Lecture 3

Random graphs I

The material in this section is mainly based on [Bol85], [Wor99].

3.1 Basic definitions

3.1.1 Graphs

Let us first fix our definition of what a graph is. There are multiple definitions available to capture the intuitive idea that a graph is a set of vertices and a set of edges between these vertices. We will want to allow multiple edges between pairs of vertices and loops, so we choose the following definition, in which we write $|X|$ for the cardinality of a set X.

Definition 3.1. A graph is a triple $G = (V, E, \mathcal{I})$ where V is a set, called the set of *vertices* of G, E is a set, called the set of *edges* of G and

$$\mathcal{I} \subset E \times V$$

is called the *incidence relation* of G and satisfies the condition that for all $e \in E$ we have

$$|\{v \in V;\ (e, v) \in \mathcal{I}\}| \in \{1, 2\}.$$

An edge that is incident to a single vertex is called a *loop*. A graph without loops in which every pair of vertices has at most one edge incident to it is called *simple*. If $v, w \in V$ and there exists and $e \in E$ so that both $(e, v) \in \mathcal{I}$ and $(e, w) \in \mathcal{I}$ we say that v and w are *adjacent* or that v and w *share an edge*.

The *degree* or *valence* of a vertex $v \in V$ is given by

$$\deg(v) = |\{e \in E;\ (e,v) \in \mathcal{I}\}| + \left| \left\{ e \in E;\ \begin{array}{c} (e,v) \in \mathcal{I} \text{ and} \\ (e,w) \notin \mathcal{I},\ \forall w \in V \text{ with } w \neq v \end{array} \right\} \right|.$$

An *isomorphism* of between graphs $G_1 = (V_1, E_1, \mathcal{I}_1)$ and $G_2 = (V_2, E_2, \mathcal{I}_1)$ is a pair of bijective maps $f_V : V_1 \to V_2,\ f_E : E_1 \to E_2$ such that

$$(e, v) \in \mathcal{I}_1 \ \Leftrightarrow\ (f_E(e), f_V(e)) \in \mathcal{I}_2.$$

An *automorphism* of a graph $G = (V, E, \mathcal{I})$ is an isomorphism between G and itself. The group formed by all automorphisms of G will be denotes $\mathrm{Aut}(G)$.

A *walk* between vertices $v, w \in V$ is sequence of vertices $(v_1, v_2, \ldots, v_r)$ with $v_1 = v$, $v_r = w$ and so that for all $i = 1, \ldots r-1$ the vertices v_i and v_{i+1} are adjacent. A *cycle* in G is a walk between v and itself for some vertex $v \in V$.

G is called *connected* is there exists a walk between every pair of vertices $v, w \in V$.

Some remarks:

- The condition on $|\{v \in V;\ (e, v) \in \mathcal{I}\}|$ guarantees that every edge connects to either one or two vertices.

- In the definition above, a loop at a vertex (an edge that connects to only that vertex) adds 2 to the degree of this vertex.

- Given a graph G, we will often write $V(G)$ and $E(G)$ for the sets of its vertices and edges respectively.

The above serves as a formal definition of what a graph is. It is however not always the easiest way to describe graphs. Often we will just think in terms of pictures. Let us give an example of a graph.

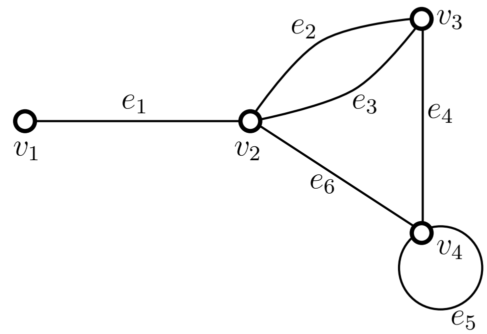

Figure 3.1: A graph.

