Causal Inference

Miguel A. Hernán, James M. Robins

Chapter 4

EFFECT MODIFICATION

- So far we have focused on the **average causal effect in an entire population** of interest. However, many causal questions are about subsets of the population.
- You will be interested in characterizing how the causal effect varies across subsets of the population when the intervention can be targeted to different subsets, or when the findings of the study need to be applied to other populations.
- This chapter emphasizes that there is not such a thing as **the causal effect** of treatment. Rather, the causal effect depends on the characteristics of the particular population under study.

4.1 Definition of effect modification

- We say that M is a modifier of the effect of A on Y when the average causal effect of A on Y varies across levels of M.
- Since the average causal effect can be measured using different effect measures (e.g., risk difference, risk ratio), the presence of effect modification depends on the effect measure being used:

Additive effect modification:

$$\mathrm{E}[Y^{a=1} - Y^{a=0} | M = 1] \neq \mathrm{E}[Y^{a=1} - Y^{a=0} | M = 0]$$

Multiplicative effect modification:

$$\frac{\mathrm{E}[Y^{a=1}|M=1]}{\mathrm{E}[Y^{a=0}|M=1]} \neq \frac{\mathrm{E}[Y^{a=1}|M=0]}{\mathrm{E}[Y^{a=0}|M=0]}$$

Table 4.1

	M	Y^0	Y^1
Rheia	1	0	1
Demeter	1	0	0
Hestia	1	0	0
Hera	1	0	0
Artemis	1	1	1
Leto	1	0	1
Athena	1	1	1
Aphrodite	1	0	1
Persephone	1	1	1
Hebe	1	1	0
Kronos	0	1	0
Hades	0	0	0
Poseidon	0	1	0
Zeus	0	0	1
Apollo	0	1	0
Ares	0	1	1
Hephaestus	0	0	1
Cyclope	0	0	1
Hermes	0	1	0
Dionysus	0	1	0

- A null average causal effect in the population does not imply a null average causal effect in a particular subset of the population.
- Heart transplant increases the risk of death in women and decreases the risk in men.

- We say that there is **qualitative effect modification** because the average causal effects in the subsets M=1 and M=0 are in the opposite direction.
- In the presence of qualitative effect modification, additive effect modification implies multiplicative effect modification, and vice versa.
- In the absence of qualitative effect modification, however, one can find effect modification on one scale (e.g., multiplicative) but not on the other (e.g., additive).

Multiplicative, but not additive, effect modification by M:

$$\Pr[Y^{a=0} = 1|M = 1] = 0.8$$
$$\Pr[Y^{a=1} = 1|M = 1] = 0.9$$
$$\Pr[Y^{a=0} = 1|M = 0] = 0.1$$
$$\Pr[Y^{a=1} = 1|M = 0] = 0.2$$

4.2 Stratification to identify effect modification

- A stratified analysis is the natural way to identify effect modification.

Stratification: the causal effect of A on Y is computed in each stratum of M. For dichotomous M, the stratified causal risk differences are:

$\Pr[Y^{a=1} = 1|M = 1] -$
$\Pr[Y^{a=0} = 1|M = 1]$

and

$\Pr[Y^{a=1} = 1|M = 0] -$
$\Pr[Y^{a=0} = 1|M = 0]$

- How does the unavailability of the counterfactual outcomes affect the use of stratification to detect effect modification?

The answer depends on the study design.

Ideal marginally randomized experiment:

we demonstrated that, leaving aside random variability, the average causal effect of treatment can be computed using the observed data. For example, the causal risk difference $\Pr[Y^{a=1} = 1] - \Pr[Y^{a=0} = 1]$ is equal to the observed associational risk difference $\Pr[Y = 1|A = 1] - \Pr[Y = 1|A = 0]$. The same reasoning can be extended to each stratum of the variable M because, if treatment assignment was random and unconditional, exchangeability is expected in every subset of the population. Thus the causal risk difference in women,

Ideal randomized experiment with conditional randomization:

The procedure to compute the conditional risks $\Pr[Y^{a=1} = 1|M = m]$ and $\Pr[Y^{a=0} = 1|M = m]$ in each stratum m has two stages: 1) stratification by M, and 2) standardization by L (or, equivalently, IP weighting). We computed

Table 2.2

	L	A	Y
Rheia	0	0	0
Kronos	0	0	1
Demeter	0	0	0
Hades	0	0	0
Hestia	0	1	0
Poseidon	0	1	0
Hera	0	1	0
Zeus	0	1	1
Artemis	1	0	1
Apollo	1	0	1
Leto	1	0	0
Ares	1	1	1
Athena	1	1	1
Hephaestus	1	1	1
Aphrodite	1	1	1
Cyclope	1	1	1
Persephone	1	1	1
Hermes	1	1	0
Hebe	1	1	0
Dionysus	1	1	0

