In [2]:
```
import numpy as np
import pandas as pd

import torch
import pyro
import pyro.distributions as dist

pyro.set_rng_seed(101)
pyro.enable_validation(True)
```

In [3]:
```
df = pd.read_csv('./driver.csv', index_col=0)
```

In [4]:
```
df.head()
```

Out[4]:

	x	y	z
1	0	1.490090	0
2	1	5.170279	1
3	1	7.434170	1
4	0	1.531446	0
5	0	0.935765	0

In [5]:
```
df.describe()
```

Out[5]:

	x	y	z
count	1000.000000	1000.000000	1000.000000
mean	0.498000	2.531169	0.491000
std	0.500246	2.535948	0.500169
min	0.000000	-3.353001	0.000000
25%	0.000000	0.244929	0.000000
50%	0.000000	2.321081	0.000000
75%	1.000000	4.864423	1.000000
max	1.000000	7.895047	1.000000

In [39]:
```
# Calculating do operation using valid adjustment set formula
def estimate_y(x):
    est_sum = 0
    for z in df['z'].unique():
        filter_df = df.query('x == @x & z == @z')
        mean_y = np.mean(filter_df['y'])

        prob_z = df.query('z == @z').size/df.size

        est_sum += (mean_y*prob_z)

    return est_sum
```

In [37]:
```
# E[Y_X=x | X=cx]
def estimate_counterfactual(x, cx):
    est_sum = 0
    for z in df['z'].unique():
        filter_df = df.query('x == @x & z == @z')
        mean_y = np.mean(filter_df['y'])

        filter_df = df.query('x == @cx')
        prob_z = filter_df.query('z == @z').size/filter_df.size

        est_sum += (mean_y*prob_z)

    return est_sum
```

In [38]:
```
print('E(Y_X=0 | X=1) = %.3f' % (estimate_counterfactual(0, 1)))
print('ETT = %.3f' % (estimate_counterfactual(1, 1) - estimate_counter
factual(0, 1)))
print('E(Y_X=1 - Y_X=0) = %.3f - %.3f = %.3f' % (estimate_y(1), estima
te_y(0), estimate_y(1) - estimate_y(0)))
```

```
E(Y_X=0 | X=1) = 3.091
ETT = 1.346
E(Y_X=1 - Y_X=0) = 3.075 - 1.861 = 1.214
```

Analysis

It is clear from the results that the ETT is slight higher than the total effect. This indicates that the training program must have had an impact on the motivation of the drivers results in higher revenues. Having done the ETT analysis, we can safely say that the training program in general is going to result in higher revenues despite the original motivation of the drivers.

In [ ]: