**109-1 迴歸分析\_期末報告**

2018年公投第14案：  
廣義線性迴歸模型探討

**R08227112 林子堯**

國立台灣大學心理學研究所

# 壹、緒論

## 一、前言

在2018年11月24日，中華民國全台各地舉行了全國性的公民選舉投票。在公投編號的第10到12案中與第14到15案分別由「下一代幸福聯盟」與「同運組織」提出，在內容上對於同志婚姻平權、立法保障和同志教育的議題上卻有完全分歧的觀點。在社會上偏向性別平權的一方，倡導要投出「兩好三壞」（14、15投案同意，10、11、12案投不同意）。相反的，若是偏向傳統觀念或是愛家聯盟觀點的人民，則是說要投成「三好兩壞」（10、11、12投案同意，14、15案投不同意）。雖然說在者兩組提案並非完全相反或互斥，只是在社會、網路社群上的意見常常是完全支持一方而排斥另一組。

對於這樣的現象，本研究以公投的第 14 案作為主要探討對象，該提案的內文為：「**您是否同意，以民法婚姻章保障同性別兩人建立婚姻關係？**」，說明的是要不要採用民法來規範同志婚姻制度。因此對此案表示同意表示著可以用民法來保障。相反的，若是表示不同意，也僅僅說明的指示不採用民法來保障同志婚姻，但不排除用其他的法律形式來保障。然而，大眾對此案的傾向上同意者通常會對同性戀議題表現出正向的態度，而對此案表示不同意者通常也會對性戀議題表現出負面的態度。因此本研究有預設可以透過第 14 案之同意率結果來與同性戀正負向態度連結，但需謹記著，本公投的結果和同性戀態度實際上是兩回事，之中可能有某種程度的相關而已。

## 二、文獻回顧

      關於同性戀與不同社會人口因子的關係在過往便有相當多研究探討。Herek (1984) 探討恐同症的文化因素時即列舉了許多對同性戀持負面態度的人較可能擁有的性質，包含年齡較大、較沒受到完好教育、住在鄉下或小城鎮（特別是那些持負面態度為其常態的區域），且男性通常比女性更容易持負面態度；Seltzer (1992) 也得到與上述相同的結果，除此之外，他的研究還顯示結過婚的人也較易持有負面態度。然而並非所有研究都得到相同的結果，Bowman (1979) 及 Glenn & Weaver (1979) 都沒有找到顯著的性別差異； West & Cowell (2013) 的研究則顯示年齡較大並不能預測較強的負面態度，甚至年齡較大者反而負面態度較小，此外他們也發現高收入者持較正面的態度，但也有其他研究顯示並沒有關係 ( see also Gelbal & Duyan, 2006；Becker & Scheufele, 2009)。另一方面，Seltzer (1992) 發現教育程度與年齡有交互作用的影響：受到大學教育以上的中年人比起老年、青年人都持有較正面的態度，王晴峰（2013）也表示處於青少年與大學年齡組的男性比中年人更可能持強烈的否定立場。

還有其他研究將焦點放在區域、宗教、種族，如Jäckle & Wenzelburger (2015) 其中包含：宗較、年紀、家中小孩人數、有無結婚、收入、教育狀態……等，都有顯著影響。其中特別是宗教信仰最能影響同性戀態度的取向。Adamczyk & Liao (2019) 也於近期進行過後設分析，將主要影響各國同性戀態度的原因歸類至經濟、民主、宗教三個主要因子。或許這些變項對同性戀態度亦有相當程度的影響，但我們並無此方面的資料且台灣並無宗教信仰的普查，故能期望後續研究得以加入。

## 三、研究目的與假設

本研究在2018 年公投第 14 案之同意率與同性戀態度有某種程度以上關聯的假設下，試以「廣義線性回歸模式」(generalized linear regression model, GLM) 找出影響同意率的主要因素為何。由於資料收集來源的限制，最終以該公投的同意率作為反應變項，而預測變項分別為：投票率、是否位於六都、性別比、年齡中位數、教育程度、結婚率以及薪水中位數，以及這七個變項之間的交互作用項來建立模型。依照文獻的結果，本研究預測在六都比起非六都、或是教育程度越高地區，對於同意率會偏向較高；而性別比越高、結婚率越高的地區，則同意率越低。至於在文獻中尚未探討過的投票率，和過去研究沒有共同結論的年齡中位數和薪水中位數對於同意率，正向、負向或是無關連都有可能。