Table 4.2
Stratum $M = 0$

	L	A	Y
Cybele	0	0	0
Saturn	0	0	1
Ceres	0	0	0
Pluto	0	0	0
Vesta	0	1	0
Neptune	0	1	0
Juno	0	1	1
Jupiter	0	1	1
Diana	1	0	0
Phoebus	1	0	1
Latona	1	0	0
Mars	1	1	1
Minerva	1	1	1
Vulcan	1	1	1
Venus	1	1	1
Seneca	1	1	1
Proserpina	1	1	1
Mercury	1	1	0
Juventas	1	1	0
Bacchus	1	1	0

- there is both additive and multiplicative effect modification by nationality M of the effect of transplant A on death Y
- In fact, it is possible that nationality is simply a marker for the factor that is truly responsible for the effect modification. For example, suppose that the quality of heart surgery is better in Greece than in Rome. We refer to nationality as a surrogate effect modifier, and to quality of care as a causal effect modifier.

- effect modification by M does not necessarily imply that M plays a causal role.

- To avoid potential confusions, some authors prefer to use the more neutral term "heterogeneity of causal effects across strata of M" rather than "effect modification by M." The next chapter introduces "interaction", a concept related to effect modification, that does attribute a causal role to the variables involved.

4.3 Reasons to care about effect modification

- First, if a factor M modifies the effect of treatment A on the outcome Y then the average causal effect will differ between populations with different prevalence of M.
- There is generally no such a thing as "the average causal effect of treatment A on outcome Y (period)", but "the average causal effect of treatment A on outcome Y in a population with a particular mix of causal effect modifiers."
- Evaluating the presence of effect modification is helpful to identify the groups of subjects that would benefit most from an intervention.
- Additive, but not multiplicative, effect modification is the appropriate scale to identify the groups that will benefit most from intervention.

4.4 Stratification as a form of adjustment

- Standardization (or IP weighting) is used to adjust for L and stratification is used to identify effect modification by M.
- But stratification is not always used to identify effect modification by M. In practice stratification is often used as an alternative to standardization (and IP weighting) to adjust for L.
- Stratification necessarily results in multiple stratum-specific effect measures (one per stratum defined by the variables L). Each of them quantifies the average causal effect in a non overlapping subset of the population but, in general, none of them quantifies the average causal effect in the entire population. Therefore, we did not consider stratification when describing methods to compute the average causal effect of treatment in the population in Chapter 2. Rather, we focused on standardization and IP weighting.
- Unlike standardization and IP weighting, adjustment via stratification requires computing the effect measures in subsets of the population defined by a combination of all variables L that are required for conditional exchangeability.

4.5 Matching as another form of adjustment

- Matching is another adjustment method. The goal of matching is to construct a subset of the population in which the variables L have the same distribution in both the treated and the untreated.
- To construct our matched population we replaced the treated in the population by a subset of the treated in which the matching factor L had the same distribution as that in the untreated.
- Often one chooses the group with fewer subjects (the untreated in our example) and uses the other group (the treated in our example) to find their matches. The chosen group defines the subpopulation on which the causal effect is being computed. In the previous paragraph we computed the effect in the untreated. In settings with fewer treated than untreated individuals across all strata of L, we generally compute the effect in the treated.

4.6 Effect modification and adjustment methods

- Standardization, IP weighting, stratification and matching are different approaches to estimate average causal effects, but they estimate different types of causal effects.
- These four approaches can be divided into two groups according to the type of effect they estimate: standardization and IP weighting can be used to compute either marginal or conditional effects, stratification and matching can only be used to compute conditional effects in certain subsets of the population.
- All four approaches require exchangeability, positivity, and well-defined interventions, but the subsets of the population in which these conditions need to hold depend on the causal effect of interest.
- We have computed four causal risk ratios and have obtained four different numbers: 0.8, 2.0, 0.5, and1.0. All of them are correct.

- Well defined causal effect and a well characterized target population are prerequisites for meaningful causal inference.
- Both prerequisites are automatically present in experiments that compare two or more interventions in a population that meets certain a priori eligibility criteria.
- However, these prerequisites cannot be taken for granted in observational studies.
- Rather, investigators conducting observational studies need to explicitly define the causal effect of interest and the subset of the population in which the effect is being computed.
- Otherwise, misunderstandings might easily arise when effect measures obtained via different methods are different.