وزارة التعليم العالي والبحث العلمي

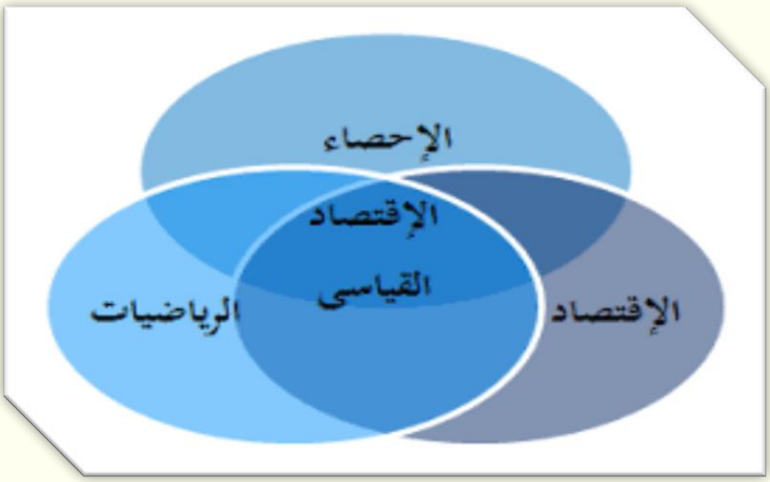
المدرسة العليا للتجارة

قسم التسيير

مطبوعة بيداغوجية في مقياس:

الاقتصاد القياسي

موجهة لطلبة السنة الأولى ماستر



إعداد: د. حمدوش عائشة

2024-2023

فهرس المحتويات

مقدمة: 5

الفصل الأول: مفاهيم أولية في الاقتصاد القياسي 6

1. تعريف الاقتصاد القياسي: 7

2. تعريف النموذج والمتغيرات: 8

3. مراحل بناء النموذج : 10

4. البيانات وطرق الحصول على البيانات 12

5. أنواع البيانات 15

الفصل الثاني: نموذج الانحدار الخطي البسيط 17

1. النموذج 17

2. فرضيات النموذج 19

3. تقدير معالم النموذج باستعمال طريقة المربعات الصغرى 20

4. توزيع المعاينة لمقدرات المربعات الصغرى 30

5.طريقة الترجيح الأعظم لتقدير $\boldsymbol{a}$ ، $\boldsymbol{b}$ ، $\boldsymbol{\sigma 2}$ 31

6. التقدير بمجال الثقة لمعالم النموذج الخطي البسيط 32

7. الاختبارات الإحصائية حول الدلالة الاحصائية (المعنوية) للنموذج المقدر 34

8. التنبؤ 37

9. تحليل للبواقي 38

10. تمارين تطبيقية 45

الفصل الثالث: نموذج الانحدار الخطي المتعدد 50

1. النموذج 50

2. فرضيات النموذج 51

3.تقدير معالم النموذج باستعمال طريقة المربعات الصغرى العادية 51

4. خواص المقدرات 52

5. توزيع المعاينة للمقدرات 55

6. الاختبارات الإحصائية للدلالة الإحصائية للنموذج المقدر 56
7.الاختبارات الخطية حول شعاع معالم النموذج: **b** 59
8. التنبؤ 60
9. استعمال المتغيرات المؤشرة في النمذجة 61
الفصل الرابع: أثار، اكتشاف وتصحيح خروقات فرضيات النموذج 71
1.عدم تجانس تباين الأخطاءHeteroscedasticity 72
7. 2.1. اختبارات اكتشاف عدم تباين الخطأ 74
8. 1.2.1. اختبار Goldfeld-Quandt.. 74
9. 2.2.1اختبار White 75
10. 3.2.1اختبار ثبات التباين الشرطي للأخطاء ARCH-LM.. 76
11. 3.1معالجة عدم ثبات تباين حد الخطأ 77
2.الارتباط الذاتي للأخطاء 79
12. 1,2. الارتباط الذاتي من الدرجة الأولى AR(1): 79
13. 2.2. خواص الارتباط الذاتي من الدرجة الأولى.. 79
14. 3.2. اختبارات الكشف عن الارتباط الذاتي 81
15. 2.3.2اختبار دربين واتسون Durbin-Watson test (1950 et 1951)........................ 81
16. 3.3.2 Breusch-Godfrey اختبار.. 83
17. 4.2طرق التقدير في حالة وجود ارتباط ذاتي بين الأخطاء 84
18. 1.4.2. تقدير ρ عن طريق إحصائية Durbin-Watson 88
19. 2.4.2تقدير ρ بطريقة Theil-Nagar 88
20. 3.4.2طريقة Cochrane-Orcutt.. 88
21. 4.4.2طريقة Hildreth-Lu 89
22. 5.2.اختبار h 92
3.مفهوم التعدد (الازدواج) الخطي Multicollinearity.. 94
23. 1.3أسباب التعدد الخطي وآثاره 94
24. 2.3اختبارات اكتشاف التعدد الخطي 95
25. 1.2.3طريقة التحليل الترافدي لـ Frisch 96
26. 2.2.3قياس التعدد الخطي أو شرط الأعداد Condition numbers.. 96
27. 3.2.3طريقة Farrar-Glauber.. 98

28. 3.3الحلول المقترحة للتعدد الخطي 99
29. تمارين تطبيقية... 102
المراجع.. 104

مقدمة:

تعتبر دراسة الاقتصاد القياسي محطة أساسية في فهم مفاهيم الحياة الاقتصادية، حيث يتجلى دورها الرئيسي في فحص وتحليل الظواهر الاقتصادية بطرق منهجية ومنطقية. إن الاقتصاد القياسي يمثل الإطار الذي يسهم في تفسير سلوك الأسواق واتخاذ القرارات الاقتصادية الاستراتيجية.

كما أنَ الاقتصاد القياسي يعتبر كأداة للتنبؤ بالتحولات الاقتصادية حيث أن فهم العلاقات بين المتغيرات الاقتصادية وتأثيرها على متغيرات أخرى يمثل خطوة أساسية نحو اتخاذ القرارات الاقتصادية الصائبة.

يعد فهم الاقتصاد القياسي أساسًا لفهم تفاعلات السوق وتحليل السياسات الاقتصادية، وهو مفتاح لاتخاذ القرارات الاستراتيجية في عالم الأعمال والتجارة. إن هذه المطبوعة تهدف إلى توجيه الطلبة نحو فهم المفاهيم والأسس الأساسية في هذا المجال.

هذه المطبوعة هي مجموعة محاضرات لمقياس الاقتصاد القياسي المقرر لطلبة السنة أولى ماستر بالمدرسة العليا للتجارة.

سنتعرض في هذه المطبوعة مواضيع مختلفة تتعلق بالاقتصاد القياسي، بدءًا من المفاهيم الأساسية مثل تعريف النموذج ومكوناته ثم النموذج الخطي البسيط والمتعدد والفرضيات اللازمة لتقديره وانتهاءً بمخالفة هذه الفرضيات وكيفية معالجتها. كل هذه الدروس مزودة بأمثلة وتمارين.

نأمل أن تكون هذه المطبوعة إضافة قيمة للتكوين الأكاديمي وتسهم في بناء أسس قوية لفهم التفاعلات الاقتصادية الحديثة.

الرجاء من مستعملي هذه المطبوعة افادتنا بأي ملاحظات تعديلية أو تحسينية على العنوان البريد الالكتروني التالي: a_hamadouche@esc-alger.dz .

الفصل الأول: مفاهيم أولية في الاقتصاد القياسي

في هذا الفصل، نهتم بوصف طبيعة دراسة الاقتصاد القياسي وفهم متطلباته وبعض مصطلحاته.

يبدأ كل شيء بنظرية من مجال دراسة الباحث - سواء كانت المحاسبة أو علم الاجتماع أو الاقتصاد – حيث يهتم حول مدى ارتباط المتغيرات المهمة ببعضها البعض. في الاقتصاد نعبر عن أفكارنا حول العلاقات بين المتغيرات الاقتصادية باستخدام المفهوم الرياضي للوظيفة. على سبيل المثال، للتعبير عن العلاقة بين الدخل (R)والاستهلاك (C)، يمكننا كتابة : $C = f(R)$ الذي يعني أن مستوى الاستهلاك هو دالة تابعة للدخل.

فمثلا يمكن التعبير عن الطلب على سلعة فردية ما (سعر نوع من السيارات مثلا) على النحو التالي:

$$Q^d = f(P, P^S, P^c, R)$$

والتي تعني أن الكمية المطلوبة، Q^d ، هي دالة لسعر السلعة P ، وسعر السلعة البديلة لها P^S وسعر السلعة المكملة P^c لها ومستوى الدخل R. يمكن كتابة المعروض من سلعة زراعية مثل اللحم البقري على الشكل:

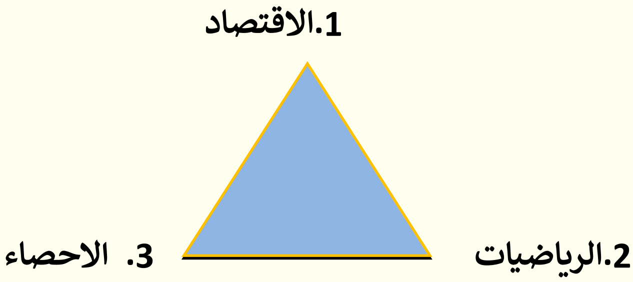
$$Q^s = f\left(P, P^c, P^f\right)$$

حيث Q^s هي الكمية المعروضة، P هو سعر اللحم البقري ، P^c هو سعر المنتجات التنافسية في الإنتاج (على سبيل المثال ، سعر الدجاج) ، و P^f هو سعر العوامل أو المدخلات (على سبيل المثال ، سعر الذرة المستخدمة في عملية الإنتاج. كل من المعادلات المذكورة أعلاه هي نموذج اقتصادي عام يصف كيف نتصور الطريقة التي تترابط بها المتغيرات الاقتصادية. النماذج الاقتصادية من هذا النوع توجه تحليلنا الاقتصادي. بالنسبة لمعظم مشاكل القرار أو الاختيار الاقتصادي، لا يكفي معرفة أن بعض المتغيرات الاقتصادية مترابطة، أو حتى اتجاه العلاقة. بل، يجب ،بالإضافة الى ذلك ،أن نفهم المقادير المعنية. أي أننا يجب أن نكون قادرين على تحديد مقدار تأثير التغيير في متغير على متغير آخر.

يدور الاقتصاد القياسي حول كيفية استخدام النظرية والبيانات من الاقتصاد والأعمال والعلوم الاجتماعية، جنبا إلى جنب مع الأدوات الإحصائية ، للإجابة على مختلف الأسئلة المطروحة.

