this writing.

Theorem 1.1. *Let $c < (d-2)!$. Then, a.a.s all the components of $G^d(N,p(N))$, where $p(N) \sim c/N^{d-1}$, are either trees or unicycles.*

Proof. See [1] for $d=2$ and [2] for the general case.

Está todo hecho en el modelo uniforme, pero la transferencia uniforme-> binomial es "sencilla"

□

We will call the **fragment** F_N to the union of the unicyclic components in $G^d(N,p(N))$.

Theorem 1.2. *Let $c < (d-2)!$ and let $p(N) \sim c/N^{d-1}$. Let δ_N be the random variable that counts the vertices in F_N. Then there is some constant C such that $\mathrm{E}[\delta_n]$ is smaller than C for any N.*

Este teorema lo he visto para grafos en el artículo de Erdos de 1960 , y más completo en el libro de Karonski, por ejemplo. Para hypergrafos no lo he encontrado, pero igual es cosa de buscar mejor. Lo demuestro de todas formas.

Proof. Let $C_d(m)$ denote the number of connected labeled d-uniform hypergraphs with m edges and $n := (d-1)m$ vertices. It is satisfied that for big values of m:

$$C_d(m) \leq \frac{n^n}{e^{(m-n)}} \frac{1}{(d-2)!^m},$$

where the constant hidden by the O-notation depends only on d. Let $X_{N,m}$ be the random variable that counts the number of unicyclic components in $G^d(N, c/N^{d-1})$ with exactly m edges. Then we have $\delta_N = \sum_{m=2}^{N/(d-1)} (d-1)m \cdot X_{N,m}$. Also,

$$\begin{aligned}
\mathrm{E}[X_{N,m}] &\leq \binom{N}{n} p(N)^m (1-p(N))^{\binom{N}{d}-\binom{N-n}{d}-m} \frac{n^n}{e^{(n-m)}} \frac{1}{(d-2)!^m} \\
&\leq \frac{N^n e^n}{n^n} e^{-\frac{n(n-1)}{2N}} p(N)^m e^{-p(N)\left[\binom{N}{d}-\binom{N-n}{d}-m\right]} \frac{n^n}{e^{(n-m)}} \frac{1}{(d-2)!^m} \\
&= \left(\frac{c}{(d-2)!}\right)^m e^{n-\frac{n(n-1)}{2N}-p(N)\left[\binom{N}{d}-\binom{N-n}{d}-m\right]+(m-n)} \qquad (1)
\end{aligned}$$

Operating on the exponent of e in the last term:

$$\begin{aligned}
&n - \frac{n(n-1)}{2N} - p(N)\left[\binom{N}{d} - \binom{N-n}{d} - m\right] + (m-n) \leq \\
&-\frac{n(n-1)}{2N} - p(N)\left[\binom{N}{d} - \binom{N-n}{d}\right] + p(N)m + m \leq \\
&m\Big(-\frac{(d-1)}{2} - \frac{c}{(d-2)!} + p(N) + 1\Big) \leq \\
&m\Big(1 - \frac{c}{(d-2)!}\Big).
\end{aligned}$$

Thus, substituting in eq. (1) we obtain

$$\mathrm{E}[X_{N,m}] \leq \left(\frac{c}{(d-2)!} e^{1-\frac{c}{(d-2)!}}\right)^m \tag{2}$$

One can easily check that xe^{1-x} grows monotonously from 0 to 1 as x goes from 0 to 1. Thus, $\frac{c}{(d-2)!}e^{1-\frac{c}{(d-2)!}} < 1$.

By definition it is satisfied

$$\mathrm{E}[\delta_N] = \sum_{m=2}^{N/(d-1)} m\mathrm{E}[X_{N,m}],$$

except for the case $d = 2$, where the sum starts at $m = 3$ instead of $m = 2$. This is because there are no connected graphs with 2 vertices and 2 edges. Because of eq. (2), for sufficiently large values of m the terms $m\mathrm{E}[X_{N,m}]$ are bounded by $m(\frac{c}{(d-2)!}e^{1-\frac{c}{(d-2)!}})^m$ uniformly for all values of N. Otherwise, for small values of m, the $m\mathrm{E}[X_{N,m}]$ are bounded as well because

$$\lim_{N\to\infty} m\mathrm{E}[X_{N,m}] = \left(ce^{\frac{c}{(d-2)!}}\right)^m \frac{C_d(m)}{(m\cdot(d-1))!}$$