The graph $G = (V, E, \mathcal{I})$ above is given by $V = \{v_1, v_2, v_3, v_4\}$, $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ and

$$\begin{aligned}\mathcal{I} &= \{(e_1, v_1),\ (e_1, v_2),\ (e_2, v_2),\ (e_2, v_3),\ (e_3, v_2),\ (e_3, v_3),\ (e_4, v_3),\ (e_4, v_4),\\ &\quad (e_5, v_4)\ (e_6, v_4),\ (e_6, v_2)\}\end{aligned}$$

3.1.2 Random graphs

There are multiple models of random graphs around. The most widely studied model is probably that of the Erdős-Rényi random graph: fix $p \in (0, 1)$, take n vertices and add each of the possible edges between these vertices to the graph with probability p and leave it out with probability $1 - p$. We will however be interested in regular graphs:

Definition 3.2. Let $k \in \mathbb{N}$. A graph $G = (V, E, \mathcal{I})$ is called k-regular if

$$\deg(v) = k$$

for all $v \in V$.

Our goal now is to pick a graph at random among all k-regular graphs on a given set of vertices. Of course, the set of k-regular graphs on n vertices up to isomorphism is a finite set for all $k, n \in \mathbb{N}$. So, we could just pick one at random from this finite set. This model however turns out to be hard to control in general. Instead we will study the *configuration model* for random regular graphs.

First let us fix once and for all disjoint sets $W_1(n), \dots, W_n(n)$ with

$$|W_i(n)| = k$$

for $i = 1, \dots, n$, for every $n, k \in \mathbb{N}_{\geq 1}$ so that $n \cdot k$ is even. Furthermore, we will write

$$W(n) = \bigsqcup_{i=1}^{n} W_i(n).$$

We can now define configurations:

Definition 3.3. Let $n, k \in \mathbb{N}$ so that $n \cdot k$ is even. Then, a k-regular *configuration* on n vertices is a set of pairs

$$C = \{\{a_i, b_i\} \subset W(n)\}_{i=1}^{n \cdot k/2}$$

so that

$$\bigcup_{i=1}^{n \cdot k/2} \{a_i, b_i\} = W(n).$$

We will write $\mathcal{G}_{n,k}$ for the (finite) set of k-regular configurations on n vertices.

Note that the last condition guarantees that every element of $W(n)$ appears exactly once in a pair of the configuration C.

Definition 3.4. Let $n, k \in \mathbb{N}$ so that $n \cdot k$ is even. Furthermore, let $C = \{\{a_i, b_i\} \subset W(n)\}_{i=1}^{n \cdot k/2}$ be a k-regular configuration on n vertices.

The graph $G(C) = (V, E, \mathcal{I})$ associated with C is given by

$$V = \{v_1, \dots, v_n\},\ E = \{e_1, \dots, e_{n \cdot k/2}\}$$

and

$$(e_i, v_j) \in \mathcal{I} \ \Leftrightarrow\ \{a_i, b_i\} \cap W_j(n) \neq \emptyset.$$

In other words, our finite sets $W_i(n)$ represent the vertices of $G(C)$ and we connect two of them if and only if two of their elements appear as a pair in the configuration. As such, we will often think of the elements in C as labels on half-edges of $G(C)$. Figure 3.2 gives an example:

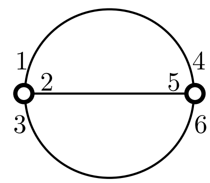

Figure 3.2: The graph $G(C)$ corresponding to the configuration $C = \{\{1,4\},\ \{2,5\},\ \{3,6\}\}$.

The number of configurations is easy to count. For $n \in \mathbb{N}$ even we write

$$n!! = (n-1)(n-3)\cdots 3 \cdot 1.$$

We have:

Lemma 3.5. *Let $n, k \in \mathbb{N}$ so that $n \cdot k$ is even. Then:*

$$|\mathcal{G}_{n,k}| = (n \cdot k)!!.$$

Proof. See Exercise 3.2. □

We can now define the configuration model for random regular graphs. Recall that, given a set X, $\mathcal{P}(X)$ denotes the power set of X.

Definition 3.6. *The configuration model.* Let $n, k \in \mathbb{N}$ so that $n \cdot k$ is even. We define a probability measure

$$\mathbb{P}_{n,k} : \mathcal{P}(\mathcal{G}_{n,k}) \to [0,1]$$

by

$$\mathbb{P}_{n,k} = \frac{|A|}{|\mathcal{G}_{n,k}|}, \text{ for all } A \subset \mathcal{G}_{n,k}.$$

Because configurations give rise to regular graphs, the definition above allows us to speak of random regular graphs. That is, if we say "the probability that a k-regular graph on n vertices has property P", we will mean the probability with respect to the probability measure $\mathbb{P}_{n,k}$.