1. تعريف الاقتصاد القياسي:

- الاقتصاد القياسي فرع من فروع الاقتصاد يعالج تقدير العلاقات الاقتصادية.
- الاقتصاد القياسي هو دراسة العلاقات والظواهر الاقتصادية بطريقة كمية و ذلك باستعمال التحليل الإحصائي و الصياغة الرياضية.
- الاقتصاد القياسي يعبر كميا عن الارتباطات الممكن أن تكون بين الظواهر الاقتصادية والتي النظرية تثبت وجودها.
- النظرية الاقتصادية تسمح بإعطاء الأفكار حول الطرق التي تحدد المتغيرات الاقتصادية و الاقتصاد القياسي يسمح بالتحقق التجريبي و يضع كميا الارتباطات التي تكون مقبولة.
- موضوع الاقتصاد القياسي هو مواجهة نموذج اقتصادي بمجموعة من المعطيات –معطيات زمنية أو مقطعية أو مختلطة- ثم التحقق من مصداقيته.
- الاقتصاد القياسي فرع من فروع الاقتصاد يعالج تقدير العلاقات الاقتصادية.
- الاقتصاد القياسي ملتقى ثلاث ميادين علمية:



تتم عمليتي النمذجة والتقدير تحت مراقبة الاقتصادي المتخصص في الميدان لإثبات قبول أو رفض النموذج، فمثلا: قد تضع النظرية الاقتصادية قيودا حول أحد معاملات النموذج المقترح كأن يكون موجب و أصغر من الواحد كما هو الحال بالنسبة للميل الحدي للاستهلاك في دالة الاستهلاك الكنيزية.

1.الاقتصادي	2.الرياضي	3.الإحصائي
متخصص في المجال، يعبر عن ظاهرة اقتصادية ما باستعمال النظرية الاقتصادية أين توضح طبيعة العلاقة والتأثير بين المتغيرات التي تشرح الظاهرة. مثال1: الاستهلاك مرتبط أو مشروح بواسطة بالدخل مثال2: الطلب على سلعة مرتبط بسعرها	يقترح علاقة جبرية للنظرية تربط بين المتغيرات تسمى النموذج، مثال: الاستهلاك=b *الدخل+ a $C = a + by$ حيث a وb تسمى معالم المجتمع أو معالم النموذج وهي مجهولة ونهدف لتقديرها	يتم تقدير معالم النموذج بإيجاد عبارات تسمح بتقدير قيم a و b ثم اعطائها قيم عددية من خلال البيانات. ثم تقييم النموذج احصائيا.

2. تعريف النموذج والمتغيرات:

1.2 تعريف النموذج

يعرف النموذج على أنه:

- تمثيل لظاهرة ما في شكل معادلة أو عدة معادلات رياضية خطية أو غير خطية وقابلة للتحويل إلى خطية أو غير خطية وغير قابلة للتحويل إلى خطية.
- و بما أن العلاقة الرياضية المقترحة ليست محددة تماما فانه يجب إدخال الخطأ في تحديد الصيغة الحقيقية للنموذج الاقتصادي.
- إن الخطأ يشمل:
  - كل المتغيرات الممكن أن تشرح الظاهرة المدروسة ولم تؤخذ بعين الاعتبار في النموذج
  - الخطأ في تحديد نوع العلاقة بين الظاهرة المدروسة و المتغيرات الشارحة لها.
  - أخطاء القياس على المتغيرات المدروسة أثناء جمع المعلومات.

النموذج يكون كما يلي :

$$\boldsymbol{V_D = f(V_{ID}) + Erreur}$$

$$\boldsymbol{Y = f(X) + \varepsilon}$$

- Y شعاع المتغير التابع ذو البعد $(n \times 1)$ حيث n حجم العينة
- X مصفوفة المتغيرات المستقلة ذات البعد $(n \times)$ حيث k يمثل عدد المتغيرات المستقلة المأخوذة لشرح Y (في حالة وجود متغير واحد $k = 1)$

- شكل الدالة $f(X)$ يتم تعيينه اما عن طريق النظرية (محدد مسبقا) أو من خلال التمثيل البياني للمعطيات الخاصة بالمتغيرات وقد يأخذ مثلا أحد الاشكال التالية:

أ- الشكل الخطي: $Y = a + bX + \varepsilon$

ب- الشكل التربيعي (غير خطي في المتغيرات): $Y = a + bX + cX^2 + \varepsilon$

ت- الشكل غير الخطي في المعاملات: $Y = a + \frac{1}{b}X + \varepsilon$

ث- الشكل الأسي (غير خطي) : $Y = e^{a+bX+\varepsilon}$

ج- الشكل نصف اللوغاريتمي : $lnY = a + bX + \varepsilon$

ح- الشكل اللوغاريتمي : $lnY = a + blnX + \varepsilon$

2.2 تعريف المتغيرات

- المتغيرات تمثل قيم لظواهر اقتصادية مشاهدة أو مقاسة مثل: الكميات المباعة من سلعة ما، سعر سلعة ما ،معدل الفائدة للاستثمار، رصيد الميزان التجاري، سعر الصرف، الناتج الداخلي الخام ،غياب الطلبة ، منطقة سكن الطلبة، تغطية الانترنيت لدى الطلبة ....
- ان تقييم النتائج يتعلق بنوعية المتغيرات.
- المشاكل الممكن مصادفتها حول المتغيرات:
  - مشكل عدم الملاءمة: مثلا دراسة مبيعات الخبز و استعمال معطيات خاصة بمبيعات البسكوت أو الفرينة
  - أخطاء القياس إما أثناء جمع المعطيات أو أثناء نقلها
  - أخطاء في تحديد وحدة القياس: مثلا كمية مبيعات الخبز مقاسة بعدد الوحدات المباعة أم برقم الأعمال
  - مشكل التمثيل: مثلا قياس فقط مبيعات المخابز دون مبيعات المساحات الكبرى-
- قد تكون المتغيرات: كمية، كيفية اسمية و كيفية ترتيبية.
- في الاقتصاد القياسي نميز اصطلاحا بين نوعين من المتغيرات:

- المتغيرات التي تمثل الظاهرة أو الظواهر التي نريد أن نشرحها، تسمى المتغيرات التابعة، المشروحة ، المفسَرَة Les Variables dépendantes(V à expliquer)
- المتغيرات التي تشرح الظاهرة التي نريد أن نشرحها، تسمى المتغيرات المستقلة، الشارحة، الخارجية، المفسرة Les Variables Indépendantes

3. مراحل بناء النموذج :

يمر بناء النموذج بأربع مراحل:

1.3 مرحلة توصيف وصياغة النموذج

تعني هذه المرحلة بتوصيف النموذج من خلال تحديد المتغيرات المستقلة والتابعة، وتوقع نظري مسبق بقيمة اشارات المعلمات وتعد هذه المرحلة من أهم مراحل بناء النموذج، لأنها تقوم على خبرة الاقتصادي وتمكنه من المعرفة الاقتصادية، فخلالها يستعين بالنظرية الاقتصادية لإيجاد العلاقات الدالية بين متغيرين أو أكثر تمهيدا لوضعها في النموذج، لذا يقوم بوضع عدد من الفرضيات حول تلك المتغيرات ويتوقف ذلك على مدى امكانية وضع العلاقة في صيغة رياضية باستخدام المعادلات والرموز.

تعد عملية تحديد المتغيرات وحصرها عدديا الخطوة الأساسية على طريق صياغة النموذج أن كل المعادلة في النموذج تفسر متغيرا واحدا بدلالة متغيرات أخرى وما يتصل بها من معاملات وثوابت.
لذا فإننا أمام نوعين من المتغيرات، الأولى وتسمى المتغيرات الخارجية التي تتحدد بقوى تقع خارج النموذج، أما النوع الثاني من المتغيرات فهي المتغيرات الداخلية التي تتحدد قيمها عن طريق النموذج، أي تقدير معادلات النموذج والتي يتم في المرحلة الثانية من مراحل بناء النموذج.

2.3 مرحلة تقدير النموذج المصاغ

في هذه المرحلة يستعان بالأدوات الرياضية لتحويل الدالة الى معادلة رياضية، تمهيدا لتقدير معلماتها بقيم عددية ومن أكثر صيغ التقدير شيوعا هي طريقة المربعات الصغرى العادية (OLS) والتي سنتعرف عليها لاحقا.

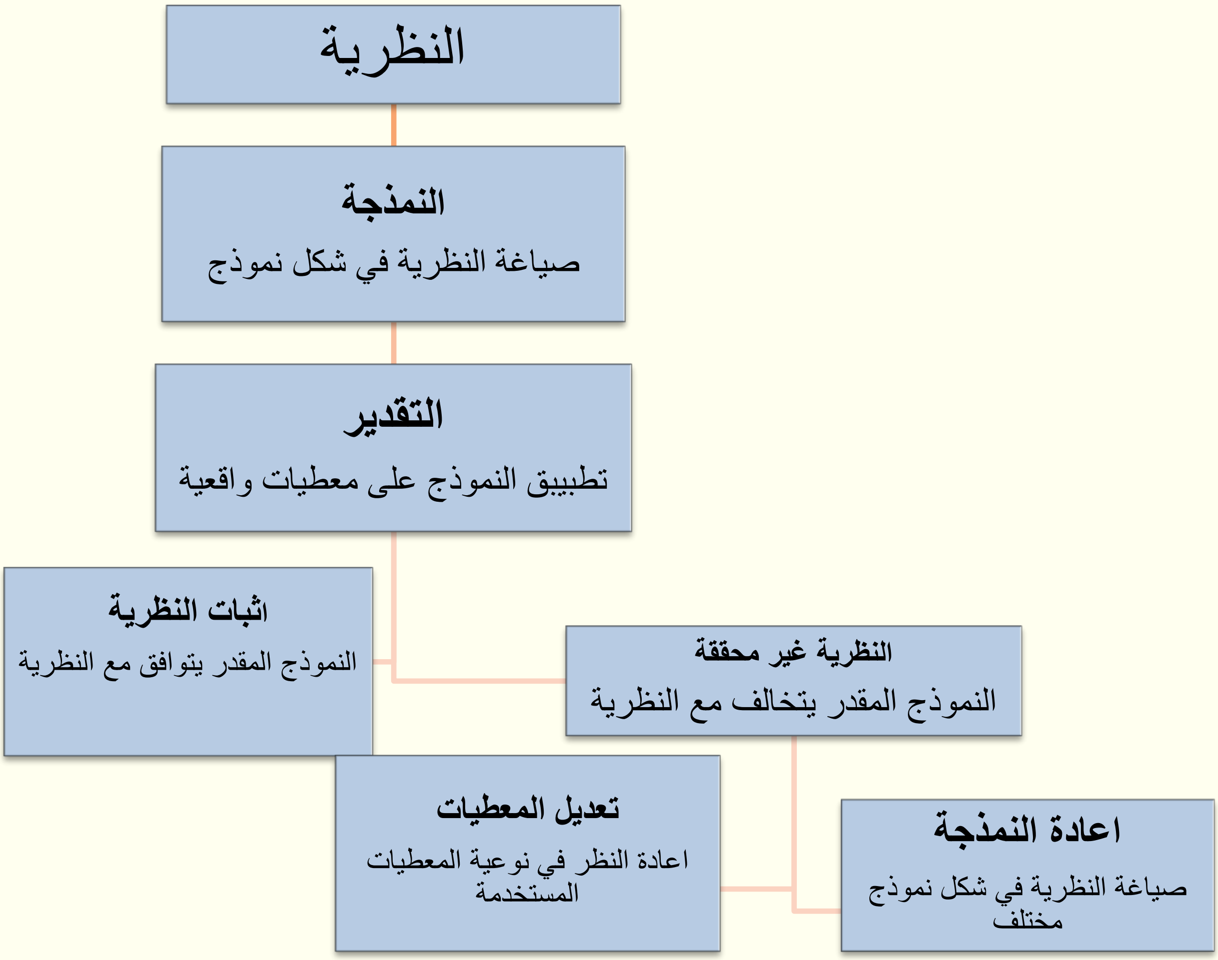
3.3 مرحلة اختبار النموذج المقدر:

بعد تقدير المعلمات لابد من تقييمها، واختبار دقة تقديرها باستخدام الأدوات والوسائل الاحصائية المناسبة.