In consequence, using the dominated converge theorem we can conclude that

$$\lim_{N\to\infty} \mathrm{E}[\delta_N] = \sum_{m=2}^{N/(d-1)} m\mathrm{E}[X_{N,m}]$$

exists and is not infinite. □

Theorem 1.3. *Let $c < (d-2)!$, and let $p(N) \sim c/N^{d-1}$. For any $k \geq 2$, let $\gamma_{N,\geq k}$ be the random variable that counts how many cycles with at least k edges lie in $G^d(N, p(N))$. Then*

$$\lim_{N\to\infty} \mathrm{E}[\gamma_{N,\geq k}] = \sum_{l=k}^{\infty} \left(\frac{c}{(d-2)!}\right)^l \frac{1}{2l}$$

In particular, if γ_N is the random variable that counts the cycles in $G^d(N, p(N))$ then, if $d > 2$:

$$\lim_{N\to\infty} \mathrm{E}[\gamma_N] = \begin{cases} \frac{c}{2(d-2)!} + \ln\left(1 - \frac{c}{2(d-2)!}\right), & \text{if } d > 2 \\ \frac{c}{2} + \frac{c^2}{4}\ln\left(1 - \frac{c}{2}\right), & \text{if } d = 2. \end{cases}$$

Esto está hecho en el artículo de Erdos para grafos, y para hypergrafos no lo he visto.

Proof. For $k \geq 2$ ($k \geq 3$ if $d = 2$), let $\gamma_{N,k}$ be the random variable that counts the k-cycles that lie in $G^d(N, p(N))$. A simple computation yields $\mathrm{E}[\gamma_{N,k}] = \frac{(N)_{k(d-1)}}{2k}\left(\frac{c}{N^{d-1}}\right)^k$, and $\lim_{N\to\infty} \mathrm{E}[\gamma_{N,k}] = \frac{c^k}{2k}$.

Not only this, but also $\mathrm{E}[\gamma_{N,k}] \leq \frac{c^k}{2k}$ for all N. In consequence, applying the dominated convergence theorem we obtain

$$\lim_{N\to\infty} \sum_{l=k}^{\infty} \mathrm{E}[\gamma_{N,l}] = \sum_{l=k}^{\infty} \frac{c^l}{2l}.$$

Using that $\gamma_{N,\geq k}$ is the sum of all the $\gamma_{N,l}$ for $l \geq k$ and the Taylor expansion of $ln(1-x)$ yields the desired results. $\square$

Given $0 < c < (d-2)!$ we define the function $B_d(c)$ as the limit as N tends to infinity of the expected number of cycles in $G^d(N, c/N^{d-1})$. Because of our last theorem

$$B_d(c) = \begin{cases} \frac{c}{2(d-2)!} + \ln\left(1 - \frac{c}{2(d-2)!}\right), & \text{if } d > 2 \\ \frac{c}{2} + \frac{c^2}{4} \ln\left(1 - \frac{c}{2}\right), & \text{if } d = 2. \end{cases}$$

Corollary 1.1. *Let $c < (d-2)!$ and $p(N) \sim c/N^{d-1}$. For each $k \in \mathbb{N}$ let $Z_{N,k}$ be the random variable that counts the unicyclic components in $G^d(N, p(N))$ with exactly k edges. Then*

$$\sum_{i=1}^{\infty} \lim_{N\to\infty} \mathrm{E}\big[Z_{N,i}\big] = \lim_{N\to\infty} \mathrm{E}\big[\sum_{i=1}^{\infty} Z_{N,i}\big] = B_d(c).$$

A simple application of the factorial moments method proves the following theorem:

Theorem 1.4. *Let $k_1, \dots, k_j \geq 2$ (≥ 3 if $d = 2$). Then, as N tends to infinity the γ_{N,l_i}'s converge in distribution to independent Poisson variables with mean values $\lambda_i := \frac{c^{l_i}}{2l_i}$ respectively. That is, for any $a_1, \dots, a_j \in \mathbb{N}$,*

$$\lim_{N\to\infty} \Pr\Big(\bigwedge_{i=1}^{r} \gamma_{N,l_i} = a_i\Big) = \prod_{i=1}^{j} e^{-\lambda_i} \frac{\lambda_i^{a_i}}{a_i!}$$

Esto está hecho en el libro de Bollobás de random graphs. Para hypergrafos no lo he encontrado, pero realmente es lo mismo. Si queréis lo hago.

