Danmarks Tekniske Universitet

Lukas Leindals, s183920
Sunniva Olsrud Punsvik, s183924

# 02463 - Aktiv machine learning og agency F20

## Project 3
## Causality - Intervention Effects

30th of April 2020
Danmarks Tekniske Universitet

# Introduction

Not being able to perform a certain experiment is not unheard of and often occurs in real life settings - this may be due to ethical considerations, not being able to make the most out of historical data, resources and expenses, etc. However, causal inference is able to provide the necessary tools to overcome those hindrances mentioned above. That being said, there is still a huge debate surrounding the nature of causality and proving causality often times is not even possible.

The objective here is to investigate intervention effects based of a DTU server to search for how variables influence different variables and/or is affected by other variables. There are 7 variables to analyse, as well as there is a hidden variable $\_H$, to do interventions on (not on the hidden variable). With the given information, one had to produce a causal model showing these relationships.

# Method & Analysis

There were given 7 variables: $I, P, S, B, K, A$ and the hidden variable $\_H$. There was a cost related to per sample, per experiment, and per incorrect guess. So for this particular set up, it was chosen to go with 100 observational data points without any interventions and 75 samples with intervention on all the variables except $\_H$ and all had a set-value of 1. To get some more samples, bootstrapping was used to obtain 300 samples in each scenario (data with different and no interventions).

To map the causal graph, one can look at the changes in the mean and variance when performing interventions. To appoint a variable as a mediator, confounder, or collider; how the variance and mean changes after an intervention is crucial. For the hidden variable, it must mean that there is not some logical explanation as of why two variable might be affected by each other, and that is the role of $\_H$.

# Results

The following results were obtained where e.g. *K1* means an intervention of $K = 1$ was made:

| | None | I1 | P1 | S1 | B1 | K1 | A1 |
|---|---|---|---|---|---|---|---|
| I | 0.0 | 0.772 | 0.319 | 0.154 | -0.355 | 0.640 | -0.346 |
| P | 0.0 | 0.261 | -0.419 | 0.207 | 0.259 | -0.031 | 0.169 |
| S | 0.0 | -0.119 | -0.208 | 0.485 | -0.069 | -0.129 | -0.153 |
| B | 0.0 | 0.128 | -0.474 | 0.221 | -0.415 | -0.098 | 0.308 |
| K | 0.0 | -0.076 | -0.051 | -0.099 | -0.147 | 0.538 | -0.112 |
| A | 0.0 | 0.173 | -0.451 | -0.667 | -0.399 | -0.124 | 0.248 |

Table 1: The percentagewise difference in means with different interventions

| | None | I1 | P1 | S1 | B1 | K1 | A1 |
|---|---|---|---|---|---|---|---|
| I | 0.0 | -1.000 | 0.332 | 0.082 | 0.201 | 0.419 | -0.191 |
| P | 0.0 | -0.268 | -1.000 | 0.064 | 0.210 | 0.369 | -0.125 |
| S | 0.0 | 0.097 | 0.132 | -1.000 | 0.064 | 0.102 | 0.114 |
| B | 0.0 | -0.242 | -0.876 | 0.167 | -1.000 | 0.482 | -0.101 |
| K | 0.0 | -0.257 | -0.243 | -0.163 | -0.000 | -1.000 | 0.098 |
| A | 0.0 | -0.403 | -0.858 | 0.526 | -0.957 | -0.557 | -1.000 |

Table 2: The percentagewise difference in variances with different interventions

| | I | P | S | B | K | A |
|---|---|---|---|---|---|---|
| I | 1.000 | 0.384 | 0.138 | 0.092 | 0.124 | -0.063 |
| P | 0.384 | 1.000 | 0.156 | 0.934 | -0.087 | -0.056 |
| S | 0.138 | 0.156 | 1.000 | 0.121 | -0.153 | 0.052 |
| B | 0.092 | 0.934 | 0.121 | 1.000 | -0.117 | -0.010 |
| K | 0.124 | -0.087 | -0.153 | -0.117 | 1.000 | -0.107 |
| A | -0.063 | -0.056 | 0.052 | -0.010 | -0.107 | 1.000 |

Table 3: The correlations when no interventions were performed

The obtained causal graph:

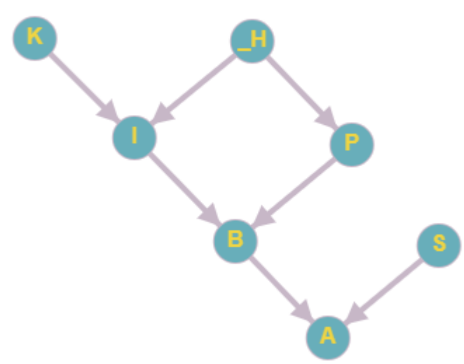


Figure 1: Causal Graph

# Discussion & Conclusion (learning outcome)

By comparing table 1 with table 2 with the threshold 0.2, one can see that $S$ affected $A$ and $K$ affected $I$, as no interventions seemed to show differences above the threshold for $S$ and $K$.

