Section 22

Lecture 7

Section 23

Dynamic regimes

Dynamic regimes

Definition (Dynamic regime)

A dynamic regime $g = (g_0, \ldots, g_k)$, where $g_k : (\overline{A}_{k-1}, \overline{L}_k) \mapsto A_k$, is a policy that assigns treatment (possibly at multiple time points) based on the measured history $(\overline{A}_{k-1}, \overline{L}_k)$.

We will restrict ourselves to settings where

$$g_k : (\overline{L}_k) \mapsto A_k$$

.

Dynamic regime SWIGs

Definition (d-SWIG from Robins and Richardson)

Given a template $\mathcal{G}(a)$ and a dynamic regime g for $\bar{a}$, the d-SWIG $\mathcal{G}(g)$ is defined by applying the following transformation:

- Replace each fixed node a_j with a random node A_j^{g+} that inherits children from a_j. Include dashed directed edges from every variable that is an input to the function g_i that determines the variable A_i^{g+}.
- Each random node V_i that is a descendant of at least one variable A_i^{g+} is relabeled as V_i^g.

Time-varying exposures (treatments) are frequent

Examples:

- Smoking status, which depends on other events in life.
- A therapeutic drug, for which the dose is adjusted according to the response over time (patients take the drug every day, every week etc)
- Cancer screening, which e.g. depends on previous diagnostic tests.
- Surgical interventions (e.g. transplants) are given at a certain time after the diagnosis.
- Expression of genes.

Running example: HIV

Consider a 5-year follow-up study of individuals infected with the human immunodeficiency virus (HIV)[30].

- A_k takes value 1 if the individual receives antiretroviral therapy in month k, and let $L = 0$ otherwise. Define $A_{-1} = 0$.
- Suppose Y measures health status at 5 years of follow-up.
- So far we have considered *deterministic* treatment rules, for example "always treat", where the outcome of interest is $Y^{a=1}$ vs "never treat", where the outcome of interest is $Y^{a=0}$.
  When $\overline{A} \equiv \overline{A}_K$, we can define 2^K such static regimes...
- However, often we want to make *dynamic* treatment decisions.
- Let $L_k \in \{0, 1\}$ be an indicator of low CD4 cell count measured at month k.
- Depending on the value of L_k, we may argue that it is good or bad to start treatment at time k.

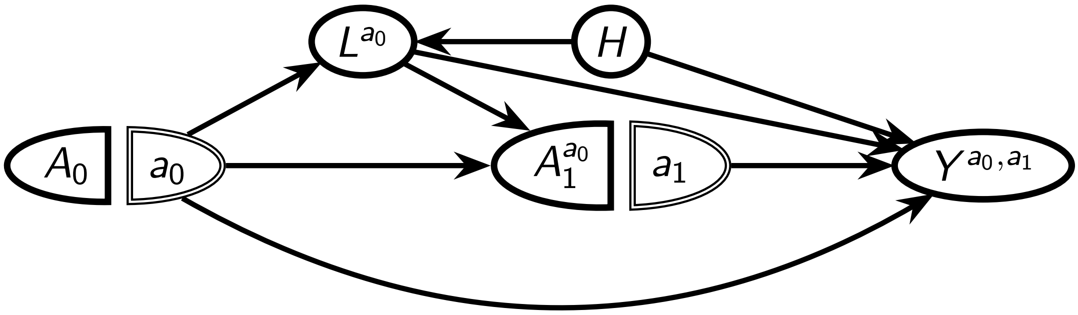
[30] Hernan and Robins, *Causal inference: What if?*

Example of Dynamic Regime

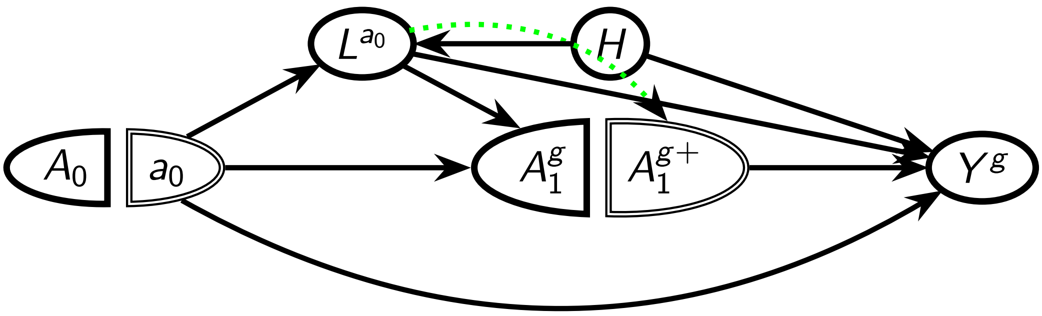
A simple example of a dynamic regime g for setting with two treatments is

- $A_0^{g+} = a_0$.
- $A_1^{g+} = L_1^{a_0}$
  In the HIV example this would mean that you are treated at time 1 if the CD4 cell count is low at that time.

