Probabilities of Causation: Bounds from Incomplete and Subpopulation Data

Yuxuan Xie
edithxie944@gmail.com

Ang Li
Department of Computer Science
Florida State University
angli@cs.fsu.edu

Abstract

Probabilities of causation (PoCs) play a central role in individual-level explanation and decision making. Tian and Pearl showed that PoCs are in general not point-identifiable and derived tight bounds for three fundamental PoCs using observational and experimental data. However, in real-world applications, such data are often incomplete; for instance, data may only be available for a subpopulation. In this paper, we show how informative bounds on fundamental PoCs can be obtained from incomplete data. We further investigate whether informative bounds on the PoCs of one subpopulation can be derived given data from another subpopulation; that is, we study the transportability of PoC bounds.

1 Introduction

Causal reasoning at the individual level plays an important role in many consequential decisions in medicine, law, and public policy [Lu and Boutilier, 2011, Li and Pearl, 2019, Mueller and Pearl, 2023]. When a patient recovers after receiving a treatment, was the treatment actually responsible for the recovery? When a harmful outcome is avoided after a policy intervention, would that outcome have happened otherwise? Questions of this kind go beyond what population-level causal effects can answer, even when those effects are identified cleanly from randomized experiments.

Probabilities of causation (PoCs) [Pearl, 1999, Galles and Pearl, 1998, Halpern, 2000, Tian and Pearl, 2000] provide a principled way to study such questions within the Structural Causal Model (SCM) framework [Pearl, 2009]. The three most commonly studied quantities are the probability of necessity (PN), the probability of sufficiency (PS), and the probability of necessity and sufficiency (PNS) [Pearl, 1999]. Each captures a different aspect of individual-level causation and has been used in applications such as treatment prioritization, legal reasoning, and personalized decision making.

A fundamental challenge in this field is that PoCs rely on joint counterfactual events that can never be observed simultaneously. Because of this, these quantities are generally not point-identifiable from data alone. In practice, deriving informative bounds is often the more realistic goal [Tian and Pearl, 2000, Mueller et al., 2022, Li and Pearl, 2024].

Tian and Pearl showed that tight bounds for PN, PS, and PNS can be derived using Balke's linear programming framework [Balke, 1995] by combining experimental and observational data without introducing assumptions on the underlying data-generating process [Tian and Pearl, 2000]. Later work demonstrated that additional structural information, such as a fully specified causal diagram, can further tighten these bounds [Mueller et al., 2022, Zhang et al., 2022]. Intuitively, structural knowledge rules out counterfactual distributions that would otherwise remain compatible with the observed data.

Despite the theoretical progress on PoCs, most existing methods still assume access to ideal data collected from the target population itself. In practice, however, this assumption is often unrealistic.

Preprint.

In many real-world applications, available data may only come from a subpopulation rather than from the full population of interest.

1.1 Motivating Example

Consider a vaccine designed to prevent a serious disease. Beyond estimating whether the vaccine lowers the disease rate on average, policymakers may also want to answer a more individual-level question: for how many people was the vaccine actually responsible for preventing the disease? This corresponds to the probability of necessity and sufficiency (PNS), namely, the proportion of individuals who would remain healthy if vaccinated but would become sick if they had not received the vaccine.

In practice, however, obtaining the right data for this question can be difficult. Suppose the goal is to estimate the PNS for the entire population, including both adults and youths. Yet the available observational data mainly come from adults because youths are substantially less likely to receive the vaccine. As a result, treatment decisions in the youth population are rarely observed, and the collected data no longer reflect the behavior of the full target population.

Under this setting, directly applying Tian and Pearl's bounds to the entire population becomes challenging. One possible approach is to ignore the observational data altogether and rely only on experimental data, but this often leads to much looser and less informative bounds. At the same time, the partial observational data from the adult population still contain useful information that should not be discarded. This naturally raises several questions. Can we still leverage observational data from adults to tighten the bounds for the overall population? What if the experimental data are also collected only from adults? More broadly, given information from one subpopulation, can we say anything meaningful about another subpopulation, such as youths for whom little or no data are available? This question is closely related to transportability, which has been extensively studied for causal effects by Elias Bareinboim and Judea Pearl [Pearl and Bareinboim, 2022, 2011, Bareinboim and Pearl, 2013, 2012].

In this paper, we develop tight bounds for probabilities of causation, including PN, PS, PNS, and the effect of treatment on the treated (ETT), another important causal quantity introduced by Shpitser and Pearl [2009], under such settings involving subpopulation and distributional mismatch.

2 Preliminaries and Related Work

In this section, we review the definitions of ETT, PN, PS, and PNS introduced by Pearl and related works [Pearl, 1999, Shpitser and Pearl, 2009]. We also review existing results on the tight bounds of these quantities, particularly the classical bounds for PN, PS, and PNS derived by Tian and Pearl [Tian and Pearl, 2000]. Readers who are already familiar with these concepts may skip this section.

Following the literature above, we adopt the language of Structural Causal Models (SCMs) [Galles and Pearl, 1998, Halpern, 2000]. In this framework, the counterfactual statement "Variable Y would have value y had X been set to x" is denoted by $Y_x = y$, and abbreviated as y_x. Throughout the paper, we consider two sources of information: experimental data, represented through interventional distributions such as $P(y_x)$, and observational data, represented through the joint distribution $P(x, y)$. Note that we assume the data accurately represent the underlying distribution; sample error is beyond the scope of this paper. Readers may refer to [Li et al., 2022] for additional requirements on sample size.