The next thing that was done was to take a look at table 3. The first thing to notice is a high correlation between $B$ and $P$. When looking at which variables these two also correlates with, it is seems likely that $P$ and $I$ seemingly has fairly high correlation. This indicates some kind of causal relationship between the three. When looking at the variance, an intervention on $P$ has a high impact on $B$, but not the other way around, which indicates $P$ affects $B$. When looking at which variables affects the variance of $B$, $K$ seems to stand out. As it is known $K$ influences $I$ and $I$ also seems to have an effect on $B$ a relationship between $I$, $B$ and $P$ where $B$ is a collider is suspected.

When looking at which interventions has an effect on $A$, all the interventions seems to have a large effect. We have already found that $S$ affects $A$, but also $B$ and $P$ seems to have a large effect. This leads to a relationship between $B$ and $A$ where $B$ affects $A$, making $A$ a collider between $B$ and $S$ and $B$ a mediator sending information from both $I$ and $P$

When looking at the results, only one relationship seemed to still be unexplained. $P$ and $I$, which seems to have a fairly high impact on each other, which cannot be explained by the graph. This leads to the theory that a confounder exists between the two, which must be the unobserved $\_H$ variable.

This leads to the final graph shown in figure 1.

To make these causal relations more clear, one could consider intervening with the largest of the boundary values (2 and -2), this could help by making the differences in mean larger. Furthermore, intervening on $S$ was unnecessary when looking back in hindsight, as this was an uniformly distributed value, which could simply just have been conditioned on to obtain the same effect.

From this project we have learned that understanding a causal relationship can be very difficult, but with simple summary statistics of mean and variance, along with the correlation, are very useful tools to understanding causal relationships. Another thing we learned was that a hidden variable can be detected by looking at relationships, that cannot be explained by the graph and relationships found via interventions.

# Appendix

## Link to code in GitHub

https://github.com/The-very-most-awesome-team-of-cool-kids/Causality

**Introduction:** It is good that you mention that the field of causality is still developing and that it is not simply possible to check causality in all cases. Maybe you could have mentioned the importance of causal analysis in comparison with correlations, as this is an important distinction which is often misinterpreted. You adequately and sufficiently explain the setup for the given problem.

**Methods and Analysis:**
You clearly show that you have thought about the number of samples, which is a crucial step in this analysis, I generally think that using a bootstrap strategy is a good approach for estimating distributions is a good idea, however I am not following the argument of increasing the number of samples to 300. The way I would use it would be to estimate the certainty of the distribution for the number of samples or to estimate if a lower number would have been enough. By upsampling by bootstrapping you actually have the over estimated your confidence in coupling which could be a dangerous practice, however in this case 100 samples appears to be sufficient so you are good.

The procedure of first getting samples without an intervention is a good way to start, as it will serve as a baseline to compare interventional distributions against, as you have done in table 1 and 2. You could have considered to plot the distributions, which would have given you a good insight in the data - e.g. in the categorical variable *S*, but also for the other variables to asses if your assumption of the data being normally distributed.

Overall, it's really great that you shortly describe the theory, and specify that you will look at mean and variance (which is sufficient for this graph), but you could have mentioned that this generally presumes that the data in normally distributed (in line with the comment above). You could consider if this assignment has learned you some general rules for how to determine if a variable is e.g. a confounder or a mediator. This could possibly make it easier for you to present the project at the exam where there is limited time.

**Results, Discussion and Conclusion**:
Generally, you have tackled the problem in a very "uniform" way, performing the same intervention for all variables and looking at the resulting mean and variance changes. Since you have also looked at the correlations without interventions (listed in Table 3), another approach could have been to perform specific interventions on the variables which show a strong correlation (since if there is no correlation, you will likely also not be able to get evidence of coupling looking by looking at mean and variance anyways). E.g.