Causal Inference Through the Structural Causal Marginal Problem

Luigi Gresele[*,1], Julius von Kügelgen[*,1,2], Jonas M. Kübler[*,1],
Elke Kirschbaum[3], Bernhard Schölkopf[1], Dominik Janzing[3]

[1] Empirical Inference Department, Max Planck Institute for Intelligent Systems, Tübingen
[2] Department of Engineering, University of Cambridge
[3] Amazon Research Tübingen
[*] Co-first authors
Part of this work was done while L.G. and J.M.K. were interning at Amazon.

MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

UNIVERSITY OF CAMBRIDGE

aws

Counterfactual questions

- A patient, Alice, is recommended a treatment X against her disease and agrees to take it.
- The effectiveness of the treatment has been rigorously established through a randomised control trial, which found a positive average causal effect (ACE).
- However, **the ACE is an average of treatment efficacy over the whole population, including some individuals who respond better and others who respond worse.**
- **Alice might wonder what *her own* chances of recovery would have been, had she not taken** X (Heckman, 1992; Shpitser and Pearl, 2009).
- This requires envisioning **consequences of a hypothetical change** (not taking the treatment), **given that the opposite happened** (in reality, she took it): a **counterfactual**.

Empirical content of counterfactuals

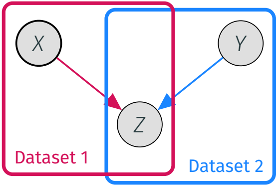
- Counterfactuals can be expressed through the framework of structural causal models (SCMs) (Pearl, 2000).
- However, we typically do not have access to an SCM but only to observational or experimental data which may be insufficient: **In general, we cannot unambiguously answer counterfactual questions based on empirical observations**.
- We simply cannot perform an experiment where the same person is both given and not given a treatment, an issue also referred to as the *fundamental problem of causal inference* (Imbens and Rubin, 2015).

Identifiability of counterfactual expressions

- Pearl (2011): Counterfactual expressions should only be evaluated when they can be estimated based on empirical observations $\rightarrow$ *Identifiability* requirement (Shpitser and Pearl, 2007, 2008; Pearl, 2001; Correa et al., 2021).
- When full identification is not achievable, informative bounds can sometimes still be given based on empirically observable quantities $\rightarrow$ *Partial Identification* (Balke and Pearl, 1994).
- In these results **all the considered variables are jointly observed** (Bareinboim and Pearl, 2016).

What if the variables are not *jointly observed?*

- What if we instead have studies involving distinct, but overlapping subsets of variables?



- In Alice's case, suppose that a separate study characterises the interventional effect of a rare condition Y on her disease.
- Since the condition is rare, and testing for it is costly, there are no studies characterising the joint effect of X and Y on recovery.
- **Could Alice nevertheless make use of the available information on the effect of Y and combine it with information on X to better answer her counterfactual question?**

Structural Causal Marginal Problem

Our goal: (Partial) identification of counterfactual models by merging information from different datasets, involving *distinct but overlapping sets of variables*.

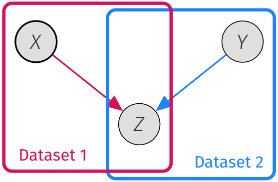
Causal reformulation of the *marginal problem* in statistics (Vorob'ev, 1962; Kellerer, 1964).

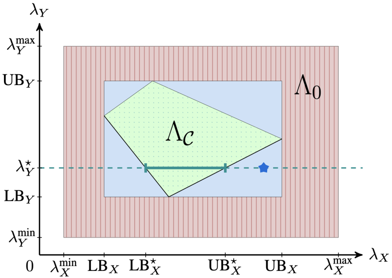
> **(Statistical) Marginal Problem:** *Given some distributions over non-identical but overlapping subsets of variables, determine existence and uniqueness of a consistent joint distribution over their union.*

***Structural causal* marginal problem**: We want to merge *marginal SCMs* s.t. the marginal and joint SCMs are *counterfactually consistent*.

We formalise counterfactual consistency in the context of categorical SCMs and assuming causal sufficiency.

2D Schematic of the Structural Causal Marginal Problem





Enforcing consistency reduces the space of admissible marginal SCMs.

- Λ_0 (outer red area): combinations of counterfactual marginal models that cannot be counterfactually consistent;
- $\Lambda_{\mathcal{C}}$ (green dotted area): (λ_X, λ_Y) that are counterfactually consistent;
- Solid blue area: (λ_X, λ_Y) that are not counterfactually consistent but cannot be falsified without additional assumptions or constraints; e.g., **knowing one of the marginal SCMs exactly further restricts the choices for the other marginal** (horizontal green line; blue star can be ruled out).

Overview

- We introduce an approach to counterfactual inference based on merging information from multiple datasets.
- This can be seen as the causal reformulation of a classic problem in statistics called the *marginal problem* (Vorob'ev, 1962; Kellerer, 1964).
- We show that counterfactuals can acquire empirical content when considered in the broader context of a joint model, *even if only observations of the marginal models are available*.
- While focusing mostly on simple examples, the present work still makes a significant conceptual point: SCMs can sometimes be falsified as interventional models over additional *variables* become available.

Thank you for your attention &
see you at the poster!

References

A. Balke and J. Pearl. Counterfactual probabilities: Computational methods, bounds and applications. In *Uncertainty Proceedings 1994*, pages 46–54. Elsevier, 1994.

E. Bareinboim and J. Pearl. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, 113(27):7345–7352, 2016.

J. D. Correa, S. Lee, and E. Bareinboim. Nested counterfactual identification from arbitrary surrogate experiments. *arXiv preprint arXiv:2107.03190*, 2021.

J. J. Heckman. Randomization and social policy evaluation. *Evaluating welfare and training programs*, 1:201–30, 1992.

G. W. Imbens and D. B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, 2015.

H. Kellerer. Maßtheoretische Marginalprobleme. *Math. Ann.*, 153:168–198, 1964. in German.

J. Pearl. *Causality: Models, reasoning, and inference*. Cambridge University Press, 2000.

References ii

J. Pearl. Direct and indirect effects. In *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 411–420, San Francisco, CA, 2001. Morgan Kaufmann.

J. Pearl. The logic of counterfactuals in causal inference. 2011.

I. Shpitser and J. Pearl. What counterfactuals can be tested. In *23rd Conference on Uncertainty in Artificial Intelligence, UAI 2007*, pages 352–359, 2007.

I. Shpitser and J. Pearl. Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research*, 9:1941–1979, 2008.

I. Shpitser and J. Pearl. Effects of treatment on the treated: Identification and generalization. In *25th Conference on Uncertainty in Artificial Intelligence*, 2009.

N. N. Vorob'ev. Consistent families of measures and their extensions. *Theory of Probability & Its Applications*, 7(2):147–163, 1962.