Exercises for Statistical analysis of network data – Sheet 1

1. This corresponds to just noting the edges:

(a)

$$A_1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This has as edge list $E_1 = \{(1,2),(1,3),(2,3),(3,4)\}$, and is known as a tadpole graph.

(b)

$$A_2 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

This has as edge list $E_2 = \{(1,2),(1,3),(2,4),(3,4),(3,5)\}$, and is known as a tadpole graph.

(c)

$$A_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This has edge list $E_3 = \{(1,2),(2,3),(3,4),\}$, and is known as a path. Paths are special cases of trees, and have fewer edges than nodes.

(d)

$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

This has edge list $E_4 = \{(1,5),(2,5),(3,5),(4,5)\}$. This is a star, and is also a special case of a tree.

(e)

$$A_5 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

This has edge list $E_5 = \{(1,2),(1,3),(2,4),(3,4)\}$. This is a 4-cycle.

(f)

$$A_6 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

This is a special case of a tree as the number of edges is less than the number of nodes. This has edge list $E_6 = \{(1,2),(2,3),(3,4),(3,5)\}$.

2. The plots are at the end of the document. Adjacency matrices are given by

(a) $E_1 = \{(1,2),(1,3),(2,3)\}$. This has adjacency matrix

$$A_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

(b) $E_2 = \{(1,5),(2,5),(3,5),(4,5)\}$. This has adjacency matrix

$$A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

(c) $E_3 = \{(1,2),(2,3),(3,4),(4,5)\}$. This has adjacency matrix

$$A_3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

(d) $E_4 = \{(1,2),(2,3),(3,4),(4,5),(5,6),(6,1)\}$. This has adjacency matrix

$$A_4 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

A good check on whether something has been tabulated correctly is to take $A - A^T$. As the adjacency matrix is symmetric, this will normally uncover errors.

3. Assume that a_{ij} is generated from an Erdős-Rényi network with edge probability ρ.

(a) Using moment generating functions we can note that the sum of $n-1$ independent Bernoullis is $\text{Bin}(n-1,\rho)$.

(b) The expectation of a binomial is $(n-1)\rho$. The variance is $(n-1)\rho(1-\rho)$.

(c) The dispersion is $\text{Var}\{d_i\}/\text{E}\{d_i\} = (n-1)\rho(1-\rho)/((n-1)\rho) = 1-\rho$. We need $\rho = o(1)$.

(d) For an inhomogeneous random graph we have what is known as a Poisson Binomial random variable, which is the distribution of $n-1$ Bernoullis with different success probabilities. The expectation is $\sum_{i\neq j} p_{ij}$ and the variance is $\sum_{i\neq j} p_{ij}(1-p_{ij})$.

4.

$$\text{E}T = \text{trace}(\text{E}(A^3))/6 = \frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\text{E}(a_{ij}a_{jk}a_{ki}).$$

To determine the expectation we note that

$$\begin{aligned} \text{E}(a_{ij}a_{jk}a_{ki}) &= \text{E}(a_{ij})\text{E}(a_{jk})\text{E}(a_{ki})\text{I}(i\neq j)\text{I}(i\neq k)\text{I}(k\neq j) && (1)\\ &+ \text{E}(a_{ij})\text{E}(a_{jk})\text{E}(a_{ki})\text{I}(i = j)\text{I}(i\neq k)\text{I}(k\neq j) && (2)\\ &+ \text{E}(a_{ij})\text{E}(a_{jk})\text{E}(a_{ki})\text{I}(i\neq j)\text{I}(i = k)\text{I}(k\neq j) && (3)\\ &+ \text{E}(a_{ij})\text{E}(a_{jk})\text{E}(a_{ki})\text{I}(i\neq j)\text{I}(i\neq k)\text{I}(k = j) && (4)\\ &= \rho^3\text{I}(i\neq j)\text{I}(i\neq k)\text{I}(k\neq j). && (5) \end{aligned}$$

Therefore

$$\begin{aligned} \text{E}T &= \frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\rho^3\text{I}(i\neq j)\text{I}(i\neq k)\text{I}(k\neq j) && (6)\\ &= \binom{n}{3}\cdot\rho^3. && (7) \end{aligned}$$

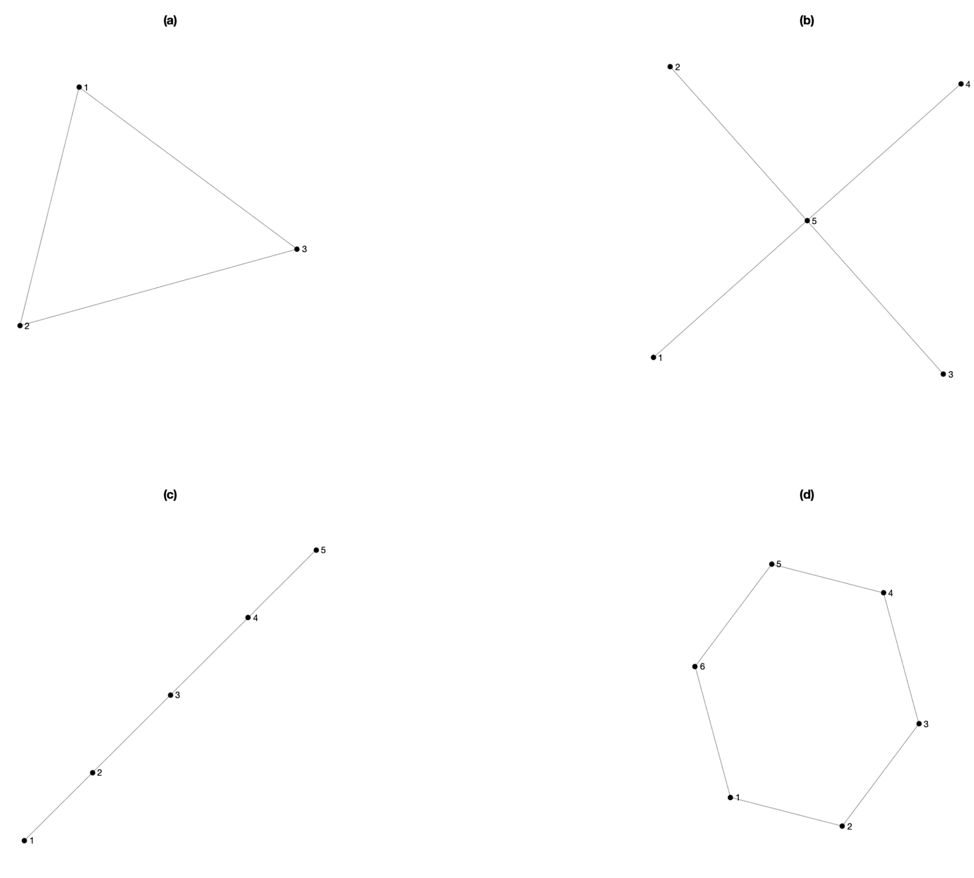

Figure 1: Plots of the graphs.