ان طريقة تقدير النموذج تتوقف على الاختبارات الاقتصادية والاحصائية والقياسية مع توفر جميع الفرضيات الخاصة بنموذج الانحدار ومن ثم تحقيق الخصائص المرغوبة بالمقدرات، وعندما تكون المقدرات سليمة تكون الاختبارات الاقتصادية والاحصائية سليمة، لأن تلك الاختبارات تعتمد في تقييمها على قيم وتباين المقدرات السليمة، فيقبل النموذج باعتباره أفضل النماذج المختارة بعد الاختبار.

4.3 مرحلة التنبؤ

أحد الأهداف الرئيسية لتطبيق بحث القياس الاقتصادي هو استخدام النموذج المقدر للتنبؤ بقيمة المتغيرات التابعة استنادا الى قيم المتغيرات المستقلة من أجل التعرف على مسار الظاهرة المدروسة في المستقبل، حيث يعرف التنبؤ بأ نه تحليل بيانات الماضي وتطبيق نتائجها على المستقبل من خلال استخدام نموذج رياضي مناسب.



الشكل رقم (1): تمثيل بياني يبين مراحل بناء نموذج اقتصاد قياسي

4. البيانات وطرق الحصول على البيانات

يعمل الباحثون في علم الاقتصاد وغيرهم من الباحثون في علم الاجتماع في عالم معقد يتم فيه "ملاحظة" البيانات المتعلقة بالمتغيرات ونادرا ما يتم الحصول عليها من تجربة مضبوطة. هذا يجعل مهمة التعرف على المعايير الاقتصادية أكثر صعوبة.

يتم الحصول على البيانات التي تسمح لنا بتقدير المعلمات غير المعروفة للعلاقات الاقتصادية اما بواسطة الطرق التجريبية أو بواسطة الطرق غير التجريبية كما يمكن أن تكون بيانات تجميعية مثل البيانات الاقتصادية.

1.4 البيانات التجريبية

من بين الطرق للحصول على معلومات حول المعلمات غير المعروفة للعلاقات الاقتصادية ملاحظة نتائج تجربة ما. ففي ميدان العلوم الفيزيائية والزراعة، من السهل تخيل التجارب الخاضعة للرقابة. يحدد العلماء قيم متغيرات التحكم الرئيسية ثم يلاحظون النتيجة. قد تُزرع قطعا متشابهة من الأرض مع مجموعة معينة من القمح، ثم تُغير كميات الأسمدة والمبيدات المطبقة على كل قطعة أرض، مع ملاحظة في نهاية موسم النمو بوشل القمح المنتج في كل قطعة. يؤدي تكرار التجربة على N من قطع الأرض إلى إنشاء عينة من N ملاحظة. مثل هذه التجارب الخاضعة للرقابة والملاحظة نادرة في ميدان الأعمال والعلوم الاجتماعية. يتمثل أحد الجوانب الرئيسية للبيانات التجريبية في أنه يمكن تثبيت قيم المتغيرات التوضيحية عند قيم محددة في التجارب المتكررة للتجربة.

أحد الأمثلة التجارية يأتي من أبحاث التسويق. لنفترض أننا نهتم بالمبيعات الأسبوعية لسلعة معينة في سوبر ماركت. عند بيع سلعة ما ، يتم تمريرها فوق وحدة مسح ضوئي لتسجيل السعر والمبلغ الذي سيظهر في فاتورة البقالة الخاصة بالمشتري. ولكن في الوقت نفسه ، يتم إنشاء سجل بيانات ، وفي كل وقت يكون سعر السلعة وأسعار جميع السلع المنافسة لها معروفا ، بالإضافة إلى عروض المتجر الحالية واستخدام القسيمة. يتم التحكم في الأسعار وبيئة التسوق من قبل إدارة المتجر، لذلك يمكن تكرار هذه "التجربة" عدة أيام أو أسابيع باستخدام نفس قيم متغيرات "التحكم".

هناك بعض الأمثلة على التجارب المخطط لها في العلوم الاجتماعية، لكنها نادرة بسبب الصعوبات في تنظيمها وتمويلها. مثال بارز على تجربة مخططة هو مشروع تينيسي ستار (Tennessee’s Project Star) اتبعت هذه التجربة مجموعة واحدة من أطفال المدارس الابتدائية من رياض الأطفال حتى الصف الثالث، بدءا من عام 1985 وتنتهي في عام 1989. في التجربة، تم تعيين الأطفال بشكل عشوائي داخل المدارس إلى ثلاثة أنواع من الفصول: فصول صغيرة تضم 13-17 طالبا، وفصول عادية الحجم تضم 22-25 طالبا، وفصول عادية الحجم مع مساعد مدرس بدوام كامل لمساعدة المعلم. كان الهدف هو تحديد تأثير الفصول الصغيرة على تعلم الطلاب، كما تم قياسه من خلال درجات الطلاب في اختبارات التحصيل. ستؤثر نتيجة هذه التجربة على السياسة العامة تجاه التعليم لسنوات قادمة.

2.4 البيانات غير التجريبية

من البيانات غير التجريبية نجد بيانات المسح. فمثلا يجري قسم التطوير والبحث في شركة اوريدو استطلاعات عبر الهاتف للعملاء. في استطلاع هاتفي، يتم اختيار الأرقام بشكل عشوائي والاتصال بها. يتم تسجيل الردود على الأسئلة وتحليلها. في مثل هذه البيئة، يتم جمع البيانات عن جميع المتغيرات في وقت واحد، والقيم ليست ثابتة ولا قابلة للتكرار.

يتم إجراء مثل هذه الدراسات الاستقصائية على نطاق واسع من قبل الحكومات الوطنية. على سبيل المثال، المسح السكاني العام الحالي (RGPH 2022)هو المسح السادس يشمل جميع الأسر الجزائرية يجريه الديوان الوطني للإحصائيات. يتم استخدام بيانات RGPH من قبل صانعي السياسات الحكومية والمشرعين كمؤشرات مهمة للوضع الاقتصادي للدولة ولتخطيط وتقييم العديد من البرامج الحكومية. كما أنها تستخدم من قبل الصحافة والطلاب والأكاديميين وعامة الناس.

3.4 البيانات التجميعية

يمكن جمع البيانات على مستويات مختلفة من التجميع:

- البيانات الجزئية (Micro-Data) التي يتم جمعها على وحدات صنع القرار الاقتصادي الفردية مثل الأفراد والأسر والشركات.
- البيانات الكلية (Macro-Data) البيانات الناتجة عن التجميع أو التجميع على الأفراد أو الأسر أو الشركات على المستوى المحلي أو مستوى الولاية أو المستوى الوطني.

وقد تمثل البيانات أيضا تدفقا أو مخزونا:

- تدفق لقياس ظاهرة ما على مدى فترة من الزمن، مثل استهلاك البنزين خلال الربع الأخير من عام 2020.
- مخزون في شكل قياس ظاهرة ما في نقطة زمنية معينة، مثل كمية النفط الخام التي تحتفظ بها سوناطراك في صهاريج التخزين في 1 نوفمبر 2008، أو قيمة أصول بنك الجزائر في 1 يوليو 2009.

قد تكون البيانات كمية أو نوعية:

- كمية - نتائج مثل الأسعار أو الدخل التي يمكن التعبير عنها كأرقام أو بعض التحولات فيها، مثل الأسعار الحقيقية أو دخل الفرد.

- نوعية - النتائج التي هي من حالة "إما أو". على سبيل المثال، قام المستهلك بشراء أو لم يشتري سلعة معينة، أو كان الشخص متزوجا أو غير متزوج.

5. أنواع البيانات

1.5 البيانات المقطعية

تتغير قيم المتغير التابع و المتغيرات المستقلة بتغير الافراد التي قد تكون أشخاص او عوائل أو بلدان أو مؤسسات كما في الشكل التالي :

الفرد i	المتغير التابع Y_i	المتغير المستقل x_i
1	y_1	x_1
2	y_2	x_2
3	y_3	x_3
.	.	.
i	Y_i	x_i
.	.	.
n	y_n	x_n

2.5 بيانات زمنية

تتغير قيم المتغير التابع والمتغيرات المستقلة بتغير الزمن الذي قد يكون يوم أو شهر أوثلاثي أو سداسي أو سنة كما هو مبين في الشكل التالي:

الزمن t	المتغير التابع Yi	المتغير المستقل xi
1	y_1	x_1
2	y_2	x_2
3	y_3	x_3
.	.	.
t	Y_t	X_t
.	.	.
T	Y_T	X_T

3.5 بينات مقطعية- زمنية(بيانات بانل):

تتغير قيم المتغير التابع و المتغيرات المستقلة بتغير الزمن و الأفراد وذلك من أجل نفس العينة من الافراد في فترة زمنية معينة كما في الشكل التالي :

x_i	المتغير التابع Y_{ti}	الفرد i	الزمن t
$x_{t1,1}$	$y_{t1,1}$	1	
.	.	.	
$x_{t1,i}$	$Y_{t1,i}$	i	$t1$
.	.	.	
$x_{t1,n}$	$y_{t1,n}$	N	
$x_{t2,1}$	$y_{t2,1}$	1	
.	.	.	$t2$
$x_{t2,i}$	$Y_{t2,i}$	i	.
.	.	.	.
$x_{t2,n}$	$y_{t2,n}$	N	.
.	.	1	.
.	.	.	.
$X_{tj,i}$	$Y_{tj,i}$	i	.
.	.	.	T_j
.	.	n	. .
$x_{tT,1}$	$y_{tT,1}$	1	.
.	.	.	.
$x_{tT,i}$	$Y_{tT,i}$	i	.
.	.	.	t_T
$x_{tT,n}$	$y_{tT,n}$	n	

الفصل الثاني: نموذج الانحدار الخطي البسيط

يعد الانحدار من أهم أدوات الاقتصاد القياسي. ان نماذج الانحدار تحاول ايجاد متوسط العلاقة بين متغير أول الذي يتمثل في المتغير التابع، ومتغير آخر (أو متغيرات) تمثل المتغير (المتغيرات) المستقل.

