

随机化模拟案例

数据来源：2016年香港人口中期普查数据

研究对象：34562名18-55岁香港本地劳动者

模拟说明：将34562名劳动者随机分成两组，其中控制组为17281名，实验组为17281名，比较两组劳动者的每月工作收入（salary）的差异。若随机化机制（randomize assignment）得以完全实现，控制组和实验组劳动者的每月工作收入均值和分布高度相似，在统计意义上没有显著差异。本次模拟重复实现500次。

程序代码：

****使用统计软件Stata 15.1**

version 15.1

****设置数据来源路径**

cd "F:\Data Storages\香港人口数据\已处理"

****读取数据**

use hk_april_new.dta, clear

****保留2016年劳动人口**

keep if wave == 2016

****设置随机种子**

set seed 12121

****进行随机分组模拟，用t-test比较组间差异**

```
forvalues i = 1/500 {  
    preserve  
    quietly: randomtag , count(17281) gen(t)  
    quietly:recode t (0 = 1 "control") (1 = 2 "experiment"), gen(group)  
    logout, save(d:/Tables/Table_`i') excel replace:ttest main_wage , by(group)  
    hist main_wage, by(group)  
    quietly:graph export "d:/Graphs/Graph_`i'.png", as(png) width(800) replace  
    graph close  
    restore  
}
```