Note however that while it is clear that every graph can be obtained from some configuration (we just label the vertices and edges), some graphs might be given a higher probability by $\mathbb{P}_{n,k}$ than others. Later on, we will see that while $\mathbb{P}_{n,k}$ indeed prefers certain graphs, the differences between the probabilities assigned are small enough so that $\mathbb{P}_{n,k}$ can still be used to make statements about graphs picked uniformly at random among isomorphism classes.

3.2 Counting regular graphs

Our main application of the configuration model is counting regular graphs. That is, we will first control the cycle counts and the number of automorphisms of a random regular graph and then use this to give asymptotic estimates on the number of regular graphs of a fixed degree on a large number of vertices.

3.2.1 Cycles

Let $r \in \mathbb{N}$. A *cycle* of length r (or r-cycle) in a graph G is sequence of vertices $(v_1, v_2, \ldots, v_r)$ so that for all $i = 1, \ldots r-1$ the vertices v_i and v_{i+1} share an edge and so do the vertices v_r and v_1. Cycles that are obtained from one another by cyclic permutation or 'reading backwards' will be considered the same. A cycle in which all the vertices are distinct is called a *circuit*.

Our first goal is to understand the number of cycles of a given length in a random regular graph. To this end, define random variables $X_{n,k,r} : \mathcal{G}_{n,k} \to \mathbb{N}$ defined by

$$X_{n,k,r}(C) = |\{r\text{-cycles in } G(C)\}|$$

We will use the Chen-Stein method to prove a Poisson limit theorem (due to Bollobás [Bol80]) for these random variables.

Let us also define Poisson distributed random variables $X_{k,r} : \Omega \to \mathbb{N}$ with means

$$\lambda_{k,r} = \frac{(k-1)^r}{2r}$$

Finally, for a finite set $R \subset \mathbb{N}$, we define vectors of random variables

$$\mathbf{X}_{n,k,R} = (X_{n,k,r})_{r \in R} \text{ and } \mathbf{X}_{k,R} = (X_{k,r})_{r \in R}$$

Theorem 3.7. *Fix $k \in \mathbb{N}_{\geq 3}$. For any finite set $R \subset \mathbb{N}$ there exists a constant $C_R > 0$ so that*

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbf{X}_{n,k,R}, \mathbf{X}_{k,R}\right) \leq C_R/n$$

for all $n \in \mathbb{N}$.

Proof. We will apply the Chen-Stein method to prove this. Let us first analyse the possible labelings of an r-cycle in the graph corresponding to a configuration. Given an r-cycle in such a graph, we can traverse it and record the $2r$ labels that appear in it in order. If we also group every pair consecutive labels corresponding to the same edge, we obtain a list of the form

$$((a_1, b_1), \ldots, (a_r, b_r)) \in \left(W(n)^2\right)^r.$$

Note however that this does not define a map from cycles to lists of labels: to obtain the list, we need to know where to start traversing the cycle and in which direction to traverse it.

Let us write $A_{n,r}$ for the set of all such lists of labels that could possibly appear as an r-cycle in a configuration. In other words, $A_{n,r}$ is the set of lists $((a_1, b_1), \ldots, (a_r, b_r)) \in (W(n)^2)^r$ so that:

- The labels form a cycle: b_i and a_{i+1} lie in the same set $W_j(n)$ for all $1 \leq i \leq r-1$, as do b_r and a_1
- The pairs of labels that appear as edges are consitent: if (a_i, b_i) appears in a pair, then neither a_i nor b_i appears in a pair with another label (a pair is however allowed to appear multiple times).

Given $\alpha \in A_{n,r}$, we write

$$X_\alpha : \mathcal{G}_{n,k} \to \{0, 1\},$$

where $X_\alpha(C)$ counts the number of appearances of α in C, which is either 0 or 1. So X_α is a Bernoulli variable for all $\alpha \in A_{n,r}$ and all $r \in W$.

However, $\sum_{\alpha \in A_{n,r}} X_\alpha$ is not equal to $X_{n,k,r}$. Indeed we over count by going througth all the $\alpha \in A_{n,r}$: in $A_{n,r}$ each cycle artificially has a starting point and direction of travel. This implies that every labeled r-cycle is counted $2r$ times in $A_{n,r}$.

