2014/3/20 الثلاثاء المحاضرة (23)

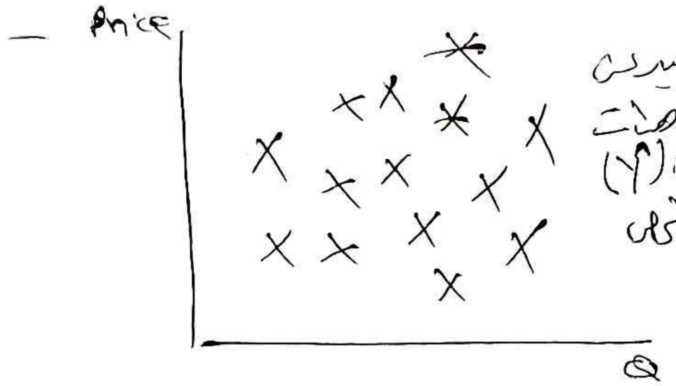
Gauss – Markov theory — نظرية كاوس – ماركوف

meaning

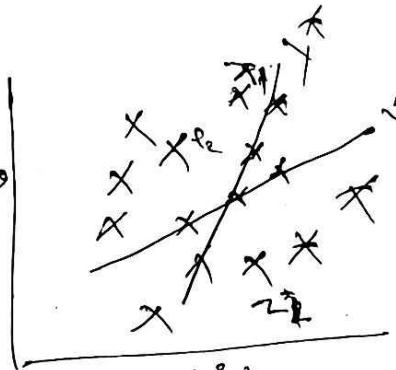
the ordinary Least square (OLS) Estimations — طريقة المربعات الصغرى الاعتيادية

ان هذه النظرية تقدم ضمانات لتطبيق (OLS) وتفضيلها عن غيرها من الطرق المستخدمة في التقدير لنموذج خطي ومعادلة منفردة.

ولتوضيح هذه الطريقة اي توضيح التقدير بواسطة المربعات الصغرى (OLS) نبدأ اولاً بمعرفة معنى المربعات الصغرى الاعتيادية ولتوضيح ذلك نفترض الشكل الاتي حيث توضح النقاط نقاط الانتشار للظاهرة المدروسة ومن الملاحظ من هذه النقاط انه يمكن تكوين العديد من الخطوط يمر عبر هذه النقاط



ولتقدير الخط الذي يمر عبر هذه المشاهدات رسمنا الخط $(\hat{Y})$ وربما يختار شكل آخر Y^*



(2)

سيكون ابتعاد نقطة هنا عن الخط هو عبارة عن (e_i) حيث ان (e_i) هو حد الخطأ العشوائي ومن متابعة الشكل اعلاه نجد ان (e_i) بالنسبة الى Y^* اصغر مما هو عليه بالنسبة $(\hat{Y})$ وذلك عند النقطة (Z) ولو نأخذ نقطة اخرى مثل (Z_1) سنجد ان ابتعاد هذه النقطة (حد الخطأ العشوائي) بالنسبة لـ $\hat{Y}$ اقل مما هو عليه بالنسبة لـ (Y^*) لذا فكل خط يعبر عن حقيقة معينة. السؤال الذي يطرح نفسه ← ما هو الخط الصحيح؟

للتوصل الى الخط الصحيح (خط الانحدار) سنعتمد على معيار المربعات الصغرى الاعتيادية والذي يقوم على تقدير خط الانحدار بحيث يقع في وسط نقاط الانتشار.

ان وقوع الخط في وسط نقاط الانتشار يعني سيكون هناك قيم متعددة لـ (e_i) (لابتعاد نقاط الانتشار عن الخط) ولكن تم التعبير عن كل هذه النقاط بمقدار مجموع (e_i) اي Σe_i وان مجموع انحرافات القيم عن وسطها الحسابي = 0 $(\Sigma e_i = 0)$ دائماً وهذه نتيجة