# 貳、資料來源與前處理

## 一、資料來源與前處理

為了研究全台灣 2018 年公投第 14 案之同意率之可能影響因子，試以建立多元線性迴歸模型來找出之間的關聯。在本研究中，將以全台 368 個「鄉、鎮、市、區」（中華民國的[直轄市](https://zh.wikipedia.org/wiki/%E7%9B%B4%E8%BD%84%E5%B8%82_(%E4%B8%AD%E8%8F%AF%E6%B0%91%E5%9C%8B))、[縣](https://zh.wikipedia.org/wiki/%E7%B8%A3_(%E4%B8%AD%E8%8F%AF%E6%B0%91%E5%9C%8B))、[市](https://zh.wikipedia.org/wiki/%E5%B8%82_(%E4%B8%AD%E8%8F%AF%E6%B0%91%E5%9C%8B))所管轄的下級[行政區](https://zh.wikipedia.org/wiki/%E8%A1%8C%E6%94%BF%E5%8D%80)）作為研究資料單位，其中反應變項（應變項）為該公同的同意率，預測變項（獨變項）分別有：投票率、是否位於六都、性別比、年齡中位數、教育程度、結婚率以及薪水中位數，共七項。其中，各鄉鎮市區之同意率與投票率從政府資料開放平臺「第14案全國性公民投票計票結果」而來；內政資料開放平臺「村里戶數、單一年齡人口」和「15歲以上現住人口按性別、年齡、婚姻狀況及教育程度分」分別提供了 2018 年 10 月的人口年齡結構分佈和婚姻狀況及教育程度；最後，財政部財政資訊中心提供了「105年度綜合所得稅申報核定統計專冊」，我們可以從中獲取各鄉鎮市區的薪水中位數，然而可惜的是目前財政部最新的資訊只提供到了民國 105 年（西元 2016 年），因此只能先以此筆資料作為薪水中位數的處理。在下方將一一介紹每個變項的定義。

公投第 14 案之同意率為該地區在該案的同意票數除以總有效票數；投票率為該地區實際對該案進行投票人數在符合資格選民中所佔的比例；是否位於六都為類別變項，若該地區位於六個直轄市地區（中華民國之第一行政區，包括：台北市、新北市、桃園市、台中市、台南市、高雄市）之內，則標記為 1，其他地區則標記為 0；性別比為該地區的符合資格選民（18 歲以上）之男性人口總數除以女性人口總數；年齡中位數如同字面上所意，然而我們選定的範圍是該地區符合資格選民（18 歲以上）的年齡中位數；結婚率則是定義為曾經有結果婚的比率，因此在計算上則是將該地區目前是有配偶、喪偶以及離婚的人數總和除以該地的 15 歲以上之人口總數 ，由於原始資料的將每五歲作為一個級距統計，因此無法切出符合資格選民的年齡（18 歲以上）人數；教育程度則是定義在具有大專生以上教育程度的比例，因此在計算上則是將該地區的大學、大專、碩士和博士以上的人數總和除以該地的 15 歲以上之人口總數；最後的薪水中位數也如同字面上所意，但此變項則是直接採財政部的以統計好的資料，因此我們並未對數據進行任何計算處理。

## 二、變數概況

在下表1中，顯示七個變項的基本描述統計概況。而在全台 368 個鄉鎮市區中，其中有 158 個地區是屬於六都，剩下的 210 個地區不屬於六都範圍內。而在下圖1中，線顯示七個變項的直方圖，同時以顏色來區分是否為六都來觀測是否不同位置在該變數上有不同的趨勢出現。從中大致上可發現是否為六都在其他變項上有不同的表現，可能暗示著這些變項中有可能有交互作用項的存在。

在圖2中，則是將所有變項的資料繪製在台灣地圖上（參考連結：<https://xup6y3ul6.github.io/ALSM_final_report/referendum14_map.html#1>），其中各個鄉鎮市區的顏色代表著對此公同案的同意率程度，當用滑數點擊該地區時（如圖2所示），則可以顯示其他變項的數值，以利在資料分析之前大致了解各個地區的狀況。值得注意的是，在六都的地區相較於非六都地區同意率相較起來較高，然而在所有地區同意率最高的地方卻是位於「連江縣東引鄉」，比起其他同樣是離島地區（普遍同意率偏低）的來說可能是比較特別的存在。這也確實在我們後續的迴歸模型建立上，造成非常具有影響力的個別資料點。

表 1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 反應變項 | | 預測變項 | | | | | |
|  | 同意率 | 投票率 | | 性別比 | 年齡  中位數 | 結婚率 | 教育程度 | 薪水中位數(千元) |
| 最小值 | 0.184 | 0.291 | | 0.833 | 40.00 | 0.590 | 0.117 | 587.0 |
| Q1 | 0.277 | 0.505 | | 0.986 | 45.00 | 0.652 | 0.214 | 691.0 |
| 中位數 | 0.305 | 0.542 | | 1.047 | 47.00 | 0.666 | 0.275 | 735.0 |
| 平均值 | 0.301 | 0.534 | | 1.053 | 47.39 | 0.668 | 0.290 | 784.9 |
| Q3 | 0.326 | 0.577 | | 1.118 | 50.00 | 0.683 | 0.356 | 823.2 |
| 最大值 | 0.487 | 0.689 | | 1.582 | 55.00 | 0.753 | 0.653 | 1722.0 |



