

Examining the effectiveness of automatic coffee machines

Shuoyu Chen

November 26, 2020

Examining the effectiveness of automatic coffee machines

Link to Github: <https://github.com/hallietiff/Examining-the-effectiveness-of-automatic-coffee-machines.git>

Shuoyu Chen

November 26, 2020

Abstract

In this report, I will analyse how automatic coffee makers more effective than manual ones by statistical analysis, in which Difference in Differences method including creating models and graphs would be selected to help the analysis. Although the automatic machines are a little effective, they can be used in mass manufacturing in some specific factories.

Keywords

Difference in Differences, Causal Inference, Observational Study, Coffee Machine

Introduction

Nowadays, people's lives are more and more convenient due to intelligence appliances. We always use statistical analysis to investigative techniques and help test whether the new version improves the performance in realm of science and technology. From distinctive aspects, we can check the practicality of a new invention. Observational data is easier to be gain ethics approval than experimental data, so making causal inference turns out to be important to define things as convenience.

It would be nice to make causal inference through Difference in Differences which has become widely popular in recent years. In this report, I will use this approach to

find whether the automatic coffee makers can produce coffee more effectively than manual ones.

I create a data set by simulation and use it with difference in differences method to investigate whether using a automatic machine would be more quick than making coffee manually, and the model used is explained in the Methodology section (Section 2). Then the results of difference in differences is indicated in the Results section (Section 3). In the Discussion section (Section 4), I will summarise the results and make conclusions of this analysis.

Methodology

Data

In order to get 2000 cases of time to make a cup of coffee, I simulated 1000 person, everyone would repeat the action twice, and divided them into two treatment groups randomly, one for which is labelled 0 and the other for using the automatic coffee maker labeling 1. The time is simulated from the normal distribution with mean $\mu = 12$ and standard deviation $\sigma = 1$ if this person makes the coffee manually for the first time. Similarly, for the second time has the time simulated from normal distribution with mean $\mu = 10$ and standard deviation $\sigma = 1$. For the situation of using automatic coffee maker, time is simulated from the same distribution but parameter $\mu = 8$ if it is the first time, and $\mu = 6$ when it is the second time. I chose the parameters of $\mu = 6, 8, 10, 12$ and $\sigma = 1$ based on the average time needed to make a cup of coffee (Brian).

Model

The general multiple regression model equation I use in this report is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon$$

where y represents time to produce a cup of coffee, x_1 is treatment group and x_2 is number of trying. β_0 is the intercept of this model, which indicates the time needed when the treatment group and trying are both 0. That is, time to make coffee if the person use manual machine for the first time is β_0 . The coefficient of treatment group being 1 turns out to be β_1 , which means the time would raise by β_1 units when the machine is automatic. Similarly, for the number of trying variable, the coefficient stands as β_2 . If trying of this case is 1, the time would increase by β_2 units. It automatically includes their interaction due to $x_1 x_2$, which is the estimate of β_3 . Therefore, if the person use the automatic coffee maker for the second time, the time would raise by β_3 units. This model is selected since I would like to find the linear relationship between the predictor time and these two variables, number of trying and whether automatic machines or not.

