Unit Selection with Nonbinary Treatment and Effect

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

Abstract
The unit selection problem aims to identify a set of individuals who are most likely to exhibit a desired mode of behavior, for example, selecting individuals who would respond one way if encouraged and a different way if not encouraged. Using a combination of experimental and observational data, Li and Pearl derived tight bounds on the “benefit function”, which is the payoff/cost associated with selecting an individual with given characteristics. This paper extends the benefit function to the general form such that the treatment and effect are not restricted to binary. We propose an algorithm to test the identifiability of the nonbinary benefit function and an algorithm to compute the bounds of the nonbinary benefit function using experimental and observational data.

Introduction
Several areas of industry, marketing, and health science face the unit selection dilemma. For example, in customer relationship management (Berson, Smith, and Thearling 1999; Lejeune 2001; Hung, Yen, and Wang 2006; Tsai and Lu 2009), it is useful to determine the customers who are going to leave but might reconsider if encouraged to stay. Due to the high expense of such initiatives, management is forced to limit inducement to customers who are most likely to exhibit the behavior of interest. As another example, companies are interested in identifying users who would click on an advertisement if and only if it is highlighted in online advertising (Yan et al. 2009; Bottou et al. 2013; Li et al. 2014; Sun et al. 2015). The challenge in identifying these users stems from the fact that the desired response pattern is not observed directly but rather is defined counterfactually in terms of what the individual would do under hypothetical unrealized conditions. For example, when we observe that a user has clicked on a highlighted advertisement, we do not know whether they would click on that same advertisement if it were not highlighted.

The binary benefit function for the unit selection problem was defined by Li and Pearl (Li and Pearl 2019) (we will call this Li-Pearl’s model), and it properly captures the nature of the desired behavior. Using a combination of experimental and observational data, Li and Pearl derived tight bounds of the benefit function. The only assumption is that the treatment has no effect on the population-specific characteristics. Inspired by the idea of Mueller, Li, and Pearl (Mueller, Li, and Pearl 2021) and Dawid et al. (Dawid, Musio, and Murtas 2017) that the bounds of probabilities of causation could be narrowed using covariates information, Li and Pearl (Li and Pearl 2022c) narrowed the bounds of the benefit function using covariates information and their causal structure. However, the abovementioned studies are based on binary treatment and effect. Recently, researchers have shown interest in developing bounds for probabilities of causation with nonbinary treatment and effect. Zhang, Tian, and Bareinboim (Zhang, Tian, and Bareinboim 2022), as well as Li and Pearl (Li and Pearl 2022a), proposed nonlinear programming-based solutions to compute the bounds of nonbinary probabilities of causation numerically. Li and Pearl (Li and Pearl 2022b) provided the theoretical bounds of nonbinary probabilities of causation. The benefit function is a linear combination of probabilities of causation; therefore, in this paper, we focus on discovering the bounds of any benefit function without restricting them to binary treatment and effect.

Consider the following motivating scenario: a clinical study is conducted to test the effectiveness of a vaccine. The treatments include vaccinated and unvaccinated. The outcomes include uninfected, asymptomatic infected, and infected in a severe condition. The benefited individuals include the following: the individual who would be infected in a severe condition if unvaccinated and would be asymptomatic infected if vaccinated, the individual who would be infected in a severe condition if unvaccinated and would be infected if vaccinated, and the individual who would be asymptomatic infected if unvaccinated and would be uninfected if vaccinated. The harmed individuals include the following: the individual who would be asymptomatic infected if unvaccinated and would be infected in a severe condition if vaccinated, the individual who would be uninfected if unvaccinated and would be infected in a severe condition if vaccinated, and the individual who would be uninfected if vaccinated and would be asymptomatic infected if vaccinated. All others are unaffected individuals. The researcher performing the clinical study has collected both experimental and observational data. The researcher then wants to know the expected difference between benefited and harmed individuals to emphasize the effectiveness of the vaccine.

We cannot apply Li-Pearl’s model because we have two treatments and three outcomes. In this paper, we extend Li-Pearl’s benefit function to general form without restricting...
them to binary treatment and effect. We will provide an algorithm to test the identifiability of the nonbinary benefit function and an algorithm to compute the bounds of the nonbinary benefit function using experimental and observational data.

Preliminaries

In this section, we review Li and Pearl’s binary benefit function of the unit selection problem (Li and Pearl 2019), and the theoretical bounds of the probabilities of causation recently proposed by Li and Pearl (Li and Pearl 2022b).

In this paper we use the language of counterfactuals in structural model semantics, as given in (Galles and Pearl 1998; Halpern 2000). We use $Y_x = y$ to denote the counterfactual sentence “Variable $Y$ would have the value $y$, had $X$ been $x$”. For simplicity purposes, in the rest of the paper, we use $Y_x, y$ to denote the event $Y_x = y$, $y_{x'}$ to denote the event $Y_{x'} = y'$, and $y_{x'}$ to denote the event $Y_{x'} = y'$. We assume that experimental data will be summarized in the form of the causal effects such as $P(y_x)$ and observational data will be summarized in the form of the joint probability function such as $P(x, y)$.

Individual behavior was classified into four response types: labeled compiler, always-taker, never-taker, and defier. Suppose the benefit of selecting one individual in each category are $\beta, \gamma, \theta, \delta$ respectively (i.e., the benefit vector is $(\beta, \gamma, \theta, \delta)$). Li and Pearl defined the objective function of the unit selection problem as the average benefit gained per individual. Suppose $x$ and $x'$ are binary treatments, $y$ and $y'$ are binary outcomes, and $c$ are population-specific characteristics, the objective function (i.e., benefit function) is following (if the goal is to evaluate the average benefit gained per individual for a specific population $c$, $\arg\max_c$ can be dropped):

$$\arg\max_c \beta P(y_x, y_{x'}|c) + \gamma P(y_x, y_{x'}|c) + \theta P(y_{x'}, y_{x'}|c) + \delta P(y_{x'}, y_{x'}|c).$$

Using a combination of experimental and observational data, Li and Pearl established the most general tight bounds on this benefit function (which we refer to as Li-Pearl’s Theorem in the rest of the paper). The only constraint is that the population-specific characteristics are not a descendant of the treatment.