圖 1

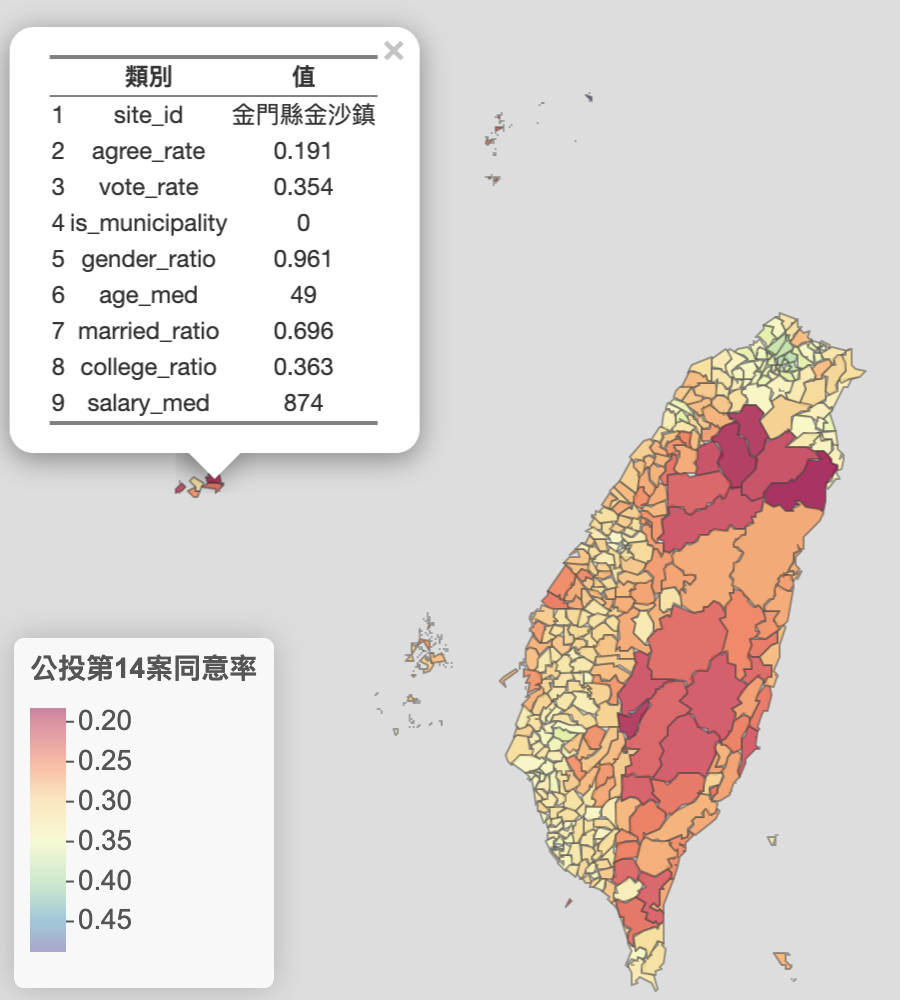


圖 2

# 參、迴歸模型

1. **基本流程概述**

本研究透過以下流程建立回歸模型：

1. 抓取資料及資料清洗
2. 建立一階模型
3. 整斷與修補模型 (變異數同質和常態分佈假設)
4. 建立其他候選模型
5. 檢查共線性、離群值、影響力指標並作修整
6. 靴拔重抽法
7. 選出最終模型

本段將沿該順序來呈現模型建立時的過程與問題，其中第1部分已於前段探討，因此本段將從第2部分開始呈現，並在最後選擇本研究最終回歸模型。在模型發展過程的第4部分將建立三個候選模型，並且每個候選模型都將進行之後的數據分析，但是呈現的數據與圖表將以最終模型的結果為主。

1. **一階模型建立**

透過先前的文獻回顧，本研究決定了影響14案公投同意率的可能影響因子，並在前段中詳述各項獨變項的情形。而在將前段中提及的變項納入回歸模型後，本研究建立以下一階回歸模型:

其中，Agree 為同意率、Ｖote 為投票率、BigCity 為是否位於六都（類別變項）、Gender 為性別比、Age 為年齡中位數、Married 為結婚率、Education 為教育程度以及 Salary 表示為薪水中位數。

之後透過最小平方法分析資料並得到一階回歸模型的參數估計值，並分別透過t檢定分別檢定該係數是否達到顯著，其中虛無假設回歸係數為0對立假設為回歸係數不為0，α =0.05自由度為360。檢定結果如下表所示：

表 2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) | Signif |
| Intercept | 0.357 | 0.056 | 6.179 | 0.000 | \*\*\* |
| vote | -0.075 | 0.028 | -2.627 | 0.009 | \*\* |
| BigCity1 | 0.014 | 0.004 | 4.021 | 0.000 | \*\*\* |
| Gender | -0.036 | 0.025 | -1.476 | 0.141 |  |
| Age | 0.003 | 0.001 | 4.373 | 0.000 | \*\*\* |
| Married | -0.278 | 0.096 | -2.891 | 0.004 | \*\* |
| Education | 0.206 | 0.041 | 5.039 | 0.000 | \*\*\* |
| Salary | 0.000 | 0.000 | 0.290 | 0.772 |  |

以上檢定結果我們能發現性別與薪水中位數在顯著性檢定中並未能拒絕迴歸係數為0的虛無假設，因此在接下來的模型分析中也將探討是否應包含這兩變項。而其他變項所估計的迴歸係數中，結婚率與投票率的迴歸係數為負，而是否為六都、年齡中位數與高等教育比率的迴歸係數為正，與前先對同性戀態度的文獻結果相符。

1. **一階模型診斷與轉換檢查**

而在建立模型後為，確保所建立的模型中隨機變數符合常態分配與變異數同質的假設，將觀察估計值-殘差散佈圖( )與常態機率點圖(normal probability plot )以確認該模型符合假設。