Unless otherwise specified, let X and Y be binary variables in a causal model M, where x and y denote the events $X = true$ and $Y = true$, respectively, while x' and y' denote their complements. For simplicity of presentation, we focus on binary treatments and outcomes, following the standard formulation in the literature. Extensions to multi-valued variables are discussed in [Pearl, 2009].

We next review the definitions of several commonly studied probabilities of causation.

Definition 2.1 (Probabilities of Causation)**.** Let X and Y be binary variables in a causal model M. The effect of treatment on the treated (ETT), probability of necessity (PN), probability of sufficiency (PS), and probability of necessity and sufficiency (PNS) are respectively defined as

$$\text{ETT} \triangleq P(y_x|x'), \text{PN} \triangleq P(y'_{x'}|x, y), \text{PS} \triangleq P(y_x|x', y'), \text{PNS} \triangleq P(y_x, y'_{x'}).$$

For subpopulations defined by a set of observed non-descendant characteristics C, the c-specific ETT, PN, PS, and PNS are defined as follows:

$$c\text{-ETT} \triangleq P(y_x|x',c), c\text{-PN} \triangleq P(y'_{x'}|x,y,c), c\text{-PS} \triangleq P(y_x|x',y',c), c\text{-PNS} \triangleq P(y_x,y'_{x'}|c). \quad (1)$$

Tian and Pearl [2000] derived tight bounds for PNS, PN, and PS using a linear programming formulation that does not impose assumptions on the underlying causal structure beyond consistency with the available experimental and observational data. The approach optimizes over all admissible counterfactual distributions, following the framework introduced by Balke [Balke, 1995]. Later, Li and Pearl [2019] provided a theoretical proof of the validity and tightness of these bounds.

The tight bounds for PNS, PN, and PS are given by

$$\max\left\{\begin{array}{c} 0, \\ P(y_x)-P(y_{x'}), \\ P(y)-P(y_{x'}), \\ P(y_x)-P(y) \end{array}\right\} \le \text{PNS} \le \min\left\{\begin{array}{c} P(y_x), \\ P(y'_{x'}), \\ P(x,y)+P(x',y'), \\ P(y_x)-P(y_{x'})+P(x,y')+P(x',y) \end{array}\right\} \quad (2)$$

$$\max\left\{\begin{array}{c} 0, \\ \frac{P(y)-P(y_{x'})}{P(x,y)} \end{array}\right\} \le \text{PN} \le \min\left\{\begin{array}{c} 1, \\ \frac{P(y'_{x'})-P(x',y')}{P(x,y)} \end{array}\right\} \quad (3)$$

$$\max\left\{\begin{array}{c} 0, \\ \frac{P(y_x)-P(y)}{P(x',y')} \end{array}\right\} \le \text{PS} \le \min\left\{\begin{array}{c} 1, \\ \frac{P(y_x)-P(x,y)}{P(x',y')} \end{array}\right\} \quad (4)$$

ETT is one of the few counterfactual quantities for which general identification results are available [Shpitser and Pearl, 2009].

Theorem 2.2. *The ETT of X on Y are identifiable if $P(Y_X)$ are identifiable and given by the following formula:*

$$P(y_x|x') = \frac{P(y_x)-P(x,y)}{P(x')}. \quad (5)$$

The results above assume access to the full experimental distribution $P(Y_x)$ and observational distribution $P(X,Y)$. In this work, we study a different setting in which a non-descendant binary covariate C (e.g., gender) is observed, but only the full distributions $P(Y_x|c)$, $P(X,Y|c)$, and $P(c)$ are available, while only partial or no information about $P(Y_x|c')$ and $P(X,Y|c')$ is known.

Under this setting, we derive tight bounds for PNS, PS, PN, and ETT, where ETT is generally no longer identifiable. In addition, we investigate under what conditions the subgroup-specific quantities for c', including PNS, PS, PN, and ETT, can still be meaningfully bounded.

3 Main Results

Again, assume the population consists of two subpopulations, c and c'. We assume that the full distributions $P(Y_x|c)$, $P(X,Y|c)$, and $P(c)$ are available, while only partial or no information about $P(Y_x|c')$ and $P(X,Y|c')$ is known. We will present our results under different levels of data availability for the subpopulation c'.

3.1 Linear Programming Formulation

We first introduce the Linear Programming (LP) formulation that leads to the following theorems. We consider four binary variables, X, Y_x, $Y_{x'}$, and C (note that Y is determined once X, Y_x, and $Y_{x'}$ are fixed due to the consistency rule). Therefore, the LP contains the following 16 counterfactual distribution parameters:

$$
\begin{aligned}
p_{1111} &= P(y_x, y_{x'}, c, x) = P(x, y, c, y_{x'}) & p_{1110} &= P(y_x, y_{x'}, c, x') = P(x', y, c, y_x) \\
p_{1101} &= P(y_x, y_{x'}, c', x) = P(x, y, c', y_{x'}) & p_{1100} &= P(y_x, y_{x'}, c', x') = P(x', y, c', y_x) \\
p_{1011} &= P(y_x, y'_{x'}, c, x) = P(x, y, c, y'_{x'}) & p_{1010} &= P(y_x, y'_{x'}, c, x') = P(x', y', c, y_x) \\
p_{1001} &= P(y_x, y'_{x'}, c', x) = P(x, y, c', y'_{x'}) & p_{1000} &= P(y_x, y'_{x'}, c', x') = P(x', y', c', y_x) \\
p_{0111} &= P(y'_x, y_{x'}, c, x) = P(x, y', c, y_{x'}) & p_{0110} &= P(y'_x, y_{x'}, c, x') = P(x', y, c, y'_x) \\
p_{0101} &= P(y'_x, y_{x'}, c', x) = P(x, y', c', y_{x'}) & p_{0100} &= P(y'_x, y_{x'}, c', x') = P(x', y, c', y'_x) \\
p_{0011} &= P(y'_x, y'_{x'}, c, x) = P(x, y', c, y'_{x'}) & p_{0010} &= P(y'_x, y'_{x'}, c, x') = P(x', y', c, y'_x) \\
p_{0001} &= P(y'_x, y'_{x'}, c', x) = P(x, y', c', y'_{x'}) & p_{0000} &= P(y'_x, y'_{x'}, c', x') = P(x', y', c', y'_x)
\end{aligned}
$$

The target quantities are defined as follows:

$$
\text{PNS} = P(y_x, y'_{x'}) = \sum_{kl} p_{10kl} \qquad c'\text{-PNS} = P(y_x, y'_{x'}|c') = \frac{\sum_l p_{100l}}{P(c')} \tag{6}
$$

$$
\text{ETT} = P(y_x|x') = \frac{\sum_{jk} p_{1jk0}}{P(x')} \qquad c'\text{-ETT} = P(y_x|x', c') = \frac{\sum_j p_{1j00}}{P(x', c')} \tag{7}
$$

$$
\text{PN} = P(y'_{x'}|x, y) = \frac{\sum_k p_{10k1}}{P(x, y)} \qquad c'\text{-PN} = P(y'_{x'}|x, y, c') = \frac{p_{1001}}{P(x, y, c')} \tag{8}
$$

$$
\text{PS} = P(y_x|x', y') = \frac{\sum_k p_{10k0}}{P(x', y')} \qquad c'\text{-PS} = P(y_x|x', y', c') = \frac{p_{1000}}{P(x', y', c')} \tag{9}
$$

Since the distributions $P(Y_x|c)$, $P(X, Y|c)$, and $P(c)$ are assumed to be available throughout the paper, the following data constraints are included in all optimization formulations presented later.

$$
\begin{array}{lll}
\text{for } i, j, k, l \in \{0, 1\} & \sum_{ijkl} p_{ijkl} = 1 & \sum_{jl} p_{1j1l} = P(y_x, c) \\
\sum_{jl} p_{0j1l} = P(y'_x, c) & \sum_{il} p_{i11l} = P(y_{x'}, c) & \sum_j p_{1j11} = P(x, y, c) \\
\sum_j p_{0j11} = P(x, y', c) & \sum_i p_{i110} = P(x', y, c) &
\end{array} \tag{10}
$$

If the observational distribution for c' is additionally available, then the following constraints can be incorporated into the optimization formulations.

$$
\text{for } i, j, k, l \in \{0, 1\} \sum_j p_{1j01} = P(x, y, c'), \sum_j p_{0j01} = P(x, y', c'), \sum_i p_{i100} = P(x', y, c') \tag{11}
$$

If further the experimental distribution for c' is available, then the following constraints can be incorporated into the optimization formulations.

$$
\text{for } i, j, k, l \in \{0, 1\} \sum_{jl} p_{1j0l} = P(y_x, c'), \sum_{il} p_{i10l} = P(y_{x'}, c') \tag{12}
$$

Moreover, since ETT (c'-ETT), PN (c'-PN), and PS (c'-PS) require $P(x')$, $P(x', y')$, and $P(x, y)$ as their respective denominators, the following constraints must be incorporated into the optimization formulations when computing ETT (c'-ETT), PN (c'-PN), and PS (c'-PS), respectively. Note that, because $P(X, Y, c)$ is always available, the quantities $P(x', c')$, $P(x', y', c')$, and $P(x, y, c')$ can be derived accordingly.

$$
\sum_{ijk} p_{ijk0} = P(x') \quad (13) \qquad \sum_{ik} p_{i0k0} = P(x', y') \quad (14) \qquad \sum_{jk} p_{1jk1} = P(x, y) \quad (15)
$$

Optimizing Equations 6 to 9 under the appropriate sets of constraints in Equations 10 to 15 yields the tight bounds of the target quantities, since the optimization is global. We apply Balke's linear programming framework [Balke, 1995] to derive the closed-form solutions for each bound. Therefore, the bounds are presented in the next section without proofs; their correctness can be verified by manually enumerating the vertices, as described in [Balke, 1995]. Formal mathematical proofs can also be established carefully using the consistency rule (i.e., $P(y_x, x) = P(y, x)$) and Fréchet inequalities.

3.2 No Data of c' is Available

Under the constraints in Equations 10, the optimization of Equations 6, Equations 7 (with the additional constraint in Equation 13), Equations 8 (with the additional constraint in Equation 14), and Equations 9 (with the additional constraint in Equation 15) yields the following theorems.

