

## Assignment 1

**Issue date:** 24 Apr 2014    **Due date:** 01 May 2014, 11:00

It is explicitly recommended to solve exercises in groups of two.

*(Tutorial on 28 Apr: Introduction to graph theory)*

### Exercise 1: Empirical Research

**6 Points**

Assume that you are given the following controversial hypothesis:

*“Obesity is socially contagious.”*

What could be the underlying theory?

Try to answer the following questions concerning the *research pipeline*:

- Which data of participants in your study would you gather?
- Which data might not be observable?
- Is it enough to gather data only once? If not, why?
- How would you design the research to obtain meaningful network data?
- How do you analyze the data?

[please turn over]

## Exercise 2: Random Networks

2+4 Points

- (a) The  $G(n, p)$  model was first introduced by Edgar Gilbert in 1959 [1]. Read the Introduction of his paper and discuss the plausibility of his exemplary application. (Link to paper)
- (b) An often used example for a *preferential attachment* mechanism is the evolution of citation networks. Assume that this evolution is driven by the following “random citation process”:

*Papers to cite are chosen at random with probability proportional to the number of citations those papers already have.*

An unrealistic assumption of this process is that every scientist writing a paper is aware of the distribution of citations for all papers ever written.

Describe a random citation process that could explain the evolution, however without the knowledge of how often papers already got cited.

**Please submit your answers electronically to teaching assistant David (david.schoch@uni-konstanz.de).**

## References

- [1] Edgar N Gilbert. Random Graphs. *The Annals of Mathematical Statistics*, 1141–1144, 1959.