للانحدار أهداف تتمثل فيما يلي:

- الحصول على تقديرات معاملات المتغيرات المستقلة؛
- فحص تأثير المتغيرات المستقلة على المتغير التابع، (على سبيل المثال ، تأثير الدخل على الاستهلاك في نظرية الاستهلاك الكنزي)؛
- التحقق من النظرية الاقتصادية (على سبيل المثال، نظرية الاستهلاك الكنزي)؛
- إجراء تحليل للسياسات (على سبيل المثال، كيفية تأثير التغييرات في سياسة الدخل على أنماط الاستهلاك)؛
- للتنبؤ أو التنبؤ بالمتغير التابع (على سبيل المثال، ما هو مستوى الاستهلاك في المستقبل

1. النموذج

نريد أن نضع علاقة إرتباط بين متغيرين X و Y ، حيث :

- ✓ Y يمثل المتغير الذي نريد شرحه و يسمى أيضا بالمتغير التابع :

Variable to be explained (Dependent variable)

- ✓ X هو المتغير الشارح و يسمى بالمتغير المستقل

Explanatory variable (Independent variable)

نموذج الانحدار الخطي البسيط يكتب على النحو التالي:

$$\mathrm{Y_i} = \mathrm{a} + \mathrm{b}\, X_i + \varepsilon_i \;\dots\dots (1)$$

- ✓ a و b هي معالم Parameters (coefficients) النموذج يراد تقديرها. (b: الميل Slope و a : الثابت Constatnt)
- ✓ ε تمثل العبارة العشوائية و التي تسمى أيضا خطأ النموذج له دور مهم في علاقة الانحدار ، حيث أنه يضم كل المعلومات التي لم تؤخذ بعين الاعتبار في العلاقة الخطية التي نريد

وضعها بين X و Y أي كل ما يتعلق بالمتغيرات الممكن أن تشرح Y و لم نأخذها بعين الاعتبار، خطأ القياس ، وأيضا خطأ التقريب بالعلاقة الخطية.

نستعمل عينة من n ملاحظة مستقلة ولها نفس التوزيع لتقدير هذه المعالم. يتم عرض تلك الملاحظات في جدول احصائي يبين الأفراد في عمود وقيم المتغير التابع في عمود ثاني وقيم المتغير المستقل في عمود أخر كما هو مبين بالجدول أدناه.

Y_i	X_i	i
Y_1	X_1	1
Y_2	X_2	2
....	....	....
Y_i	X_i	i
....	....	....
Y_n	X_n	n

ان خصائص مقدرات معالم النموذج(1)، a و b ،تعتمد بقوة على الفرضيات التي نضعها حول ε ،

مثال : نهتم بتفسير النفقات الشهرية بواسطة الدخل للعوائل ، ومن أجل ذلك نأخذ عينة من 10 عائلات ونلاحظ عليها المداخيل و النفقات الشهرية كما هو موضح في الجدول رقم () .

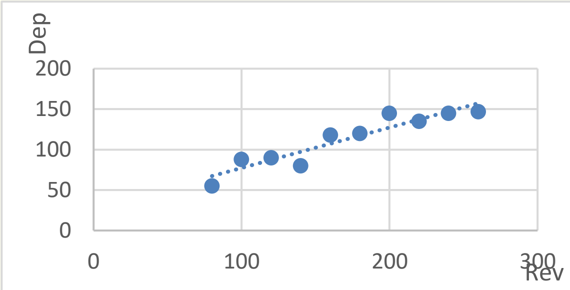
الجدول رقم(1) : مداخيل والنفقات الشهرية لـ10 عائلات

العائلة i	1	2	3	4	5	6	7	8	9	10
المداخيل: Rev	260	240	220	200	180	160	140	120	100	80
النفقات الشهرية:Dep	147	145	135	145	120	118	80	90	88	55

أرسم التمثيل البياني لهذه البيانات $(Rev_i\,, Dep_i)$ ، ثم اقترح النموذج المناسب لتمثيل العلاقة التي تشرح النفقات الشهرية بواسطة المداخيل .

الحل:

الشكل رقم(2): التمثيل البياني للثنائيات (Rev_i, Dep_i)



التمثيل البياني لمعطيات العشر عائلات يوحي بإمكانية تقريب العلاقة بين المتغيرين بالعلاقة الخطية التالية:

$$Dep_i = a + b\ Rev_i + \ \varepsilon_i \ \ \ldots..(2)$$

حيث ε تمثل كل المتغيرات التي تجعل العلاقة بين النفقات الشهرية و المداخيل تنحرف عن الخط المستقيم أي كل المتغيرات التي تؤثر في النفقات الشهرية من غير المداخيل.

ان النموذج (2) مبني ليس فقط على أساس التمثيل البياني لمعطيات العشر عائلات وانما هناك نظرية اقتصادية تدعم كتابته وهي نظرية كينز لتفسير الاستهلاك.

2. فرضيات النموذج

يتم بناء نموذج الاقتصاد القياسي تحت الفرضيات التالية :

الفرضية الأولى : يفترض أن تكون معاملات الانحدار (معلمات النموذج) ثابتة أو غير عشوائية بمعنى أن قيمها ثابتة عند أخذ عدة عينات ، كما يفترض أن نموذج الانحدار خطي في المعلمات، وقد يكون خطيا أو غير خطيا في المتغيرات Y و X.

حيث: - x_i تمثل المتغيرات المستقلة في النموذج و يفترض أنها غير عشوائية .

- y_i تمثل المتغيرات التابعة وهي متغيرات عشوائية عن طريق الخطأ ε، وذلك كون $Y = f(\varepsilon)$ أي الخطأ الوحيد على Y هو ناتج عن عدم كفاية X لشرح Y .

الفرضية الثانية: الأخطاء العشوائية لها توقع رياضي معدوم وتباين ثابت $\varepsilon_i \backsim (0, \sigma_\varepsilon^2)$

- تباين الخطأ ثابت Homoscédasticité: $\forall i\ ;\ V(\varepsilon_i) = \sigma_\varepsilon^2$
- التوقع الرياضي للأخطاء معدوم: $E(\varepsilon_i) = 0$.

الفرضية الثالثة: المتغيرات المستقلة x_i مستقلة عن المتغيرات العشوائية ε_i ،

أي أنّ $cov(x_i\ ; \varepsilon_i) = 0$

الفرضية الرابعة: استقلال الأخطاء $cov(\varepsilon_i, \varepsilon_j) = 0$

أي عدم ارتباط الأخطاء فيما بينها non autocorrélation des erreurs

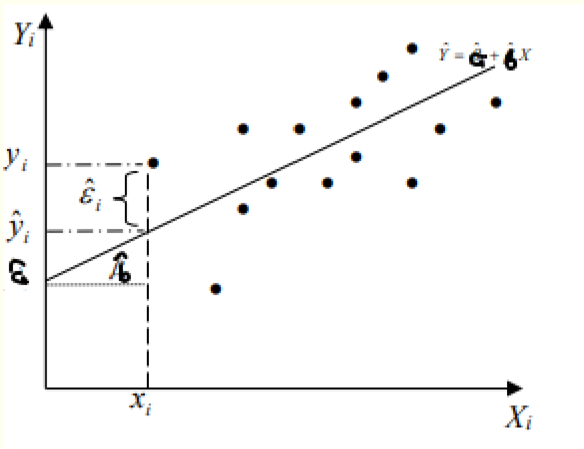
الفرضية الخامسة: الأخطاء العشوائية تتبع التوزيع الطبيعي $\varepsilon_i \backsim N(0, \sigma_\varepsilon^2)$. يمكن الاستغناء عن هذه الفرضية في حالة العينات الكبيرة واستعمال خاصية التقارب نحو التوزيع الطبيعي

3. تقدير معالم النموذج باستعمال طريقة المربعات الصغرى

نهتم بإيجاد مقدرات لمعالم النموذج $y_i = a + b\,x_i + \varepsilon_i$ (النموذج الذي نريد تقديره)، أي نهتم بإيجاد عبارتي $\hat{a}$ و $\hat{b}$ التي تعطينا أحسن خط مستقيم يمثل العلاقة بين x و y.

إن أي خط مرسوم ما بين هذه النقاط المنتشرة يمكن أن يمثل تقديرا للعلاقات المفروضة بالمعادلة، و يكون هذا الخط ممثلا بالعلاقة المقدرة التالية : $\hat{y}_i = \hat{a} + \hat{b}x_i \quad , i = \overline{1, n}$

<u>الشكل رقم(3):تمثيل الخط المستقيم المقدر للعلاقة الحقيقية بين y و x</u>



1.3 إيجاد مقدرات للمعالم a و b

إن مبدأ المربعات الصغرى لإيجاد أفضل خط مستقيم هو إيجاد قيم a و b التي تجعل مجموع مربعات البواقي(la Somme des Carrés Résiduelles SCR) أصغر ما يمكن أي :

$$SCR = \sum(y_i - \hat{y})^2 = \sum_{i=1}^{n} e_i^2 = \sum(y_i - a + bx_i)^2 = \min \dots (3)$$

حيث e_i تسمى البواقي وهي تمثل الجزء من المتغير التابع y الغير مشروح بواسطة x

إن حل المعادلة(3) هو قيم a وb التي تعدم المشتق الأول لعبارة مجموع مربعات البواقي بشرط أن يكون المشتق الثاني موجب أي:

$$\begin{cases} \frac{\partial SCR}{\partial a} = 0 \\ \frac{\partial SCR}{\partial b} = 0 \\ \frac{\partial^2 SCR}{\partial a^2} > 0 , \frac{\partial^2 SCR}{\partial b^2} > 0 \end{cases} \Rightarrow \begin{cases} -2 \sum(y_i - a - bx_i) = 0 \Rightarrow \sum e_i = 0 \\ -2\sum x_i(y_i - a + bx_i) = 0 \Rightarrow \sum x_i e_i = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \sum y_i = n\,a + b \sum x_i & \dots\dots\dots (1) \\ \sum x_i y_i = \hat{a} \sum x_i + \hat{b} \sum x_i^2 & \dots\dots\dots (2) \end{cases}$$