Li and Pearl (Li and Pearl 2022b) provided eight theorems to compute bounds for any type of probabilities of causation with nonbinary treatment and effect. Suppose variable $X$ has $m$ values and $Y$ has $n$ values, the following probabilities of causation are bounded. Besides, if the probabilities of causation are conditioned on a population-specific variable $c$ that is not affected by $X$, then all the theorems still hold (we provided the extended theorems from Li and Pearl in the appendix).

$$P(y_{i,x_j}, y_i),$$

s.t., $1 \leq i \leq n, 1 \leq j \leq m,

P(y_{i,x_j}, y_k),$$

s.t., $1 \leq i, k \leq n, 1 \leq j \leq m, i \neq k,

P(y_{i,x_j}, x_k),$$

s.t., $1 \leq i \leq n, 1 \leq j, k \leq m, j \neq k,

P(y_{i,x_j}, y_k, x_p),$$

s.t., $1 \leq i, k \leq n, 1 \leq j, p \leq m, j \neq p,

P(y_{i,x_j}, x_{i,j}, y_{k}),$$

s.t., $1 \leq i, ... j, k \leq n, 1 \leq j_1, ... j_k \leq m, j_1 \neq ... \neq j_k

P(y_{i_1,x_{j_1}}, ... , y_{i_k,x_{j_{k}}}, x_p, y_{q}),$$

s.t., 1 \leq i, ... , i_k \leq n, 1 \leq j_1, ... , j_k \leq m,

j_1 \neq ... \neq j_k

P(y_{i_1,x_{j_1}}, ... , y_{i_k,x_{j_{k}}}, x_p, y_{q}),$$

s.t., 1 \leq i_1, ... , i_k, q \leq n, 1 \leq j_1, ... , j_k \leq m,

j_1 \neq ... \neq j_k

P(y_{i_1,x_{j_1}}, ... , y_{i_k,x_{j_{k}}}, x_p, y_{q}),$$

s.t., 1 \leq i_1, ... , i_k, q \leq n, 1 \leq j_1, ... , j_k \leq m,

j_1 \neq ... \neq j_k

\arg\max_c \alpha_1 P(y_{x_1}, y_{x_2}, \ldots, y_{x_m}|c) + \alpha_2 P(y_{x_1}, y_{x_2}, \ldots, y_{x_m}|c) + \ldots + \alpha_n P(y_{x_1}, y_{x_2}, \ldots, y_{x_m}|c) + \ldots + \alpha_{n-1+1} P(y_{x_1}, y_{x_2}, \ldots, y_{x_m}|c) + \ldots + \alpha_{n^m} P(y_{x_1}, y_{x_2}, \ldots, y_{x_m}|c).

Note that $c$ can be interpreted as the population-specific variable, the only assumption is that the treatment $X$ has no effect on the population-specific variable. Recall from Li
and Pearl’s paper (Li and Pearl 2019), the benefit vector is provided by the decision-maker who uses the model.

In the next section, we will provide an algorithm that could check whether a given benefit function with the benefit vector is identifiable with purely experimental data (i.e., we can find the exact value of the benefit function rather than bounds). If it is not identifiable we will then provide an algorithm that computes the bounds of the benefit function given the benefit vector using experimental and observational data.

Main Results

Identifiability of Benefit Function

Recall that in binary case, the conditions of identifiable are gain equality (i.e., $\beta + \delta = \gamma + \theta$) or monotonicity (i.e., $P(y'_{x_1}, y'_{x_2}) = 0$) (Li and Pearl 2019). Here, it is complicated in nonbinary cases; therefore; we provide an algorithm to test whether a given benefit function with the benefit vector is identifiable with purely experimental data.

Theorem 1. Suppose variables $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$. Then the benefit function $f(c)$ is identifiable if Algorithm 1 returns $(\text{True}, \text{res})$, and res is the value of the benefit function.

$$f(c) = \alpha_1 P(y_{1x_1}, y_{1x_2}, ..., y_{1x_m} | c) + \alpha_2 P(y_{2x_1}, y_{2x_2}, ..., y_{2x_m} | c) + ... + \alpha_n P(y_{nx_1}, y_{nx_2}, ..., y_{nx_m} | c).$$

The correctness of the algorithm simply follow the fact that $\sum_{\text{terms}} P(..., y_{i}, ..., | c) = P(y_{i} | c)$. Therefore, if there exist such $n^m - 1$ terms in the benefit function, then we can obtain an equivalent benefit function by replacing one of the $n^m - 1$ terms with experimental data $P(y_{i} | c)$. We exhausted all equivalent benefit functions to check if we could replace all the counterfactual terms with experimental data (i.e., identifiable).

For example, consider $m = n = 2$ and the benefit function:

$$7P(y_{1x_1}, y_{1x_2}) + 2P(y_{1x_1}, y_{2x_2}) + 4P(y_{2x_1}, y_{1x_2}) - P(y_{2x_1}, y_{2x_2}).$$

Algorithm 1: Check identifiability of the benefit function

Input: $a$, the benefit function, where $a[i]$ is a $m + 1$ tuple that stands for $i$th term in the benefit function.