In the end we will want to deal with the set of labeled cycles and not the set of labeled directed cycles with a starting point, we could try to find a convenient description of the set $A_{n,r}/\sim$, where $\sim$ is some equivalence relation that takes care of the symmetry. Another course of action, the one we will actually pursue, is to simply divide all the quantities we need to compute by $2r$.

The rest of the proof of the theorem now progresses in a similar fashion to the proof of Theorem 2.2. For notational simplicity, we shall deal with the univariate case here. The passage to the multivariate case is analogous to the proofs of the approximation theorems.

Let us first compute the means. Like we said, we have:

$$\mathbb{E}\left[X_{n,k,r}\right] = \frac{1}{2r}\sum_{\alpha\in A_{n,r}} \mathbb{E}\left[X_\alpha\right] = \frac{1}{2r}\sum_{\alpha\in A_{n,r}} \mathbb{P}\left[C \text{ contains the pairs in } \alpha\right].$$

We note that the probability $\mathbb{P}\left[C \text{ contains the pairs in } \alpha\right]$ only depends on the number of distinct pairs in α. Indeed, if this number of pairs is e, then

$$\mathbb{P}\left[C \text{ contains the pairs in } \alpha\right] = \frac{1}{(n\cdot k-1)(n\cdot k-3)\cdots(n\cdot k--2\cdot e+1)}.$$

As such, it makes sense to divide $\mathbb{E}\left[X_{n,k,r}\right]$ into terms: each term corresponding to an isomorphism type of cycles. We can then write

$$\mathbb{E}\left[X_{n,k,r}\right] = \frac{1}{2r}\sum_{\mathcal{C}} a_{n,k}(\mathcal{C})\cdot p_{n,k}(\mathcal{C}),$$

where the sum runs over isomorphism types $\mathcal{C}$, $a_{n,k}(\mathcal{C})$ is the number of lists in $A_{n,r}$ that gives a cycle of the isomorphism type $\mathcal{C}$ and

$$p_{n,k}(\mathcal{C}) = \mathbb{P}\left[C \text{ contains the pairs in } \alpha_{\mathcal{C}}\right]$$

for any $\alpha_{\mathcal{C}} \in A_{n,r}$ that gives rise to a cycle of the isomorphism type $\mathcal{C}$.

It will turn out that $\mathbb{E}\left[X_{n,k,r}\right]$ is dominated by the term corresponding to circuits, so let us first compute that term. If $\alpha \in A_{n,r}$ corresponds to a circuit, then it contains exactly r distinct pairs, as such

$$p_{n,k}(r\text{-circuit}) = \frac{1}{(n\cdot k-1)(n\cdot k-3)\cdots(n\cdot k-2\cdot r+1)}.$$

Furthermore, to count the number of lists in $A_{n,r}$ giving rise to r-circuits, we note that all we need to choose is which distinct r vertices we use and which of the labels of these vertices to connect to each other. This gives a total of

$$a_{n,k}(r\text{-circuit}) = n \cdot (n-1) \cdots (n-r+1) \cdot (k(k-1))^r$$

options.

For the other terms, we note that by the same reasoning as above

$$a_{n,k}(\mathcal{C}) \cdot p_{n,k}(\mathcal{C}) \leq \frac{k^{2r} n^v}{(n - 2 \cdot r + 1)^e},$$

where v is the number of vertices in $\mathcal{C}$ and e the number of edges. If $\mathcal{C}$ is not a circuit, it has more edges than vertices, which implies that

$$0 \leq \mathbb{E}\left[X_{n,k,r}\right] - \frac{1}{2r} \frac{n \cdot (n-1) \cdots (n-r+1) \cdot (k(k-1))^r}{(n \cdot k - 1)(n \cdot k - 3) \cdots (n \cdot k - 2 \cdot r + 1)} \leq \frac{C}{n},$$

where $C > 0$ is a constant that depends on r and k (it for instance contains the number of isomorphism classes $\mathcal{C}$ we need to sum over) but not on n. We have

$$\lambda_{k,r} \left(\frac{n \cdot k - r \cdot k}{n \cdot k - 1} \right)^r \leq \frac{1}{2r} \frac{n \cdot (n-1) \cdots (n-r+1) \cdot (k(k-1))^r}{(n \cdot k - 1)(n \cdot k - 3) \cdots (n \cdot k - 2 \cdot r + 1)} \leq \lambda_{k,r}$$

Because

$$\left(\frac{n \cdot k - r \cdot k}{n \cdot k - 1} \right)^r = 1 + \mathcal{O}\left(n^{-1}\right)$$

as $n \to \infty$, we obtain that

$$\left|\mathbb{E}\left[X_{n,k,r}\right] - \lambda_{k,r}\right| = \mathcal{O}\left(n^{-1}\right)$$

as $n \to \infty$.

