ATE estimations from real RCT data

This notebook examines the use of Bayesian Networks for estimating Average Treatment Effects (ATE) in Randomized Controlled Trials (RCTs) within the Neyman-Rubin potential outcome framework from the STAR trial datset.

Dataset

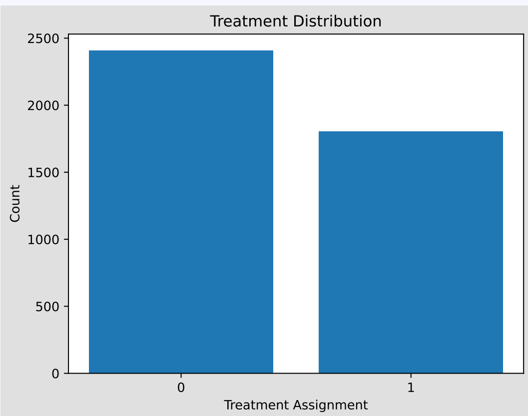
The data used in this notbook come from the Tennessee Student/Teacher Achievement Ratio (STAR) trial. This randomized controlled trial was designed to assess the effects of smaller class sizes in primary schools (T) on students' academic performance (Y).

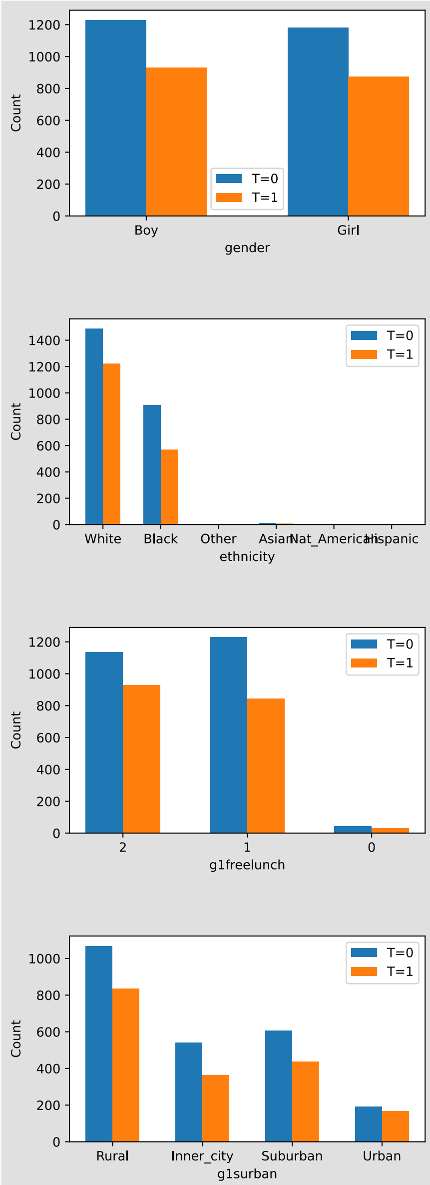
The covariates in this study include:

- `gender`
- `age`
- `g1freelunch` being the number of lunchs provided to the child per day
- `g1surban` the localisation of the school (inner city or rural)
- `ethnicity`

	Y	T	gender	ethnicity	age	g1freelunch	g1surban
0	514.000000	0	Boy	White	4.596851	2	Rural
1	512.666667	0	Girl	Black	5.694730	1	Inner_city
2	470.333333	1	Girl	Black	4.180698	1	Suburban
3	500.666667	1	Girl	White	5.963039	2	Urban
4	516.333333	0	Boy	Black	5.867214	1	Inner_city

Let's examine the distribution of the dataset. An RCT environment should induce the same distribution for both control and treatment groups.