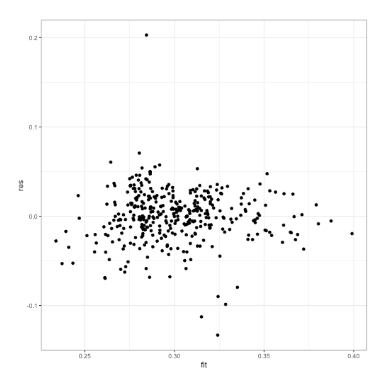


圖 3

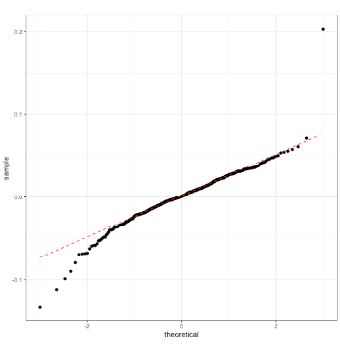


圖 4

上圖3與圖4分別為估計值-殘差散佈圖與常態機率點圖。從圖3中我們能發現殘差的分布呈現中間較分散兩側較窄的情形，可能違反變異數同質性的假設，因此對變異數同質進行進一步檢定。本研究將進行Brown-Forsythe檢定與Breusch–Pagan檢定，其中Brown-Forsythe檢定的虛無假設為該筆資料變異數同質，對立假設為該筆資料變異數不同質，並進行α =0.05自由度為366的t檢定，而臨界值= 1.97，結果tBF = 2.35 > 1.97，故拒絕虛無假設，變異數不同質。進行Breusch–Pagan檢定，虛無假設為該分配符合常態分配且變異數同質，並進行α =0.05自由度為7的卡方檢定，臨界值為2.16。結果BP=65.631>2.16，故拒絕虛無假設，該分配不為變異數同質的常態分配。

在圖4中則發現並非所有點皆落於對角線上，可能違反常態分配假設。對常態分配理論值與資料實際值的相關係數進行假設檢定，虛無假設為分配不為常態分配α =0.05，則常態分配理論值與實際值的相關係數為0.96也並未大於0.996的臨界值，因此無法拒絕虛無假設，故保留虛無假設:該分配不為常態分配。模型違反了以上兩點假設，因此可能須對模型進行轉換以符合假設，但是從上面兩圖中皆能發現存在明顯界外值，該資料點為連江縣的東引鄉，因此上述情形可能是由於界外值造成。

為了使本模型符合模型假設，因此以下將進行Box-Cox轉換，並找出最適合的λ 值使平方誤差總和(SSE)最小。而下圖5為在不同λ 下的平方誤差總和。

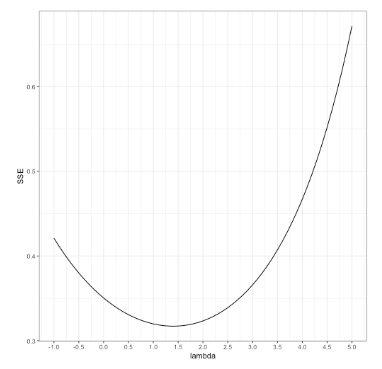


圖 5

透過圖5我們能發現在λ 位於1至1.5之間時，SSE最小，SSE最小的點位於λ =1.3時。則上述結果代表原本的回歸模型可能不需要經過轉換，因此造成模型違反假設的原因可能是因為界外值對模型的影響過大造成。

1. **候選模型探討**

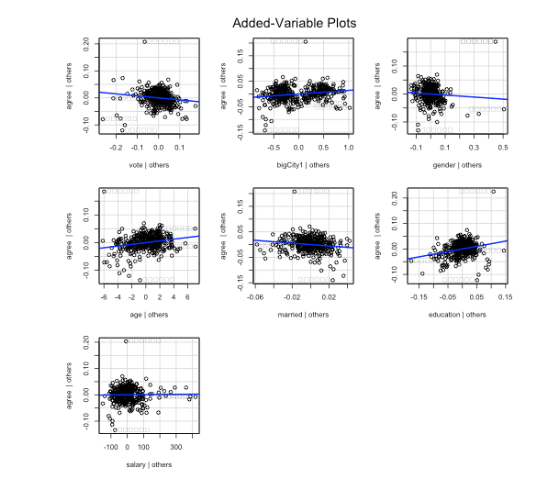


圖 6

在建立一階模型後為了發展預測力更好的模型，我們將透過加變數圖(added variable plot)來確認在去除其他變數的影響下，獨變項與依變項之見是否存在非線性的關係而需要加入高次項作修正，並將獨變項與依變項之間不具有線性關係的變項去除。

圖6為各比獨變項與依變項的加變數圖，其中能發現薪水中位數與同意率之間應該是沒有線性關係，故在之後的模型中都不納入薪水中位數的變項，而其他變項與依變項間應該是線性關係，故不考慮加入高次方項的獨變項修正模型。

在確定模型不需包含高次方項後，接下來將探討是否需要加入交互作用項，而為了選出包含交互作用項的最適模型，本研究使用最適子集演算法(Best subsets algorithms)來選出在不同參數個數下最佳的子模型，而在不同參數個數下的模型在各項指標上的作圖如圖7所示。