In order to derive a bound on the total variational distance between $X = X_{n,k,r}$ and a Poisson variable with mean $\lambda' = \mathbb{E}[X]$, we will use Theorem 1.8, in a way that will be very similar to the proofs of Theorems 2.1 and 2.2. We need to estimate

$$\left|\mathbb{E}\left[\lambda' \cdot g_A(X+1) - X \cdot g_A(X)\right]\right|.$$

First suppose that $X = \sum_{\alpha \in A} X_\alpha$ is a sum of Bernoulli variables. We can then write

$$\begin{aligned}\mathbb{E}\left[\lambda' \cdot g_A(X+1) - X \cdot g_A(X)\right] &= \sum_\alpha p_\alpha \mathbb{E}\left[g_A(X+1)\right] - \mathbb{E}\left[X_\alpha g_A(X)\right] \\ &= \sum_\alpha p_\alpha \left(\mathbb{E}\left[g_A(X+1)\right]\right. \\ &\qquad \left. - \mathbb{E}\left[g_A\left(1 + \sum_{\beta \neq \alpha} X_\beta\right) \mid X_\alpha = 1\right]\right),\end{aligned}$$

where $p_\alpha = \mathbb{E}\left[X_\alpha\right]$.

Now suppose that for every $\alpha \in A$ we can define a partition

$$A = \{\alpha\} \sqcup A_{\alpha,1} \sqcup A_{\alpha,2}$$

and random variables $X'_{\alpha,\beta} : \mathcal{G}_{n,k} \to \{0,1\}$ for all $\beta \in A \setminus \{\alpha\}$ so that

1. $X'_{\alpha,\beta}$ has the same distrbution as X_β when conditioned on $X_\alpha = 1$. That is
$$\mathbb{P}[X'_{\alpha,\beta} = 1] = \mathbb{P}[X_\beta = 1 |\, X_\alpha = 1].$$
2. $X_{\alpha,\beta}(C) \geq X_\beta(C)$ for all $C \in \mathcal{G}_{n,k}$ and all $\beta \in A_{\alpha,1}$.

If this were the case, then

$$\begin{aligned}\mathbb{E}\left[\lambda' \cdot g_A(X+1) - X \cdot g_A(X)\right] &\leq \sum_\alpha p_\alpha \left(\mathbb{E}\left[g_A(X+1)\right]\right. \\ &\qquad \left. - \mathbb{E}\left[g_A\left(1 + \sum_{\beta \neq \alpha} X'_{\alpha,\beta}\right)\right]\right) \\ &= \sum_\alpha p_\alpha \cdot \mathbb{E}\left[g_A(X+1) - g_A\left(1 + \sum_{\beta \neq \alpha} X'_{\alpha,\beta}\right)\right],\end{aligned}$$

Using the fact that $||g_A|| \leq 1$ (Proposition 1.10), we get

$$\begin{aligned}
|\mathbb{E}[\lambda' \cdot g_A(X+1) - X \cdot g_A(X)]| &\leq 2 \cdot \sum_\alpha p_\alpha \cdot \mathbb{E}\left[\left|X - \sum_{\beta \neq \alpha} X'_{\alpha,\beta}\right|\right] \\
&= 2 \cdot \sum_\alpha p_\alpha \cdot \mathbb{E}\left[\left|X_\alpha - \sum_{\beta \neq \alpha} (X'_{\alpha,\beta} - X_\beta)\right|\right] \\
&\leq 2 \cdot \sum_\alpha p_\alpha \cdot \mathbb{E}[X_\alpha] + p_\alpha \cdot \mathbb{E}\left[\sum_{\beta \in A_{\alpha,1}} (X'_{\alpha,\beta} - X_\beta)\right] \\
&\quad + \sum_\alpha p_\alpha \cdot \mathbb{E}\left[\sum_{\beta \in A_{\alpha,2}} (X'_{\alpha,\beta} + X_\beta)\right]
\end{aligned}$$

where we used property (2) to obtain the last inequality. Now note that

$$p_\alpha \cdot \mathbb{E}\left[X'_{\alpha,\beta}\right] = \mathbb{E}[X_\alpha X_\beta].$$