نهاية لانوصلنا الى شيء لذا تأخذ مجموع المربعات بدلاً من مجموع القيم وسنستفيد من هذه النقطة عند التوصل اليها بعد اي صيغة تقدير معلمات النموذج ٢٢

$$e_i = Y_i - \hat{Y}$$

$$\hat{Y} = \hat{a} + \hat{b} X_i$$

$$e_i = Y_i - \hat{a} - \hat{b} X_i$$

بتربيع الطرفين وادخال المجموع نحصل على

$$e_i^2 = (Y_i - \hat{a} - \hat{b} X_i)^2$$

$$\sum e_i^2 = \sum (Y_i - \hat{a} - \hat{b} X_i)$$

وحيث ان فكرة المربعات الصغرى هي ايجاد قيم المعلمات a و b بحيث يكون عندها مجموع مربعات الاخطاء العشوائية اقل ما يمكن

$$\sum e_i^2 \xrightarrow{\text{اقل ما يمكن}} \text{Minimization للتصغير}$$

لذا نجد المشتقة الجزئية من واحد المشتقة الجزئية الاولى لـ $\sum e_i^2$ ومساواتها بالصفر

اي بتطبيق الشرط الضروري للتدنية (التقليل)

Necessary condition for minimization

$$\frac{\partial \sum e_i^2}{\partial \hat{a}} = 2 \sum (Y - \hat{a} - \hat{b} X_i)(-1) = 0$$

بالقسمة على -2

$$0 = \sum (Y - \hat{a} - \hat{b} X_i)$$

وادخال المجموع

$$\sum Y - n\hat{a} - \hat{b} \sum X_i = 0$$

$$\boxed{\sum Y = n\hat{a} + b \sum X_i} \text{ ---- (1)}$$

$$\partial \sum e_i^2 = 2 \sum (Y - \hat{a} - \hat{b} X_i)(-X_i) = 0$$

$$= -2 \sum X_i (Y - \hat{a} - \hat{b} X_i)$$

بالقسمة

$$= \sum X_i Y - \hat{a} \sum X_i - \hat{b} \sum X_i^2 = 0$$

(25) $\sum XY = \hat{a}\sum X_i + \hat{b}\sum X_i^2$ ⟵ (2)

$$\sum Y = n\hat{a} + \hat{b}\sum X_i \quad \text{———} (1)$$

∴

$$\sum X_i Y = \hat{a}\sum X_i + \hat{b}\sum X_i^2 \quad \text{———} (2)$$

نحل المعادلتين آنيا بشكل م

مصفوفة المعاملات

$$\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} * \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum X_i Y \end{bmatrix}$$

$$\hat{a} = \frac{\begin{vmatrix} \sum Y & \sum X_i \\ \sum X_i Y & \sum X_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}} = \frac{\sum Y \sum X_i^2 - \sum X_i \sum X_i Y}{n\sum X_i^2 - (\sum X_i)^2}$$

$$\hat{b} = \frac{\begin{vmatrix} n & \sum Y \\ \sum X_i & \sum X_i Y \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}} = \frac{n\sum X_i Y - \sum Y_i \sum X_i}{n\sum X_i^2 - (\sum X_i)^2}$$

أو

$$\hat{b} = \frac{\sum X_i Y - \frac{(\sum X_i)(\sum Y_i)}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

بقسمة على n

هناك طريقة أبسط لتقدير $\hat{a}$ وهي بقسمة المعادلة رقم (1) على n

$$\frac{\sum Y}{n} = \frac{n\hat{a}}{n} + \frac{\hat{b}\sum X_i}{n}$$

$$\bar{Y} = \hat{a} + \hat{b}\bar{X}$$

∴ $\boxed{\hat{a} = \bar{Y} - \hat{b}\bar{X}}$

طريقة المصفوفات

تعتمد طريقة المصفوفات على إيجاد قيم المعلمات (a و b)
باستخدام المصفوفات وفقاً للصيغة الإحصائية

$$\beta = (X^T X)^{-1}(X^T Y)$$

β: تشير المعلمات (a و b)

X^T: مبدل المصفوفة X

Y: مصفوفة Y

X: مصفوفة X

حيث أن مصفوفة X

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ n & x_n \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & y_1 \\ 1 & y_2 \\ 1 & y_3 \\ \vdots & \vdots \\ n & y_n \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1\,1\,1 & \cdots & n \\ x_1\,x_2\,x_3 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ n & x_n \end{bmatrix}$$

$$(X^T X) = \begin{bmatrix} n & \Sigma X_i \\ \Sigma X_i & \Sigma X_i^2 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 1\,1\,1 & \cdots & n \\ x_1\,x_2\,x_3 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & y_1 \\ 1 & y_2 \\ 1 & y_3 \\ \vdots & \vdots \\ n & y_n \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \Sigma Y \\ \Sigma X_i Y \end{bmatrix}$$

$$\beta = (X'X)^{-1}(X'Y)$$

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} n & \Sigma X_i \\ \Sigma X_i & \Sigma X_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma Y \\ \Sigma Y X_i \end{bmatrix}$$