Static vs dynamic



$Y^{a_0,a_1} \perp\!\!\!\perp A_0$ and $Y^{a_0,a_1} \perp\!\!\!\perp A_1^{a_0} \mid L_0^{a_0}, A_0$.



$Y^g \perp\!\!\!\perp A_0$ and $Y^g \perp\!\!\!\perp A_1^{a_0} \mid L_0^{a_0}, A_0$.
$Y^g \perp\!\!\!\perp A_0$ and, using the graph and consistency, $Y^g \perp\!\!\!\perp A_1 \mid L_0, A_0 = a_0$.

Sufficient condition for identification (Repetition of previous slide)

Theorem (Identification of static regimes)

Consider an intervention that sets $\overline{a} = \overline{a}_K = (a_0, \dots, a_K)$. Under positivity and consistency,

$$P(Y^{\overline{a}} = y) = b_{\overline{a}}(y)$$

if and only if for $k \in \{0, \dots, K\}$

$$Y^{\overline{a}} \perp\!\!\!\perp I(A_k = a_k) \mid L_0, \dots, L_k, A_0 = a_0, \dots, A_{k-1} = a_{k-1}.$$

This theorem follows from Robins[31] and Richardson and Robins[32], and is closely related to the backdoor theorem of Pearl[33]; Indeed, we can just call it "The SWIG backdoor criterion"

~~The theorem establishes when we can~~ use the g-formula to identify causal effects.

[31] Robins, "A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect".

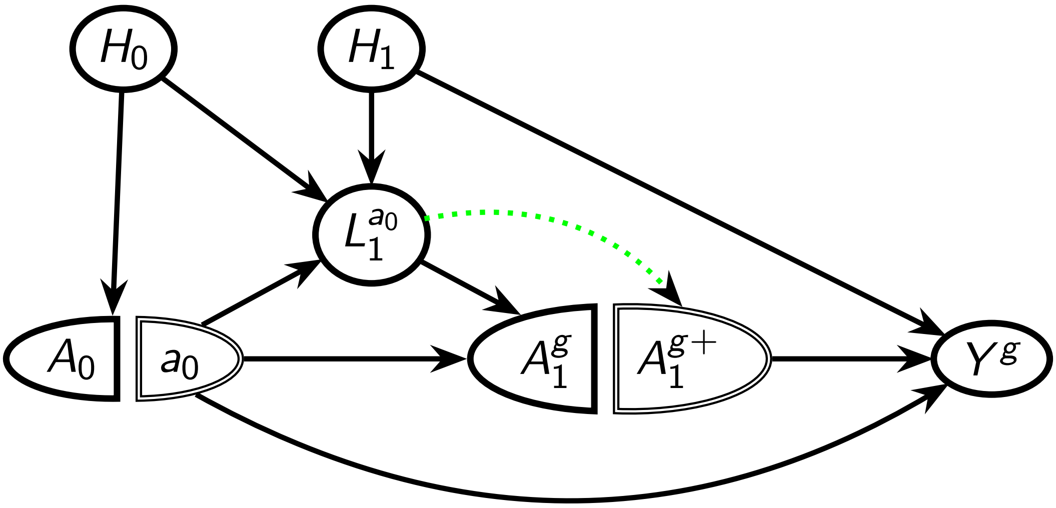
[32] Richardson and Robins, "Single world intervention graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality".

[33] Pearl, "Causal diagrams for empirical research".

Identification results for dynamic regimes

- We can use the same identification conditions (independencies in Slide 186) as for static regimes, only if A_k^{g+} **is not** a function of A_j^{g+} for $j < k$. However, we need to use the extended g-formula as the identification formula (as defined in Slide 207).
- if A_k^{g+} **is** a function of A_j^{g+} for any $j < k$, we need slightly stronger conditions (we are not presenting them now). This is e.g. the case in the graph in Slide 205 (due to the red arrow).