Set $p_{\alpha\beta} = \mathbb{E}[X_\alpha X_\beta]$. We obtain:

$$\begin{aligned}
\mathrm{d_{TV}}(X_{n,k,r}, Z_{\lambda'}) \leq\ & 2 \cdot \sum_\alpha (p_\alpha^2 + \sum_{\beta \in A_{\alpha,1}} p_{\alpha\beta} - p_\alpha p_\beta + \\
& \sum_{\beta \in A_{\alpha,2}} p_{\alpha\beta} + p_\alpha p_\beta). \qquad (3.1)
\end{aligned}$$

Note that none of the quantities depend on the variables $X'_{\alpha,\beta}$, it is only important that they exist. Furthermore, the only difference with Theorem 2.1 are the terms $p_{\alpha\beta} - p_\alpha p_\beta$. These terms measure the dependence of X_α and X_β. As such, we will want to choose $A_{\alpha,1}$ so that X_β is 'close to independent' of X_α for all $\beta \in A_{\alpha,1}$. The proof of the multivariate version of the statement above is very similar to the proof of Theorem 2.2 and we will skip it for now.

Set $A = A_{n,r}$. Our first task is now to find partitions $A = A_{\alpha,1} \sqcup A_{\alpha,2} \sqcup \{\alpha\}$ and variables $X'_{\alpha,\beta}$.

We start with the variables. Given $C \in \mathcal{G}_{n,k}$ and $\alpha \in A_{n,r}$, we obtain a new configuration $C'_\alpha \in \mathcal{G}_{n,k}$ as follows:

1. All pairs of labels in C that contain no labels from α become pairs of labels in C'_α.

2. If (i, j) appears in α but $\{i, j\} \notin C$, then that means that there are two pairs $\{i, x\}, \{j, y\} \in C$ with $x \neq j$ and $y \neq i$. We replace these pairs in C by the pairs $\{i, j\}$ and $\{x, y\}$. We do this until all the pairs in α appear in the configuration obtained.

Now set $X'_{\alpha,\beta}(C) = X_\beta(C'_\alpha)$. The partition we choose is given by:

$$A_{\alpha,1} = \{\beta \in A_{n,r};\ \beta \text{ shares no vertices with } \alpha\}.$$

We claim that these random variables satisfy the desired properties. Property (1) follows from the fact that, given α, the map $C \to C'_\alpha$ is constant to 1. This follows from symmetry: the actual labels involved play no role. As such

$$\begin{aligned}
\mathbb{P}[X'_{\alpha,\beta}(C) = 1] &= \mathbb{P}[X_\beta(C'_\alpha) = 1] \\
&= \frac{1}{|\mathcal{G}_{n,k}|} \sum_{C' \in \mathcal{G}_{n,k} : X_\alpha(C) = 1} |\{C \in \mathcal{G}_{n,k};\ C'_\alpha = C'\}| \cdot X_\beta(C')
\end{aligned}$$

Because the map $C \to C'_\alpha$ is constant to 1, we obtain

$$|\mathcal{G}_{n,k}| = |\{C' \in \mathcal{G}_{n,k};\ X_\alpha(C) = 1\}| \cdot |\{C \in \mathcal{G}_{n,k};\ C'_\alpha = C'\}|,$$

for any $C' \in \{C' \in \mathcal{G}_{n,k};\ X_\alpha(C) = 1\}$. Hence

$$\begin{aligned}
\mathbb{P}[X'_{\alpha,\beta}(C) = 1] &= \frac{1}{|\{C' \in \mathcal{G}_{n,k};\ X_\alpha(C) = 1\}|} \sum_{C' \in \mathcal{G}_{n,k} : X_\alpha(C) = 1} X_\beta(C') \\
&= \mathbb{P}[X_\beta = 1 |\ X_\alpha = 1].
\end{aligned}$$

Property (2) follows directly from the definition of $A_{\alpha,1}$. Indeed if $\beta \in A_{\alpha,1}$ and $X_\beta(C) = 1$ then $X_\beta(C'_\alpha) = 1$, just because β has no labels in common with α, so C'_α still contains the pairs in β.

