

Udacity 教學網站註冊效果之 A/B Test 分析 (進階篇)(4/27更新)


[課程閱讀](#)

[本日作業](#)

[問題討論](#)

[學習心得\(完成\)](#)
[重要知識點](#)

[A/B Test \(Udacity網頁改善專題\)](#)

[知識點回顧](#)

[延伸閱讀 - A/B Test \(Udacity網頁改善專題\)](#)


重要知識點

標準差與樣本數的計算

- 標準差：詳見範例程式中 4.2.2 Estimating Analytically
- 樣本數：詳見範例程式中 5. Experiment Sizing

假設檢驗的計算與判斷

- 假設最小間距 (Minimum distance) GC : 0.01 · NC : 0.075
- 計算信賴區間(Confidence Interval) : $\text{mean} \pm Z\text{-score} * \text{Stdev}$
- 判斷技巧：基礎數值(baseline value) 是否落在 信賴區間(CI) 外，是則現實上(practically)顯著

Udacity 希望了解，在免費 14 天 試學網頁上，除了要信用卡資訊外，還了解學生願意花多少小時學。如果少於某門檻(5 小時)，就建議學生不要註冊，免費聽聽影音就好，免得浪費資源，降低學習成功率。我們的題目是，增加這個頁面，是否對 Gross Conversion(GC) 和 (Net Conversion) NC 在統計學上 ($\alpha=0.05$ ， $\text{Power}=0.8$) 有幫助 ($d=0.01/0.0075$)，亦即統計上的顯著 (Significant)。

- CI = click 數目
- GC = 註冊數/CI (聽了建議仍然註冊的比例)
- NC = 繳費數/CI (14 天之後繳費且繼續的比例)
- 我們期待 GC 比原來下降，但 NC 不降，這表示省去資源但收入不降。
- 程式檔名：ab-tests-with-python.ipynb
- 數據：
experiment_data.csv, control_data.csv

-
- 多用在網頁效能增進 Click Thru Rate (CTR) 的比較：改進是否顯著
 - 每一個 Click 都是 Bernoulli 實驗，多個就是二項式分布
 - 數量大我們用或然率表達，且因中央極限定理我們可以常態分佈之 Z-Score 計算
 - 假設檢定：A 與 B 的平均值是否相同？
 - 在蒐集資料前需要
 - 算出 Std (Standard Deviation)
 - 要決定樣本數 (Sample Size) (type I error : 0.05, Power = 0.8, $d = 0.01/0.0075$)

Python 函數計算各個標準差與樣本數

知識點回顧

標準差與樣本數的計算

- 標準差：詳見範例程式中 4.2.2 Estimating Analytically
- 樣本數：詳見範例程式中 5. Experiment Sizing

假設檢驗的計算與判斷

- 假設最小間距 (Minimum distance) GC : 0.01 , NC : 0.075
- 計算信賴區間(C Confidence Interval) : $\text{mean} \pm Z\text{-score} * \text{Stdev}$
- 判斷技巧：基礎數值(baseline value) 是否落在 信賴區間(CI) 外，是則現實上(practically)顯著

延伸閱讀 - A/B Test (Udacity網頁改善專題)

A/B Test 迷思解惑：

網站：[A/B testing 常見迷思與問題大解析！. 常有人會誤會 AB Testing... | by Nana Chiang | 3PM LAB | Medium](#)

網站：[A/B testing 你真的做對了嗎？產品開發團隊必懂的五大方法論 | by Lucas Hsieh | AsiaYo Engineering | Medium](#)

[下一步：閱讀範例與完成作業](#)

