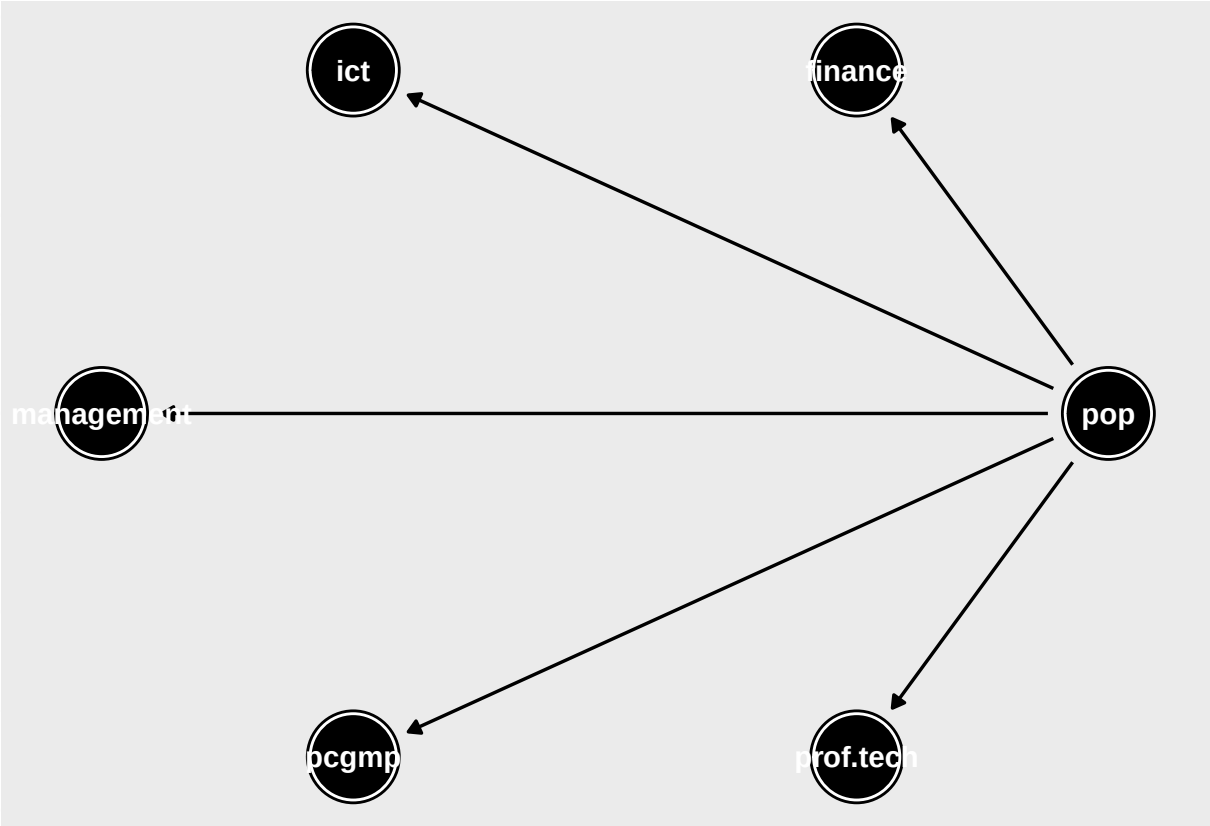
36-402 Homework 12

Eu Jing Chua
eujingc

May 19, 2019

Question 1

Q1 a)



However, multiple other graphs are possible as the theory says nothing about the relationships between per-capita output and the four industries, in which case any set of relationships would be compatible with the theory.