To bound the sums in (3.1) we use similar observations as in the computation of $\mathbb{E}[X_{n,k,r}]$. If α forms a cycle of e edges and $v \leq e$ vertices, then

$$p_\alpha = \mathcal{O}\left(n^{-e}\right)$$

However, the number of terms with the same isomorphism type as α is $\mathcal{O}(n^v)$. So the first sum is $\mathcal{O}(n^{v-2e}) = \mathcal{O}(n^{-1})$ (using that the number of isomorphism classes we are considering is finite and depends on R only). Similar arguments work for the sums corresponding to $A_{\alpha,2}$.

If α and β are vertex-disjoint, the probabilities $p_{\alpha\beta}$ and $p_\alpha p_\beta$ are readily computed. It follows that

$$p_{\alpha\beta} - p_\alpha p_\beta \leq \frac{C}{n} \cdot p_\alpha p_\beta.$$

This means that the sums corresponding to the sets $A_{\alpha,1}$ contribute at most $C \cdot \lambda'^2/n$.

All in all, we obtain that

$$\mathrm{d_{TV}}\left(X_{n,k,r}, Z_{\lambda'}\right) \leq C/n.$$

Using the triangle inequality, we see that

$$\mathrm{d_{TV}}\left(X_{n,k,r}, X_{k,r}\right) \leq \mathrm{d_{TV}}\left(X_{n,k,r}, Z_{\lambda'}\right) + \mathrm{d_{TV}}\left(Z_{\lambda'}, X_{k,r}\right).$$

The above controls the first term, Exercise 3.3 controls the second. □

As an immediate consequence we obtain:

Corollary 3.8. *Fix $k \in \mathbb{N}_{\geq 3}$. For any finite set $R \subset \mathbb{N}$ we have*

$$\mathbf{X}_{n,k,R} \xrightarrow{\mathrm{TV}} \mathbf{X}_{k,R}$$

as $n \to \infty$.

3.3 Exercises

Exercise 3.1. Let $G = (V, E, \mathcal{I})$ be a graph. Show that

$$\sum_{v \in V} \deg(v) = 2\left|E\right|.$$

Exercise 3.2. Prove Lemma 3.5.

Exercise 3.3.

(a) Let $X_1, X_2 : \Omega \to \mathbb{N}^r$ be random variables. Show that

$$\mathrm{d_{TV}}\left(X_1, X_2\right) = \frac{1}{2} \sum_{k \in \mathbb{N}^r} \left|\mathbb{P}\left[X_1 = k\right] - \mathbb{P}\left[X_2 = k\right]\right|.$$

(b) Let $r \in \mathbb{N}$ and $\lambda_1, \lambda_2 \in (0,\infty)^r$ and let $\mathbf{X}_1 = (X_{1,1}, \ldots, X_{1,r}), \mathbf{X}_2 = (X_{2,1}, \ldots, X_{2,r}) : \Omega \to \mathbb{N}^r$ be random variables so that $X_{i,j}$ are Poisson distributed with mean $\lambda_{i,j}$ and pairwise independent. Show that

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbf{X}_1, \mathbf{X}_2\right) = \mathcal{O}\left(\sum_{i=1}^{r} |\lambda_{1,i} - \lambda_{2,i}|\right)$$

as $\sum_{i=1}^{r} |\lambda_{1,i} - \lambda_{2,i}| \to 0$.

Exercise 3.4. Let $k \geq 3$. Show that

$$\lim_{n\to\infty} \mathbb{P}_{n,k}\left[\text{The graph is connected}\right] = 1.$$

Hint: try to estimate $\mathbb{E}_{n,k}\left[X\right]$, where $X : \mathcal{G}_{n,k} \to \mathbb{N}$ counts the number of connected compondents of at most $n/2$ vertices.

Bibliography

[Bol80] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[Bol85] Béla Bollobás. *Random graphs.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1985.

[Wor99] N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.