STATISTICAL HORIZONS

Marginal Structural Models

Daniel Westreich, Ph.D.

Upcoming Seminar:
March 18-20, 2021, Remote Seminar

Introduction to time-fixed marginal structural models

Day 1, Part 3.

Overview

Concepts: what IS a marginal structural model?

Inverse probability weights

G-formula

Goals

Our goal is to estimate a marginal contrast in potential outcomes – that is, a causal effect.

Recall our 2x2 table of potential outcomes from the previous lecture.

Causal types

	$Y^{x=1} = 1$	$Y^{x=1} = 0$
$Y^{x=0} = 1$	A Doomed	B Protected
$Y^{x=0} = 0$	C Harmed	D Immune

Recall the fundamental problem of causal inference

If you observe one potential outcome, you can't observe any others.

So, can you ever identify where someone sits in this table? (Spoiler: no.)

	$Y^{x=1} = 1$	$Y^{x=1} = 0$
$Y^{x=0} = 1$	A Doomed	B Protected
$Y^{x=0} = 0$	C Harmed	D Immune

The best you can do is to identify which row or column they are in. Consider a trial. If I'm assigned x=0, and you observe my Y, you know whether I'm in row 1 (Yes outcome) or row 2 (No outcome). Same is true of x=1 and the columns.

That is, we can classify people into the margins of this table.

Causal types, extended to the margins

	$Y^{x=1} = 1$	$Y^{x=1} = 0$	Total
$Y^{x=0} = 1$	A Doomed	B Protected	A+B
$Y^{x=0} = 0$	C Harmed	D Immune	C+D
	A+C	B+D	N

So we can (at least theoretically) observe A+B, or A+C, etc. From here we can calculate the risk under x=0, and the risk under x=1.

Marginal risks and contrasts

	$Y^{x=1} = 1$	$Y^{x=1} = 0$	Total
$Y^{x=0} = 1$	Doomed	Protected	A+B
$Y^{x=0} = 0$	Harmed	Immune	C+D
	A+C	B+D	N

Specifically, the risk under x=1 is (A+C)/((A+C)+(B+D)) = (A + C) / N

The risk under x=0 is (similarly) = (A + B) / N

Causal risk difference is (A+C)/N – (A+B)/N = (C - B) / N

Causal risk ratio = (A+C) / (A+B)

These last two are examples of contrasts between marginal potential outcomes.

The problem

Now, in particular, these risks are risks over all of N – the whole population.

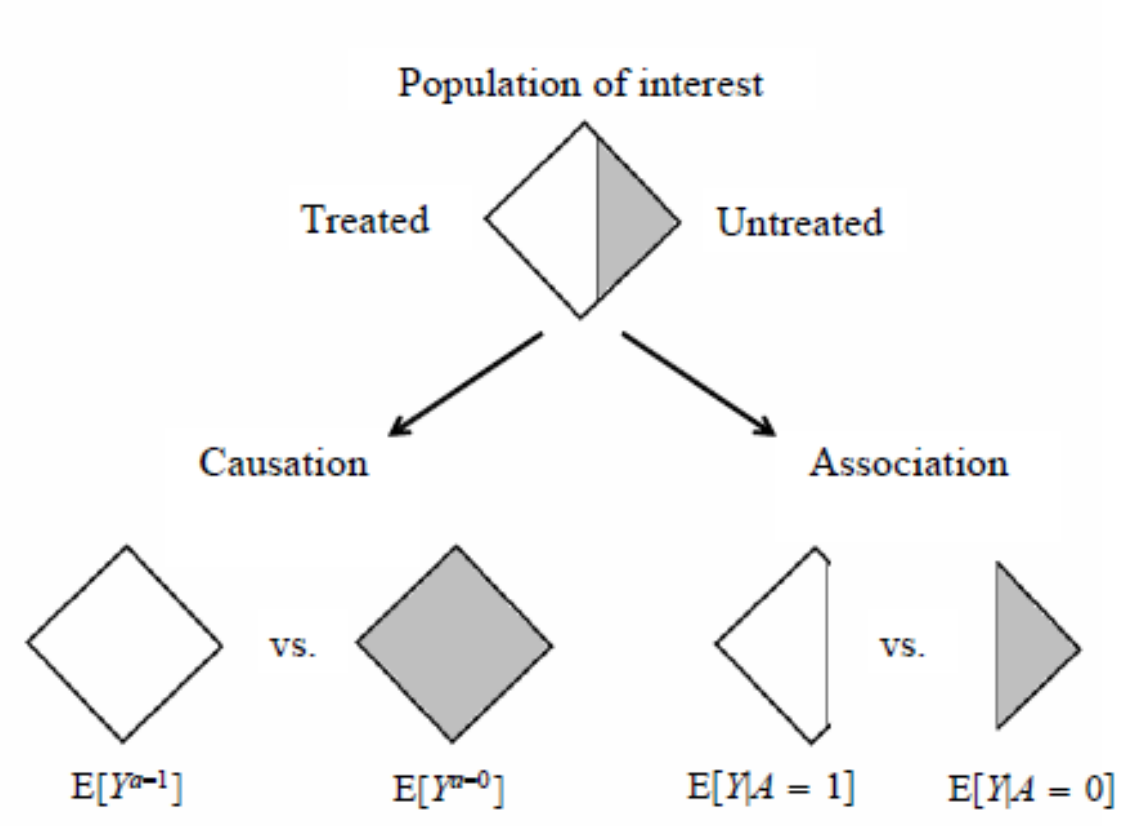
In a real study, you don't observe the whole population under x=0 and also under x=1. Take a trial. We might observe half the population under each value of x, if assignment of x is a coinflip.

