

EITM Turin: Formal Models and Bayesian Nets

Macartan Humphreys

3-5 July 2017

Go [here](#) for the most up-to-date version of this syllabus.

1 Content

We will focus on some Bayesian methods for translating from theories to empirical causal models. You will pick up some general tools but also hopefully a broader framework for thinking about the link between theoretical and empirical models. A key insight we will develop is that well developed theories have multiple implications regarding what constellations of outcomes one should observe in the world. Bayesian models are particularly well suited to making inferences from multiple disparate bits of information, even in a single case. You will be able to make use of the types of ML and Bayesian modelling you learned about in James Lo's modules; in addition you will also get a basic introduction to Stan.

I am hoping that over the three days you will work individually or in groups to translate a theoretical model into a structural model, define estimands, and generate code to make inferences based on observations from one or more cases. This will be hard enough conceptually so lets focus on simple models and make sure we get the analytic structure right.

2 Readings

All readings are linked from this document.

A lot of the material of the course is gathered together in a draft book manuscript, [Integrated Inferences](#), that Alan Jacobs and I are working on right now (referred to below as "HJ"). You will be among the first readers of this text; I hope getting to think through the new ideas in here will make up for the pain of working through draft text. This draft book is not for wider circulation and is also liable to change!

3 Prerequisites

You should already have background in formal modelling and causal inference.

In addition you should know some R. Really, the more you can invest on getting on top of R before the class the better. Take James's sessions seriously, take the EITM refreshers, or swot up yourself.

- Resources for learning R: <http://www.r-bloggers.com/how-to-learn-r-2/>

Please make sure your R is up-to-date and that you are working in R `studio`. Then make sure you have the following packages installed.

```
if (!require("rstan")) install.packages("rstan")
if (!require("dagitty")) install.packages("dagitty")
```

Those are both quite stable package. We will also use an alpha version of a package that Alan Jacobs and I are working on for our book. This is not particularly stable but will be useful for us. Apologies in advance for bugs. To install this you need `devtools`.

```
if (!require("devtools")) install.packages("devtools")
if (!require("biqq")) devtools::install_github("macartan/biqq")
```

3.1 File Sharing via Dropbox

If we have a good connection I think it will be useful for us to all work in a single dropbox. That makes it easy to look at each other's work easily and help each other out. Please make sure you have dropbox installed on your machines.

3.2 Writing with Rmd

Please plan to do drafting in Rmarkdown. Rmarkdown is just a very simple markup language that lets you integrate writing and coding. This document is written in Rmarkdown and the slides will be also.

The key thing is that you can insert code chunks like this.

```
# Define a random number
x <- rnorm(1)
```

which are run as the document compiles, and can be accessed as needed, like this: we just sampled the random number $x = -0.7072583$.

For this you will use Rstudio as an editor. More information here: rmarkdown.rstudio.com.

4 Modules

Here is the plan:

4.1 Monday 3 July: Tools

We will focus on the basics:

- 09:00 Introductions, outlines, expectations
- 09:00 - 10:30 Lecture: Causal Inference and Causal Estimands. Read: [HJ](#) Ch 2.1, 2.2
- 11:00 - 11:30 Lecture: Basics of Bayesian Inference. Read: [HJ](#) Ch 3
- 11:30 - 12:30 Exercises: Bayesian updating by hand.
- 12:30 - 13:30 Lunch
- 13:30 - 15:00 Lecture: DAGS, Structural Models, daggity. Read Pearl: [Causal diagrams](#); [HJ](#) Ch 2.3, 2.4; [d-separation without tears](#)
- 15:30 - 16:30 Lecture: Defining estimands on a DAG
- 16:30 - 17:30 Exercises in DAGs, d-separation.

4.2 Tuesday 4 July: Theories and the probative value of data

We will focus on connecting DAGs and theories

- 9:00 - 10:00 Lecture: What's a theory? Theories of moderation and mediation. Mapping from a game to a DAG. Read: [HJ](#) Ch 4; [Clarke and Primo](#)
- 10:00 - 11:00 Exercises. Game to DAG mappings.

- 11:00 - 12:30 Lecture: Probative value and case level process tracing. Illustrations. Read: [HJ](#) Chs 5 & 7
- 12:30 - 13:30 Lunch
- 13:30 - 15:30 Groupwork: Your task is to take a game, represent it as a DAG, define estimands of interest, describe an inferential strategy. This may involve using data from one or more cases. Final presentations will be tomorrow so you will likely want to keep working on this this evening.
- 15:30 - 17:00 Lecture: Bayesian inference in Stan. Look at one of these intros: [Viefers slides](#) ; [Andy Gelman lecture](#)

4.3 Wednesday 5 July: Scaling up / Group Presentations

We will focus on scaling up to larger data problems.

- 09:00 - 10:00 Mixed methods inference in stan. Read: [HJ](#) Ch 8
- 10:00 - 11:00 Implications for design. Read: [HJ](#) Ch 9 & 10
- 11:00 - 11:30 Liberia illustration (structural model with experimental data, quant only)
- 11:30 - 12:30 Exercises: Run and assess mixed methods stan models
- 12:30 - 13:30 Lunch
- 13:30 - 14:30 Lecture: Strategies for evaluating models
- 14:30 - 15:30 Final presentation prep time
- 15:30 - 17:30 Project presentations. Each presentation should have:
 1. Game description
 2. DAG of game
 3. Estimands and inference strategy
 4. Conditional inferences given possible data observations
 5. Reflections on model assessment