

CH4 作業檢討

陳家威¹

OCTOBER 29, 2022

¹R10323045@ntu.edu.tw

題目 4-11

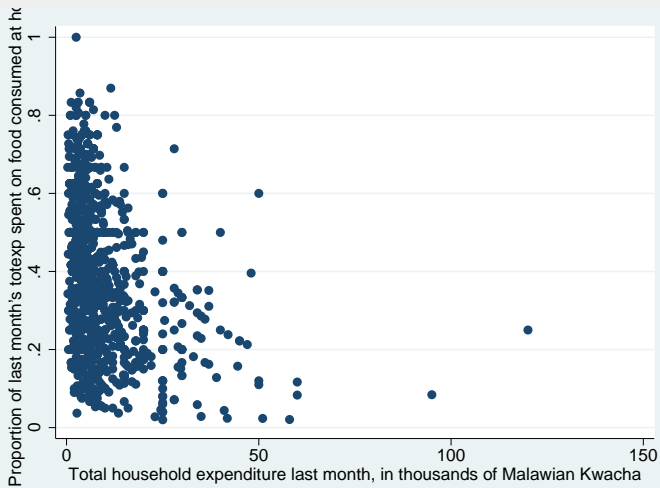
A 地理作業

自己査

尋找正確的模型有一些作法

1. 從既有文獻。適合短時間畢業。
2. 觀察散佈圖，看圖說故事。數據驅動 (Data driven) 的分析。適合初步了解數據。
3. 從一些結構性的關係 (structural relation) 推倒變數之間的關聯。較適合處理有嚴重內生性的變數關係。

scatter pfood totexp



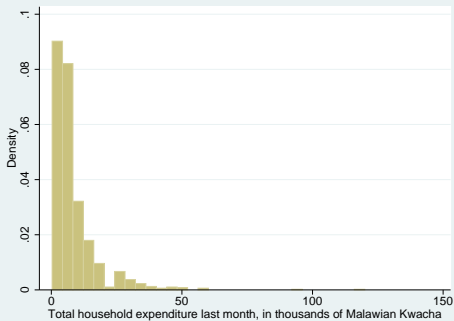
TOTEXP 變數好像不太「正常」

繪製長條圖

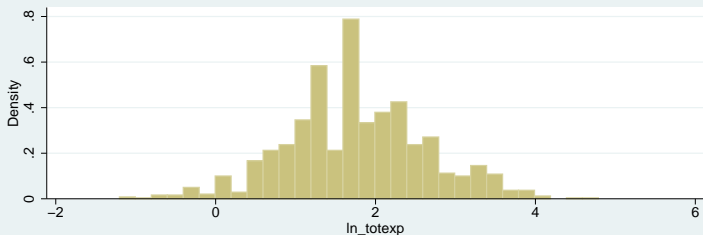
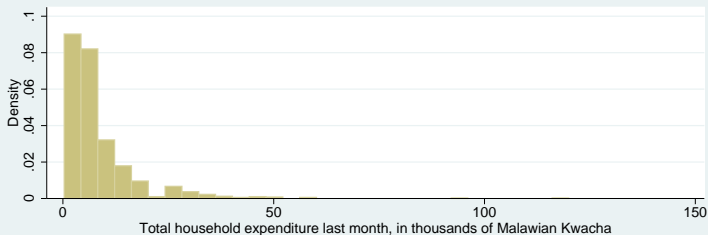
```
hist totex, name(hist_of_totexp)
```

匯出圖檔

```
graph export "Q4_11_hist_of_totexp.eps",  
name(hist_of_totexp)
```



如果取對數？



我們決定以 $TOTEXP$ 的對數來處理資料

b, c 小提上次助教課已做過，不再重複

```
1 eststo est_b: reg pfood ln_totexp
2 esttab, ci
3
```

```
1         capture program drop elas
2
3 // program 函數名稱
4 program elas
5     args total_ex est_name
6
7     qui estimates restore `est_name'
8     nlcom Elasticity: (_b[_cons] + _b[ln_totexp]*(log(`
9     total_ex') +1) ) ///
10         /(_b[_cons] + _b[ln_totexp]*log(`total_ex') )    ///
11         , noheader
12 end
```



```
13 quietly sum totexp, detail
14 scalar total_ex_p5 = r(p5)
15 scalar total_ex_p75 = r(p75)
16
17 elas total_ex_p5 est_b
18 elas total_ex_p75 est_b
19
```