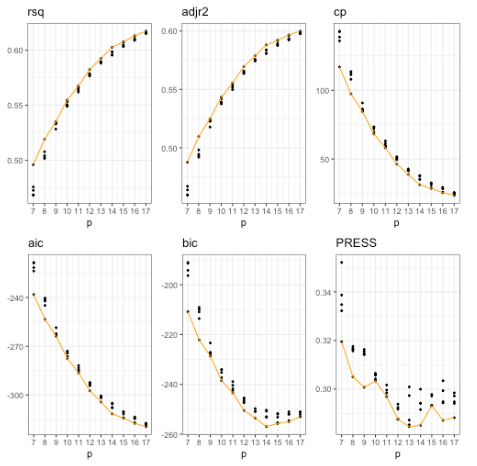


圖 7

上圖中包含的指標包括決定係數、修正決定係數、、赤池信息準則、貝氏信息準則與 PRESS。而在各項指標的表現上，皆有隨模型參數個數越多而表現越好的趨勢，其中表現最好的模型出現在參數個數15至17的區間。但是考量到包含過多的參數個數不利於我們解釋模型所代表的意義，故我們根據PRESS的指標在參數個數為9時相較於參數個數為10的時候低，故選擇參數個數為9的模型來發展包含交互作用項的模型。而該模型相較於原本的一階模型多包含投票率與六都、投票率與教育程度及六都與結婚率的交互作用，但是卻不包含性別比與薪水中位數。

在先前考慮高次項時，雖然我們透過加變數圖決定去除薪水中位數，但是卻不確定是否依然需要將性別比納入考量，而過去文獻中對於信別比是否影響同性戀態度也沒有明確結論，故我們將性別比再加入模型中發展新的候選模型。最後得到了以下三個候選模型:第一個是包含性別比的一階模型，第二個是包含交互作用項但不包含性別比的模型，第三個則是包含性別比與交互作用項的模型。

模型一：

模型二：

模型三：

由於我們在最後選擇模型三為最終模型，因此在接下來的模型分析中將呈現模型三的分析結果，並概述模型一與模型二的結果。

1. **模型共線性與離群值診斷**

在前一段的變數概況中，我們發現模型的各項變數之間存在高相關，故為確定模型是否存在共線性的問題，我們將透過模型變數的VIF指標進行進一步的分析，並由於模型加入了交互作用項，故在之後的分析中都先將變數進行標準化後分析，而各變數的VIF指標如下表3所示。

表 3

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Vote | BigCity | Age | Married | Edication | Gender | Vote  Education | Vote  BigCity | BigCityMarried |
| 1.47 | 1.25 | 1.96 | 2.25 | 2.23 | 2.59 | 1.25 | 1.54 | 1.31 |

透過表3我們能發現所有變數的VIF指標都約在1到2左右，因此不具有明顯的共線性，所以在之後的分析中並不需要修正共線性的問題。而模型一與模型二野的分析中也得到相同的結論。

在先前的分析中，我們發現了資料中可能存在離群值，並可能因此使模型不符合常態分配與變異數同質的假設。所以接下來我們將透過化student化去除後殘差(studentized deleted residuals)來看是否存在一變項的離群值，並透過槓桿量(Leverage) 來看在X變項上是否存在離群值。

表 4

|  |  |  |  |
| --- | --- | --- | --- |
| 資料 | 殘差 | 槓桿量 | Studentized 後除去殘差 |
| 連江縣莒光鄉 | -1.05 | 0.2 | -1.69 |
| 金門縣烏坵鄉 | -0.97 | 0.18 | -1.54 |
| 連江縣東引鄉 | 4.95 | 0.16 | 8.5 |
| 金門縣金城鎮 | -0.66 | 0.14 | -1.02 |
| … | … | … | … |
| 屏東縣春日鄉 | 0.81 | 0.13 | 1.25 |

上表4為在student化去除後殘差與槓桿量上表現較高的前幾筆資料，根據Beferroni校正後得到的臨界值為3.67，如果student化去除後殘差大於該值則判斷為離群值。而槓桿量的臨界值為0.5，如果大於該值則被判斷為離群值。從以上結果我們發現並沒有槓桿量被判斷為離群值，而有一筆資料的依變項被判斷為離群值，該筆資料為連江縣的東引鄉。而在模型一與模型二的分析中也得到相似的結論。

1. **資料點影響力檢查與修整**

由於資料中存在離群值，因此可能對模型估計產生問題，但是即使被判斷為離群值，該筆資料也不一定會對模型產生過大的影響力。故接下來將判斷離群值是否對模型有影響力。而在接下來的判斷中將從Cook距離(Cook’s distance)、DFFITS與DFBETAS這三項指標中判斷離群值的影響力。

上表5為在三項影響力指標中教大的前六筆數值，而其中標示為紅色的數值則是被判斷為影響力較大，可能嚴重影響模型的資料。從上面結果來看連江縣東引鄉除了是離群值也可能會嚴重影響模型，這結果和對模型一與模型二進行分析時相同。除此之外也有其他會對模型產生影響的資料，因此模型需要對根據離群值資料進行修整。

而本研究將透果迭代再加權最小平方法(IRLS)對資料的比重進行調整，以降低離群值的影響，其中使用的權重函數為Huber 權重函數。而對模型三進行反復迭代後到了第10次迭代後，模型的迴歸係數就不再變動。則以該迭代後得到的迴歸係數為模型三的新迴歸係數。而在模型一與模型二上也以相同方式進行迭代，並分別在地10次與第14次迭代後達到穩定。