إن المعادلتين 1 و 2 تسميان بالمعادلتين الطبيعيتين لطريقة المربعات الصغرى (les equations normales de **MCO**)

$$(1) \Rightarrow \hat{a} = \frac{\sum y_i}{n} - \hat{b}\frac{\sum x_i}{n} = \bar{y} - \hat{b}\,\bar{x} \quad \dots\dots\dots (3)$$

بتعويض (3) في (2)نحصل على :

$$\sum x_i y_i = (\bar{y} - \hat{b}\bar{x}) \sum x_i + \hat{b} \sum x_i^2$$

$$\sum x_i y_i = \left(\sum x_i^2 - n\bar{x}^2\right) + n\bar{y}\bar{x}$$

$$\hat{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \dots\dots\dots (4) \Rightarrow$$

وبما أن المشتق الثاني لمجموع مربعات البواقي بالنسبة للمعالم أي:

$$\frac{\partial^2 SCR}{\partial a^2} = 2 > 0 \ ; \ \frac{\partial^2 SCR}{\partial a^2} = 2\sum_{i=1}^{n} x_i^2 > 0$$

فإن المعادلتين 3 و 4 تمثل عبارتي مقدرات المربعات الصغرى للنموذج : $y_i = a + b\, x_i + \varepsilon_i$

$$\begin{cases} \hat{a} = \bar{y} - \hat{b}\,\bar{x} \\ \hat{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \end{cases} :$$

ملاحظة : يمكن كتابة عبارة مقدر b كما يلي :

1. $\hat{b} = \frac{cov\,(x,Y)}{V(x)} = r_{x,Y}\frac{\sigma_Y}{\sigma_X}$ حيث $r_{X,Y}$: معامل الارتباط بين المتغيرين Y و X

2. $\hat{b} = \frac{\sum x_i y_i}{\sum x_i^2}$ بحيث : $x_i = X_i - \bar{X}$ ، $y_i = Y_i - \bar{Y}$ ، x_iوy_i تسمى الانحرافات

مثال تطبيقي: باعتبار معطيات المثال السابق، أحسب مقدرات النموذج المقترح لتمثيل العلاقة بين النفقات الشهرية والمداخيل باستعمال طريقة المربعات الصغرى.

الجدول رقم(4):الحسابات اللازمة لتطبيق طريقة المربعات الصغرى

$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$(y_i - \bar{y})$	المداخيل Rev(x)	النفقات الشهرية Dep(y)	االعائلة
8100	3123	34,7	260	147	1
4900	2289	32,7	240	145	2
2500	1135	22,7	220	135	3
900	981	32,7	200	145	4
100	77	7,7	180	120	5
100	-57	5,7	160	118	6
900	969	-32,3	140	80	7
2500	1115	-22,3	120	90	8
4900	1701	-24,3	100	88	9
8100	5157	-57,3	80	55	10
33000	16490	0	170	112,3	**المجموع**

باستعمال عبارتي مقدرات المربعات الصغرى لمعالم النموذج الخطي البسيط المبينة في المعادلتين (3) و(4) أعلاه, وباستعمال نتائج الحسابات اللازمة والمبينة في الجدول رقم() نجد ما يلي:

$$\begin{cases} \hat{a} = \bar{y} - \hat{b}\,\bar{x} \quad = 11{,}23 - 0{,}499 * 17 = 2{,}747 \\ \hat{b} = \dfrac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \dfrac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \dfrac{16490}{33000} = 0{,}499 \end{cases}$$

ومنه يمكن كتابة النموذج المقدر كما يلي :

$$\hat{y}_i = \hat{a} + \hat{b}x_i = 2{,}747 + 0{,}499 * x_i$$

2.3 نتائج المربعات الصغرى :

عند تطبيق المربعات الصغرى ، يمكننا استخلاص النتائج الهامة التالية:

1. إن خط الانحدار يمر على النقطة $(\bar{x}, \bar{y})$:

$$\sum e_i = 0 \Rightarrow \sum\left(y_i - \hat{a} - \hat{b}x_i\right) = 0 \Rightarrow \bar{y} = \hat{a} + \hat{b}\bar{x}$$

2. البواقي غير مرتبطة بالمتغير المستقل: $Cov(x_i, e_i) = 0$

$$Cov(x_i, e_i) = \frac{1}{n} \sum(x_i - \bar{x})(e_i - \bar{e}) = \frac{1}{n}\sum(x_i - \bar{x})\, e_i$$

$$= \frac{1}{n}\sum x_i\, e_i - \frac{1}{n}\bar{x}\sum e_i = \frac{1}{n}\sum x_i e_i = 0$$

$$\Longrightarrow cov(x_i, e_i) = 0$$

3. البواقي مستقلة عن مقدر المتغير التابع $Cov(\hat{y}_i, e) = 0$

$$cov(\hat{y}_I, e_i) = \frac{1}{n}(\hat{y}_i - \bar{y})(e_i - \bar{e})\frac{1}{n}\sum \hat{y}_i\, e_i$$

$$= \frac{1}{n}\sum(\hat{a} + \hat{b}x_i)e_i = 0$$

$$= \frac{1}{n}\left(\hat{a}\sum e_i + \hat{b}\sum x_i\, e_i\right) = 0$$

4. $SCT = SCE + SCR$ و $SCE = \hat{b}\sum(x_i - \bar{x})(y_i - \bar{y}) = \hat{b}^2\sum(x_i - \bar{x})^2$

حيث :

SCT مجموع المربعات الكلية أو مجموع الانحرافات الكلية	مجموع المربعات الإجمالي للتغيرات التي تحدث في المتغير التابعY وتسمى أيضا بالتغيرات الكلية: $$SCT = \sum_{i=1}^{n}(y_i - \bar{Y})^2$$
SCEمجموع مربعات الانحدار أو مجموع الانحرافات المفسرة	يسمى بمجموع مربعات الانحدار يعني جزء من تباين Y الذي تم تفسيره بواسطة الانحدار. إي الجزء من المتغيرات التي تحدث في المتغير التابع والذي تم تفسيرها بواسطة النموذج المقدر: $SCE = \sum_{i=1}^{n}(\hat{y}_i - \bar{Y})^2$
SCR مجموع مربعات البواقي أو مجموع الانحرافات غير المفسرة	مجموع مربعات البواقي، , وهذا مؤشر للجزء الذي لم يفسر بواسطة نموذج الانحدار: $SCR = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

البرهان:

$$SCE = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n}\left(\hat{a} + \hat{b}x_i - \hat{a} - \hat{b}\bar{x}\right)^2 = \sum_{i=1}^{n}\left(\hat{b}x_i - \hat{b}\bar{x}\right)^2$$

$$= \sum_{i=1}^{n}\hat{b}^2(x_i - \bar{x})^2 = \sum_{i=1}^{n}\hat{b}\,\frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}(x_i - \bar{x})^2$$

لدينا : $SCR = \sum_{i=1}^{n}\left(y_i - \hat{a} - \hat{b}x_i\right)^2$

بتعوض $\hat{a}$ بـ $\bar{y}-\hat{b}\bar{x}$ ، نحصل على :

$$SCR = \sum_{i=1}^{n}\left(y_i-\bar{y}+\hat{b}\bar{x}-\hat{b}x_i\right)^2$$

$$SCR = \sum\left(y_i-\bar{y}-\hat{b}(x_i-\bar{x})\right)^2$$

$$= \sum(y_i-\bar{y})^2+\hat{b}^2\sum(x_i-\bar{x})^2-2\hat{b}\sum(y_i-\bar{y})\,(x_i-\bar{x})$$

$$= \sum(y_i-\bar{y})^2+\hat{b}\frac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}\sum(x_i-\bar{x})^2-2\hat{b}\sum(y_i-\bar{y})\,(x_i-\bar{x})$$

$$= \sum(y_i-\bar{y})^2-\hat{b}\sum(y_i-\bar{y})\,(x_i-\bar{x}) = SCT-SCE$$

3.3 خصائص مقدرات المربعات الصغرى

أ. خاصية عدم التحيز:

النظرية 1 : ليكن $\hat{a}$ و $\hat{b}$ مقدري المربعات الصغرى لمعالم النموذج الخطي البسيط: a و b ، أنَ

$\hat{a}$ و $\hat{b}$ مقدرات غير متحيزان لـ a و b

البرهان : نقول أن المقدر $\hat{b}$ مقدر غير متحيز لـ b اذا تحقق ما يلي:

$$\boldsymbol{E(\hat{b}) = b}$$

نفس الشيء بالنسبة للمقدر $\hat{a}$ ، نقول أنه مقدر غير متحيز للمعلمة a اذا تحقق مايلي:

$$\boldsymbol{E(\hat{a}) = a}$$

لدينا

$$\begin{cases} y_i = a+bx_i+\varepsilon_i \\ \bar{y} = a+b\bar{x}+\bar{\varepsilon} \end{cases} \Longrightarrow y_i-\bar{y} = b(x_i-\bar{x})+(\varepsilon_i-\bar{\varepsilon})$$

نعوض $y_i-\bar{y}$ في عبارة $\hat{b}$:

$$\hat{b} = \frac{\sum(y_i-\bar{y})\,(x_i-\bar{x})}{\sum(x_i-\bar{x})^2} = \frac{\sum(x_i-\bar{x})[b(x_i-\bar{x})+(\varepsilon_i-\bar{\varepsilon})]}{\sum(x_i-\bar{x})^2}$$

$$= b + \frac{\sum(x_i-\bar{x})(\varepsilon_i-\bar{\varepsilon})}{\sum(x_i-\bar{x})^2} = b + \frac{\sum(x_i-\bar{x})\varepsilon_i}{\sum(x_i-\bar{x})^2}$$

$$\Rightarrow E(\hat{b}) = b + \frac{\sum(x_i-\bar{x})E(\mathcal{E}_i)}{\sum(x_i-\bar{x})^2} = b$$

$$\Rightarrow E(\hat{b}) = b$$

لدينا:

$$\hat{a} = \hat{y} - \hat{b}\bar{x} = a + b\bar{x} + \bar{\mathcal{E}} - \hat{b}\bar{x} = a + (b-\hat{b})\bar{x} + \bar{\mathcal{E}}$$

$$= a + (b-\hat{b})\bar{x} + \bar{\mathcal{E}} = a - (\hat{b}-b)\bar{x} + \bar{\mathcal{E}} \Rightarrow \hat{a} - a = (\hat{b}-b)\bar{x} + \bar{\mathcal{E}}$$

ومنه:

$$\boxed{E(\hat{a}) = a}$$

ب. خاصية أصغر تباين التباين

نظرية 2 : خاصية BLUE

إن مقدري المربعات الصغرى b وâ هما أفضل مقدرين خطيين غير متحيزين حيث أن لهما أصغر تباين ممكن مقارنة مع بقية المقدرات الخطية و غير المتحيزة الأخرى. (Best Linear Unbiased Estimator BLUE)

ت. مصفوفة التباينات لمقدرات معالم النموذج الخطي البسيط

النظرية 3 :