1: $m = \text{length}(a);$
2: // Base case, if all benefit vector equals to $(0, ..., 0)$, then the input benefit function is identifiable, and its value equals to the adjusted value.
3: for $i = 1$ to $l$ do
4:   if $a[i][1] \neq 0$ then
5:     $m = \text{false};$
6:     break;
7: end if
8: end for
9: if $m == \text{true}$ then
10:   $\text{Return}(\text{true}, e);$
11: end if
12: // Build an equivalent benefit function by the fact that if $\exists l \leq r \leq (m + 1)s.t., a[j_{1}]r = ... = a[j_{n-1}]r$, then the sum of these $n^m - 1$ terms without coefficients is equal to $P(y_{a[j_1]r} | c)$, then we recursively solve the equivalent benefit function.
13: for every $n^m - 1$ pair in $a, a[j_{1}], ..., a[j_{n-1}]$, s.t., there $\exists l \leq r \leq (m + 1)s.t., a[j_{1}]r = ... = a[j_{n-1}]r$ do
14:   for $k = 1$ to $n^m - 1$ do
15:     $na = a;$
16:     $nc = e + a[j_{1}]r * d[1][r] - a[j_{1}]r;$
17:     for $t = 1$ to $n^m - 1$ do
18:       if $t \neq k$ then
19:         $na[j_{1}]r = na[j_{1}]r - na[j_{1}]r;$
20:       end if
21:     end for
22:     end for
23:   end for
24:   $\text{Remove}(na[j_{1}]r);$  
25:   $\text{res} = \text{IBF}(na, d, nc);$
26:   if $\text{res}[0] == \text{true}$ then
27:     $\text{Return} \text{res};$
28:   end if
29: end for
30: Return $(\text{false}, e);$
|                | Vaccinated | Unvaccinated |
|----------------|------------|--------------|
| Uninfected     | 52 People  | 329 People   |
| Asymptomatic   | 512 People | 58 People    |
| Severe Condition| 36 People | 213 People   |
| Overall        | 600 People | 600 People   |

Table 1: Experimental data of the clinical study. Here, 600 people were forced to take the vaccine and 600 people were forced to take no vaccine.

\[ P(y_{1x_1}, y_{1x_2}, \ldots, y_{nx_m}|c), \ldots, P(y_{2x_1}, y_{1x_2}, \ldots, y_{nx_m}|c), \ldots, P(y_{nx_1}, y_{nx_2}, \ldots, y_{nx_m}|c), \text{ by Li and Pearl’s theorems (Li and Pearl 2022b).} \]

We then have the following theorem.

**Theorem 2.** Suppose variables \( X \) has \( m \) values \( x_1, \ldots, x_m \) and \( Y \) has \( n \) values \( y_1, \ldots, y_n \). Then the bounds of the benefit function \( f(c) \) is obtained by Algorithm 2.

\[
f(c) = \alpha_1 P(y_{1x_1}, y_{1x_2}, \ldots, y_{nx_m}|c) + \alpha_2 P(y_{1x_1}, y_{1x_2}, \ldots, y_{nx_m}|c) + \ldots + \alpha_n P(y_{1x_1}, y_{1x_2}, \ldots, y_{nx_m}|c) + \alpha_{n+1} P(y_{2x_1}, y_{1x_2}, \ldots, y_{nx_m}|c) + \ldots + \alpha_{n+m} P(y_{nx_1}, y_{nx_2}, \ldots, y_{nx_m}|c).
\]

Again, the correctness of the algorithm simply follow the fact that \( \sum_{i=1}^{n} \alpha_i P(y_{ix_1}, y_{ix_2}, \ldots, y_{ix_m}|c) = P(y_{ix_1}|c) \). We exhausted all equivalent benefit functions and take the maximum of all the lower bounds and take the minimum of all the upper bounds of equivalent benefit functions.

**Example: Effectiveness of a Vaccine**

Recall the motivating example at the beginning, a clinical study is conducted to test the effectiveness of a vaccine. The treatments include vaccinated and unvaccinated. The outcomes include uninfected, asymptomatic infected, and infected in a severe condition. The researcher of the clinical study has collected both experimental and observational data.

**Task 1**

The researcher wants to know the expected difference between benefited and harmed individuals to emphasize the effectiveness of the vaccine.

Let \( X \) denotes vaccination with \( x_1 \) being vaccinated and \( x_2 \) being unvaccinated and \( Y \) denotes outcome, where \( y_1 \) denotes uninfected, \( y_2 \) denotes asymptomatic infected, and \( y_3 \) denotes infected in a severe condition. The experimental and observational data of the clinical study are summarized in Tables 1 and 2.

Based on the clinical study, the researcher of the vaccine claimed that the vaccine is effective in controlling the severe condition, the number of severe condition patients dropped from 213 to only 36.

Algorithm 2: Compute the bounds of the benefit function

**Input:** \( a, \) the benefit function, where \( a[i] \) is a \( m+1 \) tuple that stands for \( i \)th term in the benefit function.

If the \( i \)th term is \( \alpha_i P(y_{ix_1}, y_{ix_2}, \ldots, y_{ix_m}|c) \), then \( a[i] = (\alpha_i, i_1, i_2, \ldots, i_m) \).

\( lb \), the lower bound of all possible terms obtained from Li-Pearl’s theorems, where \( lb([i_1, i_2, \ldots, i_m]) \) is the lower bound of \( P(y_{ix_1}, y_{ix_2}, \ldots, y_{ix_m}|c) \).

\( ub \), the upper bound of all possible terms obtained from Li-Pearl’s theorems, where \( ub([i_1, i_2, \ldots, i_m]) \) is the upper bound of \( P(y_{ix_1}, y_{ix_2}, \ldots, y_{ix_m}|c) \).

\( e \), the adjusted value of the benefit function.

The initial call of the algorithm is \( BBF(a[1], \ldots, a[m], lb, ub, 0) \), where \( a[1], \ldots, a[m] \) corresponding to the original benefit function.

All lists in this algorithm start with index 1.