Theorem 3.1. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, c)$ and $P(X, Y, c)$ that are compatible with G. Then, PNS = $P(y_x, y'_{x'})$ is tightly bounded as follows:*

$$\max\left\{\begin{array}{c} 0, \\ P(y_x, c) - P(y_{x'}, c), \\ P(y, c) - P(y_{x'}, c), \\ P(y_x, c) - P(y, c) \end{array}\right\} \leq \textit{PNS} \leq \min\left\{\begin{array}{c} 1 - P(y'_x, c), \\ 1 - P(y_{x'}, c), \\ 1 - P(x, y', c) - P(x', y, c), \\ 1 - P(y'_x, c) - P(y_{x'}, c) + \\ + P(x, y', c) + P(x', y, c) \end{array}\right\} \tag{16}$$

Besides, there are no informative bounds (i.e., no transportable information from c) for the c'-specific PNS= $P(y_x, y'_{x'}|c')$. That is, the tight bounds are given by $0 \leq c'$-PNS ≤ 1.

Theorem 3.2. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, c)$, $P(X, Y, c)$, and $P(X)$ that are compatible with G. Then, ETT = $P(y_x|x')$ is tightly bounded as follows:*

$$\frac{P(y_x, c) - P(x, y, c)}{P(x')} \leq \textit{ETT} \leq \frac{P(x') - P(y'_x, c) + P(x, y', c)}{P(x')} \tag{17}$$

Besides, there are no informative bounds (i.e., no transportable information from c) for the c'-specific ETT= $P(y_x|x', c')$. That is, the tight bounds are given by $0 \leq c'$-ETT ≤ 1.

Theorem 3.3. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, c)$, $P(X, Y, c)$, and $P(x, y)$ that are compatible with G. Then, PN = $P(y'_{x'}|x, y)$ is tightly bounded as follows:*

$$\max\left\{\begin{array}{c} 0, \\ \frac{P(y,c) - P(y_{x'},c)}{P(x,y)} \end{array}\right\} \leq \textit{PN} \leq \min\left\{\begin{array}{c} 1, \\ \frac{P(x,y) + P(x,y',c) + P(x',y,c) - P(y_{x'},c)}{P(x,y)} \end{array}\right\} \tag{18}$$

Besides, there are no informative bounds (i.e., no transportable information from c) for the c'-specific PN= $P(y'_{x'}|x, y, c')$. That is, the tight bounds are given by $0 \leq c'$-PN ≤ 1.

Theorem 3.4. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, c)$, $P(X, Y, c)$, and $P(x', y')$ that are compatible with G. Then, PS = $P(y_x|x', y')$ is tightly bounded as follows:*

$$\max\left\{\begin{array}{c} 0, \\ \frac{P(y',c) - P(y'_x,c)}{P(x',y')} \end{array}\right\} \leq \textit{PS} \leq \min\left\{\begin{array}{c} 1, \\ \frac{P(x',y') + P(x',y,c) + P(x,y',c) - P(y'_x,c)}{P(x',y')} \end{array}\right\} \tag{19}$$

Besides, there are no informative bounds (i.e., no transportable information from c) for the c'-specific PS= $P(y_x|x', y', c')$. That is, the tight bounds are given by $0 \leq c'$-PS ≤ 1.

Note that the theorems are stated with respect to a causal diagram G. This notation is mainly adopted for clarity of presentation and to facilitate future extensions involving additional graphical constraints. In the current paper, however, the results do not require any specific assumptions on G and hold for arbitrary causal diagrams, as long as C is not a descendant of X. Otherwise, $P(Y_x|c)$ would no longer correspond to experimental data and would instead become a counterfactual quantity.

3.3 With Additional Observational Data for c'

Under the constraints in Equations 10 and 11, optimizing Equations 6 – 9 yields the following theorems. No additional constraints are needed in this setting since the full observational distribution is assumed to be available.

Theorem 3.5. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, c)$ and $P(X, Y, C)$ that are compatible with G. Then, PNS = $P(y_x, y'_{x'})$ and c'-PNS= $P(y_x, y'_{x'}|c')$ are tightly bounded as follows:*

$$\max\left\{\begin{array}{c} 0, \\ P(y_x, c) - P(y_{x'}, c), \\ P(y, c) - P(y_{x'}, c), \\ P(y_x, c) - P(y, c) \end{array}\right\} \leq \textit{PNS} \leq \min\left\{\begin{array}{c} 1 - P(y'_x, c)- \\ -P(x, y', c') - P(x', y, c'), \\ 1 - P(y_{x'}, c)- \\ -P(x, y', c') - P(x', y, c'), \\ P(x, y) + P(x', y'), \\ 1 - P(y'_x, c) - P(y_{x'}, c)+ \\ +P(x, y', c) + P(x', y, c)- \\ -P(x, y', c') - P(x', y, c') \end{array}\right\} \tag{20}$$

$$0 \leq c'\text{-}\textit{PNS} \leq \frac{P(c') - P(x, y', c') - P(x', y, c')}{P(c')} \tag{21}$$

The additional observational data for c' does not actually result in tighter bounds for ETT, c'-ETT, PN, c'-PN, PS, and c'-PS, as expected from Li and Pearl [2023], where observational data was shown to play a less significant role in bounding PoCs.

Note that additional observational data do introduce bounds for c'-PNS, but these bounds do not involve any data from the c subpopulation. In other words, the result does not provide any advantage over directly applying the Tian–Pearl bounds to c' using only the additional observational data.

3.4 With Additional Experimental Data for c'

Under the constraints in Equations 10 and 12, the optimization of Equations 6, Equations 7 (with the additional constraint in Equation 13), Equations 8 (with the additional constraint in Equation 14), and Equations 9 (with the additional constraint in Equation 15) yields the following theorems.