27

طريقة الانحرافات (طريقة التقدير حول نقطة المتوسط)

نظراً لكبر حجم الأرقام المستخدمة من التقدير $\hat{a}$ و $\hat{b}$ عند استخدام طريقة القيم الأصلية (التقدير حول نقطة الأصل)، يتم اللجوء الى تبسيط هذه الأرقام باستخدام الانحرافات عن الوسط الحسابي فإذا فرضنا:

أن: تمثل الانحرافات قيم المتغير عن الوسط الحسابي للمتغير Y_i

x_i: " " " " " " X_i

$\therefore \quad y_i = Y_i - \bar{Y} \;,\; Y_i = y_i + \bar{Y} \longrightarrow \hat{y} = \hat{Y}_i - \bar{Y} \rightarrow \hat{Y}_i = \hat{y}_i + \bar{Y}$

$\therefore \quad x_i = X_i - \bar{X} \longrightarrow X_i = x_i + \bar{X}$

$$e_i = Y_i - \hat{Y}_i = y_i + \cancel{\bar{Y}} - \hat{y}_i - \cancel{\bar{Y}}$$

$$e_i = y_i - \hat{y}_i$$

وحيث أن

$\hat{Y} = \hat{a} + \hat{b} X_i$ ——— (1)

$\bar{Y} = \hat{a} + \hat{b} \bar{X}$ ——— (2)

بالطرح

$\hat{Y}_i - \bar{Y} = \hat{b}(X_i - \bar{X})$

$\hat{y} = \hat{b} x_i$

بالتعويض عن $\hat{y}$ في e_i

$$e_i = y_i - \hat{b} x$$

ونربع الطرفين واخذ المجموع Σ

$$\Sigma e_i^2 = \Sigma (y_i - \hat{b} x_i)^2$$

ومن اجل جعل Σe_i^2 أقل ما يمكن نأخذ المشتقة الاولى لـ Σe_i^2 بالنسبة لـ $\hat{b}$ ونساويها بالصفر

$$\Sigma e_i^2 = 2\Sigma (y_i - \hat{b} x_i)(-x_i) = 0$$

$$-2\left(\Sigma y_i x_i - \hat{b} \Sigma x_i^2\right) = 0$$

بالقسمة على -2

$$\Sigma y x_i - \hat{b} \Sigma x_i^2 = 0$$

$$\Sigma y x_i = \hat{b} \Sigma x_i^2$$

$$\boxed{\hat{b} = \frac{\Sigma y x_i}{\Sigma x_i^2}}$$

(28) ELASTICITY elasticity — المرونة

تقدير المرونة من معلمة الانحدار

تعرف المرونة بأنها درجة استجابة المتغير المعتمد التابع للمتغير المستقل وتحسب وفقاً للصيغة التالية

$$E_p = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}}$$

$$\eta = \frac{\Delta Y}{Y} \times \frac{X}{\Delta X}$$

$$\eta = \frac{\Delta Y}{\Delta X} \times \frac{X}{Y}$$

حيث أن

$$\frac{\Delta Y}{\Delta X} = \hat{b}$$

$$\boxed{\eta = \hat{b} \times \frac{X}{Y}}$$

ويجب ان نلاحظ ان الوسط الحسابي للقيم الفعلية المعروفة لـ Y تساوي الوسط الحسابي للقيم المقدرة لـ $\hat{Y}$ حيث نلاحظ ان

$$\bar{Y} = \bar{\hat{Y}}$$

$$\hat{Y} = \hat{a} + \hat{b} X \qquad \hat{a} = \bar{Y} - \hat{b} \bar{X}$$

$$\bar{\hat{Y}} = \hat{a} + \hat{b} \bar{X}$$

$$\bar{\hat{Y}} = (\bar{Y} - \hat{b} \bar{X}) + \hat{b} \bar{X}$$

$$= \bar{Y} - \cancel{\hat{b} \bar{X}} + \cancel{\hat{b} \bar{X}}$$

$$\bar{\hat{Y}} = \bar{Y}$$

correlation coefficient معامل الارتباط

لقياس قوة العلاقة بين متغيرين او اكثر يستخدم معامل الارتباط الذي طوره كارل بيرسون (Carl Pearson) ويرمز له بالرمز r وتتراوح القيم بين -1 و +1 اي ان

$$-1 \le r \le 1$$

في حالة (r=-1) فالعلاقة بين المتغيرين X و Y علاقة عكسية تامة حيث جميع المشاهدات تقع على الخط المستقيم سالب الميل اي ان القيم في متغيري Y و X تكون عكس الاتجاه