Let H be an unicycle and let $m = |E(H)|$, $n = |V(H)| = m(d-1)$. Let $X_{N,H}$ be the random variable that counts the number of connected components in $G^d(N, p(N))$ isomorphic to H. Then

$$\mathrm{E}[X_{N,H}] = \frac{(N)_n}{|Aut(H)|} p(N)^m (1 - p(N))^{\binom{N}{d} - \binom{N-n}{d} - m}.$$

And if $p(N) \sim c/N$,

$$\lim_{N\to\infty} \mathrm{E}[X_{N,H}] = \frac{c^m}{|Aut(H)|} e^{\lim_{N\to\infty} c \frac{\binom{N}{d} - \binom{N-n}{d} - m}{N^{d-1}}} = \frac{c^m}{|Aut(H)|} e^{c \frac{m}{(d-2)!}}.$$

For convenience's sake we will often use the auxiliary variable $s = \frac{c}{(d-2)!} e^{\frac{c}{(d-2)!}}$. We can rewrite last limit in terms of s as:

$$\lim_{N\to\infty} \mathrm{E}[X_{N,H}] = (s)^m \frac{(d-2)!^m}{|Aut(H)|}.$$

Another application of the factorial moments method proves the next theorem:

Theorem 1.5. *Let $H_1, \dots, H_j$ be unicycles, and let $P(N) \sim c/N$. Then, as N tends to infinity, the X_{N,H_i}'s converge in distribution to independent Poisson variables with means $\lambda_i = s^{|E(H_i)|} \frac{(d-2)!^{|E(H_i)|}}{|Aut(H_i)}$ respectively. That is, for any $a_1, \dots, a_j \in \mathbb{N}$,*

$$\lim_{N\to\infty} \Pr\Big(\bigwedge_{i=1}^{r} X_{N,H_i} = a_i\Big) = \prod_{i=1}^{j} e^{-\lambda_i} \frac{\lambda_i^{a_i}}{a_i!}$$

Next we show some stuff

Theorem 1.6. *Let $0 < c < (d-2)!$ and $p(N) \sim c/N$. For $i \in \mathbb{N}$ let $\lambda_i = \left(\frac{c}{(d-2)!}\right)^i \frac{1}{2i!}$. Let A_N be the event that $G^d(n, p(N))$ contains no cycles. Define $F_d := F_d(c)$ as*

$$F_d(c) = \begin{cases} e^{\sum_{i=2}^{\infty} \lambda_i} = e^{\frac{c}{2(d-2)!}} \sqrt{1 - \frac{c}{(d-2)!}} & \text{if } d > 2 \\ e^{\sum_{i=3}^{\infty} \lambda_i} = e^{\frac{c}{2} + \frac{c^2}{4}} \sqrt{1-c} & \text{if } d = 2. \end{cases}$$

Then it is satisfied

$$\lim_{N\to\infty} \Pr\big(A_N\big) = F_d$$

Esto está hecho para grafos en [1].

Proof. For each $k \in \mathbb{N}$ define $F_{d,k} := F_{d,k}(c)$ as

$$F_{d,k}(c) = \begin{cases} e^{\sum_{i=2}^{k} -\lambda_i} & \text{if } d > 2 \\ e^{\sum_{i=3}^{k} -\lambda_i} & \text{if } d = 3. \end{cases}$$

A simple computation using the Taylor expansion of $\ln(1-x)$ shows that $\lim_{k\to\infty} F_{d,k} = F_d$. Fix an arbitrary $\epsilon > 0$. We show that there exists a constant k satisfying

$$\left| \lim_{N\to\infty} \Pr\big(A_N\big) - F_{d,j} \right| \leq \epsilon \text{ for any } j \geq k.$$

For each $k \in N$ let $A_{N,k}$ be the event that $G^d(N, p(N))$ contains no cycle with length at most k. Using theorem 1.4, we obtain

$$\lim_{N\to\infty} \Pr\big(A_{N,k}\big) = F_{d,k}.$$

For any $k \in \mathbb{N}$ let $\gamma_{N,\geq k}$ be the random variable that counts the number of cycles in $G^d(N, p(N))$ with length at least k. Then, using theorem 1.3 we obtain that $\lim_{k\to\infty} \lim_{N\to\infty} \mathrm{E}[\gamma_{N,\geq k}] = 0$.
Fix k such that for any $j \geq k$

$$\left| \lim_{N\to\infty} \mathrm{E}[\gamma_{N,\geq k}] \right| \leq \epsilon.$$