ان مصفوفة التباينات-التباينات المشتركة لـ مقدرات المربعات الصغرى $\hat{a}$ و $\hat{b}$ معطاة كما يلي:

$$V\left((\hat{a},\hat{b})\right) = \begin{pmatrix} V(\hat{a}) & \mathrm{cov}(\hat{a},\hat{b}) \\ \mathrm{cov}(\hat{a},\hat{b}) & V(\hat{b}) \end{pmatrix} = \begin{pmatrix} \frac{\sigma^2}{n}\frac{\sum x_i^2}{\sum(x_i-\bar{x})^2} & -\frac{\bar{x}\sigma^2}{\sum(x_i-\bar{x})^2} \\ -\frac{\bar{x}\sigma^2}{\sum(x_i-\bar{x})^2} & \frac{\sigma^2}{\sum(x_i-\bar{x})^2} \end{pmatrix}$$

حيث:

$$V(\hat{a}) = \sigma_{\hat{a}}^2 = \frac{\sigma^2}{n} \frac{\sum x_i^2}{\sum(x_i-\bar{x})^2} \quad ;$$

$$V(\hat{b}) = \sigma_{\hat{b}}^2 = \frac{\sigma^2}{\sum(x_i-\bar{x})^2} \,;$$

$$\mathrm{Cov}(\hat{a},\hat{b}) = \sigma_{a.b} = -\frac{\bar{x}\sigma^2}{\sum(x_i-\bar{x})^2}$$

البرهان :

$$V(\hat{\mathrm{b}}) = \mathrm{E}(\hat{\mathrm{b}}-\mathrm{b})^2 = \mathrm{E}\left[\frac{\sum(\mathrm{x_i}-\bar{\mathrm{x}})\varepsilon_\mathrm{i}}{\sum(\mathrm{x_i}-\bar{\mathrm{x}})^2}\right]^2$$

$$= \frac{\sum(\mathrm{x_i}-\bar{\mathrm{x}})^2\mathrm{E}(\varepsilon_\mathrm{i}^2)}{[\sum(\mathrm{x_i}-\bar{\mathrm{x}})]^2} = \frac{\sigma^2}{\sum(x_i-\bar{x})^2}$$

$$V(\hat{a}) = V\left(\hat{a}-(\hat{b}-b)\bar{x}+\bar{\varepsilon}\right) = V(\hat{b})\bar{x}^2 + \mathcal{V}(\bar{\varepsilon}) - 2\bar{x}cov(\hat{b}-b,\bar{\varepsilon})$$

$$V(\hat{a}) = \frac{\sigma^2\bar{x}^2}{\sum(x_i-\bar{x})^2} + \frac{\sigma^2}{n} \qquad ;$$

$Cov(\hat{b}-b,\bar{\varepsilon}) = 0$ لنبين ان:

$$COV(\hat{b}-b,\bar{\varepsilon}) = \mathrm{E}\left(\hat{b}-b,\bar{\varepsilon}\right) = E\left[\frac{\sum(\mathrm{x_i}-\bar{\mathrm{x}})\varepsilon_\mathrm{i}}{\sum(\mathrm{x_i}-\bar{\mathrm{x}})^2}\bar{\varepsilon}\right]^2 - E(\hat{b}-b)E(\bar{\varepsilon})$$

$$\Longrightarrow V(\hat{a}) = \frac{\sigma^2[\bar{x}^2n+\sum(x_i-\bar{x})^2]}{n\sum(x_i-\bar{x})^2} = \frac{\sigma^2\sum x_i^2}{n\sum(x_i-\bar{x})^2} = \frac{\sigma^2}{n}\left[1+\frac{n\bar{x}^2}{\sum(x_i-\bar{x})^2}\right] = \frac{\sigma^2}{n}\left[1+\frac{\bar{x}^2}{\mathcal{V}(x)}\right]$$

$$V(\hat{a}) = \frac{\sigma^2\sum x_i^2}{n^2\mathcal{V}(x)}$$

$$COV(\hat{a},\hat{b}) = E\left[(\hat{a}-a)(\hat{b}-b)\right]$$

$$= E\left[\left(-\bar{x}(\hat{b}-b)+\bar{\varepsilon}\right)(\hat{b}-b)\right]$$

$$= -\bar{x}E(\hat{b}-b)^2 + E\left[(\hat{b}-b)\bar{\varepsilon}\right]$$

$$= -\ \bar{x}E(\hat{b}-b)^2 + E\left[(\hat{b}-b)\bar{\varepsilon}\right]$$

$$= -\ \bar{x}\,\mathcal{V}(\hat{b}) = -\bar{x}\frac{\sigma^2}{\sum(x_i-\bar{x})^2}$$

4.3 مقدر غير متحيز ل $\hat{\sigma}^2$:

مقدر المربعات الصغرى الغير متحيز لـ σ^2 هو $\hat{\sigma}^2$ بحيث : $\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2} = \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n-2}$

البرهان:

لدينا: $SCR = \sum e_i^2 = \sum(y_i - \bar{y})^2 - \hat{b}^2 \sum(x_i - \bar{x})^2$

$$E\,(SCR) = E \sum(y_i - \bar{y})^2 - E(\hat{b})^2 \sum(x_i - \bar{x})^2$$

كما نعلم أن :

$$y_i - \bar{y} = b(x_i - \bar{x}) + (\varepsilon_i - \bar{\varepsilon})$$

$$\sum(y_i - \bar{y})^2 = b^2 \sum(x_i - \bar{x})^2 + \sum(\varepsilon_i - \bar{\varepsilon})^2 + 2b \sum(x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon})$$

$$E[\sum(y_i - \bar{y})^2] = b^2 \sum(x_i - \bar{x})^2 + E[\sum(\varepsilon_i - \bar{\varepsilon})^2] + 2b \sum(x_i - \bar{x}) E(\varepsilon_i - \bar{\varepsilon})$$

$$= b^2 \sum(x_i - \bar{x})^2 + E\left(\sum \varepsilon_i^2 - n\bar{\varepsilon}^2\right) + n\sigma^2 - n[v(\varepsilon) + E^2(\bar{\varepsilon})]\sigma^2 + 0$$

$$= b^2 \sum(x_i - \bar{x})^2 + n\sigma^2 - \sigma^2$$

$$\Rightarrow E[\sum(y_i - \bar{y})^2] = b^2 \sum(x_i - \bar{x}) + (n-1)\sigma^2$$

$$E\,(SCR) = b^2 \sum(x_i - \bar{x}) + (n-1)\sigma^2$$

$$E\,(SCR) = E(e_i^2) = b^2 \sum(x_i - \bar{x}) + (n-1)\sigma^2 - \sum(x_i - \bar{x})^2 E(\hat{b}^2)$$

$$E(\hat{b}^2) = V(\hat{b}) + E^2(\hat{b}) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2} + b^2$$

$$E(e_i^2) = b^2 \sum(x_i - \bar{x})^2 + (n-1)\delta^2 - \delta^2 - b^2 \sum(x_i - \bar{x})^2$$

$$= (n-2)\sigma^2$$

$$\Rightarrow E\left(\frac{e_i^2}{n-2}\right) = \sigma^2$$

و منه نحصل على :

$$\boxed{\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2}}$$ مقدرا غير متحيزا لـ σ^2 (تباين الأخطاء)

مثال تطبيقي:

نفترض أن العلاقة بين المتغيرين X وY هي من الشكل:

$$Y = \beta_0 + \beta_1 X + u \,.....(1)$$

وبافتراض أن فرضيات تطبيق المربعات الصغرى محققة وبتوفر المعلومات التالية :

$$\bar{X} = 70 \quad ; \quad \sum(X_i - \bar{X})(Y_i - \bar{Y}) = -35550 \;\; ; \;\; n = 12$$

$$\bar{Y} = 100 \quad ; \quad \sum(X_i - \bar{X})^{\wedge}2 = 2250 \; ; \; \sum(Y_i - \hat{Y}_i)^2 = 698$$

- أحسب مقدرات المعالم β_0 و β_1
- أحسب مقدر تباين الأخطاء ثم مقدر تباين كل من $\widehat{\beta_0}$ و $\widehat{\beta_1}$

الحل:

- حساب $\widehat{\beta_0}$ و $\widehat{\beta_1}$

$$\widehat{\beta_1} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^{\wedge}2} = \frac{-3550}{2250} = -1{,}578$$

$$\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1}\,\bar{X} = 100 - (-1.578)70 = 210.460$$

العلاقة المقدرة للنموذج (1) هي من الشكل :

$$\hat{Y} = 210.46 - 1.578X$$

- حساب مقدر تباين الأخطاء :

$$\hat{\sigma}^2 = \frac{\sum e^2}{n-2} = \frac{698}{12-2} = 69.8$$

- حساب مقدر تباين كل من : $\widehat{\beta_0}$ و $\widehat{\beta_1}$

$$\hat{V}(\widehat{\beta_0}) = \hat{\sigma}^2(\frac{1}{n} + \frac{\bar{X}^2}{\sum x^2}) = 69.8(\frac{1}{12} + \frac{(70)^2}{2250}) = 157.82$$

$$\hat{V}(\widehat{\beta_1}) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{n} x_i^2} = \frac{69,8}{2250} = 0.0310$$

4. توزيع المعاينة لمقدرات المربعات الصغرى

لدينا :

$$\begin{cases} \hat{a} = \bar{y} - \hat{b}\bar{x} = a + \left(b - \hat{b}\right)\bar{\varepsilon} \\ \hat{b} = b + \dfrac{\sum(x_i - \bar{x})\varepsilon_i}{\sum(x_i - \bar{x})^2} \\ \varepsilon_i \sim \mathbb{N}(0, \sigma^2) \end{cases}$$

بما أن التوزيع الطبيعي مستقر بالنسبة للمزج الخطي فإن :

$$\hat{b} \sim \mathbb{N}\left(b\,;\, \sigma^2 / \sum(x_i - \bar{x})^2\right) \Rightarrow z_{\hat{b}} = \frac{\left(\hat{b} - b\right)}{\sigma\sqrt{\dfrac{1}{\sum(x_i - \bar{x})^2}}} \sim \mathbb{N}(0; 1)$$

$$\hat{a} \sim \mathbb{N}\left(a\,;\, \frac{\sigma^2}{n}\,\frac{\sum x_i^2}{\sum(x_i - \bar{x})^2}\right) \Rightarrow z_a = \frac{(\hat{a} - a)}{\sigma\sqrt{\dfrac{\sum x_i^2}{n\sum(x_i - \bar{x})^2}}} \sim \mathbb{N}(0; 1)$$