從 (b) 小題中的模型計算殘差。建造這些殘差的直方圖，並針對 $\ln(TOTEXP)$ 繪製他們。有沒有明顯的模式？找到殘差的樣本偏態與峰態，以 1% 顯著水準進行 Jarque-Bera 檢定。

B, D, E, F, G, H 在做什麼？

- b, e, g 提出各種模型
- d, f, h 在進行殘差項診斷 (Residual Diagnosis)

線性迴歸模型要能正確地假設，包含

OLS 其中三項假設

1. 外生性假設： $E[e_i | X_i] = 0$
2. 獨立： $Cov(e_i, e_j | X_i) = 0$
3. (選擇性) 殘差分佈符合常態： $e_i | X_i \sim \mathcal{N}(0, \sigma^2)$

如果樣本數不夠大，要做統計檢定就一定需要透過殘差項符合常態分佈的假設，否則出來的不會是 t 檢定，此時判斷 p-value 也就毫無意義。

採用不同的模型

$$\text{b } PFOOD = \beta_1 + \beta_2 \ln(TOTEXP) + e$$

$$\text{e } \ln(FOOD) = \alpha_1 + \alpha_2 \ln(TOTEXP) + v$$

$$\text{g } FOOD = \gamma_1 + \gamma_2 \ln(TOTEXP) + u$$

同時估計彈性

$$\text{b } \varepsilon_b = \frac{\beta_1 + \beta_2 [\ln(TOTEXP) + 1]}{\beta_1 + \beta_2 \ln(TOTEXP)}$$

$$\text{e } \varepsilon_e = \frac{d \ln(FOOD)}{d \ln(TOTEXP)} = \alpha_2$$

$$\text{g } \varepsilon_g = \frac{\gamma_2}{\gamma_1 + \gamma_2 \ln(TOTEXP)}$$

- 計算殘差
- 繪製殘差直方圖
- 針對 $\ln(TOTEXP)$ 繪製
- 進行 Jarque-Bera 檢定

G,H 為例

從 $FOOD = PFOOD \times TOTEXP$ 得到 $FOOD$ 並估計

```
1 gen food = totexp*pfood
2 eststo est_g: reg food ln_totexp
```

建立計算彈性的函式（選擇性，也可以慢慢打）

```
1 capture program drop elas_g
2 program elas_g
3     args total_ex est_name
4     qui estimates restore `est_name'
5     nlcom Elasticity: _b[ln_totexp] ///
6         / (_b[_cons] + _b[ln_totexp]*log(`total_ex')) )///
7         , noheader
8 end
9
10 elas_g total_ex_p50 est_g
11 elas_g total_ex_p75 est_g
```

慢慢打版本

```
1
2 qui estimates restore est_g
3 nlcom Elasticity: _b[ln_totexp] ///
4     / (_b[_cons] + _b[ln_totexp]*log(total_ex_p50) )///
5     , noheader
6 nlcom Elasticity: _b[ln_totexp] ///
7     / (_b[_cons] + _b[ln_totexp]*log(total_ex_p75) )///
8     , noheader
```

上面的 total_ex_p50 是先透過 sum totexp, detail ，再從中取得 50, 70 百分位數 scalar total_ex_p75=r(p75)