Theorem 3.6. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, C)$ and $P(X, Y, c)$ that are compatible with G. Then, PNS = $P(y_x, y'_{x'})$ and c'-PNS= $P(y_x, y'_{x'}|c')$ are tightly bounded as follows:*

$$\max\left\{\begin{array}{c} 0, \\ P(y_x) - P(y_{x'}), \\ P(y_x, c) - P(y_{x'}, c), \\ P(y_x, c') - P(y_{x'}, c'), \\ P(y, c) - P(y_{x'}, c), \\ P(y_x, c) - P(y, c), \\ P(y_x, c') - P(y_{x'})+ \\ +P(y, c), \\ P(y_x) - P(y_{x'}, c')- \\ -P(y, c) \end{array}\right\} \leq \textit{PNS} \leq \min\left\{\begin{array}{c} P(y_x), \\ P(y'_{x'}), \\ 1 - P(y'_x, c) - P(y_{x'}, c'), \\ 1 - P(y_{x'}, c')- \\ -P(x, y', c) - P(x', y, c), \\ P(y_x) + P(y'_x, c)- \\ -P(x, y', c) - P(x', y, c), \\ P(y_x) + P(y'_x, c) - P(y_{x'}, c), \\ P(y_x) - P(y_{x'}, c)+ \\ +P(x, y', c) + P(x', y, c), \\ 1 - P(y'_x, c) - P(y_{x'})+ \\ +P(x, y', c) + P(x', y, c) \end{array}\right\} \tag{22}$$

$$\max\left\{\begin{array}{c} 0, \\ \frac{P(y_x, c') - P(y_{x'}, c')}{P(c')} \end{array}\right\} \leq c'\text{-}\textit{PNS} \leq \min\left\{\begin{array}{c} \frac{P(y_x, c')}{P(c')}, \\ \frac{P(c') - P(y_{x'}, c')}{P(c')} \end{array}\right\} \tag{23}$$

Theorem 3.7. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, C)$, $P(X, Y, c)$, and $P(X)$ that are compatible with G. Then, ETT = $P(y_x|x')$ and c'-ETT= $P(y_x|x', c')$ are tightly bounded as follows:*

$$\max\left\{\begin{array}{c} \frac{P(y_x, c) - P(x, y, c)}{P(x')}, \\ \frac{P(y_x) + P(x, y', c) + P(x') - 1}{P(x')} \end{array}\right\} \leq \textit{ETT} \leq \min\left\{\begin{array}{c} \frac{P(x, y', c) + P(x') - P(y'_x, c)}{P(x')}, \\ \frac{P(y_x) - P(x, y, c)}{P(x')} \end{array}\right\} \tag{24}$$

Table 1: Observational study: 1000 adults were given access to the vaccine and chose freely whether to take it. Experimental study: 500 adults were randomly assigned to receive the vaccine and 500 were assigned not to.

Observational Study (1000 adults, free choice)		Experimental Study (500 per group)	
Vaccinated	Unvaccinated	Vaccinated	Unvaccinated
30 out of 248 prevented disease (12.1%)	251 out of 752 prevented disease (33.4%)	381 out of 500 prevented disease (76.2%)	238 out of 500 prevented disease (47.7%)

$$\max\left\{\begin{array}{c}0,\\ \frac{P(y_x,c')-P(x,c')}{P(x',c')}\end{array}\right\} \le c'\text{-}ETT \le \min\left\{\begin{array}{c}1,\\ \frac{P(y_x,c')}{P(x',c')}\end{array}\right\} \tag{25}$$

Theorem 3.8. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, C)$, $P(X, Y, c)$, and $P(x, y)$ that are compatible with G. Then, PN = $P(y'_{x'}|x, y)$ and c'-PN = $P(y'_{x'}|x, y, c')$ are tightly bounded as follows:*

$$\max\left\{\begin{array}{c}0,\\ \frac{P(x,y)-P(x,y,c)-P(y_{x'},c')}{P(x,y)},\\ \frac{P(y,c)-P(y_{x'},c)}{P(x,y)},\\ \frac{P(x,y)+P(x',y,c)-P(y_{x'})}{P(x,y)}\end{array}\right\} \le PN \le \min\left\{\begin{array}{c}1,\\ \frac{1+P(x,y,c)-P(c)-P(y_{x'},c')}{P(x,y)},\\ \frac{P(x,y',c)+P(x',y,c)+P(x,y)-P(y_{x'},c)}{P(x,y)},\\ \frac{1+P(x,c)+P(x',y,c)-P(c)-P(y_{x'})}{P(x,y)}\end{array}\right\} \tag{26}$$

$$\max\left\{\begin{array}{c}0,\\ \frac{P(y_{x'},c')-P(x,y,c')}{P(x,y,c')}\end{array}\right\} \le c'\text{-}PN \le \min\left\{\begin{array}{c}1,\\ \frac{P(c')-P(y_{x'},c')}{P(x,y,c')}\end{array}\right\} \tag{27}$$

