المحور الثاني:

نموذج الانحدار الخطي البسيط

يعتبر نموذج الانحدار الخطي البسيط أبسط النماذج القياسية التي يعتمد عليها في دراسة العلاقات والتأثيرات بين المتغيرات، حيث نهتم بالتأثيرات فقط من خلال متغير تابع ومتغير مستقل (متغير مفسَّر ومتغير مفسِّر) في إطار علاقة رياضية خطية.

مثل علاقة الانفاق الاستهلاكي والدخل المتاح

1) صياغة وتقدير النموذج الخطي البسيط

أول من أطلق كلمة الانحدار هو الاقتصادي الانجليزي فرنسيس جالتون

⇐ قام بدراسة العلاقة بين طول الأبناء وطول آبائهم، وقام بصياغة النموذج التالي:

$$Chil_i = \alpha_0 + \alpha_1 fath_i$$

عدد أفراد قيد الدراسة

طول الأبناء متغير تابع — الحد الثابت — الميل الحدي للطول — طول الآباء متغير مستقل

⇓ يمكن صياغة العلاقة بمعادلة رياضية من الدرجة الأولى

$$Y = a + bx$$

مثال 2: ندرس على سبيل المثال العلاقة بين الدخل والاستهلاك لمجموعة من الأفراد ⇐

$$Y = \alpha_0 + \alpha_1 X$$

الاستهلاك، انفاق عائلي — الحد الثابت، أو الاستهلاك الحدي — الميل الحدي — الدخل

إذا كان $\alpha_1 = 0.8$ ⇐ يعني أن الأفراد يستهلكون 80% من دخلهم في حين يدخرون 20% المتبقية

+ هل في الحقيقة كامل الأفراد يستهلكون 80% من الدخل؟

+ هل الاستهلاك يتأثر بالدخل فقط؟ لا يوجد أذواق، الذوق، الأسعار، النوعية.

⇐ يستوجب إضافة عنصر اسمه الخطأ العشوائي أو الحد العشوائي e_t / ε_t إلى النموذج الذي يتعامل مع الاختلافات والمؤثرات التي لم تؤخذ بعين الاعتبار، ليصبح النموذج كالتالي

$$\boxed{Y_i = \alpha_0 + \alpha_1 X_i + \varepsilon_i}$$ النموذج القياسي

أو باستعمال عنصر الزمن

$$Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t$$

حيث:

Y_t: المتغير التابع أو المفسَّر

X_t: المتغير المستقل أو المفسِّر

α_0, α_1: معلمات النموذج

ε_t: الخطأ العشوائي، الحد العشوائي، الخطأ في تفسير Y_t ⇐

$$\varepsilon_t = Y_t - \alpha_0 - \alpha_1 X_t$$

1) أسباب ظهور أو وجود الخطأ العشوائي:

* إهمال بعض المتغيرات المستقلة التي يمكن أن تؤثر على المتغير التابع في النموذج (خطأ التحديد)
* الصياغة الرياضية غير السليمة للنموذج
* حدوث خطأ في كل من تجميع البيانات وقياس المتغيرات الإحصائية (خطأ القياس)

ويترتب عن الأسباب السابقة ما يلي:

* تحديد خاطئ للمتغيرات المستقلة: يتمثل ذلك في إغفال متغيرات مستقلة مهمة في نموذج الانحدار المراد تقديره أو احتواء هذا النموذج على متغيرات مستقلة غير هامة
* العلاقة الحقيقية بين المتغير التابع والمتغير المستقل قد تكون غير خطية

2) فرضيات نموذج الانحدار الخطي البسيط

1) **الفرضية الأولى:** الأمل أو التوقع الرياضي للأخطاء معدوم $E(\varepsilon_t) = 0$

تعني هذه النظرية أن الأخطاء العشوائية لا
تدخل في تفسير المتغير التابع Y، إذ أنها تعبر
عن حدود عشوائية تأخذ قيما سالبة، موجبة أو
معدومة لا يمكن قياسها أو تحديدها بدقة وتخضع
لقوانين الاحتمال (يعبر الخطأ العشوائي عن
ابتعاد القيم الحقيقية عن القيم النظرية للمتغير
التابع)

الفرضية 2: تجانس (ثبات) تباين الأخطاء ε

أي أن تباين الخطأ ثابت عبر الزمن ونكتب

$$Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma^2_{\varepsilon}$$

وهو يعني أن تشتت الأخطاء حول متوسطها
الحسابي ثابت عبر الزمن

الفرضية الثالثة: عدم وجود ارتباط ذاتي بين الأخطاء

أي أن الأخطاء العشوائية مستقلة عن بعضها البعض
وهذا يعني أن التباينات المشتركة للأخطاء
تكون معدومة ونكتب $Cov(\varepsilon_t, \varepsilon_{t'}) = 0$