G,H 為例（續）

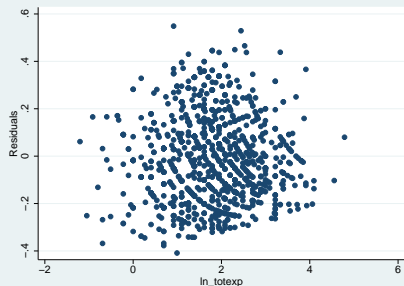
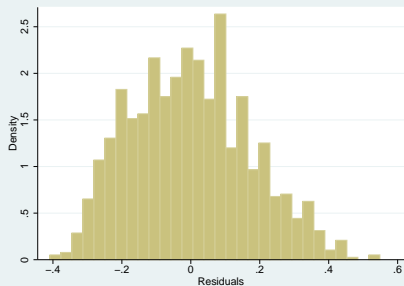
建立殘差

```
1 estimates restore est_g
2 predict res_h, residual
```

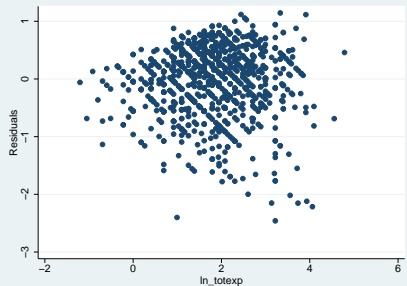
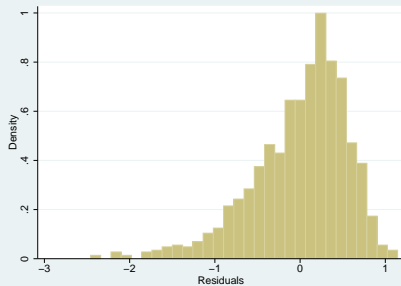
繪製殘差直方圖

```
1 histogram res_h, name("Q4_11_hist_3", replace)
2 graph export "Q4_11_hist_3.eps", name(Q4_11_hist_3) replace
3
4 twoway (scatter res_h ln_totexp ), name("Q4_11_scatter_3",
    replace)
5 graph export "Q4_11_scatter_3.eps", name("Q4_11_scatter_3")
    replace
```

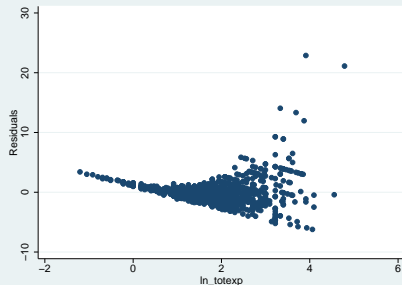
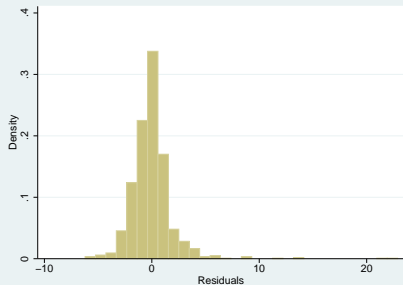

作圖結果-B,D



作圖結果-E,F



作圖結果—G,H



題目 4-13

資料檔 *collegetown* 包含 2009 至 2013 年在 Baton Rouge 售出的 500 棟房屋資料，

a $\ln(PRICE) = \beta_1 + \beta_2 SQFT + e$

b $\ln(PRICE) = \alpha_1 + \alpha_2 \ln(SQFT) + e$

c $PRICE = \delta_1 + \delta_2 SQFT + e$

分別計算

1. 模型參數
2. 樣本平均值的斜率
3. 樣本平均值的彈性

同時比較 R^2 值

對於上面 a,b,c 三種模型，

d 建構殘差直方圖，並進行 JB 檢定

e 針對 SQFT 畫殘差

f 預測 2700 平方英尺房屋的價格

g 預測 2700 平方英尺房屋的 95% 預測區間

被解釋變數為對數時的預測區間

如果回歸為

$$\ln(Y) = X\beta + e$$

因為進行 OLS 時的假設為 $e \sim \mathcal{N}(0, \sigma_e^2)$
所以會成常態分佈的，不是 Y ，而是 $\ln(Y)$

此時 Y 的對數是常態分佈，所以 Y 的分佈叫做「對數常態分佈」
對數常態分佈的期望值為

$$E[Y] = e^{(\mu + \frac{\sigma^2}{2})}$$