Theorem 3.9. *Given a causal diagram G, let C be a binary variable that is not a descendant of X in G. Given the distributions $P(Y_X, C)$, $P(X, Y, c)$, and $P(x', y')$ that are compatible with G. Then, PS = $P(y_x|x', y')$ and c'-PS = $P(y_x|x', y', c')$ are tightly bounded as follows:*

$$\max\left\{\begin{array}{c}0,\\ \frac{P(x',y')-P(x',y',c)-P(y'_x,c')}{P(x',y')},\\ \frac{P(y',c)-P(y'_x,c)}{P(x',y')},\\ \frac{P(x',y')+P(x,y',c)-P(y'_x)}{P(x',y')}\end{array}\right\} \le PS \le \min\left\{\begin{array}{c}1,\\ \frac{1+P(x',y',c)-P(c)-P(y'_x,c')}{P(x',y')},\\ \frac{P(x',y,c)+P(x,y',c)+P(x',y')-P(y'_x,c)}{P(x',y')},\\ \frac{1+P(x',c)+P(x,y',c)-P(c)-P(y'_x)}{P(x',y')}\end{array}\right\} \tag{28}$$

$$\max\left\{\begin{array}{c}0,\\ \frac{P(y'_x,c')-P(x',y',c')}{P(x',y',c')}\end{array}\right\} \le c'\text{-}PS \le \min\left\{\begin{array}{c}1,\\ \frac{P(c')-P(y'_x,c')}{P(x',y',c')}\end{array}\right\} \tag{29}$$

Again, note that additional experimental data do introduce bounds for c'-PNS, c'-ETT, c'-PN, and c'-PS, but these bounds do not involve any data from the c subpopulation. In other words, the results do not provide any advantage over directly applying the Tian–Pearl bounds to c' using only the additional experimental data.

4 Example

Let us consider the motivating example again of a vaccine designed to prevent a serious disease. The government aims to estimate the probability of necessity and sufficiency (PNS), namely, the proportion of individuals who would remain healthy if vaccinated but would not otherwise. Let x denote taking the vaccine and x' denote not taking the vaccine; let y denote remaining healthy and y' denote having the disease; and let c denote adults and c' denote youth. Based on regional demographics, adults account for 63.6% of the population, i.e., $P(c) = 0.636$. The government designed both observational and experimental studies for the adult population, as shown in Table 1.

Table 2: Performance comparison of the proposed bounds with baseline.

		↑ LB	↓ UB	gap(Base)	gap(Prop.)	# imp.
No data for c'	PNS	0.0824	0.3327	1	0.5850	10,000
	ETT	0.2477	0.2513	1	0.5010	10,000
	PN	0.1456	0.1432	1	0.7112	7,493
	PS	0.1420	0.1466	1	0.7113	7,466
With obs. data of c'	PNS	0.0824	0.0832	0.4998	0.3342	9,336
With exp. data of c'	PNS	0.0525	0.0531	0.3105	0.2049	9,371
	ETT	0.3559	0.3628	1	0.2813	10,000
	PN	0.1942	0.1924	1	0.6133	8,362
	PS	0.1904	0.1963	1	0.6133	8,373

The table gives $P(y_x|c) = 0.762$, $P(y_{x'}|c) = 0.477$, $P(x, y|c) = 30/1000$, $P(x|c) = 248/1000$, $P(x', y|c) = 251/1000$, and $P(x'|c) = 752/1000$, together with the prior $P(c) = 0.636$. Applying Theorem 3.1, we obtain $0.306 \leq \text{PNS} \leq 0.697$, which indicates that at least 30.6% of the overall population would benefit from the vaccine, even though the observational study suggests that the disease prevention rate decreases among individuals who take the vaccine.

5 Simulation Results

We further show that the proposed theorems yield informative bounds in general. For the setting without additional data from c', the baseline bounds are $[0, 1]$, since no informative bounds are available. For the setting with additional observational (experimental) data for c' of PNS, the baseline corresponds to the Tian–Pearl bounds in Equations 2, applied solely to the observational (experimental) distributions, since the additional data make the corresponding distributions available.

We randomly generated $10,000$ samples of distribution tuples $(P(X, Y, C), P(Y_x, C), P(Y_{x'}, C))$ (see the appendix for the data generation process) and applied the proposed bounds and baseline bounds using only the corresponding available data.

For each sampled distribution tuple i, let $[a_i, b_i]$ denote the bounds obtained using the proposed theorems, and let $[c_i, d_i]$ denote the baseline bounds. We compare the following criteria, with results reported in Table 2:

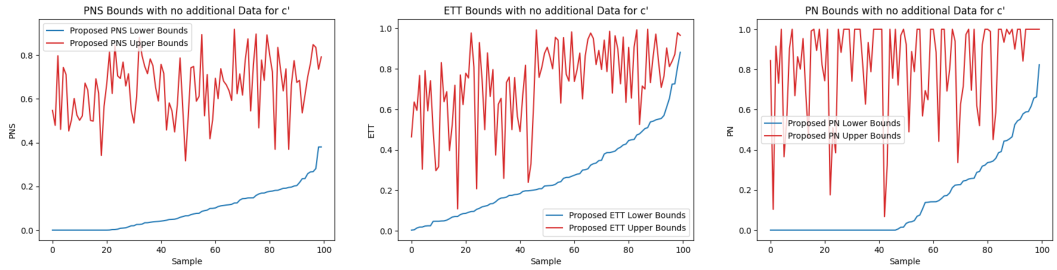
- Average increase in the lower bound (↑LB): $\frac{1}{10000} \sum(a_i - c_i)$
- Average decrease in the upper bound (↓UB): $\frac{1}{10000} \sum(d_i - b_i)$
- Average gap of the baseline bounds (gap(Base)): $\frac{1}{10000} \sum(d_i - c_i)$
- Average gap of the proposed bounds (gap(Prop.)): $\frac{1}{10000} \sum(b_i - a_i)$
- Number of sampled distributions for which the proposed bounds improved upon the baseline bounds (# imp.): $\sum g_i$, where $g_i = 1$ if $(a_i > c_i)$ or $(b_i < d_i)$, and $g_i = 0$ otherwise

As shown in Table 2, for PNS with any type of additional data from c', the proposed bounds are significantly more informative than the existing Tian–Pearl bounds. In all other settings, where no existing informative bounds are available, the proposed method still provides informative bounds.