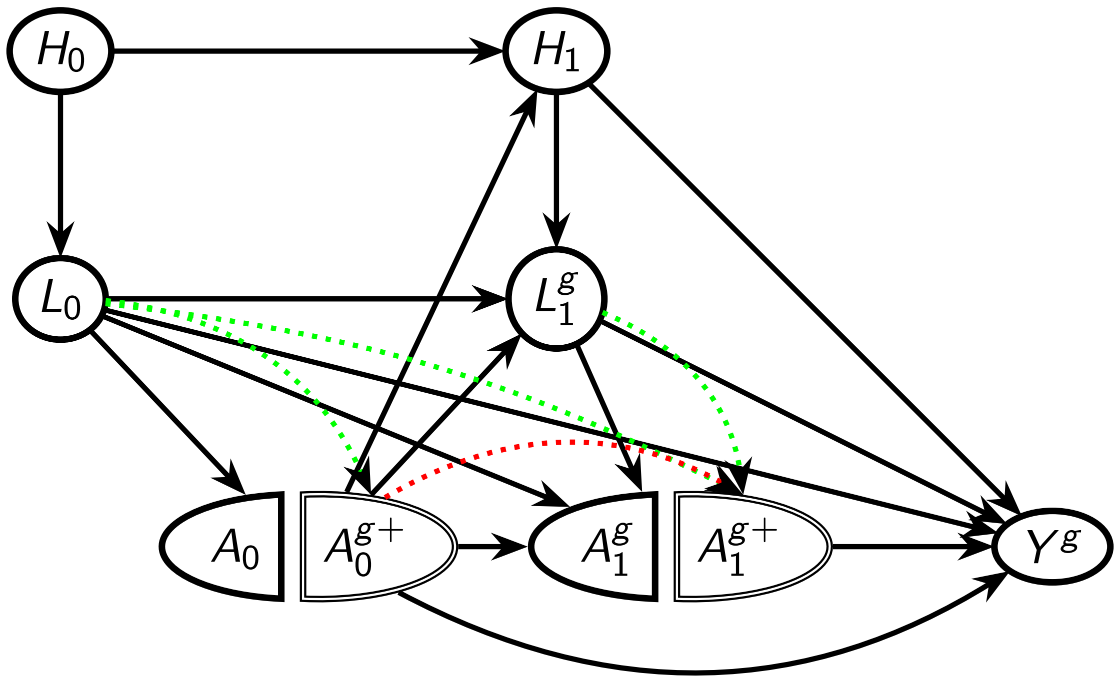
Does the identification conditions hold in the following Dynamic SWIG?



$Y^g \not\perp\!\!\!\perp A_0$ because $A_0 \leftarrow H_0 \rightarrow L_1^{a_0} \rightarrow A_1^{g+} \rightarrow Y^g$ is open. However, we would have identification in a static SWIG where $A_1^{g+} \equiv a_1$. So, in that sense, dynamic regimes require stronger conditions for identification, even though the independencies are stated in the same way.

HIV SWIG

A (busy) graph illustrating a conditional RCT, where H_0 and H_1 are hidden variables (e.g. the actual immune status of the patient).



Consistency for dynamic regimes

Now we generalize the consistency conditions such that it is valid for time-varying dynamic regimes. Indeed, it can simply be expressed as

$$\text{if } \overline{A}_K = \overline{A}_K^{g+}, \text{ then } Y^g = Y.$$

A special case for static regimes is if $\overline{A}_K = \overline{a}_K, \text{ then } Y^{\overline{a}_K} = Y$.

Marginal extended g-formula under interventions that depend on $\overline{L}_k$

Suppose that g_k is only a function of $\overline{L}_k$. Then, the marginal extended g-formula is defined as the following function of observed random variables $\overline{A}_K, \overline{L}_k, Y$.

Definition (Marginal extended g-formula)

$$b_g(y) = \sum_{\bar{a}_K} \sum_{\bar{l}_K} p(y \mid \bar{l}_K, \bar{a}_K) \prod_{j=0}^{K} p(l_j \mid \bar{l}_{j-1}, \bar{a}_{j-1}) p^g(a_j \mid \bar{l}_j),$$

where $\bar{l}_k = (l_0, \dots, l_k)$, $k \leq K$, are instantiations of **observed** variables and $p^g(a_j \mid \bar{l}_j)$ is the density of A_k^{g+} given $\overline{L}_k^g$, which is determined by g_k.