الفرضية الرابعة: قيم الخطأ العشوائي ε_t مستقلة
عن قيم المتغير المستقل X_t، أي أن التباين
المشترك بينهما معدوم: $Cov(X_t, \varepsilon_t) = 0$

الفرضية الخامسة: الخطأ العشوائي ε_t يتبع
التوزيع الطبيعي ونكتب $\varepsilon_t \sim N(0, \sigma^2_{\varepsilon})$

له متوسط حسابي يساوي الصفر وانحراف معياري ثابت
لا يتغير مع الزمن) * هذه الفرضية تجمع الفرضيتين 1 و 2

$Y_t = \alpha_0 + \alpha_1 X_t$

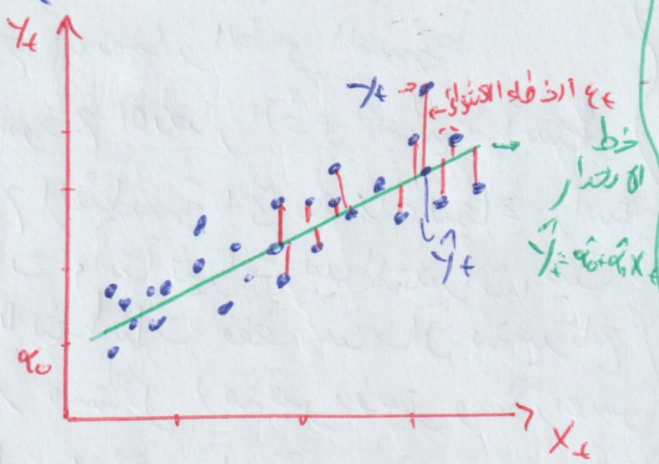
Y

تقدير معلمات النموذج الخطي البسيط باستعمال طريقة المربعات الصغرى العادية

نفرض النموذج القياسي التالي

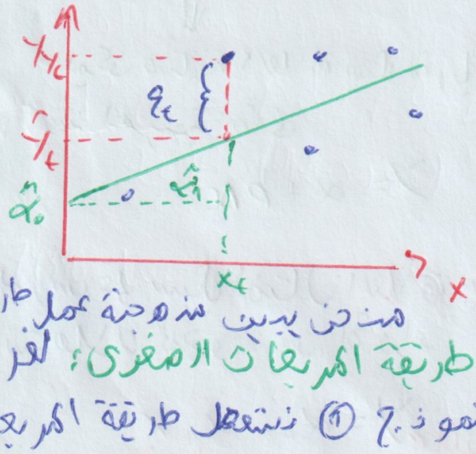
$$Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t \quad \ldots (1)$$

منحنى يبين سحابة النقاط للمتغيرين X و Y



=> نسعى إلى تقدير المعلمات α_0 و α_1 عن طريق
تحديد أمثل خط انحدار يمر عبر سحابة النقاط
وتقليص الأخطاء العشوائية إلى أقل حد ممكن

* الهدف هو إيجاد أمثل خط بحيث يتم تقليل الأخطاء
* الخطأ هو الفجوة بين النقطة (المشاهدة) وإسقاطها على خط الانحدار
* نسمي النقاط بـ Y: القيم الحقيقية ونسمي إسقاطاتها على خط الانحدار بـ $\hat{Y}$: القيم المقدرة، ومنه يكون الخطأ $\varepsilon_t = Y_t - \hat{Y}_t$



منحنى يبين منهجية عمل طريقة المربعات الصغرى

طريقة المربعات الصغرى: لغرض تقدير المعلمات
للنموذج (1) نستعمل طريقة المربعات الصغرى

moindres carrés ordinaires MCO => F

Ordinary least Squares OLS : En

وتهدف هذه الطريقة إلى تدنية وتقليص مربع
الأخطاء ε_t^2 إلى أقل حد ممكن، والهدف
من العمل على مربع الأخطاء هو كون الأخطاء من
خلال إشارتها السالبة والموجبة قد تنعدم كلها
مما يصعب من عملية التقدير لذا يتم اللجوء
للمربعات

* طريقة حساب المقدرات كالتالي:

$$\min \sum_{t=1}^{n} \varepsilon_t^2 = \min \sum_{t=1}^{n} (y_t - \hat{y}_t)^2$$

ولدينا القيم المقدرة $\hat{y}_t$: $\hat{y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 x_t$