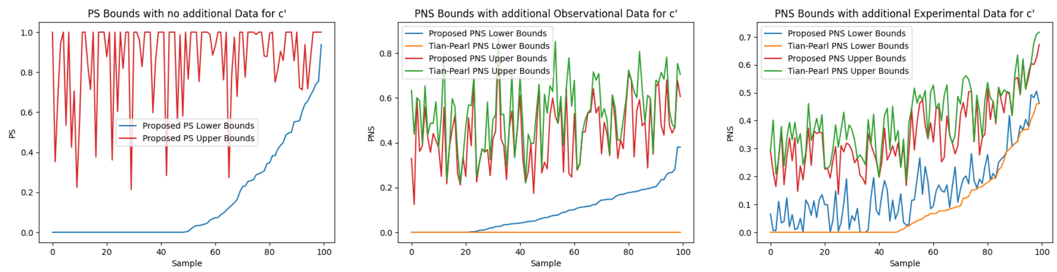
In order to visualize the informativeness of the proposed bounds, for each setting, 100 distributions were randomly selected from the 10,000 sampled distributions, sorted according to the proposed lower bounds, and used to plot both the proposed bounds and the baseline bounds, as shown in Figures 1.

6 Conclusion

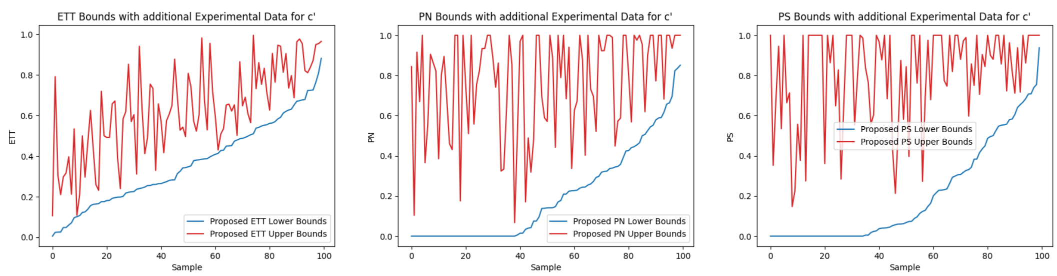
In this paper, we investigate the bounds of PoCs in settings where data are only available for a subpopulation. We derive tight bounds for PNS, ETT, PN, and PS, both with and without additional data availability. Toy examples and simulation results are provided to illustrate the informativeness and significance of the proposed bounds. We further investigate whether data from one subpopulation can provide tighter bounds for another subpopulation. However, the results show that either no



(a) PNS w/o add. Data. (b) ETT w/o add. Data. (c) PN w/o add. Data.



(d) PS w/o add. Data. (e) PNS w/ Obs. Data for c'. (f) PNS w/ Exp. Data for c'.



(g) ETT w/ Exp. Data for c'. (h) PN w/ Exp. Data for c'. (i) PS w/ Exp. Data for c'.

Figure 1: Visualization of Performance Comparison for the proposed bounds.

informative bounds can be obtained, or the resulting bounds are equivalent to directly applying the Tian–Pearl bounds to the new subpopulation.

The results in this paper are limited to the binary case. However, the framework can be extended naturally to non-binary treatments, outcomes, and covariates C by introducing additional counterfactual distribution parameters into the linear program. For example, if X, Y, and C each take three values, the corresponding linear program for $P(Y_x, Y_{x'}, Y_{x''}, C, X)$ would involve 3^5 parameters. The main difficulty lies in deriving closed-form solutions using Balke's linear programming approach [Balke, 1995], due to the rapidly increasing complexity. Nevertheless, the bounds can still be computed numerically. Deriving closed-form solutions for higher-order cases would be an appealing direction for future work, similar to the general PoC characterization developed in [Shu et al., 2025]. In addition, this paper focuses on the four fundamental PoCs, which are sufficient to illustrate the general framework, as they closely correspond to the four representative forms of PoCs defined in [Shu et al., 2025].

This paper showed that the transportability of PoC bounds cannot, in general, be achieved without additional assumptions. Therefore, a natural future direction is to investigate the transportability of PoC bounds under additional causal structural assumptions. Moreover, incorporating causal structure may further tighten the general bounds derived in this paper, similarly to the results in [Mueller et al., 2022].

Overall, this paper not only introduces, for the first time, bounds on PoCs under subpopulation data-limited settings commonly encountered in real-world applications, but also opens several appealing directions for future research.

References

Alexander Abraham Balke. *Probabilistic counterfactuals: semantics, computation, and applications*. University of California, Los Angeles, 1995.

Elias Bareinboim and Judea Pearl. Transportability of causal effects: Completeness results. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 26, pages 698–704, 2012.

