Causal models
Motivation

► As discussed before, the motivation to carry out randomized trials is to obtain evidence on causal effects.
► Causality as such is not something observable (a phenomenon), but rather, something entirely conceptual (a noumenon), that needs to be inferred from phenomenal patterns (Miettinen & Karp 2012, p. 43).
► Consequently, causal concepts need to be formalized in terms of *causal models*.
► Directed acyclic graph (DAG) is one example of a causal model.
A non-experimental setting

- We are interested in the causal relationship $Z \rightarrow Y$.
- In addition to the treatment assignment $Z$, $X$ are observed and $U$ are unobserved other determinants of the outcome $Y$.

- Is the treatment assignment independent of the unmeasured factors $U$ conditional on the measured ones $X$, that is, do we have $Z \perp U \mid X$?
Moralization criterion

► Originates from Lauritzen (1996), similar to the d-separation criterion of Pearl (2000).
► We are interested in the set of variables \((X, Z, U)\).
► Form a ‘moralized’ undirected subgraph through the following three steps:
  1. Include the variables of interest and their ‘parents’.
  2. ‘Marry’ the unmarried parents with a common ‘child’.
  3. Remove the directions of the arrows.
► Any conditional independences can now be read from the moralized graph:

![Diagram of moralization criterion]

- \(X\)
- \(Z\)
- \(U\)
Applying the criterion

- The path between $Z$ and $U$ was blocked by $X$, so we indeed have that $Z \perp \perp U \mid X$.
- But how do we make sure that there is no arrow $U \rightarrow Z$?
- A randomized trial:

Now we have $Z \perp \perp (X, U)$ (show using the criterion).
Selection diagrams

- This did not yet specify the causal effect of interest; in these graphs everything was ‘causal’.
- We can be more specific by extending the DAG to a selection diagram.
- Suppose the treatment variable $Z$ can be intervened upon (‘selected’), and indicate this with a selection node $\sigma_Z$:

![Diagram](image_url)
Identifying a causal effect using selection diagrams

- Under the non-experimental setting, we can identify the causal effect $\sigma_Z \rightarrow Y$ given the observed data $(X, Y, Z)$ if $Y \perp \perp \sigma_Z \mid (X, Z)$.

- The conditional independence condition can be read from the moralized graph:

- This means that, given $X$, observing the value of $Z$ is sufficient to estimate the relationship $\sigma_Z \rightarrow Y$. 
The role of confounder adjustment

- Note that under the non-experimental setting this does not apply if we do not condition on the observed confounders $X$.
- This is because we do not have $Y \perp \sigma_Z \mid Z$. (Read from the moralized graph on the previous slide).
- However, under the experimental setting we have

```
X \rightarrow U
\sigma_Z \rightarrow Z \rightarrow Y
```

Note that under the non-experimental setting this does not apply if we do not condition on the observed confounders $X$. This is because we do not have $Y \perp \sigma_Z \mid Z$. (Read from the moralized graph on the previous slide). However, under the experimental setting we have
Adjustment for $X$ is not needed under the experimental setting

- The moralized version is:

```
σZ

X

U

Z

Y
```

- Now we indeed have that $Y \perp \sigma Z \mid Z$.
- Now adjustment for $X$ is not necessary to control for confounding, but may be desirable for other reasons (cf. the first assignment).
Potential outcomes

- Alternatively, the causal effect and the conditions required for identifying this can be specified using what is known as potential outcomes notation.
- This framework is commonly referred to as the *Rubin’s causal model* (Rubin 1978; Rosenbaum & Rubin 1983).
- Suppose that each individual has two possible outcomes, \( Y_0 \) corresponding to the outcome without treatment, and \( Y_1 \) to the outcome with treatment.
- Depending on the value of the treatment indicator \( Z \), one of these outcomes is *factual* and the other is *counterfactual*.
- The observed outcome is then given by

\[
Y = ZY_1 + (1 - Z)Y_0.
\]
In terms of causal graphs, we can understand this by replacing the latent individual-level characteristics $U$ with the pair of potential outcomes $(Y_0, Y_1)$:

The pair $(Y_0, Y_1)$ is to be interpreted as all determinants of the outcome $Y$, short of the treatment assignment $Z$. 