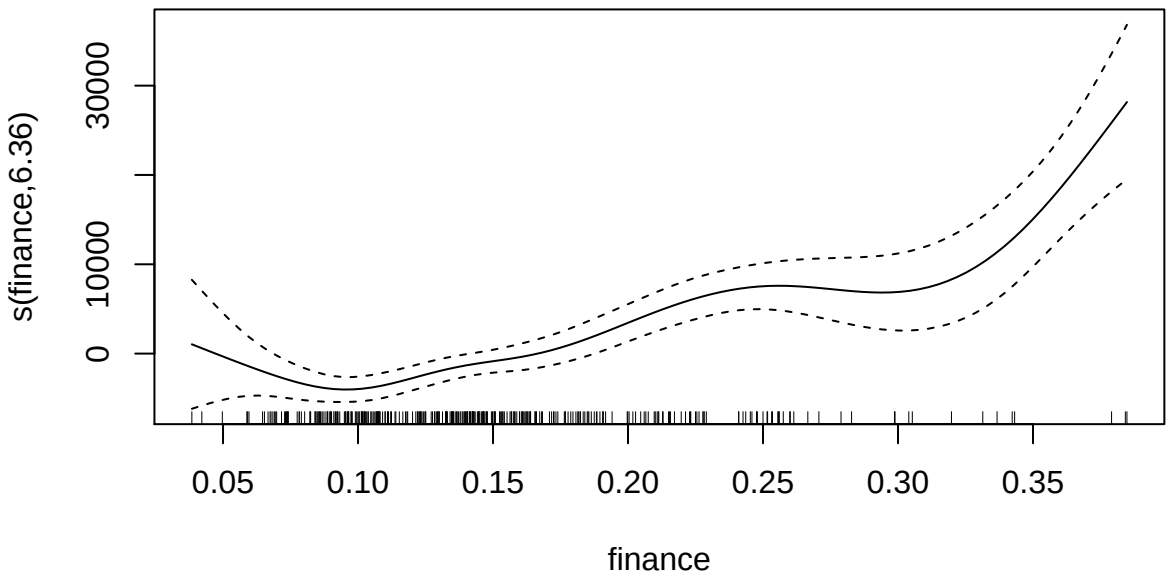
Q1 b)

`pcgmp` $\not\perp\!\!\!\perp$ `finance` as there is an open path through population.
`pcgmp` $\perp\!\!\!\perp$ `finance` | `pop` as all paths are closed.
`pcgmp` $\perp\!\!\!\perp$ `finance` | `pop`, `management` as all paths are closed.

Q1 c)

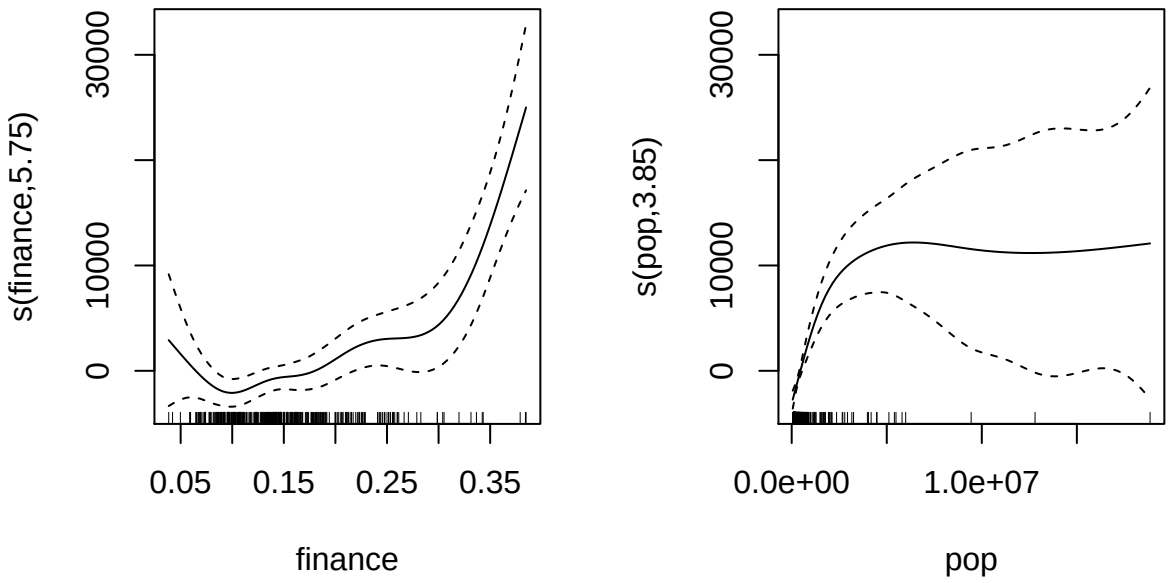
We can roughly test for the dependence between variables by assuming an additive mixture model using spline smoothing for each variable.