However, under exchangeability, then the experiences of those with x=0 can stand-in for the experiences of x=1 if, counter to fact, those with x=1 had truly had x=0. And vice-versa.

That is,

We begin with some treated and some untreated. Exchangeability allows us to interpret the association (right) as the causal effect (left). The causal effect (left) is the comparison of the marginal risks from our potential outcomes table, previously.

Image from Hernán & Robins, *Causal Inference*. To be published.



A marginal structural model

A model (aka a summary of data, in the way a risk difference is a summary of survival curves) for the marginal potential outcomes.

The structural refers to the potential outcomes: “They are structural models, because they model the probabilities of counterfactual variables and in the econometric and social science literature models for counterfactual variables are often referred to as structural”

– Robins, Hernán, Brumback *Epidemiology* 2000.

Thus, these are explicitly causal models.

Robins et al. Epidemiology 2000

Just discussed.

Briefly, Robins introduces the things on the left as marginal structural models for a point treatment.

$$pr[Y_{a_0} = 1] = \psi_0 + \psi_1 a_0 \quad (1)$$

$$\log pr[Y_{a_0} = 1] = \theta_0 + \theta_1 a_0 \quad (2)$$

$$logit\ pr[Y_{a_0} = 1] = \beta_0 + \beta_1 a_0 \quad (3)$$

These things on the right are what we can estimate from observed data.

Robins says, “The parameters of the associational models 4–6 will differ from the parameters of the MSMs 1–3, except when treatment is unconfounded.”

$$pr[Y = 1|A_0 = a_0] = \psi_0' + \psi_1' a_0 \quad (4)$$

$$\log pr[Y = 1|A_0 = a_0] = \theta_0' + \theta_1' a_0 \quad (5)$$

$$logit\ pr[Y = 1|A_0 = a_0] = \beta_0' + \beta_1' a_0. \quad (6)$$

Associational models can be estimated from data

That is, we can estimate parameters of models (4), (5), and (6) from data.

If there is no confounding (that is, if exchangeability holds) then for example, $e^{\theta_1'}$ (the associational risk ratio) can be interpreted as e^{θ_1} (the causal risk ratio).

When will there be no confounding? Generally, never – except in a randomized trial.

Thus, a randomized trial is one way to estimate the parameters of a marginal structural model.

MSMs & trials

In fact, the goal of a trial is to estimate the causal effect of some treatment assignment!

There is some confusion in the literature about this point: people seem to think, broadly, that if you're fitting a marginal structural model, you are necessarily using some "advanced" technique like inverse probability weights or the parametric g-formula. This is not true!

In fact, so long as the target of our estimation is a contrast in potential outcomes, it is reasonable to call what we're doing a marginal structural model.

Fancy methods? No: again, Robins says this in his paper: "to fit models 4–6, one could use the general-purpose SAS program Proc Genmod."

Estimating marginal causal contrasts

Obviously, use a trial if you can. But you often can't.

So how do we estimate in the presence of confounding?

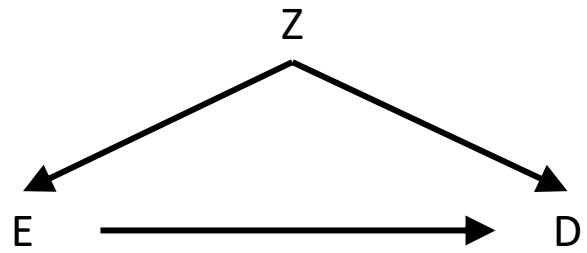
Again, regression will do! But that's not what you're here to learn. And also, while regression will do in the simple cases (time-fixed exposure), we want something that will scale to complex (time-varying exposure) cases.

We'll discuss two methods in the remainder of this lecture: inverse probability weights, and the g-formula. We'll introduce them in time-fixed settings.

Example

Before we introduce methods, however, we give you an example.

Consider the following causal diagram:



E for exposure, D for disease outcome, Z is a covariate (confounder).