وفي حالة (r=1) فالعلاقة بين المتغيرين X و Y علاقة موجبة تامة اي ان جميع المشاهدات تقع على الخط المستقيم موجب الميل اي ان القيم في متغيري Y و X تكون نفس الاتجاه

ويحسب معامل الارتباط وفق الصيغة التالية

$$r = \frac{\sum yx}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}}$$

معامل الارتباط يعطينا فقط نسبة القابلية عن x لا اكثر لتوضيح التغيرات في متغير Y . لذا مهم لقياس قوة العلاقة بين Y و X

لا توجد علاقة r=0 | قوي موجب r= | قوي سالب r=- | تام موجب r=1 | تام سالب r=-1

Q/ Estimate the correlation coefficient between ~~Price~~ income and food expenditure.

Exercises (30)

A random sample of ten families had the following income and food expenditure

Family	Family Y expenditure	Family income X	Y_iX_i	X_i^2	$y_i(Y_i-\bar{Y})$	$x_i=(X_i-\bar{X})$	yx	x_i^2
A	110	140	15400	19600	−55	−50	2750	2500
B	140	160	22400	25600	−25	−30	750	900
C	120	140	16800	19600	−45	−50	2250	2500
D	130	160	20800	25600	−35	−30	1050	900
E	170	200	34000	40000	5	10	50	100
F	230	240	55200	57600	65	50	3250	2500
G	150	150	22500	22500	−15	−40	600	1600
H	140	160	22400	25600	−25	−30	750	900
I	210	250	52500	62500	45	60	2700	3600
J	250	300	75000	90000	85	110	9350	12100
	$\Sigma Y = 1650$	$\Sigma X_i = 1900$	$\Sigma Y_iX_i = 337000$	$\Sigma X_i^2 = 388600$	$\Sigma y_i = 0$ صفر	$\Sigma x_i = 0$ صفر	$\Sigma yx = 23500$	$\Sigma x_i^2 = 27600$

$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{1900}{10} = 190$$

$$\bar{Y} = \frac{\Sigma Y}{n} = \frac{1650}{10} = 165$$

$$\hat{b} = \frac{n\Sigma YX_i - \Sigma Y \Sigma X_i}{n\Sigma X_i^2 - \Sigma X_i^2} = \frac{10(337000) - (1650)(1900)}{10(388600) - (1900)^2} = \boxed{0.85}$$

$$\hat{b} = \frac{235000}{276000} = \boxed{0.85}$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}$$

$$= 165 - (0.85)(190)$$

$$= 165 - 161.5 = 3.5$$

$$\boxed{\hat{Y}_i = 3.5 + 0.85X_i}$$

دالة الإنفاق الغذائي

① Estimate the regression line of food expenditure on income and interpret your results.

or as

② Interpret economically this estimated function

Given

③ Estimate the price elasticity of food expenditur

مميزات معادلة خط الانحدار:

ومن أهم السمات التي تتميز معادلة خط الانحدار البسيط هي ما يلي

$$\hat{Y} = \hat{a} + bX_i$$

① أن النقطة $\bar{X}$ و $\bar{Y}$ يجب أن تقع على خط الانحدار المقدر أو بمعنى آخر أن يمر خط الانحدار المقدر على هذه النقطة.

② مجموع الانحرافات عن خط الانحدار المقدر يجب أن يساوي الصفر

$$\sum e_i = 0$$

$$\sum (Y_i - \hat{Y}_i) = \sum e_i = 0$$

③ مجموع مربعات الانحرافات عن خط الانحدار أقل ما يمكن

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 \text{ أقل ما يمكن} = \text{Minmum}$$

④ يمكن استخدام معادلة خط الانحدار في التنبؤ عن قيمة $\hat{Y}_i$ لقيمة معينة من X_i غير موجودة في العينة.