In terms of causal graphs, we can understand this by replacing the latent individual-level characteristics $U$ with the pair of potential outcomes $(Y_0, Y_1)$:
Identifying condition

▶ As before, we can form the moralized graph

\[ X \rightarrow Z \rightarrow Y_0, Y_1 \]

to determine that \((Y_0, Y_1) \perp \perp Z \mid X\).

▶ How to interpret this conditional independence condition?

▶ Does it mean that the outcome is independent of the treatment?
Experimental setting

Similarly, under the experimental setting we would have

\[ X \rightarrow Y_0, Y_1 \]
\[ Z \rightarrow Y \]
\[ Y_0, Y_1 \rightarrow Y \]

and the moralized version

\[ Y_0, Y_1 \]
\[ Z \]

to determine that \((Y_0, Y_1) \perp \perp Z\).
Strong ignorability

- If \((Y_0, Y_1) \perp Z \mid X\) and \(0 < P(Z = 1 \mid X) < 1\), we say that the treatment assignment mechanism is *strongly ignorable*.

- Under the experimental setting we have further that \((Y_0, Y_1) \perp Z\) and \(0 < P(Z = 1) < 1\).

- Strong ignorability is an identifying condition for causal effects.

- Previously we have not been very specific about the causal effect of interest; now we can.

- Ideally, the *causal contrast* we would like to estimate could be, say, \(Y_1 - Y_0\).

- However, the individual level causal effects are never identifiable (why?)

- This has been called the ‘fundamental problem of causal inference’ (Holland 1986).
Average causal effects

- Fortunately, average causal effects such as
  \[ E[Y_1] - E[Y_0] \]
  are identifiable.
- To see this, recall that by definition \( Y = ZY_1 + (1 - Z)Y_0 \)
  and under the experimental setting \((Y_0, Y_1) \perp \perp Z\).
- Now
  \[
  E[Y_1] - E[Y_0] = E[Y_1 | Z = 1] - E[Y_0 | Z = 0] \\
  = E[Y | Z = 1] - E[Y | Z = 0].
  \]
- These are the mean outcomes in the two treatment arms.
Non-experimental setting

- If we only have that \((Y_0, Y_1) \perp \perp Z \mid X\), we would also have to condition on \(X\):

\[
E[Y_1] - E[Y_0] = E_X\{E[Y_1 \mid X]\} - E_X\{E[Y_0 \mid X]\} = E_X\{E[Y_1 \mid Z = 1, X]\} - E_X\{E[Y_0 \mid Z = 0, X]\} = E_X\{E[Y \mid Z = 1, X]\} - E_X\{E[Y \mid Z = 0, X]\}.
\]

- This is known in epidemiology as the *direct standardization formula*.

- It averages the stratum-specific means \(E[Y \mid Z, X]\) over the distribution of \(X\) in the standard population.
Marginal vs. conditional effects

- Standardization is not needed if instead of the marginal/average effect $E[Y_1] - E[Y_0]$ we are interested in the conditional causal contrast

$$E[Y_1 | X] - E[Y_0 | X] = E[Y_1 | Z = 1, X] - E[Y_0 | Z = 0, X] = E[Y | Z = 1, X] - E[Y | Z = 0, X].$$

- Such effects can be obtained directly from regression models.

