

Heterogeneous Treatment Effects, Machine Learning, and Experiment Design

導讀

1 What is the main question(s) raised in the paper?

不同的實驗者間即使面對相同的處理,也會產生不同的效果,因此不宜採用完全隨機抽樣的方法進行處理。現有的finite fixed mixture model,雖然同樣都是用來處理族群間,處理效果不相同的問題,但卻有著需要大量資料,與無法處理panel data 兩大缺點。因此,作者希望透過新的方法,能處理不同實驗族群面對處理效果的不同反應,與解決fixed mixture model 的痛點。

2 Why should we care about it?

並非所有實驗參與者面對同一處理都會產生相同效果,不適合進行完全的隨機抽樣,因此我們必須將不同的效果考慮進去。

3 What is the author's answer?

作者提供了一套two-step alternative 方法,透過對實驗者進行反覆抽樣,便不需要大量的資料,就可以同時進行分類處理,又不會產生data hungry 的問題。

4 How did the author get there?

作者結合了機器學習與傳統的統計學方法,透過兩階段的方法,先對受試者反覆進行實驗,透過local regression 找出會受時間影響的變數,接著利用二元樹,將實驗者按篩選出來的變數進行分類,便可以輕易的控制random effect 和mixed effect.