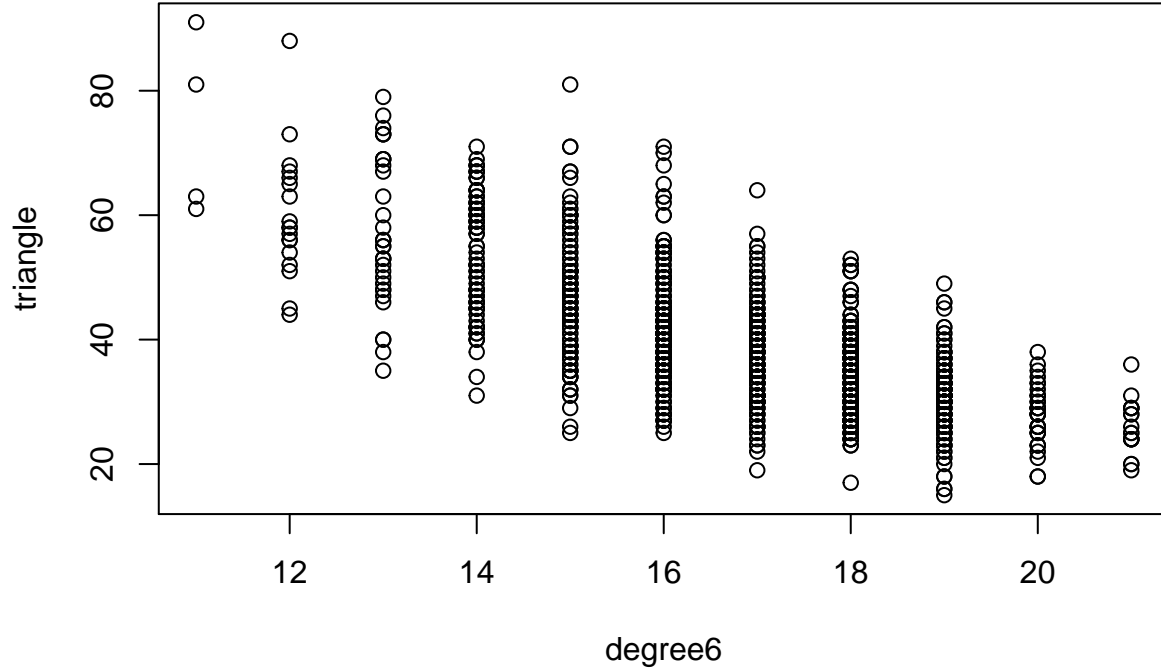


# ABC sampler test

I test the parameters of the BA model as a mechanistic model against the ERGM shown in the meeting yesterday. The ERGM (21 nodes) has two summary statistics in its likelihood: number of 6-stars and number of triangles, its true parameters are  $\theta = (3.5, 0.02)$ . The BA model parameter is  $m$ , the number of nodes added at each iteration. The KL divergence computation told us that  $m = 6$  is the optimum parameter.

I drew 1000 realisations from this ERGM and used the combination of statistics which occur most frequently as the data, 17 6stars and 44 triangles.



## [1] 17 44

The first distance I used was a Euclidean distance on the two summary statistics. The second distance is a weighted difference between the summary statistics (to approximate what our KL divergence procedure targets). As you can see below, the best estimator for the first distance is 3, and the best estimator for the second distance is 6 (which is what we found with the KL divergence). The parameter we found from the KL procedure,  $m = 6$  doesn't minimise the difference between summary statistics, it minimises the difference in the sum of parameters and statistics used in the ERGM model.

$$\text{distance}_1 := \sqrt{(S_{\text{obs},1} - S_{\text{sim},1})^2 + (S_{\text{obs},2} - S_{\text{sim},2})^2}$$

$$\text{distance}_2 := \theta_1(S_{\text{obs},1} - S_{\text{sim},1}) + \theta_2(S_{\text{obs},2} - S_{\text{sim},2})$$

