UNIVERSITY OF KONSTANZ
ALGORITHMICS GROUP
V. Amati / J. Lerner

Network Modeling
Winter Term 2014/2015

Solution for Assignments $\mathcal{N}^{\underline{o}}$ 1, Task 3

released: - **due: -**

Task 3: $\mathcal{G}(n,p)$ **characterization** **4 points**

Proof the following three properties of the $\mathcal{G}(n,p)$ model:

(1) The edge probability of every dyad is equal to p.

(2) The model is fully independent.

(3) There is just one model satisfying properties (1) and (2).

Solution:

Let $D' = D_1 \uplus D_2 \subseteq D$ be a subset of dyads partitioned into two disjoint sets D_1 and D_2; both sets D_1 and D_2 may be empty. Define $m_1 = |D_1|$ and $m_2 = |D_2|$ to be the numbers of dyads in the two sets. Let

$$\mathcal{G}' = \{(V,E) \in \mathcal{G}\,;\ D_1 \subseteq E \wedge D_2 \cap E = \emptyset\}$$

be the set of all graphs in which the dyads in D_1 are constrained to be edges and the dyads in D_2 are constrained to be non-edges. We first show that

$$P(\mathcal{G}') = p^{m_1} \cdot (1-p)^{m_2} \quad . \tag{1}$$

Let $M = |D|$ denote the number of all dyads and for any graph $G = (V,E)$ let $m(G) = |E|$ denote the number of edges of G. It is

$$
\begin{aligned}
P(\mathcal{G}') &= \sum_{G\in\mathcal{G}'} P(G) && (2)\\
&= \sum_{G\in\mathcal{G}'} p^{m(G)}\cdot(1-p)^{M-m(G)} && (3)\\
&= \sum_{m=m_1}^{M-m_2} \binom{M-m_1-m_2}{m-m_1}\cdot p^{m}\cdot(1-p)^{M-m} && (4)\\
&= \sum_{m'=0}^{M-m_1-m_2} \binom{M-m_1-m_2}{m'}\cdot p^{m'+m_1}\cdot(1-p)^{M-(m'+m_1)} && (5)\\
&= p^{m_1}\cdot \sum_{m'=0}^{M-m_1-m_2} \binom{M-m_1-m_2}{m'}\cdot p^{m'}\cdot(1-p)^{M-m_1-m'} && (6)\\
&= p^{m_1}\cdot(1-p)^{m_2}\cdot \sum_{m'=0}^{M-m_1-m_2} \binom{M-m_1-m_2}{m'}\cdot p^{m'}\cdot(1-p)^{M-m_1-m_2-m'} && (7)\\
&= p^{m_1}\cdot(1-p)^{m_2}\cdot(p+(1-p))^{M-m_1-m_2} && (8)\\
&= p^{m_1}\cdot(1-p)^{m_2}\cdot(1)^{M-m_1-m_2} && (9)\\
&= p^{m_1}\cdot(1-p)^{m_2}\quad , && (10)
\end{aligned}
$$

which is Eq (1).

Some comments on the chain of equations above:

(2) is the definition of the probability of a set of graphs.

(3) puts in the probability function of the $\mathcal{G}(n,p)$ model.

(4) changes the order of summation. The summation-index now runs over all possible numbers of edges of graphs in $\mathcal{G}'$; note that the minimum number of edges is m_1 (since all dyads in D_1 have to be edges) and the maximum number of edges is $M-m_2$ (since all dyads in D_2 have to be non-edges). The probability of any graph with m edges is $p^m\cdot(1-p)^{M-m}$ and the number of graphs with m edges in $\mathcal{G}'$ is $\binom{M-m_1-m_2}{m-m_1}$. To see that the last claim is correct note that only $m-m_1$ edges can be freely chosen (the m_1 edges in D_1 are fixed) and these $m-m_1$ edges can be chosen out of $M-m_1-m_2$ unconstrained dyads.

(5) substitutes $m'=m-m_1$; note that m runs from m_1 to $M-m_2$ if and only if m' runs from zero to $M-m_1-m_2$.

(6) pulls out the factor p^{m_1} from every term of the sum.

(7) multiplies the whole sum with $(1-p)^{m_2}$ and divides every term of the sum by $(1-p)^{m_2}$ (note that we subtract m_2 from the exponents of $(1-p)$ in all terms) which yields equality.

(8) is the binomial formula.

(9) and (10) are obvious.

Now the proof of Task 3 (1) and (2) is quite simple.

(1) To show that the edge probability of every dyad is equal to p, let $e \in D$ be a dyad and define a set of constrained dyads $D' = D_1 \uplus D_2$ by $D_1 := \{e\}$ and $D_2 = \emptyset$. It is $m_1 = |D_1| = 1$ and $m_2 = |D_2| = 0$ and thus, by Eq. (1), it is

$$P(e \in E) = P(\mathcal{G}_e) = p \ .$$

(2) To show that the model is fully independent, let $e \in D$ be a dyad, $D' = D_1 \uplus D_2 \subseteq D \setminus \{e\}$ be a set of constrained dyads not containing e, and

$$\mathcal{G}' = \{(V, E) \in \mathcal{G}\,;\ D_1 \subseteq E \wedge D_2 \cap E = \emptyset\} \ .$$

Define $m_1 = |D_1|$ and $m_2 = |D_2|$. Then it is, by applying Eq. (1) three times,

$$P(\mathcal{G}_e \cap \mathcal{G}') = p^{1+m_1} \cdot (1-p)^{m_2} = p \cdot p^{m_1} \cdot (1-p)^{m_2} = P(\mathcal{G}_e) \cdot P(\mathcal{G}') \ ,$$

which shows that $\mathcal{G}_e$ is independent of $\mathcal{G}'$.

Claim (3) of Task 3 follows directly from the lecture slides where we noted that the probability function of a fully independent model is uniquely determined by the edge probabilities of all dyads.