Introduction to Causal Discovery

Course outline
Causal discovery is the science of inferring causal models from observational data. In this interactive workshop, we will introduce the basic ideas behind causal discovery, and participants will apply causal discovery algorithms to real data in R.

We focus on constraint-based algorithms used for exclusively observational data. We will present two foundational algorithms for causal discovery in this setting, namely the PC algorithm and the FCI algorithm, the latter of which allows for the presence of unobserved confounding. We provide insights into how, why, and when the algorithms work, but without going into mathematical details or rigorous proofs. We will discuss how and when external background information, in particular temporal information, can aid causal discovery.

The short course will be organized as a mix of lectures and interactive exercise sessions where participants try out the algorithms in practice and interpret their results. Participants are welcome to bring their own data, or they can analyze example datasets provided by the instructors.

Finally, we will also provide further perspectives on other causal discovery algorithms and tools in R, and guide participants to where they can learn more.

Learning objectives
Participants will gain knowledge of:

- The basic principles of constraint-based causal discovery algorithms
- Equivalence classes of causal models with and without the assumption of sufficiency (completed partially directed acyclic graphs (CPDAGs) and partial ancestral graphs (PAGs))
- Assumptions and limitations of causal discovery
- The role of external (temporal) background knowledge in causal discovery

Moreover, after the course, participants will have practical experience with:

- Applying the PC and FCI algorithms to real data in R, both with and without external background knowledge
- Interpreting the outputs of these algorithms

Intended audience
We expect participants to have basic knowledge of directed acyclic graphs and basic statistical concepts such as tests and (conditional) independence, but no previous experience with causal discovery is required. We will be working with hands-on exercises in R, so basic experience with R is a prerequisite. Participants should bring their own computers and may also bring their own data.

Instructors
Anne Helby Petersen, Assistant Professor, University of Copenhagen
Claus Thorn Ekstrøm, Professor, University of Copenhagen
Christine Winther Bang, PhD Student, Leibniz Institute for Prevention Research and Epidemiology - BIPS