**Output:** \( (lo, up) \), lower and upper bound of the benefit function.

**Function** \( BBF(a, lb, ub, e) \):

1. \( l = length(a) \);
2. // Base case, compute the bounds.
3. \( up = e, lo = e; \)
4. for \( i = 1 \) to \( l \) do
5. \( \quad \text{if } a[i][1] < 0 \text{ then} \)
6. \( \quad \quad lo = lo + a[i][1] * ub([a[i][2], \ldots, a[i][m+1]]); \)
7. \( \quad \quad up = up + a[i][1] * lb([a[i][2], \ldots, a[i][m+1]]); \)
8. \( \quad \quad \text{else} \)
9. \( \quad \quad lo = lo + a[i][1] * lb([a[i][2], \ldots, a[i][m+1]]); \)
10. \( \quad \quad up = up + a[i][1] * ub([a[i][2], \ldots, a[i][m+1]]); \)
11. \( \quad \text{end if} \)
12. \( \text{end for} \)
13. // Build an equivalent benefit function by the fact that if \( 32 \leq r \leq (m+1) s.t., a[j][r] = \ldots = a[j[n-m-1]][r], \) then the sum of these \( n^{m-1} \) terms without coefficients is equal to \( P(y_{ja[j][r]}|c) \), we then recursively solve the equivalent benefit function.
14. for every \( n^{m-1} \) pair in \( a, [a[j][1], \ldots, a[j[n-m-1]]] \), s.t., there \( 32 \leq r \leq (m+1) s.t., a[j][r] = \ldots = a[j[n-m-1]][r] \) do
15. \( \quad \text{for } k = 1 \) to \( n^{m-1} \) do
16. \( \quad \quad na = a; \)
17. \( \quad \quad nc = e + a[j][1] * d[r-1][a[j][r]]; \)
18. \( \quad \quad \text{for } t = 1 \) to \( n^{m-1} \) do
19. \( \quad \quad \quad \text{if } t \neq k \text{ then} \)
20. \( \quad \quad \quad \quad na[j][1] = na[j][1] - na[j][1]; \)
21. \( \quad \quad \quad \text{end if} \)
22. \( \quad \quad \text{end for} \)
23. \( \quad \quad \text{Remove}(na[j][k]); \)
24. \( \quad \quad \text{res} = BBF(na, lb, ub, nc); \)
25. \( \quad \quad lo = \max\{lo, res[0]\} \)
26. \( \quad \quad up = \min\{up, res[1]\} \)
27. \( \quad \text{end for} \)
28. \( \text{end for} \)
29. Return \( (lo, up) \);
### Table 2: Observational data of the clinical study

|                | Vaccinated | Unvaccinated |
|----------------|------------|--------------|
| Uninfected     | 14 People  | 121 People   |
| Asymptomatic   | 933 People | 65 People    |
| Severe Condition| 6 People   | 61 People    |
| Overall        | 953 People | 247 People   |

Now consider the expected difference between benefited and harmed individuals. Recall the benefited individuals include the individual who would be infected in a severe condition if unvaccinated and would be asymptomatic infected if vaccinated, the individual who would be infected in a severe condition if unvaccinated and would be uninfected if vaccinated, and the individual who would be asymptomatic infected if unvaccinated and would be uninfected if vaccinated. The harmed individuals include the individual who would be asymptomatic infected if unvaccinated and would be infected in a severe condition if vaccinated, the individual who would be uninfected if unvaccinated and would be infected in a severe condition if vaccinated, and the individual who would be uninfected if unvaccinated and would be asymptomatic infected if vaccinated. All others are unaffected individuals. In order to maximize the difference between benefited and harmed individuals; therefore, we assign 1 to benefited individuals, assign -1 to harmed individuals, and 0 to all others in the benefit vector. The objective function (i.e., benefit function) is then

\[
f(c) = 0P(y_{1,x_1}, y_{1,x_2} | c) + P(y_{1,x_1}, y_{2,x_2} | c) + \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \qu
Simulated Results

In this section, we show the quality of the bounds of the benefit function obtained by Theorem 2 using four common benefit vectors.

First, we set \( m = 2 \) (i.e., \( Y \) has two values) and \( n = 3 \) (i.e., \( Y \) has three values). We set the benefit vector to one of the most common ones, \((0, 1, 1, -1, 0, 1, -1, -1, 0)\), which is to evaluate the expected difference between benefited and harmed individuals. We randomly generated 1000 populations where each population consists of different fractions of nine response types of individuals. For each population, we then generated sample distributions (observational data and experimental data) compatible with the fractions of response types (see the appendix for the generating algorithm). The advantage of this generating process is that we have the real benefit value (because we know the fractions of the response types) for comparison. Each sample population represents a different instance of the population-specific characteristics \( C \) in the model. The generating algorithm ensures that the experimental data and observational data satisfy the general relation (i.e., \( P(x, y|c) \leq P(y|x|c) \leq P(x, y|c) + 1 - P(x|c) \)).

For a sample population \( i \), let \( a_i, b_i \) be the bounds of the benefit function from the proposed theorem. We summarized the following criteria for each population as illustrated in Figure 1:

- lower bound : \( a_i \);
- upper bound : \( b_i \);
- midpoint : \( (a_i + b_i)/2 \);
- real benefit : dot product of the benefit vector and the fractions of response types;

From Figure 1, it is clear that the proposed bounds obtained from Theorem 2 are a good estimation of the real benefit. The lower and upper bounds are closely around the real benefit and the midpoints are almost identified with the real benefit. Besides, the average gap of the bounds, \( \bar{\sum (b_i - a_i)}/1000 \), is 0.330, which is also small compared to the largest possible gap of 6.

Second, we set the benefit vector to another common one, \((-1, 1, 1, -1, 1, -1, -1, -1, -1)\), which is to evaluate the expected difference between benefited and unbenefited (i.e., unaffected and harmed) individuals. We again randomly generated 1000 populations where each population consists of different fractions of nine response types. The data generating process and all other factors remain the same. We summarized the same criteria for each population as illustrated in Figure 2.

From Figure 2, it is clear that the proposed bounds obtained from Theorem 2 are a good estimation of the real benefit. The lower and upper bounds are closely around the real benefit and the midpoints are almost identified with the real benefit. Besides, the average gap of the bounds, \( \bar{\sum (b_i - a_i)}/1000 \), is 0.6520, which is also small compared to the largest possible gap of 9.

Third, we set the benefit vector to another common one, \((0, 1, 1, 0, 0, 1, 0, 0, 0)\), which is to evaluate the expected benefited individuals. We again randomly generated 1000 populations where each population consists of different fractions of nine response types. The data generating process and all other factors still remain the same. We summarized the same criteria for each population as illustrated in Figure 3.