بتعويض σ بمقدره ، يصبح توزيع كل من كما يلي :

$$t_{\hat{b}} = \frac{\left(\hat{b} - b\right)}{\hat{\sigma}\sqrt{\frac{1}{\sum(x_i - \bar{x})^2}}} \sim t_{n-2}$$

$$t_a = \frac{(\hat{a} - a)}{\hat{\sigma}\sqrt{\frac{\sum x_i^2}{n\sum(x_i - \bar{x})^2}}} \sim t_{n-2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2} = \frac{\sum(y_i - \hat{y})^2}{n-2}$$

$$\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-2}$$

5. طريقة الترجيح الأعظم لتقدير a ، b ، σ^2

لدينا : $y_i = a + bx_i + \varepsilon_i$ و $\varepsilon_i \sim N(0, \delta^2)$

$$\mathrm{E}(y_i) = a + bx_i \; ; \; v(y_i) = v(\varepsilon_i) = \sigma^2$$

وعليه فانً:

$$y \sim \mathbb{N}(a + bx_i \,; \sigma^2)$$

إن مقدرات a ، b ، σ^2 بطريقة الترجيح الأعظم هي حل جملة المعادلات :

$$\begin{cases} \dfrac{\partial lnl(y, a, b, \sigma^2)}{\partial a} = 0 \\ \dfrac{\partial lnl(y, a, b, \sigma^2)}{\partial b} = 0 \\ \dfrac{\partial lnl(y, a, b, \sigma^2)}{\partial \sigma^2} = 0 \\ \dfrac{\partial^2 lnl}{\partial a^2} < 0 \,, \dfrac{\partial^2 lnl}{\partial b^2} < 0 \,, \dfrac{\partial^2 lnl}{\partial \sigma^2} < 0 \end{cases}$$

$$L = \pi\, f(y, a, b, \sigma^2) = \pi\, \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - a - bx_i)^2}\right]$$

$$=\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n} e^{-\frac{1}{2\sigma^2}(y_i-a-bx_i)^2}$$

$$\text{Ln L} = -\frac{n}{2}Ln\,2\pi - \frac{n}{2}Ln\sigma^2 - \frac{1}{2\sigma^2}\sum(y_i - a - bx_i)$$

$$(*) \quad \begin{cases} \frac{\partial lnl}{\partial a} = \frac{1}{\sigma^2}\sum(y_i - a - bx_i) = 0 \\ \frac{\partial lnl}{\partial b} = \frac{1}{\sigma^2}\sum x_i(y_i - a - bx_i) = 0 \\ \frac{\partial lnl}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(y_i - a - bx_i)^2 = 0 \end{cases}$$

$$-n\sigma^2 + \sum(y_i - a - bx_i)^2 = 0 \quad \Rightarrow \quad \sigma^2 = \frac{\sum(y_i-a-bx_i)}{n}$$

بحل جملة المعادلات $(*)$ نجد :

$$\hat{a} = \bar{y} - \hat{\beta}\bar{x} \ldots\ldots\ldots\ldots\ldots\ldots.(1)$$

$$\hat{b} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \ldots\ldots\ldots\ldots\ldots.(2)$$

$$\boxed{\hat{\hat{\sigma}}^2 = \frac{\sum\left(y_i - \hat{a} - \hat{b}x_i\right)^2}{n} = \frac{\sum e_i^2}{n}}$$

نلاحظ : أن طريقة الترجيح الأعظم تعطي نفس المقدرات ل a و b في حين أنها تعطي مقدرا مختلفا لـσ^2 و هو مقدر متحيز حيث :

$$E\left(\hat{\hat{\sigma}}^2\right) = E\left(\frac{\sum e_i^2}{n}\right) = \frac{1}{n}E\left(\sum e_i^2\right) = \frac{1}{n}(n-2)\sigma^2$$

$$= \sigma^2 - \frac{2}{n}\sigma^2$$

إذن $\hat{\hat{\sigma}}^2$ هو مقدر متحيز و لكنه غير متحيز بالتقارب .

6. التقدير بمجال الثقة لمعالم النموذج الخطي البسيط

- <u>**مجال الثقة حول المعلمة الحقيقية b**</u>

لدينا : $\frac{\hat{b}-b}{\sigma/\sqrt{\sum(x_i-\bar{x})^2}} \backsim N(0,1)$

وبما أن σ مجهول نعوضه بمقدرة $\hat{\sigma}^2$: $T = \frac{\hat{b}-b}{\hat{\sigma}/\sqrt{\sum(x_i-\bar{x})^2}} \backsim t_{n-2}$

ومنه يمكن أن نكتب:

$$P\left(t_{\frac{\alpha}{2}} < \frac{\hat{b}-b}{\hat{\sigma}/\sqrt{\sum(x_i-\bar{x})^2}} < t_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Rightarrow P\left(\hat{b} - t_{1-\frac{\alpha}{2}}.\hat{\sigma}/\sqrt{\sum(x_i-\bar{x})^2} < \hat{b} < \hat{b} - t_{\frac{\alpha}{2}}.\hat{\sigma}/\sqrt{\sum(x_i-\bar{x})^2}\right) = 1-\alpha$$

$$I.C_{1-\alpha}(b) = \left[\hat{b} \pm t_{\frac{\alpha}{2}}.\hat{\sigma}/\sqrt{\sum(x_i-\bar{x})^2}\right]$$

- **<u>مجال الثقة حول المعلمة الحقيقية a (الثابت):</u>**

وباتباع نفس الخطوات التي تم بها بناء مجال الثقة حول الميل b نحصل على:

$$I.C_{1-\alpha}(a) = \left[\hat{a} \pm t_{\frac{\alpha}{2}}.\hat{\sigma}\Big/\sqrt{\frac{\sum x_i{}^2}{\sum(x_i-\bar{x})^2}}\right]$$

- **<u>مجال الثقة حول القيمة الحقيقية لتباين الأخطاء العشوائية σ^2 :</u>**

لدينا:

$$K = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \backsim \chi^2_{n-2}$$

$$\Rightarrow \left(\chi^2_{\frac{\alpha}{2}} < K < \chi^2_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$.\Longrightarrow I.C_{1-\alpha}(\sigma^2) = \left[\frac{(n-2)\hat{\delta}^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-2)\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}}}\right]$$

7. الاختبارات الإحصائية حول الدلالة الاحصائية (المعنوية) للنموذج المقدر

1.7 اختبار تحليل التباين

نهتم بتحليل التغيرات الكلية لـ y (المتغير الهدف أو المتغير الذي نريد تفسيره) الى جزء مشروح بواسطة النموذج والجزء غير المشروح بواسطة النموذج

لدينا:

$$SCT = \sum(y_i - \bar{y})^2 = \sum(y_i - \bar{y} + \hat{y}_i - \hat{y}_i)$$

$$= \sum(y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2 = \sum(\hat{y} - \bar{y} + e_i)^2$$

$$= \sum(\hat{y} - \bar{y})^2 + \sum(e_i)^2 + 2\sum(\hat{y}_i - \bar{y})\, e_i$$

$$= \sum(\hat{y} - \bar{y})^2 + \sum(e_i)^2 = SCE+SCR$$

علما أن:

$$\begin{cases} \sum \hat{y}_i e_i = 0 \\ \sum \bar{y} e_i = 0 \\ \bar{\hat{y}} = \bar{y} \end{cases}$$

La variabilité totale (مجموع التغيرات الكلية) مجموع مربع الإنحرافات الكلية : $\sum(y_i - \bar{y})^2$

La variabilité Expliquée مجموع مربع الإنحرافات المشروحة : $\sum\left(\hat{y} - \bar{y}\right)^2$

La variabilité Résiduelleمجموع مربع البواقي : $\sum(y_i - y)^2$

SCT = SCE + SCR ....(2)

بقسمة الطرفين على n نحصل على :

$$\frac{1}{n}\sum(y_i-\bar{y})^2=\frac{1}{n}\sum(\hat{y}-\bar{\hat{y}})^2+\frac{1}{n}\sum(y_i-\hat{y})^2$$

$\frac{1}{n}\sum(y_i-\bar{y})^2$: التباين الكلي VT

$\frac{1}{n}\sum(\hat{y}-\bar{\hat{y}})^2$: التباين المشروح VE

$\frac{1}{n}\sum(y_i-\hat{y})^2$: التباين غير المشروح VNE

و عليه يمكن بناء جدول تحليل التباين للنموذج البسيط

مصدر التباين (مصدر التغيرات)	مجموع المربعات	ddl (درجات الجرية)	متوسط مجموع المربعات
X	SCE=$\sum(\hat{y}-\bar{\hat{y}})^2=\hat{b}^2\sum(x_i-\bar{x})^2$	1	SCE/1
البواقي (الأخطاء)	SCR= $\sum(e_i)^2=\sum(y_i-\hat{y})^2$	n-2	SCR/n-2
المجموع	SCT= $\sum(y_i-\bar{y})^2$	n-1	SCT/n-1

إن اختبار معنوية النموذج باستعمال تحليل التباين تعني أن نهتم باختبار ما إذا كان X يشرح Y أي هل التباين المشروح (V.E) له دلالة احصائية أي اجراء الاختبار التالي اختبار :

$$\begin{cases} H_0: SCE=0 & \text{التابع المتغير كفاية لايشرح النموذج} \\ H_1: SCE\neq 0 & \text{النموذج يشرح كفاية المتغير التابع} \end{cases}$$

و من أجل ذلك نستعمل إحصائية الاختبار:

$$F=\frac{SCE/1}{SCR/n-2}\ \sim\ F_{1,n-2}$$

و يتم رفض $_0$ بمستوى دلالة α إذا كان:

$F_{1,n-2}$ < F القيمة المحسوبة

ملاحظة: إن أهمية تحليل التباين تظهر أكثر في حالة النموذج المتعدد.

2.7 اختبار جودة التوفيق R^2 :

لدينا :

SCT= SCE+ SCR

بقسمة الطرفين على SCT :

$$1 = \frac{\mathrm{SCE}}{\mathrm{SCT}} + \frac{\mathrm{SCR}}{\mathrm{SCT}}$$

إن $\frac{\mathrm{SCE}}{\mathrm{SCT}}$ تمثل النسبة المشروحة من y بواسطة x

$\frac{\mathrm{SCR}}{\mathrm{SCT}}$ تمثل النسبة غير المشروحة من y بواسطة x

تعريف : إن معامل التحديد ^{2} يقيس لنا النسبة المشروحة من y بواسطة x و يعطى بالعبارة التالية :

$$R^2 = \frac{\mathrm{SCE}}{\mathrm{SCT}} = 1 - \frac{\mathrm{SCR}}{\mathrm{SCT}}$$

و $$0 < R^2 \leq 1$$

كلما اقتربت قيمة ^{2} من " 1 " دل ذلك على أن x يشرح جيدا y و العكس صحيح.

$$R^2 = \frac{\mathrm{SCE}}{\mathrm{SCT}} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{\hat{b}^2 \sum(x_i - \bar{x})^2}{\sum\left(y^{22}{}_i - \bar{y}\right)^2} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})^2}{\sum(x_i - \bar{x})^2} = \frac{\left[\sum(x_i - \bar{x})(y_i - \bar{y})\right]^2}{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2} =$$

$$\left[\frac{cov\,(x,y)}{\delta_x \delta_y}\right]^2 = (r)^2$$

ملاحظة: $R^2 = (r)^2$ فقط في حالة النموذج البسيط

3.7 اختبار الفرضيات حول المعالم

ان اختبار الفرضيات حول معالم النموذج الذي نهتم بتقديره وعلى الخصوص الاختبار حول b معامل المتغير المستقل يسمح لنا بقبول أو رفض تفسير المتغير المستقل المفترض لتفسير المتغير التابع.