ومنه

$$\min \sum_{t=1}^{n} \varepsilon_t^2 = \min \sum_{t=1}^{n} (y_t - \hat{\alpha}_0 - \hat{\alpha}_1 x_t)^2 = S$$

بهدف البحث عن تعظيم أو تدنية معادلة نقوم باشتقاقها ونساوي المشتقة إلى 0 ومنه

$$\begin{cases} \dfrac{\partial S}{\partial \hat{\alpha}_0} = -2 \sum_{t=1}^{n} (y_t - \hat{\alpha}_0 - \hat{\alpha}_1 x_t) = 0 \quad \text{...(1)} \\ \dfrac{\partial S}{\partial \hat{\alpha}_1} = -2 \sum_{t=1}^{n} x_t (y_t - \hat{\alpha}_0 - \hat{\alpha}_1 x_t) = 0 \quad \text{...(2)} \end{cases}$$

$$(f^n)' = n f' f^{n-1}$$

ومنه نتخلص من 2- و ننشر المجموع لنحصل على ما يلي

$$\begin{cases} \sum y_t - \sum \hat{\alpha}_0 - \sum \hat{\alpha}_1 x_t = 0 \quad \text{...(3)} \\ \sum y_t x_t - \sum \hat{\alpha}_0 x_t - \sum \hat{\alpha}_1 x_t^2 = 0 \quad \text{...(4)} \end{cases}$$

ولدينا $\sum \alpha = n\alpha$ (مجموع عدد هو نفس العدد مضروب في n مرة)

ومنه نحصل على ما يلي

$$\begin{cases} \sum_{t=1}^{n} y_t - n\hat{\alpha}_0 - \hat{\alpha}_1 \sum_{t=1}^{n} x_t = 0 \quad \text{...(5)} \\ \sum_{t=1}^{n} y_t x_t - \hat{\alpha}_0 \sum_{t=1}^{n} x_t - \hat{\alpha}_1 \sum x_t^2 = 0 \quad \text{...(6)} \end{cases}$$

ولدينا $\bar{x} = \dfrac{\sum x_t}{n} \Rightarrow \sum x_t = n\bar{x}$ ومنه نعوض في المعادلتين فنجد

$$\begin{cases} n\bar{y} - n\hat{\alpha}_0 - \hat{\alpha}_1 n\bar{x} = 0 \quad \text{...(7)} \\ \sum x_t y_t - \hat{\alpha}_0 n\bar{x} - \hat{\alpha}_1 \sum x_t^2 = 0 \quad \text{...(8)} \end{cases}$$

من المعادلة (7) نجد

$$\cancel{n}\bar{y} - \cancel{n}\hat{\alpha}_0 - \hat{\alpha}_1 \cancel{n}\bar{x} = 0$$

$$\bar{y} - \hat{\alpha}_0 - \hat{\alpha}_1 \bar{x} = 0$$

$$-\hat{\alpha}_0 = -\bar{y} + \hat{\alpha}_1 \bar{x}$$

المعامل الثابت

$$\boxed{\hat{\alpha}_0 = \bar{y} - \hat{\alpha}_1 \bar{x}} \quad \text{.....(9)}$$

نعوض (9) في (8) فنجد

$$\sum y_t x_t - \alpha_0 n\bar{x} - \hat{\alpha}_1 \sum x_t^2 = 0$$

$$\sum y_t x_t - (\bar{y} - \hat{\alpha}_1 \bar{x}) n\bar{x} - \hat{\alpha}_1 \sum x_t^2 = 0$$

$$\sum x_t y_t - n\bar{y}\bar{x} + \hat{\alpha}_1 n\bar{x}^2 - \hat{\alpha}_1 \sum x_t^2 = 0$$

$$\sum y_t x_t - n\bar{y}\bar{x} = \hat{\alpha}_1 \sum x_t^2 - \hat{\alpha}_1 n\bar{x}^2$$

$$\sum y_t x_t - n\bar{y}\bar{x} = \hat{\alpha}_1 \left(\sum x_t^2 - n\bar{x}^2\right)$$

الميل الحدي

$$\boxed{\hat{\alpha}_1 = \frac{\sum_{t=1}^{n} y_t x_t - n\bar{y}\bar{x}}{\sum x_t^2 - n\bar{x}^2} = \frac{n\sum x_t y_t - \sum x_t \sum y_t}{n\sum x_t^2 - (\sum x_t)^2} = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2} = \frac{Cov(x_t, y_t)}{V(x)}}$$

لا تحفظ

ملاحظات:

$y_t = \alpha_0 + \alpha_1 x_t$ .... النموذج الاقتصادي

$y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t$ النموذج القياسي

$\hat{y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 x_t$ النموذج المقدر

~~[illegible]~~

مثال: يبين الجدول التالي تطور كل من الاستهلاك والدخل الوطني في الجزائر خلال الفترة 2001-2007

السنة	الاستهلاك الإجمالي Y_t	الدخل المتاح X_t
2001	18.47	39.25
2002	19.89	41.84
2003	21.26	49.06
2004	23.71	57.31
2005	25.53	69.89
2006	26.95	78.64
2007	29.48	88.00

* أحسب المعالم $\hat{\alpha}$ و $\hat{\beta}$ حيث

$$\hat{Y} = \hat{\alpha} + \hat{\beta} X_t$$

معالم الانحدار

1) الحل: (أكمل الحل باستعمال باقي القوانين)

نستعمل القانون التالي:

$$\hat{\beta} = \hat{\alpha}_1 = \frac{\sum_{t=1}^{7}(X_t-\bar{X})(Y_t-\bar{Y})}{\sum_{t=1}^{7}(X_t-\bar{X})^2} = \frac{445,24}{2114,25} = 0,21$$

الميل الحدي

$$\hat{\alpha} = \hat{\alpha}_0 = \bar{Y} - \hat{\beta}\bar{X} = 23,61 - (0,21)(60,921) = 10,89$$

ومنه النموذج المقدر يكتب كما يلي:

$$\hat{Y}_t = 10,89 + 0,21 X_t$$

حيث: 10,89: الحد الثابت أو الاستهلاك الإجمالي الثابت

0,21: الميل الحدي للاستهلاك (استهلاك 0,21 من الدخل)

وكلما تغير الدخل بوحدة واحدة، تغير معه الاستهلاك الإجمالي بـ 0,21 وحدة

~~2 خصائص مقدرات المربعات الصغرى~~

ملاحظات:

1) النموذج $Y_t = \alpha_0 + \alpha_1 X_t + e_t$ ليس هو نفسه النموذج $X_t = \alpha_0' + \alpha_1' Y_t + e_t$، العلاقة التي تربط بينهما هي

$$\hat{\alpha}_1 \hat{\alpha}_1' = \rho^2 = r^2_{xy}$$

2) في حالة $\alpha_0 = 0$ نجد $\hat{\alpha}_1 = \frac{\sum x_t y_t}{\sum x_t^2}$

2 خصائص مقدرات المربعات الصغرى:

1) خاصية عدم التحيز:

التحيز هو ذلك الفرق بين مقدرة ما ووسط توزيعها فإذا كان هذا الفرق يختلف عن الصفر نقول عن ذلك المقدر بأنه غير متحيز، ومنه نكتب

$$E(\hat{\alpha}_0) = \alpha_0$$

نقول أن $\hat{\alpha}_0$ غير متحيزة لـ α_0

$$E(\hat{\alpha}_1) = \alpha_1$$

نقول أن $\hat{\alpha}_1$ غير متحيزة لـ α_1

2) خاصية التقارب:

ويقصد بها أن تباين المقدرات سوف يتجه إلى الصفر كلما ارتفع عدد المشاهدات، حيث أنه كلما أضفنا بيانات جديدة لعملية التقدير، فإن المجتمع المدروس يقترب من المجتمع الكلي، وهذا يعني أن قيمة المقدرة تصبح أكثر تمثيلا للواقع، ونكتب:

$$\lim_{n\to+\infty} V(\hat{\alpha}_1) = 0 \quad \text{و} \quad \lim_{n\to+\infty} V(\hat{\alpha}_0) = 0$$

3) خاصية أفضل مقدر خطي غير متحيز BLUE ومتسق

حيث توضح هذه ~~الطريقة~~ الخاصية بأن من بين المقدرات الخطية وغير متحيزة، تكون مقدرتا المربعات الصغرى العادية $\hat{\alpha}_0, \hat{\alpha}_1$ أفضل مقدرتين خطيتين وغير متحيزتين، حيث أن لهما أصغر تباين ممكن مقارنة مع بقية المقدرات الخطية وغير المتحيزة الأخرى، وذلك راجع لصغر تباينهما بدرجة كبيرة ونكتب

$$\lim_{n\to+\infty} E(\hat{\alpha}_1) = \lim_{n\to+\infty} \hat{\alpha}_1 = \alpha_1$$

$$\lim_{n\to+\infty} Var(\hat{\alpha}_1) = \lim_{n\to+\infty} Var(\hat{\alpha}_1) = 0$$