From Figure 3, it is clear that the proposed bounds obtained from Theorem 2 are a good estimation of the real benefit. The lower and upper bounds are closely around the real benefit and the midpoints are almost identified with the real benefit. Besides, the average gap of the bounds, \( \bar{\sum (b_i - a_i)}/1000 \), is 0.3284, which is also small compared to the largest possible gap of 3.

Lastly, we set the benefit vector to the last common one, \((0, 0, 0, -1, 0, 0, -1, -1, 0)\), which is to evaluate the expected harmed individuals (we set the benefit vector to \(-1 \) because we want to minimize the harmed individuals). We again randomly generated 1000 populations where each
population consists of different fractions of nine response
types. The data generating process and all other factors still
remain the same. We summarized the same criteria for each
population as illustrated in Figure 4.

From Figure 4, it is clear that the proposed bounds obtained
from Theorem 2 are a good estimation of the real benefit. The
lower and upper bounds are closely around the real benefit
and the midpoints are almost identified with the real benefit.
Besides, the average gap of the bounds, \( \frac{1}{1000} \sum (b_i - a_i) \), is 0.3266,
which is also small compared to the largest possible gap of 3.

Discussion

We have shown that the proposed theorems are a good esti-
mation of the non-binary benefit function using examples
and simulated studies. One may concern about the computa-
tion complexity of Algorithms 1 and 2. They are for sure in
exponential time. However, the \( m \) and \( n \) (i.e., values of \( X \)
and \( Y \)) are usually small constant, therefore, we do not need
to worry about too much.

Conclusion and Future Work

We demonstrated the formalization of the general benefit
function with nonbinary treatment and effect. We provided
the algorithm to compute the bounds of the general benefit
function and the algorithm to check whether the benefit func-
tion is identifiable with purely experimental data. Examples
and simulation results are provided to support the proposed
theorems.

Future studies could assess the statistical properties of the
proposed bounds. How tight would the bounds be? Does it
sufficient to make decisions? Which data, experimental or
observational, would affect the bounds more? How would the
number of values in treatment and effect affect the quality of
the bounds?

Acknowledgements

This research was supported in parts by grants from the Na-
tional Science Foundation [#IIS-2106908], Office of Naval
Research [#N00014-17-S-12091 and #N00014-21-1-2351],
and Toyota Research Institute of North America [#PO-
000897].

References

Berson, A.; Smith, S.; and Thearling, K. 1999. Building data
mining applications for CRM. McGraw-Hill Professional.
Bottou, L.; Peters, J.; Quiñonero-Candela, J.; Charles, D. X.;
Chickering, D. M.; Portugaly, E.; Ray, D.; Simard, P.; and
Snelson, E. 2013. Counterfactual reasoning and learning
systems: The example of computational advertising. The
Journal of Machine Learning Research, 14(1): 3207–3260.
Dawid, P.; Musio, M.; and Murtas, R. 2017. The Probability
of Causation. Law, Probability and Risk, (16): 163–179.
Galles, D.; and Pearl, J. 1998. An axiomatic characterization
of causal counterfactuals. Foundations of Science, 3(1): 151–
182.
Halpern, J. Y. 2000. Axiomatizing causal reasoning. Journal
of Artificial Intelligence Research, 12: 317–337.
Hung, S.-Y.; Yen, D. C.; and Wang, H.-Y. 2006. Applying
data mining to telecom churn management. Expert Systems
with Applications, 31(3): 515–524.
Lejene, M. A. 2001. Measuring the impact of data mining
on churn management. Internet Research, 11(5): 375–387.
Li, A.; and Pearl, J. 2019. Unit Selection Based on Counter-
factual Logic. In Proceedings of the Twenty-Eighth Inter-
ternational Joint Conference on Artificial Intelligence, IJCAI-19,
Li, A.; and Pearl, J. 2022a. Bounds on causal effects and application to high dimensional data. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, 5773–5780.

Li, A.; and Pearl, J. 2022b. Probabilities of Causation with Non-binary Treatment and Effect. Technical Report R-516, Department of Computer Science, University of California, Los Angeles, CA.

Li, A.; and Pearl, J. 2022c. Unit selection with causal diagram. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, 5765–5772.

Li, L.; Chen, S.; Kleban, J.; and Gupta, A. 2014. Counterfactual estimation and optimization of click metrics for search engines. arXiv preprint arXiv:1403.1891.

Mueller, S.; Li, A.; and Pearl, J. 2021. Causes of effects: Learning individual responses from population data. Technical Report R-505, Department of Computer Science, University of California, Los Angeles, CA.

Sun, W.; Wang, P.; Yin, D.; Yang, J.; and Chang, Y. 2015. Causal inference via sparse additive models with application to online advertising. In AAAI, 297–303.

Tsai, C.-F.; and Lu, Y.-H. 2009. Customer churn prediction by hybrid neural networks. Expert Systems with Applications, 36(10): 12547–12553.

Yan, J.; Liu, N.; Wang, G.; Zhang, W.; Jiang, Y.; and Chen, Z. 2009. How much can behavioral targeting help online advertising? In Proceedings of the 18th international conference on World Wide Web, 261–270. ACM.

Zhang, J.; Tian, J.; and Bareinboim, E. 2022. Partial counterfactual identification from observational and experimental data. In International Conference on Machine Learning, 26548–26558. PMLR.
**Appendix**

**Proof of Theorems**

**Theorem 1.** Suppose variables $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$. Then the benefit function $f(c)$ is identifiable if Algorithm 1 returns (True, res), and $res$ is the value of the benefit function.

$$f(c) = \alpha_1 P(y_{1x_1}, y_{1x_2}, ..., y_{1x_m} | c) + \alpha_2 P(y_{2x_1}, y_{2x_2}, ..., y_{2x_m} | c) + ... + \alpha_n P(y_{nx_1}, y_{nx_2}, ..., y_{nx_m} | c)$$

**Proof.** The proof is simple. Lines 1 to 12 in Algorithm 1 simply check whether the given benefit function $f$ (encoded as $a$ in the algorithm) is identifiable.