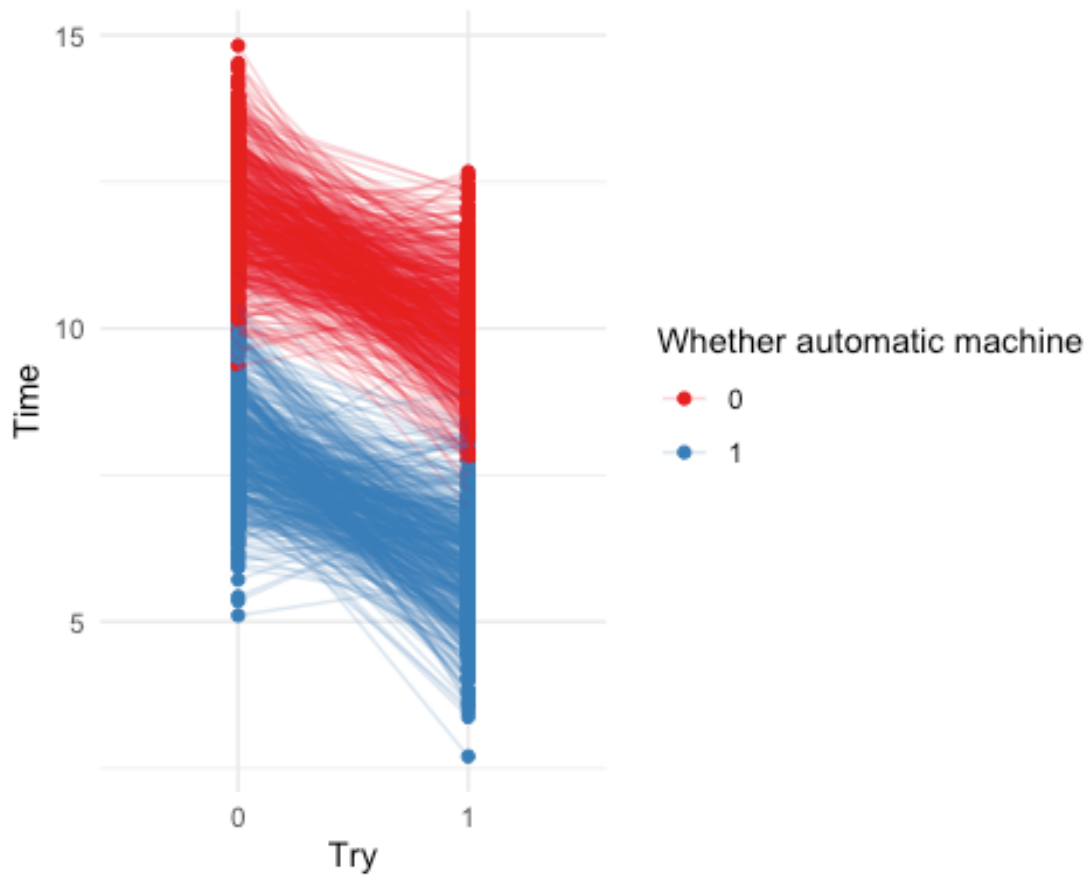
Results

term	estimate
(intercept)	12.0068
automatic coffee machine	-3.9729
the second time	-1.9800
automatic machine for the second time	-0.1035

From the summary table of the model I create above, we see that the estimate of intercept is 12.0068. And for coefficient of treatment group being 1 which represents the automatic machine, the estimate turns out to be -3.9729. That is, the time of producing a cup of coffee would decrease by 3.9729 units using automatic coffee makers. Similarly, the time would drop down by 1.9800 units if the person making it for the second time. In addition, the time reduces by 0.1035 more when the person makes coffee with automatic machines for the second time.

treatment group	average difference
manual coffee maker	-1.9800
automatic coffee maker	-2.0835

The mean difference of time needed to make a cup of coffee manually is -1.9800, and that using automatic coffee maker is -2.0835. Then we compare the absolute value of these two differences and find that the value of automatic machines is a little bit larger than that of manual ones. The average difference between their differences is 0.1035.



It is obvious that blue points representing the automatic coffee machine locate higher part of the plot in general which means they can finish a cup of coffee more quickly than manually, no matter what the number of trying is. Also, there is a little coincidence between these two situations. Apart from this, the time needed of both two approaches decreases gradually for the second time, which means people get used to the machines.

Discussion

Summary

The difference in differences equals to 0.1035 which is the estimate of the effect of the automatic coffee machine, and the multiple regression model is:

$$y = 12.0068 - 3.9729x_1 - 1.9800x_2 - 0.1035x_1x_2 + \epsilon$$

For example, the time of producing a cup of coffee is 8.0339 if the person uses the automatic coffee maker for the first time.

Conclusions

The difference in differences analysis indicated that there is a little reduction in time of making coffee with automatic machines, 5.23%. Although we can save just 5.23% of time to make a cup of coffee, total time period would shrink dramatically if factories or workshops use automatic machines to make thousands of cups of coffee.

Weaknesses and Next Steps

The data set was produced by simulation, which means it is likely to give deviation with the true value of each case. Besides, I just think of the time to make coffee but not the coffee quality, and there are many other factors affecting it like bean choices, water temperature, and so on. The quality of production should be added into consideration of effectiveness.

Next, I would add other factors mentioned above such as grinder type, bean roast date, type of beans and water temperature to analyse more aspects and degrees that they affect the quality of coffee. Then, integrated the quality and time to define the effectiveness of coffee machines.

References

1. Brian. How Long Does A Coffee Maker Take To Brew Coffee? [Online]. Retrieved from <https://www.topoffmycoffee.com/how-long-does-coffee-maker-take-brew-coffee/>
2. Code help: Diff-In-Diff-Tennis.R
3. Course material helps analysis: slide "Difference In Differences"