In the first model, we test for '*pcgmp*' $\not\perp\!\!\!\perp$ '*finance*', in which we only model `pcgmp` against `finance`. The partial response is as follows:



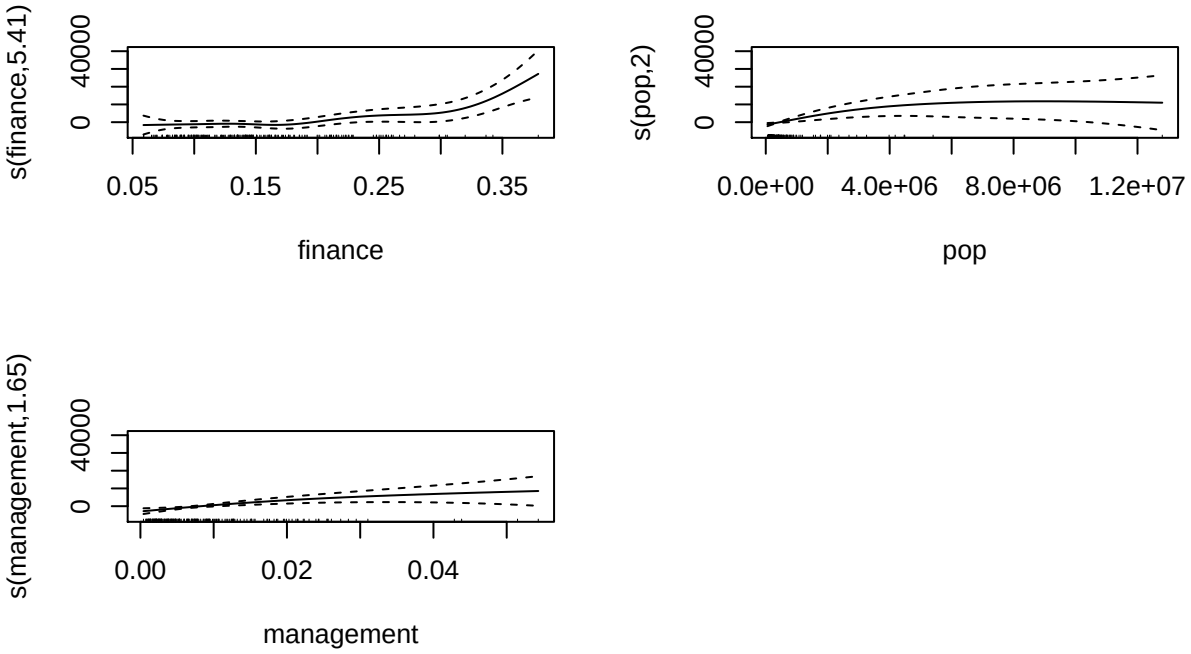
We can see a relationship exists between `pcgmp` and `finance`, where in general as `finance` increases, so does `pcgmp`. Thus, this seems to provide evidence that the dependence exists.

In the second model, we test for '$pcgmp$' ⫫ '$finance$' | 'pop', in which we model `pcgmp` against `finance`, controlling for `pop`. The partial responses are as follows:



However, the additive model shows that after controlling for `pop`, the partial response of `finance` still has a similar relationship as before, where higher values of `finance` are related to higher values of `pcgmp`. Thus, the data does not support this conditional independence.

In the third model, we test for '*pcgmp*' ⫫ '*finance*' | '*pop*', '*management*', in which we model `pcgmp` against `finance`, controlling for `pop` and `management`. The partial responses are as follows:



Similar to above, we can see that the partial response of `finance` again shows an increasing relationship, even after controlling for both variables. However, once again the data does not support this conditional independence.