On line 24 of the algorithm, we recursively call on another benefit function $f'$ (encoded as $na$ in the algorithm). Now let's consider how we obtain $f'$.

if there exist $n^{m-1}$ terms in $a$ and $r$, s.t., $a[j_1][r] = ... = a[j_{n-1}][r]$, these $n^{m-1}$ terms are $P(\ldots, y_{a[j_1]}[r], \ldots, y_{a[j_{n-1}]}[r])$, and the sum of these $n^{m-1}$ terms is equal to $P(y_{a[j_1]}[r], x_{r-1} | c)$.

We obtain $f'$ by eliminating $k$th of the $n^{m-1}$ terms in $f$, replacing $k$th term by other $n^{m-1} - 1$ terms and their sum $P(y_{a[j_1]}[r], x_{r-1} | c)$.

Therefore, $f = f' + nc$ where $nc = a[j_k][1] * P(y_{a[j_1]}[r], x_{r-1} | c)$. Thus, $f$ is identifiable if and only if $f'$ is identifiable and $f = f' + nc$.

**Theorem 2.** Suppose variables $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$. Then the bounds of the benefit function $f(c)$ is obtained by Algorithm 2.

$$f(c) = \alpha_1 P(y_{1x_1}, y_{1x_2}, ..., y_{1x_m} | c) + \alpha_2 P(y_{2x_1}, y_{2x_2}, ..., y_{2x_m} | c) + ... + \alpha_n P(y_{nx_1}, y_{nx_2}, ..., y_{nx_m} | c)$$

**Proof.** Similarly to Theorem 1, lines 1 to 12 in Algorithm 2 simply compute the bounds of the given benefit function $f$ (encoded as $a$ in the algorithm).

On line 24 of the algorithm, we recursively call on another benefit function $f'$ (encoded as $na$ in the algorithm). Now let's consider how we obtain $f'$.

if there exist $n^{m-1}$ terms in $a$ and $r$, s.t., $a[j_1][r] = ... = a[j_{n-1}][r]$, these $n^{m-1}$ terms are $P(\ldots, y_{a[j_1]}[r], \ldots, y_{a[j_{n-1}]}[r])$, and the sum of these $n^{m-1}$ terms is equal to $P(y_{a[j_1]}[r], x_{r-1} | c)$.

We obtain $f'$ by eliminating $k$th of the $n^{m-1}$ terms in $f$, replacing $k$th term by other $n^{m-1} - 1$ terms and their sum $P(y_{a[j_1]}[r], x_{r-1} | c)$.

Therefore, $f = f' + nc$ where $nc = a[j_k][1] * P(y_{a[j_1]}[r], x_{r-1} | c)$. Thus, the bounds of $f' + nc$ is the bounds of $f$.

**Li-Pearl’s Bounds of Probabilities of Causation**

The input, $lb$, $ub$, in Algorithm 2 depends on the bounds of probabilities of causation. The bounds of probabilities of causation recently proposed by Li and Pearl (Li and Pearl 2022b) is not conditional $C$. However, nothing is changed if conditioning on a variable $C$ that is not affected by $X$. We listed the conditional version of the eight theorems proposed by Li and Pearl. The proof of eight theorems is exactly the same, except every probability should be conditioned on $C$.

**Theorem 3.** Suppose variable $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{ix_j}, y_{i|c})$, where $1 \leq i \leq n, 1 \leq j \leq m$, is bounded as following:

$$\max \left\{ \frac{P(x_j, y_{i|c})}{P(y_{ix_j} | c) + P(y_{i|c}) - 1} \right\} \leq P(y_{ix_j}, y_{i|c})$$

**Theorem 4.** Suppose variable $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{ix_j}, y_{i|c})$, where $1 \leq i, k \leq n, 1 \leq j \leq m, i \neq k$, is bounded as following:

$$\max \left\{ \left[ \left[ 0, \frac{P(x_j, y_{i|c})}{P(y_{ix_j} | c) + P(y_{i|c}) - 1 + P(x_j, y_{i|c})} \right] \right] \right\} \leq P(y_{ix_j}, y_{i|c})$$

**Theorem 5.** Suppose variable $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{ix_j}, x_k | c)$, where $1 \leq i \leq n, 1 \leq j, k \leq m, j \neq k$, is bounded as following:

$$\max \left\{ \left[ \left[ 0, \frac{P(x_j, y_{i|c})}{P(y_{ix_j} | c) + P(x_j, y_{i|c})} \right] \right] \right\} \leq P(y_{ix_j}, x_k | c)$$

**Theorem 6.** Suppose variable $X$ has $m$ values $x_1, ..., x_m$ and $Y$ has $n$ values $y_1, ..., y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{ix_j}, y_k, x_p | c)$,
where $1 \leq i, k \leq n, 1 \leq j, p \leq m, j \neq p$, is bounded as following:

$$
\max \left\{ \begin{array}{c}
P(y_{i|x_j}|c) + P(x_p, y_k|c) \\
-1 + P(x_j|c) - P(y_{i|x_j}, x_p|c)
\end{array} \right\} \leq P(y_{i|x_j}, y_k, x_p|c)
$$

$P(y_{i|x_j}, y_k, x_p|c) \leq \min \left\{ \begin{array}{c}
P(y_{i|x_j}|c) - P(x_j, y_k|c) \\
P(x_p, y_k|c)
\end{array} \right\}$

**Theorem 7.** Suppose variable $X$ has $m$ values $x_1, \ldots, x_m$ and $Y$ has $n$ values $y_1, \ldots, y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{i|x_j}, \ldots, y_{k|x_{j_k}}|c)$, where $1 \leq i_1, \ldots, i_k \leq n, 1 \leq j_1, \ldots, j_k \leq m, j_1 \neq \ldots \neq j_k$, is bounded as following:

$$
\max \left\{ \begin{array}{c}
\sum_{1 \leq i \leq k} P(y_{i|x_j}|c) - k + 1, \\
m_{\geq 1} \geq \sum_{1 \leq i \leq k} \left( LB(P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, \\
y_{i+1|x_{j_i+1}}, \ldots, y_{i|x_j}|c)) + \\
\sum_{1 \leq j \leq m, s \neq j \neq j_k} \left( LB(P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, \\
y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)) + \\
\sum_{1 \leq j \leq m, s \neq j \neq j_k} \right) \right)
\end{array} \right\}
$$

