Causes of Effects: Learning individual responses from population data

Scott Mueller , **Ang Li** and **Judea Pearl**
University of California Los Angeles
Computer Science Department
Cognitive Systems Laboratory
{scott, angli, judea}@cs.ucla.edu

Supplementary Material

A Proof of Theorem 4

Proof.

$$\begin{aligned}\text{PNS} &= P(y_x, y'_{x'}) \\ &= \Sigma_z P(y_x, y'_{x'}|z) \times P(z) \end{aligned} \quad (1)$$

From [Li and Pearl, 2019], we have the z-specific PNS as follows:

$$\max\left\{\begin{array}{c} 0, \\ P(y_x|z) - P(y_{x'}|z), \\ P(y|z) - P(y_{x'}|z), \\ P(y_x|z) - P(y|z) \end{array}\right\} \leq z\text{-PNS} \quad (2)$$

$$\min\left\{\begin{array}{c} P(y_x|z), \\ P(y'_{x'}|z), \\ P(y,x|z) + P(y',x'|z), \\ P(y_x|z) - P(y_{x'}|z) + \\ + P(y,x'|z) + P(y',x|z) \end{array}\right\} \geq z\text{-PNS} \quad (3)$$

Substituting 2 and 3 into 1, theorem 4 holds.
Note that since we have,

$$\begin{aligned} & \sum_z \max\{0, P(y_x|z) - P(y_{x'}|z), \\ & P(y|z) - P(y_{x'}|z), P(y_x|z) - P(y|z)\} \times P(z) \\ \geq & \sum_z 0 \times P(z) \\ = & \ 0, \\ & \sum_z \max\{0, P(y_x|z) - P(y_{x'}|z), \\ & P(y|z) - P(y_{x'}|z), P(y_x|z) - P(y|z)\} \times P(z) \\ \geq & \sum_z [P(y_x|z) - P(y_{x'}|z)] \times P(z) \\ = & \ P(y_x) - P(y_{x'}), \\ & \sum_z \max\{0, P(y_x|z) - P(y_{x'}|z), \\ & P(y|z) - P(y_{x'}|z), P(y_x|z) - P(y|z)\} \times P(z) \\ \geq & \sum_z [P(y|z) - P(y_{x'}|z)] \times P(z) \\ = & \ P(y) - P(y_{x'}), \\ & \sum_z \max\{0, P(y_x|z) - P(y_{x'}|z), \\ & P(y|z) - P(y_{x'}|z), P(y_x|z) - P(y|z)\} \times P(z) \\ \geq & \sum_z [P(y_x|z) - P(y|z)] \times P(z) \\ = & \ P(y_x) - P(y), \end{aligned}$$

then the lower bound in theorem 4 is guaranteed to be no worse than the Tian-Pearl lower bound in equation 4. Similarly, the upper bound in theorem 4 is guaranteed to be no worse than the Tian-Pearl upper bound in equation 5. Also note that, since Z does not contain a descendant of X, the term $P(y_x|z)$ refers to experimental data under population z.

□

B Proof of Theorem 5

Proof. Since Z satisfies the back-door criterion, then equations 8 and 9 still hold and $P(y_x|z) = P(y|x,z)$, $P(y_{x'}|z) = P(y|x',z)$, and $P(y'_{x'}|z) = P(y'|x',z)$. We further have,

$$
\begin{aligned}
& P(y_x|z) - P(y_{x'}|z) \\
= \;& P(y|x,z) - P(y|x',z) \\
\geq \;& [P(y|x,z) - P(y|x',z)] \times P(x|z) \\
= \;& P(y|x,z) \times P(x|z) - P(y|x',z) \times (1 - P(x'|z)) \\
= \;& P(y,x|z) + P(y,x'|z) - P(y|x',z) \\
= \;& P(y|z) - P(y|x',z) \\
= \;& P(y|z) - P(y_{x'}|z) \qquad (4)
\end{aligned}
$$

and

$$
\begin{aligned}
& P(y_x|z) - P(y_{x'}|z) \\
= \;& P(y|x,z) - P(y|x',z) \\
\geq \;& [P(y|x,z) - P(y|x',z)] \times P(x'|z), \\
= \;& P(y|x,z) \times (1 - P(x|z)) - P(y|x',z) \times P(x'|z) \\
= \;& P(y|x,z) - P(y,x|z) - P(y,x'|z) \\
= \;& P(y|x,z) - P(y|z) \\
= \;& P(y_x|z) - P(y|z). \qquad (5)
\end{aligned}
$$

With equations 4 and 5, equation 8 reduces to equation 10 in theorem 5.