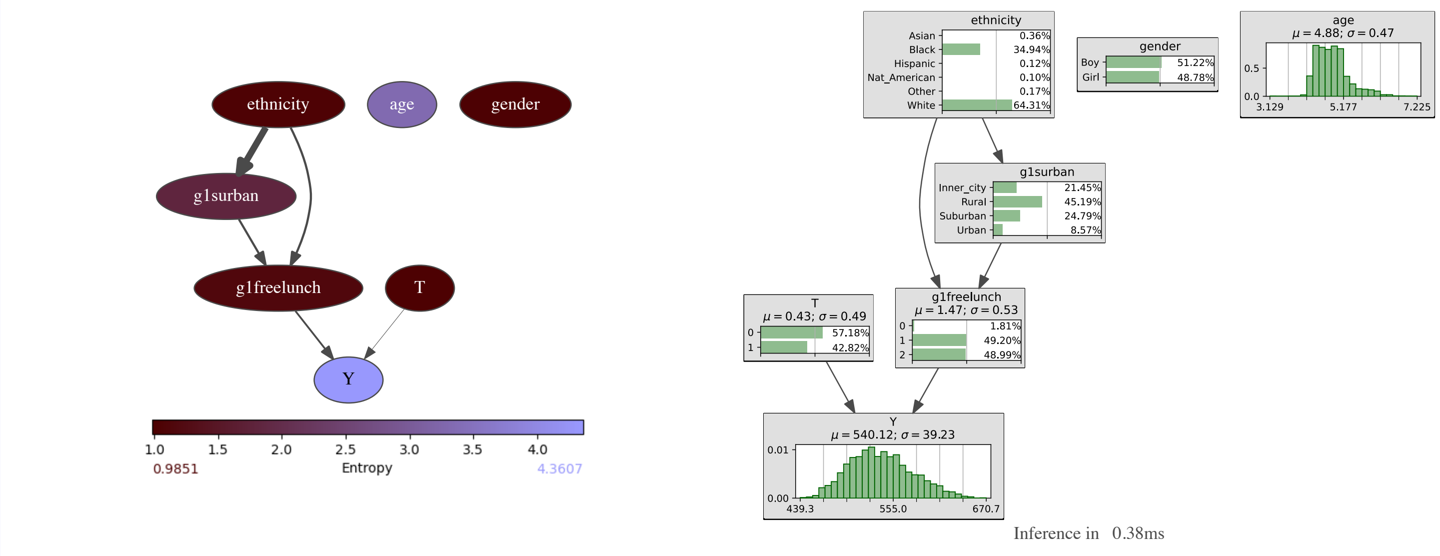
It appears that there are more units in the control group. However, the control and treatment groups appear to be similar in distribution, indicating that the ignorability assumption is likely satisfied.

We proceed with Structure Learning and Parameter Learning to estimate the Average Treatment Effect (ATE) of reduced class sizes (T) on academic performance (Y).

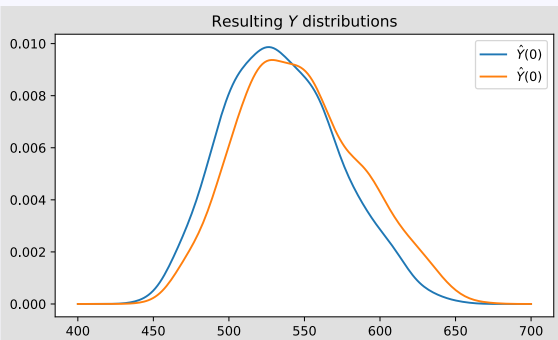
1 - Structure Learning

In the absence of prior knowledge regarding the distributions of the variables and their relationships, structure learning aims to identify the network's underlying structure. To assist the learning process, we will impose a slice order on the variables.

```
Filename               : /tmp/tmpvd9qasfr.csv
Size                   : (4215,7)
Variables              : Y[30], T[2], gender[2], ethnicity[6], age[24], g1freelunch[3], g1surban[4]
Induced types          : False
Missing values         : False
Algorithm              : MIIC
Score                  : BDeu  (Not used for constraint-based algorithms)
Correction             : NML  (Not used for score-based algorithms)
Prior                  : Smoothing
Prior weight           : 0.000001
Constraint Slice Order : {ethnicity:0, T:0, g1surban:1, age:0, gender:0, g1freelunch:1, Y:2}
```



This initial approach appears promising, as the inferred causal relationships are somewhat consistent with what might be expected from an non-expert perspective.