$P(y_{i|x_j}, \ldots, y_{k|x_{j_k}}|c) \leq \min_{1 \leq i \leq k} P(y_{i|x_j}|c)$

$P(y_{i|x_j}, \ldots, y_{k|x_{j_k}}|c) \leq \min_{1 \leq i \leq k} UB(P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, \\
y_{i+1|x_{j_i+1}}, \ldots, y_{i|x_j}|c))$.

where, $LB(f)$ denotes the lower bound of a function $f$ and $UB(f)$ denotes the upper bound of a function $f$. The bounds of $P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)$ are given by Theorem 6 or 10, the bounds of $P(y_{i|x_j}, \ldots, y_{k|x_{j_k}}|c)$ are given by Theorem 5 or 8, and the bounds of $P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)$ are given by Theorem 7 or experimental data if $k = 2$.

**Theorem 8.** Suppose variable $X$ has $m$ values $x_1, \ldots, x_m$ and $Y$ has $n$ values $y_1, \ldots, y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{i|x_j}, \ldots, y_{i|1x_{j_i-1}}, y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)$, where $1 \leq i_1, \ldots, i_k \leq n, 1 \leq j_1, \ldots, j_k, p \leq m, j_1 \neq \ldots \neq j_k \neq p$, is bounded as following:

$$
\max \left\{ \begin{array}{c}
\sum_{1 \leq i \leq k} P(y_{i|x_j}|c) + P(x_p|c) - k, \\
m_{\geq 1} \geq \sum_{1 \leq i \leq k} \left( LB(P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, \\
y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)) + \\
LB(P(y_{i|x_j}, x_{p}|c)) - 1) \\
\right)
\end{array} \right\}
$$

$P(y_{i|x_j}, \ldots, y_{i|x_{j_i}}|c) \leq \min_{1 \leq i \leq k} P(y_{i|x_j}, x_{p}|c)$

where, $LB(f)$ denotes the lower bound of a function $f$ and $UB(f)$ denotes the upper bound of a function $f$. The bounds of $P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)$ are given by Theorem 6 or 10, the bounds of $P(y_{i|x_j}, \ldots, y_{k|x_{j_k}}|c)$ are given by Theorem 5 or 8, and the bounds of $P(y_{i|x_j}, \ldots, y_{i-1|x_{j_i-1}}, y_{i+1|x_{j_i+1}}, \ldots, y_{k|x_{j_k}}|c)$ are given by Theorem 7 or experimental data if $k = 2$. 
Theorem 10. Suppose variable $X$ has $m$ values $x_1, \ldots, x_m$ and $Y$ has $n$ values $y_1, \ldots, y_n$, and variable $C$ is not affected by $X$, then the probability of causation $P(y_{i_1,x_{j_1}}, \ldots, y_{i_k,x_{j_k}}, x_p, y_q | c)$, where $1 \leq i_1, \ldots, i_k, q \leq n, 1 \leq j_1, \ldots, j_k, p \leq m, j_1 \neq \ldots \neq j_k \neq p$, is bounded as follows:

$$
\min_{1 \leq t \leq k} UB(P(y_{i_1,x_{j_1}}, \ldots, y_{i_{t-1},x_{j_{t-1}}}, y_{i_{t+1},x_{j_{t+1}}}, \ldots, y_{i_k,x_{j_k}} | c)) + \sum_{1 \leq t \leq k} UB(P(y_{i_1,x_{j_1}}, \ldots, y_{i_{t-1},x_{j_{t-1}}}, y_{i_{t+1},x_{j_{t+1}}}, \ldots, y_{i_k,x_{j_k}}, x_{j_t}, y_{r_t} | c)) + \sum_{1 \leq t \leq k} UB(P(y_{i_1,x_{j_1}}, \ldots, y_{i_{t-1},x_{j_{t-1}}}, y_{i_{t+1},x_{j_{t+1}}}, \ldots, y_{i_k,x_{j_k}} | c))
$$

where,
$LB(f)$ denotes the lower bound of a function $f$ and $UB(f)$ denotes the upper bound of a function $f$. The bounds of $P(y_{i_1,x_{j_1}}, \ldots, y_{i_{t-1},x_{j_{t-1}}}, y_{i_{t+1},x_{j_{t+1}}}, \ldots, y_{i_k,x_{j_k}} | c)$ are given by Theorem 7 or experimental data if $k = 2$ and the bounds of $P(y_{i_1,x_{j_1}}, x_p, y_q | c)$ are given by Theorem 6.

Calculation in the Example

Task 1 First by Theorem 7, we have,

$$
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.087,
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.066,
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.063,
0.431 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.523,
0.298 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.355,
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.059,
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.056.
$$

By Algorithm 2, the lower bound came from the following steps,

$$
f(c) = \begin{cases} 0, & 0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.087, \\
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.066, \\
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.063, \\
0.431 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.523, \\
0.287 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.355, \\
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.059, \\
0 \leq P(y_{i_1,x_1}, y_{i_2,x_2} | c) \leq 0.056. \\
\end{cases}
$$
and the upper bounds came from the following steps,

\[ f(c) = 0P(y_{1x_1}, y_{1x_2} | c) + P(y_{1x_1}, y_{2x_2} | c) + P(y_{1x_1}, y_{1x_2} | c) - P(y_{2x_1}, y_{1x_2} | c) + 0P(y_{2x_1}, y_{2x_2} | c) - P(y_{2x_1}, y_{1x_2} | c) - P(y_{3x_1}, y_{1x_2} | c) + 0P(y_{3x_1}, y_{3x_2} | c) = P(y_{1x_1}, y_{2x_2} | c) + P(y_{1x_1}, y_{1x_2} | c) + 0P(y_{2x_1}, y_{2x_2} | c) + P(y_{2x_1}, y_{1x_2} | c) - P(y_{3x_1}, y_{1x_2} | c) - P(y_{3x_1}, y_{2x_2} | c) + 0P(y_{3x_1}, y_{3x_2} | c) = P(y_{1x_1}, y_{1x_2} | c) - P(y_{2x_1}, y_{1x_2} | c) = P(y_{2x_1}, y_{2x_2} | c) + P(y_{2x_1}, y_{3x_2} | c) - P(y_{3x_1}, y_{2x_2} | c) - P(y_{3x_1}, y_{3x_2} | c) + P(y_{3x_1}, y_{2x_2} | c) + P(y_{3x_1}, y_{3x_2} | c) \]