We also have,

$$
\begin{aligned}
& \min\{P(y_x|z), P(y'_{x'}|z)\} \\
= \;& \min\{P(y|x,z), P(y'|x',z)\} \\
\leq \;& P(y|x,z) \times P(x|z) + P(y'|x',z) \times (1 - P(x|z)) \\
= \;& P(y|x,z) \times P(x|z) + P(y'|x',z) \times P(x'|z) \\
= \;& P(y,x|z) + P(y',x'|z) \qquad (6)
\end{aligned}
$$

and

$$
\begin{aligned}
& \min\{P(y_x|z), P(y'_{x'}|z)\} \\
= \;& \min\{P(y|x,z), P(y'|x',z)\} \\
\leq \;& P(y|x,z) \times (1 - P(x|z)) + P(y'|x',z) \times P(x|z) \\
= \;& P(y|x,z) \times (1 - P(x|z)) + P(y'|x',z) \times (1 - P(x'|z)) \\
= \;& P(y|x,z) - P(y,x|z) + P(y'|x',z) - P(y',x'|z) \\
= \;& P(y|x,z) - P(y|x',z) + P(y,x'|z) + P(y',x|z) \\
= \;& P(y_x|z) - P(y_{x'}|z) + P(y,x'|z) + P(y',x|z). \qquad (7)
\end{aligned}
$$

With equations 6 and 7, equation 9 reduces to equation 11 in theorem 5. □

C Proof of Theorem 6

Proof.

$$
\begin{aligned}
& \text{PNS} \\
= \;& P(y_x, y'_{x'}) \\
= \;& \Sigma_z \Sigma_{z'} P(y_x, y'_{x'}, z_x, z'_{x'}) \\
= \;& \Sigma_z \Sigma_{z'} P(y_x, y'_{x'}|z_x, z'_{x'}) \times P(z_x, z'_{x'}) \\
\leq \;& \Sigma_z \Sigma_{z'} \min\{P(y_x|z_x, z'_{x'}), P(y'_{x'}|z_x, z'_{x'})\} \times \\
& \quad \min\{P(z_x), P(z'_{x'})\} \\
= \;& \Sigma_z \Sigma_{z'} \min\{P(y_x|z_x), P(y'_{x'}|z'_{x'})\} \times \qquad (8) \\
& \quad \min\{P(z_x), P(z'_{x'})\} \\
= \;& \Sigma_z \Sigma_{z'} \min\{P(y|z_x, x), P(y'|z'_{x'}, x')\} \times \\
& \quad \min\{P(z_x), P(z'_{x'})\} \qquad (9) \\
= \;& \Sigma_z \Sigma_{z'} \min\{P(y|z,x), P(y'|z',x')\} \times \\
& \quad \min\{P(z_x), P(z'_{x'})\}.
\end{aligned}
$$

Combined with the Tian-Pearl bounds in equations 4 and 5, theorem 6 holds. Note that equation 8 is due to $Y_x \perp\!\!\!\perp Z_{x'} \mid Z_x$ and $Y_{x'} \perp\!\!\!\perp Z_x \mid Z_{x'}$. Equation 9 is due to $\forall x, Y_x \perp\!\!\!\perp X \mid Z_x$. □

D Proof of Theorem 7

Proof. First we show that in graph G, if an individual is a complier from X to Y, then Z_x and $Z_{x'}$ must have the different values. This is because the structural equations for Y and Z are $f_y(z, u_y)$ and $f_z(x, u_z)$, respectively. If an individual has the same Z_x and $Z_{x'}$ value, then $f_z(x, u_z) = f_z(x', u_z)$. This means $f_y(f_z(x, u_z), u_y) = f_y(f_z(x', u_z), u_y)$, i.e., Y_x and $Y_{x'}$ must have the same value. Thus this individual is not a complier. Therefore,

$$
\begin{aligned}
& \text{PNS} \\
= \;& P(y_x, y'_{x'}) \\
= \;& \Sigma_z \Sigma_{z' \neq z} P(y_z, y'_{z'}) \times P(z_x, z'_{x'}) \\
\leq \;& \Sigma_z \Sigma_{z' \neq z} \min\{P(y_z), P(y'_{z'})\} \times \\
& \min\{P(z_x), P(z'_{x'})\} \\
= \;& \Sigma_z \Sigma_{z' \neq z} \min\{P(y|z), P(y'|z')\} \times \\
& \min\{P(z|x), P(z'|x')\}
\end{aligned}
$$

Combined with the Tian-Pearl bounds in equations 4 and 5, theorem 7 holds. □

E Simulation Algorithm

We used the following algorithm to generate samples and conduct the simulations in section 5 (Note that):

Algorithm 1 Generate PNS simulation data

```
input  : Number of output samples n
         Causal diagram G
         Covariates to condition on Z
output : List of 4-tuples consisting of general lower bound,
         lower bound with causal graph, upper bound with
         causal graph, and general upper bound
begin
    for i ← 1 to n do
        cpt ← generate-cpt(G,random-uniform)
        // Lower/upper Tian-Pearl bounds
        lb, ub ← pns-bounds(cpt)
        // Lower/upper bounds with graph
        lb_graph, ub_graph ← pns-graph(cpt, Z)
        append-result(lb, lb_graph, ub_graph, ub)
    end
end
```

Procedure generate-cpt

```
input  : n causal diagram nodes (X_1, ..., X_n)
         Distribution D
output : n conditional probability tables for
         P(X_i|Parents(X_i))
begin
    for i ← 1 to n do
        s ← num-instantiates(X_i)
        p ← num-instantiates(Parents(X_i))
        for k ← 1 to p do
            sum ← 0
            for j ← 1 to s do
                a_j ← sample(D)
                sum ← sum + a_j
            end
            for j ← 1 to s do
                P(x_{i_j}|Parents(X_i)_k) ← a_j/sum
            end
        end
    end
end
```

References

[Li and Pearl, 2019] Ang Li and Judea Pearl. Unit selection based on counterfactual logic. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 1793–1799. AAAI Press, 2019.