Notice that for any j, "A_N is true and $A_{N,j}$ is false" if and only if $\Pr\big(\gamma_{N,\geq j} \geq 1\big)$. Also, by Markov inequality, $\Pr\big(\gamma_{N,\geq j} \geq 1\big) \leq \mathrm{E}\big[\gamma_{N,\geq j}\big]$. In consequence, for any $j \geq k$

$$\lim_{N\to\infty} \left| \Pr\big(A_N\big) - \Pr\big(A_{N,j}\big) \right| \leq \epsilon.$$

In consequence, for any $j \geq k$

$$\left| \lim_{N\to\infty} \Pr\big(A_N\big) - F_{d,j} \right| \leq \left| \lim_{N\to\infty} \Pr\big(A_{N,j}\big) - F_{d,j} \right| + \epsilon = \epsilon,$$

and we are finished. $\square$

Theorem 1.7. *Let $0 < c < (d-2)!$, and $p(N) \sim c/N^{d-1}$. Let H be an hypergraph whose components are all unicycles, and let $F_d := F_d(c)$ be as in last theorem. Then*

$$\lim_{N\to\infty} \Pr\Big(F_N \simeq H\Big) = F_d(c) \left(\frac{c}{(d-2)!} e^{\frac{c}{(d-2)!}} \right)^{e(H)} \frac{(d-2)!^{e(H)}}{|Aut(H)|}$$

Proof. Fix such H. Let $U_1, U_2, \ldots, U_i, \ldots$ be an enumeration of all unicycles in a way such that $e(U_i) \leq e(U_j)$ if $i \leq j$. For each $i \in \mathbb{N}$ let a_i be the number of connected components of H that are isomorphic to U_i, and let $X_{N,i}$ be the random variable that counts the number of connected components in $G^d(N, P(N))$ that are isomorphic to U_i. Clearly, $F_N \simeq H$ if and only if $X_{N,i} = a_i$ for all i. Thus,

$$\lim_{N\to\infty} \Pr\big(F_N \simeq H\big) = \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{\infty} X_{N,i} = a_i\big).$$

First, we are going to show that

$$\lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{\infty} X_{N,i} = a_i\big) = \lim_{j\to\infty} \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big).$$

Fix $\epsilon > 0$ an arbitrarily small real constant. We need to prove that there exists some $j_0 \in \mathbb{N}$ satisfying

$$\left| \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big) - \lim_{j\to\infty} \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big) \right| \leq \epsilon \text{ for all } j \geq j_0.$$

For each $k \in \mathbb{N}$ let $Y_{N,k}$ be the random variable that counts the uni-cyclic connected components of $G^d(N, p(N))$ with exactly k edges. By theorem 1.2 we have that for some $k_0 \in \mathbb{N}$

$$\lim_{N\to\infty} \sum_{l=k_0}^{\infty} \mathrm{E}\big[Y_{N,l}\big] \leq \epsilon$$

Let k_1 be the maximum number of edges in a connected component of H, and let $k = \max(k_0, k_1 + 1)$. Finally, fix j_0 such that $e(U_j) > k_1$ for any $j \geq j_0$.

Given any $j \geq j_0$, $F_N \simeq H$ if and only if

$$\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big) \wedge \big(\sum_{l=k}^{\infty} Y_{N,l} = 0\big)$$

Using Markov inequality we get

$$\lim_{N\to\infty} \Pr\big(\sum_{l=k}^{\infty} Y_{N,l} \geq 1\big) \leq \epsilon$$

Thus,

$$\left| \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big) - \lim_{j\to\infty} \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big) \right| \leq \epsilon,$$

as we wanted to prove.

Using theorem 1.5 we get

$$\lim_{j\to\infty} \lim_{N\to\infty} \Pr\big(\bigwedge_{i=1}^{j} X_{N,i} = a_i\big) = \prod_{i=1}^{\infty} e^{-\lambda_i} \frac{\lambda_i^{a_i}}{a_i!},$$

where $\lambda_i = s^{e(U_i)} \frac{(d-2)!^{e(U_i)}}{|Aut(U_i)|} = \lim_{N\to\infty} \mathrm{E}[X_{N,i}]$, and $s = s(c)$ is defined as in the statement of the theorem. Using corollary 1.1 we obtain

$$\sum_{i=1}^{\infty} \lambda_i = B_d(c),$$

and in consequence

$$\prod_{i=1}^{\infty} e^{-\lambda_i} = e^{-B_d(c)} = A_d(c).$$

The following identities hold:

$$\sum_{i=1}^{\infty} e(U_i)a_i = e(H) \qquad \prod_{i=1}^{\infty} |Aut(U_i)^{a_i} a_i! = |Aut(H)|.$$

As a consequence we get

$$\prod_{i=1}^{\infty} \frac{\lambda_i^{a_i}}{a_i!} = \prod_{i=1}^{\infty} (s(d-2)!)^{e(U_i)a_i} \frac{1}{|aut(U_i)|^{a_i} a_i!} = s^{e(H)} \frac{(d-2)!^{e(H)}}{|Aut(H)|},$$

and the theorem follows. □

Let us denote by $\mathcal{U}$ the class of hypergraphs whose connected components are unicycles. For the case $d = 2$ the asymptotic distribution of the fragment F_N coincides with the Boltzmann-Poisson distribution of random graphs from $\mathcal{U}$ described in [3].

Tobias de una forma un poco críptica da a entender que usar el teorema 1.3 de [3] sirve para algo. Yo no he sabido cómo o para qué.

For completeness sake we give an argument of why the asymptotic distribution of F_N is indeed a probability distribution.

Theorem 1.8. *Let $0 < c < (d-2)!$ and let $p(N) \sim c/N^{d-1}$. Let $H_1, \dots H_k, \dots$ be an enumeration of all hypergraphs in $\mathcal{U}$. For each $i \in \mathbb{N}$ denote by p_i the limit $\lim_{N\to\infty} \Pr(F_N \simeq H_i)$. Then*

$$\sum_{i=1}^{\infty} p_i = 1.$$

Proof. Let $\epsilon > 0$ be an arbitrarily small real constant. We show that there exists some $j_0 \in \mathbb{N}$ such that

$$1 - \sum_{i=1}^{j} p_i \leq \epsilon \qquad \text{for all } j \geq j_0.$$

Let $m = \lim_{N\to\infty} \mathrm{E}[e(F_N)]$. Notice that m exists by theorem 1.2. Define $M = m/\epsilon$. Then

$$\lim_{N\to\infty} \Pr\big(e(F_N) \geq M\big) \leq \epsilon.$$

Let j_0 be such that $e(H_i) \geq M$ for all $i \geq j_0$. Then, given any $j \geq j_0$,

$$1 - \sum_{i=1}^{j} p_i = \lim_{N\to\infty} \Pr\Big(\bigwedge_{i=1}^{j} F_N \not\simeq H_i\Big) \leq \lim_{N\to\infty} \Pr\big(e(F_N) \geq M\big) \leq \epsilon,$$

and the result follows. □

Theorem 1.9. *Let $0 < c < (d-2)!$, and $p(N) \sim c/N^{d-1}$. Let $H_1, \dots, H_i, \dots$ be an enumeration of all hypergraphs in $\mathcal{U}$ and for each $i \in \mathbb{N}$ let $p_i = \lim_{N\to\infty} \Pr(F_N \simeq H_i)$. Consider the sets*

$$L_c := \{ \lim_{N\to\infty} \Pr(G^d(N, p(N)) \textit{ satisfies } P) \,|\, P \textit{ FO property}\},$$

and

$$S_c := \{ \sum_{H_i \in T} p_i \,|\, T \subset \mathcal{U} \}.$$

Then it is satisfied that $\overline{L_c} = S_c$.

Proof. We will prove the statement by showing both $S_c \subseteq \overline{L_c}$ and $\overline{L_c} \subseteq S_c$. We begin with $S_c \subseteq \overline{L_c}$. Let $T \subset \mathcal{U}$, and let $\epsilon > 0$ be an arbitrarily small real number. We show that there exists a first order property P such that

$$\Big|\lim_{N\to\infty} \Pr\big(G^d(N, p(N)) \text{ satisfies } P\big) - \sum_{H_i \in T} p_i\Big| \leq \epsilon$$

. For each $k \in \mathbb{N}$ let $F_{N,k}$ be the union of connected components in $G^d(N, p(N))$ that are unicycles with no more than k edges, and let $\mathcal{U}_k$ be the class of hypergraphs whose connected components are unicycles with at most k edges. Given any $H \in \mathcal{U}_{\|}$, the property $F_{N,k} \simeq H$ is expressible in FO logic. $\square$

References

[1] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[2] Michał Karoński and Tomasz Łuczak. The phase transition in a random hypergraph. *Journal of Computational and Applied Mathematics*, 142(1):125–135, 2002.

[3] Colin Mcdiarmid. Random graphs from a minor-closed class. *Combinatorics, Probability and Computing*, 18(4):583–599, 2009.