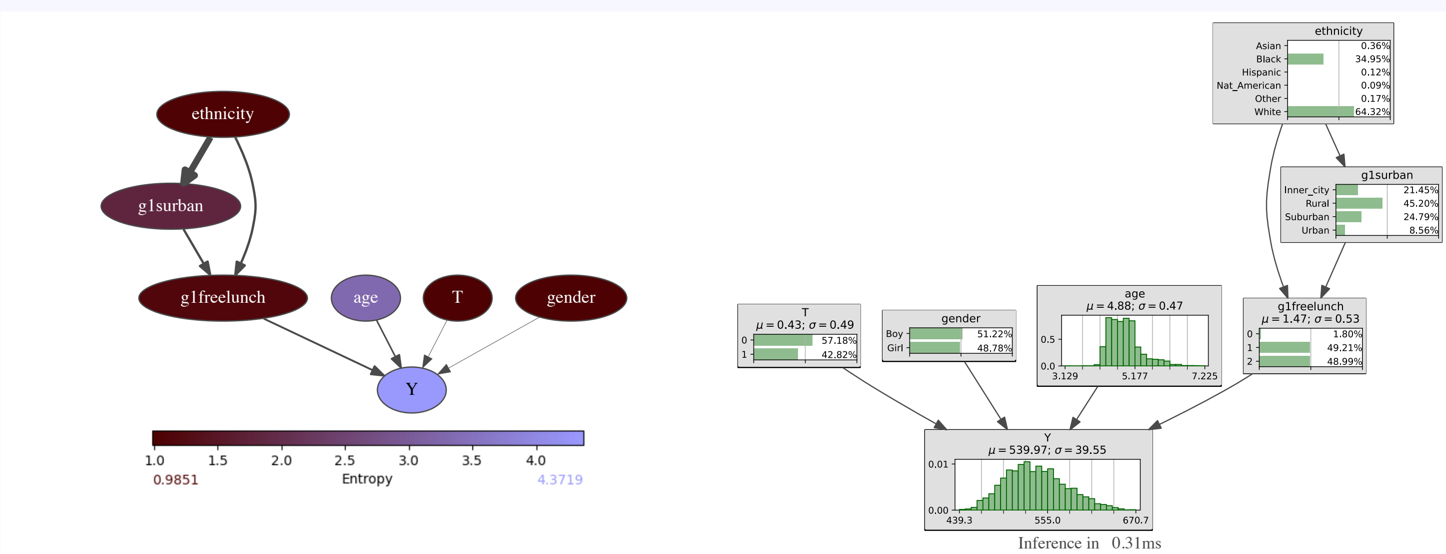

```
Estimated ATE : 11.51375626395145
```

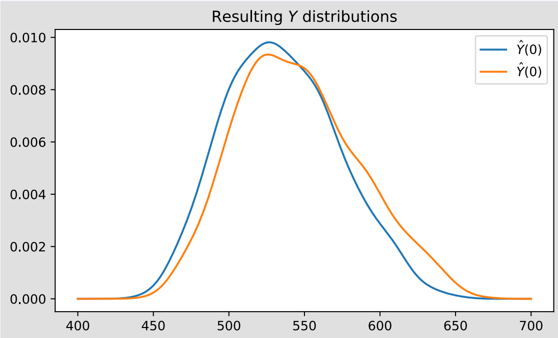
We observe a slight shift in the outcome distribution. However, because the outcome values are in the hundreds, this leads to a significant impact on the treatment effect, considering that the outcome is defined as the average of the students' three grades.

2 - Parameter Learning

Using different structures when conducting parameter learning can yield varying results. For the sake of illustration, we will examine how the estimation performs when arcs from the `age` and `gender` covariates are added to the outcome.

```
Filename       : /tmp/tmpud45te2q.csv
Size           : (4215,7)
Variables      : Y[30], T[2], gender[2], ethnicity[6], age[24], g1freelunch[3], g1surban[4]
Induced types  : False
Missing values : False
Algorithm      : MIIC
Score          : BDeu  (Not used for constraint-based algorithms)
Correction     : NML  (Not used for score-based algorithms)
Prior          : Smoothing
Prior weight   : 0.000001
```






```
Estimated ATE : 10.344241933416356
```

As anticipated, there are observable differences between the parameter learning method and the structure learning method. When compared to direct estimation methods, such as the Difference in Means (DM) estimator and the T-Learner using Ordinary Least Squares (OLS) regression, which yield average treatment effects of 12.81 and 10.77, respectively, our findings remain largely consistent. These results suggest that incorporating age and gender variables into the outcome model may deteriorate the final estimation accuracy.