Thus,

\[ f(c) = 0 + 0.355 - 0 + 0 - 0.462 = -0.107. \]

Thus, \( -0.228 \leq f(c) \leq -0.107. \)

\textbf{Task 2} By Algorithm 1, the result came from the following steps,

\[ f(c) = 0P(y_{1x_1}, y_{1x_2} | c) + P(y_{1x_1}, y_{2x_2} | c) + 2P(y_{1x_1}, y_{3x_2} | c) - P(y_{2x_1}, y_{1x_2} | c) + 0P(y_{2x_1}, y_{2x_2} | c) + P(y_{2x_1}, y_{3x_2} | c) - 2P(y_{3x_1}, y_{1x_2} | c) - P(y_{3x_1}, y_{2x_2} | c) + 0P(y_{3x_1}, y_{3x_2} | c) = P(y_{1x_1}, y_{2x_2} | c) + 2P(y_{1x_1}, y_{3x_2} | c) - P(y_{2x_1}, y_{1x_2} | c) + 0P(y_{2x_1}, y_{2x_2} | c) + P(y_{2x_1}, y_{3x_2} | c) - 2P(y_{3x_1}, y_{1x_2} | c) - 2P(y_{3x_1}, y_{2x_2} | c) + 0P(y_{3x_1}, y_{3x_2} | c) = 2P(y_{2x_1}, y_{2x_2} | c) + P(y_{2x_1}, y_{3x_2} | c) - P(y_{3x_1}, y_{2x_2} | c) - 2P(y_{3x_1}, y_{3x_2} | c) + P(y_{3x_1}, y_{2x_2} | c) + P(y_{3x_1}, y_{3x_2} | c) \]

\[ \leq 0 + UB(P(y_{2x_1}, y_{3x_2} | c)) - LB(P(y_{3x_1}, y_{2x_2} | c)) + 0 + +0.58/0.60 + 0.23/0.60 - 0.52/0.60 - 0.36/0.60 \]

\[ = 0 + 0.355 = 0.355 \]

\[ = 0.107. \]

\[ \]

\textbf{Distribution Generating Algorithm}

Here, the sample distribution generating algorithm in the simulated studies is presented. It generated both experimental and observational data compatible with the fractions of response types of individuals. The data satisfy the general relation between experimental and observational data. Note that all four simulated studies shared the same distribution generating algorithm but with different benefit vectors.
Algorithm 3: Generate sample distributions for simulated studies

Input: num, number of samples needed.
Output: num sample distributions (observational data and experimental data).

1: count = 0;
2: while count < num do
3:   //rand(0, 1) is the function that random uniformly generate a number from 0 to 1.
4:   a = [];
5:   for i = 1 to 8 do
6:       a.append(rand(0, 1));
7:   end for
8:   a.append(1.0);
9:   a.sort();
10:  // Each c_k corresponding to a sample distribution.
11:  k = count;
12:  // f is the fractions of response types of individuals, f[0] = P(y_{1x_1}, y_{1x_2}|c_k), ..., f[8] = P(y_{3x_1}, y_{3x_2}|c_k).
13:  f = [];
14:  f[0] = a[0];
15:  for i = 1 to 8 do
16:      f[i] = a[i] - a[i - 1];
17:  end for
18:  // Generate experimental data.
19:  P(y_{1x_1}|c_k) = f[0] + f[1] + f[2];
20:  P(y_{2x_1}|c_k) = f[3] + f[4] + f[5];
21:  P(y_{1x_1}|c_k) = f[6] + f[7] + f[8];
22:  P(y_{1x_2}|c_k) = f[1] + f[4] + f[7];
23:  P(y_{1x_2}|c_k) = f[2] + f[5] + f[8];
24:  // Generate observational data.
25:  P(x_1, y_1|c_k) = rand(0, P(y_{1x_1}|c_k));
26:  P(x_1, y_2|c_k) = rand(0, P(y_{2x_1}|c_k));
27:  P(x_1|c_k) = rand(P(x_1, y_1|c_k) + P(x_1, y_2|c_k), min\{P(x_1, y_1|c_k) + 1 - P(y_{1x_1}|c_k), P(x_1, y_2|c_k) + 1 - P(y_{1x_2}|c_k)\});
28:  P(x_1, y_1|c_k) = P(x_1|c_k) - P(x_1, y_1|c_k) - P(x_1, y_2|c_k);
29:  P(x_2|c_k) = 1 - P(x_1|c_k)
30:  P(x_2, y_1|c_k) = rand(0, min\{P(y_{1x_1}|c_k), P(x_2|c_k)\});
31:  P(x_2, y_2|c_k) = rand(0, min\{P(y_{2x_1}|c_k), P(x_2|c_k)\});
32:  P(x_2, y_2|c_k) = P(x_2|c_k) - P(x_2, y_1|c_k) - P(x_2, y_2|c_k);
33:  //Validate the data, the experimental data and observational data should satisfies the following:
34:  P(x, y|c_k) ≤ P(y_{1x}|c_k) ≤ P(x, y|c_k) + 1 - P(x|c_k).
35:  mark = True
36:  for i = 1 to 3 do
37:      for j = 1 to 2 do
38:         if P(y_{1x_1}|c_k) < P(x_j, y_i|c_k) or P(y_{1x_2}|c_k) > P(x_j, y_i|c_k) + 1 - P(x_j|c_k) then
39:            mark = False;
40:      end if
41:  end for
42:  end for
43:  if mark == False then
44:      continue;
45:  end if
46:  count = count + 1;
47: end while