- Say, if

$$E[Y | Z, X] = \alpha + \beta Z + \gamma X,$$

then

$$E[Y | Z = 1, X] - E[Y | Z = 0, X] = \beta.$$
Non-compliance

- Under the experimental setting the variable $Z$ represents random assignment into one of the treatment arms.
- However, this does not mean that everyone assigned to a particular treatment will receive that treatment.
- For example, a randomized screening trial might involve inviting women to mammography screening, but not everyone invited will participate, and some of those not invited will get screened anyway.
- Thus, a causal contrast such as $E[Y_1] - E[Y_0]$ represents an intention-to-treat effect rather than the treatment effect as such.
- Rubin’s causal model can be extended to deal with non-compliance (Angrist, Imbens & Rubin, 1996).
Causal graphs
Rubin's causal model
Non-compliance and the IV estimator

A DAG allowing for non-compliance

- Suppose we have no measured covariates $X$, and the treatment assignment $Z$ is randomized.
- The compliance to the assigned treatment can depend on the latent individual-level characteristics $U$: 

$$
\begin{align*}
Y_{D} & \leftarrow U \\
Z & \rightarrow D \\
D & \rightarrow Y \\
Z & \rightarrow Y
\end{align*}
$$

- The variable $D$ indicates the treatment actually received.
- Note that there is no direct arrow $Z \rightarrow Y$ (why?)
A DAG allowing for non-compliance

- We need to introduce potential received treatment variables \((D_0, D_1)\), corresponding to the possible treatment assignments:
  - \(D_0\) is the received treatment when assigned to the control arm \((Z = 0)\).
  - \(D_1\) is the received treatment when assigned to the treatment arm \((Z = 1)\).
- Further, instead of the two potential outcomes \((Y_0, Y_1)\), we now have four, indexed by both the assigned and received treatment:
  - \(Y_{00}\) is the outcome under assignment to control arm and compliance \((D_0 = 0)\).
  - \(Y_{01}\) is the outcome under assignment to control arm and non-compliance \((D_0 = 1)\).
  - \(Y_{11}\) is the outcome under assignment to treatment arm and compliance \((D_1 = 1)\).
  - \(Y_{10}\) is the outcome under assignment to treatment arm and non-compliance \((D_1 = 0)\).
Exclusion restriction

- Fortunately, this can be simplified by noting that the treatment assignment in itself does not have an effect on the outcome, beyond the received treatment.
- This is expressed by the exclusion restriction assumption:

  \[ Y_{00} = Y_{10} \equiv Y_0^* \]

  and

  \[ Y_{01} = Y_{11} \equiv Y_1^*. \]

- We can define the causal effect of interest in terms of the received treatment alone, as

  \[ E[Y_1^*] - E[Y_0^*]. \]

- Compare this to the earlier intention-to-treat effect

  \[ E[Y_1] - E[Y_0] = E[Y_{1D_1}] - E[Y_{0D_0}]. \]
Different compliance types

- Problem: the received treatment is not randomized, even if the assignment is.
- We can now identify four different types of individuals, depending on the assigned and received treatment:

<table>
<thead>
<tr>
<th></th>
<th>$D_0 = 0$</th>
<th>$D_0 = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1 = 0$</td>
<td>Never-taker</td>
<td>Defier</td>
</tr>
<tr>
<td>$D_1 = 1$</td>
<td>Complier</td>
<td>Always-taker</td>
</tr>
</tbody>
</table>
In addition to strong ignorability, for this we need further two assumptions:

1. $Z$ has a causal effect on $D$, that is, $E[D_1] - E[D_0]$ is nonzero.
2. There are no defiers, that is, $D_1 \geq D_0$.

The IV estimator is given by

$$\hat{\beta}_{IV} = \frac{\text{cov}(Z, Y)}{\text{cov}(Z, D)}.$$
The IV estimand

- What does this estimate?
- The IV estimator has the form of causal effect $Z \rightarrow Y$ divided by the causal effect $Z \rightarrow D$, that is,

$$\frac{E[Y_1] - E[Y_0]}{E[D_1] - E[D_0]}.$$

- Angrist, Imbens & Rubin (1996) show that the IV estimator estimates the average causal effect in the subpopulation of compliers, that is,

$$E[Y^*_1 \mid D_1 - D_0 = 1] - E[Y^*_0 \mid D_1 - D_0 = 1].$$
References