Elias Bareinboim and Judea Pearl. A general algorithm for deciding transportability of experimental results. *Journal of causal Inference*, 1(1):107–134, 2013.

David Galles and Judea Pearl. An axiomatic characterization of causal counterfactuals. *Foundations of Science*, 3(1):151–182, 1998.

Joseph Y Halpern. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research*, 12: 317–337, 2000.

Ang Li and Judea Pearl. Unit selection based on counterfactual logic. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 2019.

Ang Li and Judea Pearl. Probabilities of causation: Role of observational data. In *International Conference on Artificial Intelligence and Statistics*, pages 10012–10027. PMLR, 2023.

Ang Li and Judea Pearl. Probabilities of causation with nonbinary treatment and effect. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38(18), pages 20465–20472, 2024.

Ang Li, Ruirui Mao, and Judea Pearl. Probabilities of causation: Adequate size of experimental and observational samples. Technical report, Department of Computer Science, University of California, Los Angeles, CA, 2022.

Tyler Lu and Craig Boutilier. Budgeted social choice: From consensus to personalized decision making. In *IJCAI*, volume 11, pages 280–286, 2011.

S. Mueller, A. Li, and J. Pearl. Causes of effects: Learning individual responses from population data. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)*, pages 2712–2718, 2022.

Scott Mueller and Judea Pearl. Personalized decision making–a conceptual introduction. *Journal of Causal Inference*, 11(1):20220050, 2023.

Judea Pearl. Probabilities of causation: Three counterfactual interpretations and their identification. *Synthese*, pages 93–149, 1999.

Judea Pearl. *Causality*. Cambridge university press, 2nd edition, 2009.

Judea Pearl and Elias Bareinboim. Transportability of causal and statistical relations: A formal approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 25, pages 247–254, 2011.

Judea Pearl and Elias Bareinboim. External validity: From do-calculus to transportability across populations. In *Probabilistic and causal inference: The works of Judea Pearl*, pages 451–482. 2022.

I. Shpitser and J Pearl. Effects of treatment on the treated: Identification and generalization. Technical report, Department of Computer Science, University of California, Los Angeles, CA, 2009. Extended version of paper that appeared in UAI-09.

Xin Shu, Shuai Wang, and Ang Li. Identification of probabilities of causation: from recursive to closed-form bounds. *arXiv preprint arXiv:2505.15274*, 2025.

Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. *Annals of Mathematics and Artificial Intelligence*, 28(1-4):287–313, 2000.

Junzhe Zhang, Jin Tian, and Elias Bareinboim. Partial counterfactual identification from observational and experimental data. In *International Conference on Machine Learning*, pages 26548–26558. PMLR, 2022.

A Technical appendices and supplementary material

A.1 Data Generating Process of the Simulation Study

The following algorithm ensures that the generated data satisfy the relationship between the experimental and observational data described in [Tian and Pearl, 2000].

Algorithm 1 Data Generating Process for Counterfactual Bounds Simulation

Require: Number of simulations N
for $i = 1$ to N **do**
  Step 1: Generate joint distribution $P(X, Y, C)$
  Draw $(p_1, \ldots, p_8) \sim$ Dirichlet($\mathbf{1}_8$)
  Assign $P(x, y, c),\ P(x, y, c'),\ P(x, y', c),\ P(x, y', c')$
    $P(x', y, c),\ P(x', y, c'),\ P(x', y', c),\ P(x', y', c') \leftarrow p_1, \ldots, p_8$
  Step 2: Compute marginals
  $P(x, c) \leftarrow P(x, y, c) + P(x, y', c)$
  $P(x, c') \leftarrow P(x, y, c') + P(x, y', c')$
  $P(x', c) \leftarrow P(x', y, c) + P(x', y', c)$
  $P(x', c') \leftarrow P(x', y, c') + P(x', y', c')$
  $P(c) \leftarrow P(x, c) + P(x', c)$
  $P(c') \leftarrow P(x, c') + P(x', c')$
  Step 3: Sample counterfactuals uniformly within bounds
  Draw $P(y_x, c) \sim \text{Uniform}\big[P(x, y, c),\ P(x, y, c) + P(x', c)\big]$
  Draw $P(y_{x'}, c) \sim \text{Uniform}\big[P(x', y, c),\ P(x', y, c) + P(x, c)\big]$
  Draw $P(y_x, c') \sim \text{Uniform}\big[P(x, y, c'),\ P(x, y, c') + P(x', c')\big]$
  Draw $P(y_{x'}, c') \sim \text{Uniform}\big[P(x', y, c'),\ P(x', y, c') + P(x, c')\big]$
  Step 4: Derive complements via total probability
  $P(y'_x, c) \leftarrow P(c) - P(y_x, c)$
  $P(y'_x, c') \leftarrow P(c') - P(y_x, c')$
  $P(y'_{x'}, c) \leftarrow P(c) - P(y_{x'}, c)$
  $P(y'_{x'}, c') \leftarrow P(c') - P(y_{x'}, c')$
  Step 5: Derive marginal counterfactuals
  $P(y_x) \leftarrow P(y_x, c) + P(y_x, c')$
  $P(y_{x'}) \leftarrow P(y_{x'}, c) + P(y_{x'}, c')$
  $P(y'_x) \leftarrow P(y'_x, c) + P(y'_x, c')$
  $P(y'_{x'}) \leftarrow P(y'_{x'}, c) + P(y'_{x'}, c')$
  Store all quantities for simulation i
end for