We let variables indexed by subscript -1, e.g. L_{-1} be empty.
Note that $p^g(a_k \mid \bar{l}_k)$ is a known function. It is determined by the investigator (even if it has a superscript g). If g_k is a deterministic function of $\bar{l}_k$, then

$$p^g(a'_k \mid \bar{l}_k) = \begin{cases} 1 \text{ if } a'_k = g_k(\bar{l}_k), \\ 0 \text{ if } a'_k \neq g_k(\bar{l}_k), \ k \in \{0, \dots, K\}. \end{cases}$$

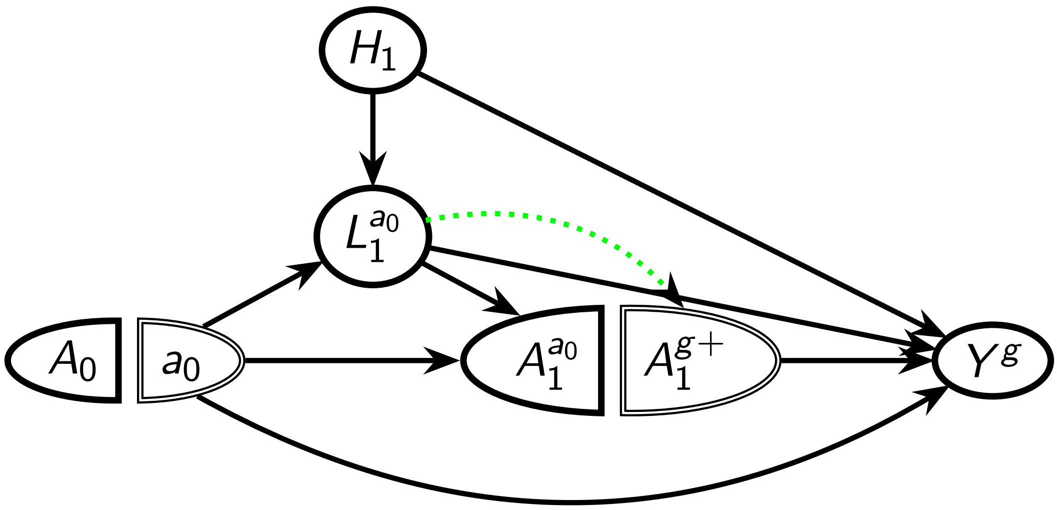
Relation to the g-formula for static regimes

The dynamic extended g-formula density generalizes the marginal g-formula from slide 186, because for a static intervention that sets $\overline{a} = (a_0, \ldots, a_K)$ we have that for $k \in \{0, \ldots, K\}$,

$$p^g(a'_k \mid \bar{l}_k) = \begin{cases} 1 \text{ if } a'_k = a_k, \\ 0 \text{ if } a'_k \neq a_k. \end{cases}$$

HIV example

Consider the example in Slide 199, and suppose the following SWIG:



let the dynamic regime g be

- $A_0^{g+} = a_0$.
- $A_1^{g+} = L_1^{a_0}$

That is, a patient is treated at time 1 if the CD4 cell count is low at that time.

HIV Example cont.

Then the g-formula reduces to

$$
\begin{aligned}
& b_g(y) \\
& = \sum_{\bar{a}'_1, l_1} p(y \mid A_1 = a'_1, L_1 = l_1, A_0 = a'_0) I(a'_1 = l_1) p(l_1 \mid A_0 = a_0) I(a'_0 = a_0), \\
& = \sum_{l_1} p(y \mid A_1 = l_1, L_1 = l_1, A_0 = a_0) p(l_1 \mid A_0 = a_0).
\end{aligned}
$$

because

$$
p^g(a'_1 \mid \bar{l}_1) = \begin{cases} 1 \text{ if } a'_1 = l_1, \\ 0 \text{ if } a'_1 \neq l_1. \end{cases}
$$

SWIG criterion to identify effects of dynamic regimes (you do not need to understand the extended g-formula density)

Definition (extended g-formula density)

The *dynamic extended g-formula density* for $Y \equiv Y_K$ under treatment regime g given by the functions $g_0, \dots, g_K$ that determine $\overline{A}_K = (A_0, \dots, A_K)$ is

$$f^g(y, \bar{l}_K, \bar{a}_K, \bar{a}_K^+) = p(y \mid \bar{l}_K, \bar{a}_K^+) \prod_{j=0}^{K} p(l_j, a_j \mid \bar{l}_{j-1}, \bar{a}_{j-1}^+) \prod_{t=0}^{K} p^g(a_t^+ \mid pa_{A_t^{g+}}),$$

where $\bar{l}_k = (l_0, \dots, l_k)$, $k \leq K$, are **observed** variables, $p^g(a_t^+ \mid pa_{A_t^{g+}})$ is the density of A_t^{g+} given $PA_{A_t^{g+}}$ is the input to g_t, for $t \in \{0, K\}$.

James M Robins. "A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect". In: *Mathematical modelling* 7.9-12 (1986), pp. 1393–1512; Thomas S Richardson and James M Robins. "Single world intervention graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality". In: *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper* 128.30 (2013).