表 5

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 地點 | DFFITS | Cook’s  Distance | DFBETA | | | | | | | | | |
| b1 | b2 | b3 | b4 | b5 | b6 | b7 | b8 | b9 |
| 連江縣東引鄉 | 3.681 | 1.257 | -0.002 | 0.005 | -0.062 | -0.027 | 0.116 | 0.184 | 0.007 | -0.002 | 0.004 |
| 金門縣金沙鎮 | -1.15 | 0.142 | 0.023 | 0.007 | -0.007 | -0.016 | -0.011 | 0.019 | 0.027 | -0.002 | 0.002 |
| 金門縣烈嶼鄉 | -0.905 | 0.089 | 0.018 | 0.006 | 0.001 | -0.019 | -0.008 | 0.016 | 0.018 | -0.015 | 0.003 |
| 連江縣莒光鄉 | -0.837 | 0.078 | 0.002 | -0.002 | 0.008 | 0.004 | -0.03 | -0.046 | -0.001 | 0.003 | 0.006 |
| 金門縣烏坵鄉 | -0.732 | 0.059 | 0.016 | 0.001 | 0 | -0.004 | -0.013 | -0.001 | 0.02 | -0.014 | -0.003 |
| 連江縣南竿鄉 | -0.664 | 0.049 | 0.005 | 0.001 | 0.006 | 0.008 | -0.027 | -0.036 | 0.001 | 0 | -0.001 |

表 6

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coef 1 | Coef 2 | … | Coef 9 | Coef 10 | Coef 11 |
| Vote | -0.1854 | -0.1948 | … | -0.2057 | -0.2058 | -.02058 |
| BigCity1 | 0.129 | 0.1237 | … | 0.1205 | 0.1205 | 0.1205 |
| Age | 0.1801 | 0.1882 | … | 0.1914 | 0.1914 | 0.1914 |
| Married | -0.1779 | -0.1652 | ... | -0.1547 | -0.1547 | -0.1547 |
| Education | 0.4564 | 0.4447 | ... | 0.4482 | 0.4483 | 0.4483 |
| Gender | -0.1997 | -0.2343 | ... | -0.2485 | -0.2485 | -0.2485 |
| Vote:Education | 0.1872 | 0.1729 | ... | 0.1541 | 0.1540 | 0.1540 |
| Vote:BigCity | -0.2150 | -0.2038 | ... | -0.1895 | -0.1895 | -0.1895 |
| BigCity:Married | -0.1109 | -0.1152 | ... | -0.1166 | -0.1166 | -0.1166 |

而對每一筆資料的權重如下圖8所示，從圖中能發現有一筆資料的權重被大幅降低，該筆資料為連江縣東引鄉，除此之外也有其他資料的權重經過調整。

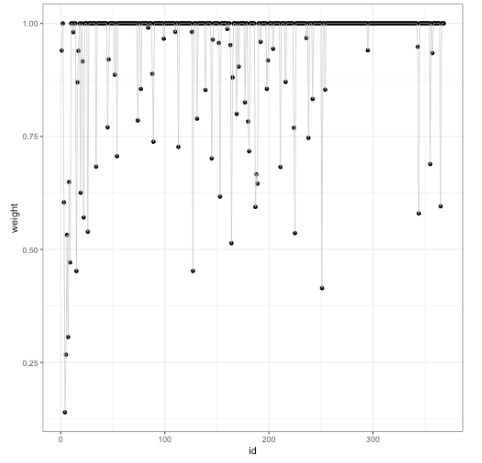


圖 8

而重新調整過的離群值在影響力指標上的表現就明顯下降，並不會嚴重影響模型。

1. **靴拔重抽法（Bootstrap）**

在對資料點進行權重後，本研究將確認迴歸係數是否達到顯著。由於本筆資料不符線性模型假設且經過權重，在無法使用統計檢定的情況下，本研究使用拔靴重抽法的方式，重複對殘差進行抽樣並建立新的資料，在用新的資料重新估計參數。透過重複上述步驟多次後得到參數的分配，並建立該分配下95%的信賴區間，如果區間不包含0則我們推迴歸係數應是顯著的。

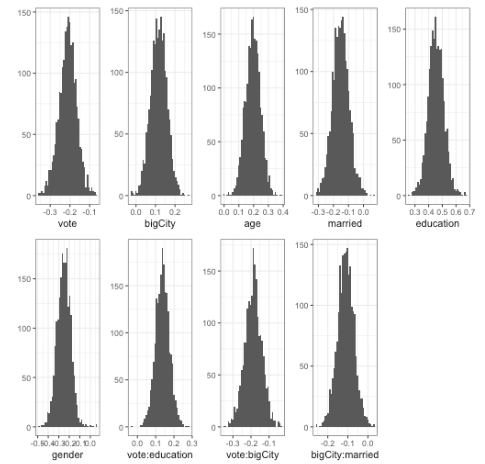


圖 9

圖9為各迴歸係數靴拔重抽法後的分布，透過上圖能發現大多數的分布都不包含零，而之後透過上圖建立各分配95%的信賴區間如下表7所示:

表 7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean | Sd | CI lower | CI upper |
| Vote | -0.211 | 0.045 | -0.289 | -0.108 |
| BigCity1 | 0.116 | 0.042 | 0.043 | 0.207 |
| Age | 0.195 | 0.051 | 0.089 | 0.286 |
| Married | -0.151 | 0.056 | -0.273 | -0.050 |
| Education | 0.455 | 0.057 | 0.328 | 0.552 |
| Gender | -0.250 | 0.066 | -0.372 | -0.118 |
| Vote:Education | 0.137 | 0.038 | 0.094 | 0.247 |
| Vote:BigCity | -0.187 | 0.047 | -0.282 | -0.096 |
| BigCity:Married | -0.110 | 0.041 | -0.205 | -0.042 |

則在表7能發現所有的參數的信賴區間皆不包含0，故我們能說所有的迴歸係數應都達到顯著。而在模型一中，透過靴拔法建立的信賴區間中，除了結婚率的區間包含0以外，其他區間皆不包含零。而模型二則是所有區間都不包含 0。

1. **最終模型**

透過上述靴拔法的結果我們得知模型一可能不夠穩定，因此先將模型一排除考量，而模型二與模型三的迴歸係數皆達顯著，但是就模型的解釋力上來看，模型三的表現較好，故我們選擇模型三為最終模型。

# 

# 肆、結果與討論

## ㄧ、結果與討論

本研究最終的迴歸模型為：  
其中，整體的。在主要效果的部分，對於各個預測變項與反應變項的關係為：當投票率越高的地區、結婚率越高或是性別比越高（男性越多），在此公投第 14 案的同意率上也會越低；相反的，若是位於六都（相較於非六都）地區、年齡中位數越高或是教育程度越高的話，則同意率也會較高。在此部分的研究結果其實可以呼應到先前的文獻發現，如男性、鄉村地區、結過婚者或是較少接觸高等教育者，較容易對同性戀議題抱持著負面的態度，而女性、都市地區、接受較高的教育程度反而比較有正向的對同性戀態度。另一方面，在文獻中尚未有共識的變項，如年齡和薪資收入，在本研究發現整體鄉鎮市區地區的年齡中位數越高的話，則同意率越高。而薪資收入並未納入最終的模型中，主要原因是該變項在統計上並沒有顯著的影響，可能說明的是薪資收入與同意率之間沒有顯著的關係。

在交互作用項的部分，可發現如果投票率越高的地區且教育程度越高，反而同意率越高，相比如果單單只有投票率越高的話，則是同意率越低；此外在六都（相較於非六都）投票率越高者，以及在六都（相較於非六都）結婚率越高者，這兩個交互作用項都造成同意率降低。而在過去文獻尚未有這方面的相關研究，特別是對於投票率如何對其他因子的影響是很少會有的。

另外，年齡的主要效果是個值得注意的議題。一般大眾認為，對於同性戀相關的知識在年紀較長者會不容易接受，因此部分的年輕人會有種想法會認為：「都是老年人（都會不同意且都是他們在投票）太多了，所以婚姻平權公投案都沒通過。」然而，在此研究中卻發現，年齡中位數越高的地區，表示者該區的人口較為老化的地區，反而同意率卻是較高的，與我們平常認知的產生衝突。在既往的文獻中，也顯示說不一定年齡較長者對同性戀態度一定都是較為負面。對此我們額外將年齡拆成三組，分為該地區年輕人（18-30 歲）人口比率、中年人（31-64 歲）人口比率以及老年人（65 歲以上）人口比率，分別對同意率做散布圖。從圖XX中可發現，年輕人與同意率呈負相關（r = -0.14\*），中年人則呈現正相關（r = 0.21\*），反而老年人是無相關的（r = -0.04）（參考圖10）。對於年齡與同性戀態度或許容易支持和反對的族群，不一定是我們一開始所想像的那樣，年輕人族群可能更抱有負面的態度。

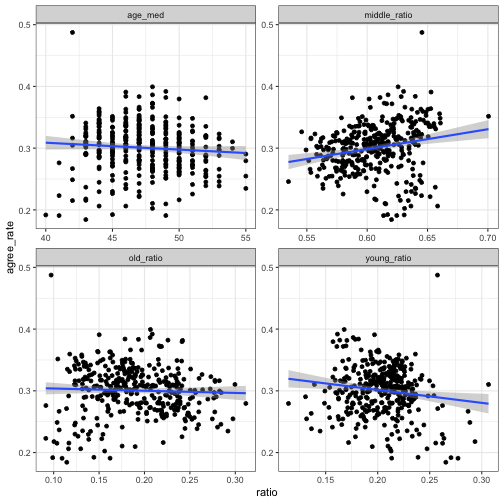


圖 10

然而，在本迴歸模型的建立並不是說明哪些變項一定會造成同意率高地或是同性戀態度的正面還是負面。「相關不代表因果」，這些變項特質雖然與公投同意率結果有關聯，但可能背後有其他因素才是真正的原因。而且同樣的投票結果不代表有相同的原因，每個人選擇支持或不支持公投內容，背後都有著不同的因素。從個人價值、成長經驗、重要他人影響、政黨傾向、到有沒有時間和能力去投票，都可能是影響投票行為的原因。因此想要了解本次第 14 案公投背後的動機與態度，必須進一步田野調查，無法單純從資料就能完全離解全貌。

