tworavens使用笔记

界面布局

左边是 -- -- 变量选择、subset选择

中间是 --- 显示界面, 主要是变量之间的关系结构

右边是 -- -- 选择评估模型

具体部分

正上方

数据集基本信息

中间的正上方是数据集的名字,鼠标放上去会显示一些详细信息

repdata

Acharya, Avidit, Blackwell, Matthew, Sen, Maya, 2015,

"Replication Data for: "Explaining Causal Findings
Without Bias: Detecting and Assessing Direct Effects"",
doi:10.7910/DVN/VNXEM6, Harvard Dataverse, V1,
UNF:6:3OeRtVBgwh7ZPT6weqWvPQ==

Variable transformation (变量转换)

右上角的栏,选择一个现在的变量,然后可以继续选择一些对当前变量的操作,比如取对数、开方(前提是能够做这样的操作,如果一个变量不能做这样的操作,会报错,然后创建失败)也可以写R语句创建一个新的变量,具体操作如下:

- 选择一个变量
- 写R语句, 比如赋值语句, 回车即可实现这条语句
- 一个例子 ccode^2



重制按钮

按了之后,相当于刷新页面,重新加载一次。之前的所有操作的撤销



Esitimation按钮

- 按下去就会执行和创建左边的那个变量
- 回车同理
- 创建失败, 会弹出错误警告窗口



左边

- 左边会列出数据集中当前的所有数据。
- 使用者首先需要理解这些数据是什么意义,然后选择其感兴趣的数据做分析。
- 除了基本数据, 还有observations 和 population等。

Variable Tab (变量列表)

- 变量列表,列出了所有的变量。你可以选择哪些显示在中间、做分析。
- 自变量和因变量的颜色会有不同,作为区分。
 - 。 自变量 橘红色
 - 。 因变量 浅蓝色



Variable List (变量信息列表)

把鼠标移动到某一个特定的变量上面,就会显示一些与该变量相关的信息、值。列出了部分变量值,全部内容可见截图。

- Stand Dev 标准差
- Uniques不同取值个数
- Herfindahl 赫芬达尔—赫希曼指数



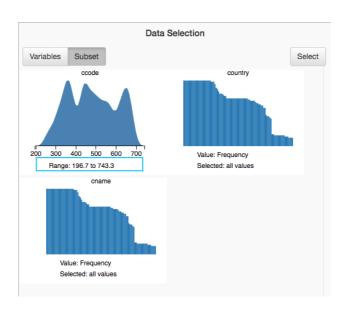
Subset Tab (已选择数据列表)

显示当前已经选出来的数据,可以在直方图或者密度图上,选择一块,然后点击Select按钮,之后中间的数据就只有这一部分了。

Select前:



Select后:



可以看到, 值的范围明显变小了。

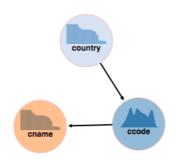
中间

- 就是显示各个变量的关系的,变量用pebbles形式表现。
- 主要是通过线的连接,表现出一个 变量之间的关系的网络图,其名字叫 directed graph

Model Builder (模型生成器)

Pebble 每一个都是对应的一个变量,它中间的那个图,表明了现在是在对当前的数据进行哪一种分析,图形和subset的是一样的

然后鼠标在移上去的时候,数据的周围会出现一些选项,就可以把数据变成一些不同类型的变量(因变量or 自变量)不同的种类的变量,会用不同的颜色轮廓来区分,右下角会给标注。 Pebble 可以用箭头连起来,右键点变量就可以开始画一个箭头。 用箭头来代表可能的因果关系。A指向B,就说明A是因、B是果。可以双向。点一下箭头,就可以删除了。



比如这个例子

- country指向ccode,则country是因,是自变量;ccode是果,是因变量
- ccode, cname同理
- 注意! 一定要确认因变量、自变量之间正确的逻辑,否则之后计算可能出现错误

Pin Button (布局按钮)

就是右上角那个大头针按钮,不按的话,Pebble是自动控制布局,你没法控制;按了之后,就可以自己控制。



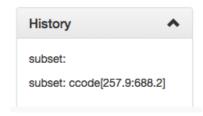
Wipe Button

清空中间布局。



History Log(历史记录)

左下角, 会记录你有哪些操作。

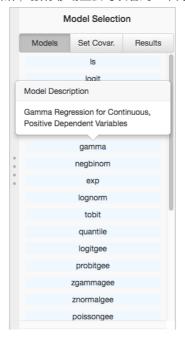


右边

主要是对数据做一些分析学上的操作,得到结果。

Models Tab (模型界面)

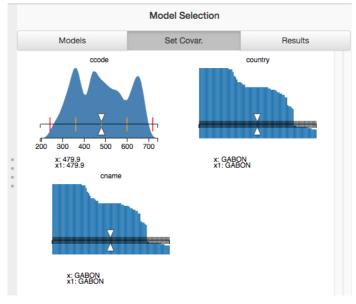
列出了TwoRavens可以进行的分析操作,鼠标移动上去可以看到一个简单介绍。点击一下就可以做选择。



Set Covariates Tab (数据交互)

主要是对选出来的数据做一些交互

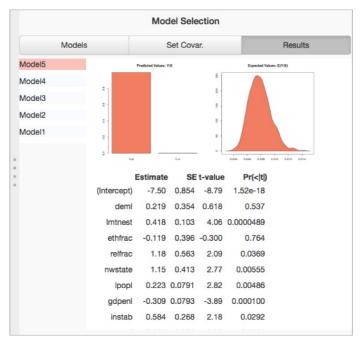
- 上面那个三角形,代表的是你选择的变量的baseline(起点),下面是想变化到的部分,就是上面那个三角形指的值,变化到下面那个三角形指的值,然后继续处理。
- 不同颜色的线代表含义不同
 - 。 绿色的是平均值;
 - 。 橙色是偏差一个标准差;
 - 。 红色是偏差两个标准差、且不断递推。
 - 。如果是离散变量则没有这些区别。



Results Tab

分析出来的结果。

具体操作可参考: 《Teaching with TwoRavens: Replicating Fearon and Laitin (2003)》,本教程在附录中可见。



备注

本文撰写时参考了如下指南,限于本人英语水平有限,可能存在部分错误,如有发现,请联系foxerlee1@gmail.com指出,十分感谢!

- 《A User Guide to TwoRavens: An overview of features and capabilities》, Vito D'Orazio,
 James Honaker
- 《Teaching with TwoRavens: Replicating Fearon and Laitin (2003)》, Vito D'Orazio