## 二、研究限制

首先，本研究是針對公投第 14 案的同意率，投過迴歸模型的建立找出相關重要的影響變項。然而在本案中的議題是針對「以民法的方式保障同性別兩人的婚姻建立關係」，並非如過往研究探討的是「同性戀態度」以及相關的因子。即便這兩個議題是有高度的相關，但本質上還是兩個完全不同的議題，因此在結論的連結上必須謹慎小心，兩者是不能互相推論的。舉例來說，對於同性戀態度正向者不一定會在此公投投下同意票，他的考量可能會認為修該民法的成本過高，而對此次公投中的第 12 案，「以民法婚姻規定以外之其他形式來保障同性別兩人...」表示贊同。

第二，迴歸模型是以鄉鎮市區作為每筆資料的單位，來看預測變項與反應變項之間的關聯。雖然說我們一開始預設可以以一個地區來預設該區的投票者整體的傾象，但這真的具有代表性嗎？以及以鄉鎮市區所帶表的意義性又會是什麼？畢竟，我們最終獲得的是整體群族資料，並非是投票人確切的行為和態度。因此在此研究結果，或許可以說在某變數上，整體的選民是偏向於某個面向（同意或不同意），但絕對不能推論到個體實際的想法和投票行為。

第三，在資料來源上有諸多的限制與困難。由於在民主社會中是採以不記名投票，我們無法追朔每個投票人的投票行為的，也因此造就上述第二點我們採以地區整體的表現作為模型資料的基本單位。另一方面，在預測變項的選擇上很有可能忽略掉重要的選擇變項，如宗教信仰、家庭人數、政黨傾向...等，特別是宗教信仰上在先前文獻中特別提及它大大的影響對同性戀的態度。然而在台灣目前沒有對宗教信仰進行普查，因此在目前無法得知他們對於公投同意率的影響，實為可惜。

第四，在一開始迴歸模型的建立上，我們將每筆鄉鎮市區資料點採相同的權重來分析，直到做「迭代再加權最小平方」（IRLS），才透過迴歸模型與原始資料的殘差，給予資料點不同的權重而進行下一次迭代迴歸模型的建立。實際上每個鄉鎮市區的人口數皆是不同的，理應將人口數納入模型建立的考量之中，特別是人口多的地區應給予較大的重視，相反的人口少的地區可能權重就會較小。但這方面的權重大小該如何拿捏，以及該如何與 IRLS 中的權重相互協調，值得後續研究中加以討論。

最後，在本次資料中，可發現離島地區（特別是連江縣東引鄉）在同意率或其他變項上，就純數值來說會是離群值且對模型具有很大的影響力，但實際上它還是具有異議性的，可能相對於其他資料點來說較具有其他特別的性質（如：人口數特少或是有人特別對此議題進行宣導等），是值得後續研究中探討的對象。因此在此澄清用離群值形容僅就普遍的同意率值來說有較不一樣的數值而已，並非說明他就是跟其他資料點不一樣而應該被忽略。

## 三、未來改進方向

期望在後續的研究中，能納入潛在的重要變項（如宗教信仰、家庭人數...等）做探討，並驗證使否跟先前文獻的結果是否有出入。有可能在這些變項加入後，會使得本研究最終選出的迴歸模型有所改變。由於本研究是為了對嘗試使用多元迴歸線性模型來做模型的建立，以符合課程的要求，然而就此筆研究資料的特性，在公投同意率實際上的範圍會介於 0 到 1 以內，或許採取羅吉斯迴歸分析 （Logistic regression）是個不錯的選擇。

# 

# 參考文獻

王晴锋. (2013). “恐同症” 的根源——基于宗教, 现代性和文化的阐释. 吉首大学学报: 社会科学版, 34(1), 59-64.

Adamczyk, A., &amp; Liao, Y.-C. (2019). Examining Public Opinion About LGBTQ-Related Issues in the United States and Across Multiple Nations. *Annual Review of Sociology, 45*(1), null. doi:10.1146/annurev-soc-073018-022332

Becker, A. B., &amp; Scheufele, D. A. (2009). Moral Politicking: Public Attitudes toward Gay Marriage in an Election Context. *The International Journal of Press/Politics, 14(*2), 186–211. https://doi.org/10.1177/1940161208330905

Bowman, R.(1979). Public Attitudes Toward Homosexuality In New Zealand. *International Review of Modern Sociology, 9*, 229-238.

Gelbal, S. &amp; Duyan, V. (2006). Attitudes of University Students toward Lesbians and Gay Men in Turkey. *Sex Roles, 55*, 573-579.

Glenn, N. D., &amp; Weaver, C. N. (1979). Attitudes toward premarital, extramarital, and homosexual relations in the U.S. in the 1970s. *Journal of Sex Research, 15*(2), 108-118.

Herek, G. M. (1984). Beyond" homophobia": A social psychological perspective on attitudes toward lesbians and gay men. *Journal of homosexuality, 10*(1-2), 1-21.

Jäckle, S., & Wenzelburger, G. (2015). Religion, religiosity, and the attitudes toward homosexuality—A multilevel analysis of 79 countries. *Journal of homosexuality*, *62*(2), 207-241.

Seltzer, R. (1992). The social location of those holding antihomosexual attitudes. *Sex Roles, 26*, 391-398.

# West, K., & Cowell, N. M. (2015). Predictors of prejudice against lesbians and gay men in Jamaica. *The Journal of Sex Research, 52*(3), 296-305