أن شكل اختبار الفرضية حول المعالم يكون كما يلي:

أ. الاختبار حول b: $H_0 : b = 0$ ضد $H_1 : b \neq 0$

ان قبول H_1 يعني أن النموذج المفترض لتفسير المتغير التابع y مقبول احصائيا (أي أن x يشرح y)

ان قاعدة القرار بمستوى دلالة α هي كما يلي:

$$d(T) = \begin{cases} a_0 : |T| < t_{n-2}(\alpha/2) \\ a_1 : |T| \geq t_{n-2}(\alpha/2) \end{cases}$$

حيث :

- $T = \frac{\hat{b}}{\hat{\sigma}_{\hat{b}}}$ تتبع توزيع ستودنت بالمعلمة $n-2$ و تسمى بـ T المحسوبة تحت(أو بفرض) H_0 محققة
- $t_{n-2}(\alpha/2)$ تسمى بقيمة T المجدولة

ب. الاختبار حول a: $H_0 : a = 0$ ضد $H_1 : a \neq 0$

بنفس الطريقة المستعملة بالنسبة للاختبار حول المعلمة b نختبر a الا أن رفض H_1 لا يعني رفض النموذج وفي أغلب الحالات نهمل الاختبار حول a

8. التنبؤ

ان هدف وضع نموذج لتفسير ظاهرة ما ممثلة بالمتغير y هو تفسير تغيرات هذه الظاهرة باستعمال متغيرات أخرى وبعد اختبار النموذج والتأكد من صحته يمكن استعمال النموذج المقدر من أجل التنبؤ بقيم الظاهرة المستقبلية اذا توفرت معلومات حول المتغيرات التي تشرحها .

اذا افترضنا أن :

- النموذج المقدر هو: $y_i = \hat{a} + \hat{b} \quad _i + e_i \,; \qquad i = 1,2, \ldots n$
- قيمة x من أجل $n+1$ معلومة و لتكن $_{n+1}$

عندئذ القيمة المقدرة لـ y عند $n+1$ معطاة كما يلي: $\hat{y}_{n+1} = \hat{a} + \hat{b}x_{n+1}$ وعندها يمكن حساب خطأ التنبؤ عند $n+1$ و هو: $e_{n+1} = y_{n+1} - \hat{y}_{n+1}$ حيث y_{n+1} هي القيمة الحقيقية لـ y عند $n+1$

يمكننا التأكد أن :

$$E(e_{n+1}) = 0$$

$$V(e_{n+1}) = \sigma^2{}_{e_{n+1}} = \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}\right]$$

لدينا: $e_{n+1} \sim N(0,\ \sigma^2{}_e\ )$ و بتعويض σ^2 بمقدره $\hat{\sigma}^2$ فان: $\frac{e_{n+1}}{\hat{\sigma}_{e_{n+1}}} \sim t_{n-2}$

ومنه يمكن أن نستنتج مجال التنبؤ حول $_{n+1}$ بدرجة ثقة $1-\alpha$ كما يلي :

$$\left[\hat{y}_{n+1} \pm t_{n-2}\left(\frac{\alpha}{2}\right)\hat{\sigma}_e\right] = \left[\hat{y}_{n+1} \pm t_{n-2}\left(\frac{\alpha}{2}\right)\hat{\sigma}\sqrt{\left[1 + \frac{1}{n} + \frac{(x_{n+1}-\bar{x})^2}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}\right]}\right]$$

نلاحظ أنه كلما كانت $\hat{\sigma}$ صغيرة ⟵ الإنحدار جيد أي كلما :كانت

- n كبير أي حجم العينة المستعمل في بناء النموذج كبير.
- $(x_0 - \bar{x})$ صغيرة أي أن الملاحظة قريبة من $\bar{}$.

فان الانحدار يكون جيد

يمكن الحصول على مجال الثقة حول التوقع E $(y_{n+1}$) بدرجة ثقة $1-\alpha$

$$E\,(y_{n+1}) \in \left[\hat{y}_{n+1} \pm\ \hat{\sigma}_{(\hat{y}_{n+1})} t_{n-1\left(1-\frac{\alpha}{2}\right)}\right]$$

$$\hat{\sigma}^2_{\hat{y}_{n+1}} = \hat{\sigma}^2_{\varepsilon}$$

9. تحليل للبواقي

لقد تم تعريف البواقي أو "الأخطاء الملاحظة" على أنها الفرق بين القيم الملاحظة والقيم المقدرة بنموذج الانحدار، ولها خصوصية تمثيل الجزء الذي لم تفسره معادلة الانحدار.

الغرض من تحليل البواقي هو اختبار صلاحية نموذج الانحدار. إنه يجعل من الممكن اكتشاف حالات الفشل في النموذج، ولهذا السبب من الضروري القيام به قبل أي تحليل للانحدار.

طرق تحليل البواقي أو البواقي المعيارية هي أساسا مبنية على التحليل البياني.

يتم اجراء التحليل البياني للبواقي بتمثيلين بيانيين ممكنين ، حيث أنهما يسمحان باكتشاف وجود فشل في النموذج قيد النظر:

رسم بياني للبواقي والقيم المقدرة لـ y

رسم بياني لقيم البواقي وقيم متغيرات الانحدار x

تحليل هذه البيانات بسيط للغاية. في الواقع، يتعلق الأمر بالتحقق من أن تمثيل البواقي لا يمثل أي شكل معين، ويمكن حصر النقاط التي نهتم بملاحظتها على التمثيل البياني للبواقي فيما يلي :

- ✓ الكشف عن الحالات الشاذة: تحديد القيم المتطرفة أو أخطاء التنبؤ.
- ✓ توزيع البواقي: فحص ما اذا كانت البواقي تتبع التوزيع الطبيعي .
- ✓ تجانس تباينات الأخطاء: التحقق من ثبات تباين الأخطاء للقيم المتنبئ بها .
- ✓ التغيير الهيكلي: فحص الشكل العام للتمثيل البياني للأخطاء الذي قد يشير الى علاقات غير خطية لم يتم أخذها بعين الاعتبار بواسطة النموذج.
- ✓ غياب متغير مهم: يمكن أن يؤثر عدم الاخذ بعين الاعتبار احد المتغيرات على توزيع البواقي ، مما يتطلب إعادة تقدير وتقييم النموذج

وبالتالي فإن تحليل البواقي يوفر رؤى أساسية لتحسين متانة الانحدار الخطي.

ومن أجل توضيح أهمية التحليل البياني للبواقي نأخذ المثال الموجود بالمقال العلمي المعنون بـ Graphs in Statistical Analysis لكاتبه F. J. Anscombe والمنشور بــ The American Statistician, Vol. 27, No. 1. (Feb., 1973) حيث يقترح عينة من 11ملاحظة لأربع ثنائيات من المتغيرات كما هو مبين في الجدول رقم () ثم يبين أن تقدير نموذج الانحدار الخطي البسيط للثنائيات الاربعة يعطي نفس النتائج الا أنه عند فحص التمثيل البياني تبين أنه نموذج واحد فقط مقبول ، أما الثلاثة الاخرين فكل منها يعاني من مشكل معين.

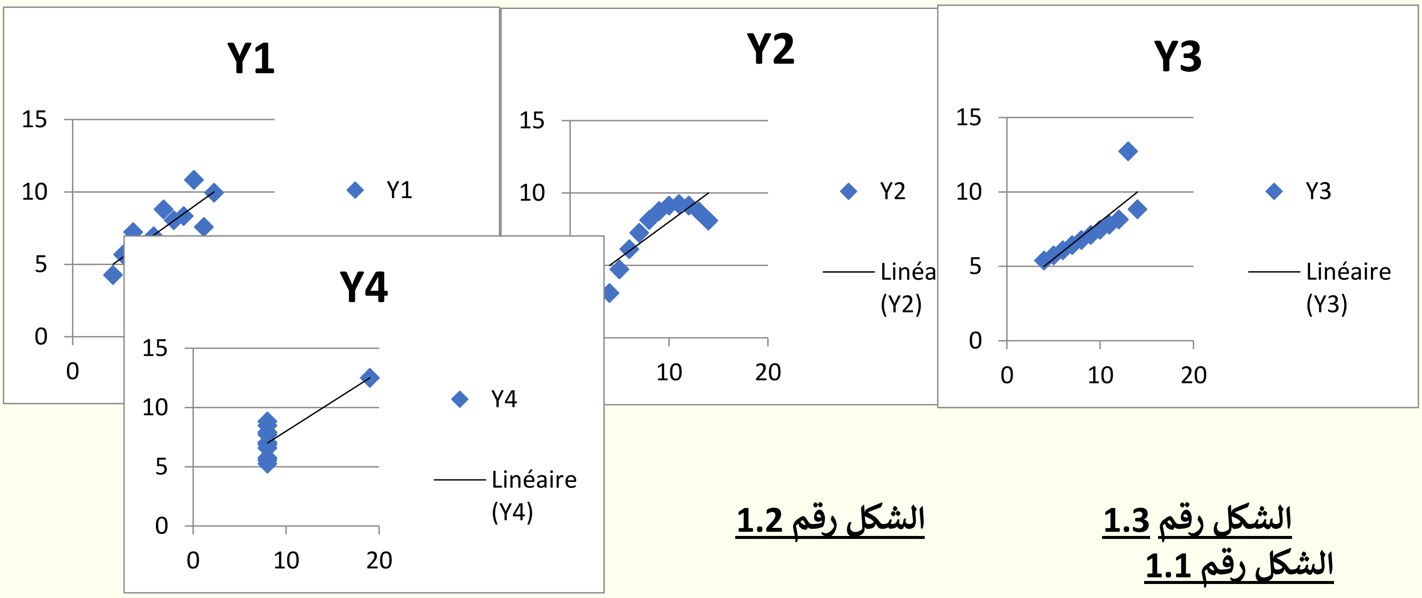
6. نص المثال التوضيحي

الجدول رقم (): بيانات لعينة من 11 ملاحظة حول ثماني متغيرات(أربع ثنائيات)

Y4	X4	X3	Y3	X2	Y2	Y1	x1	i
6,6	8	10	7,46	10	9,14	8,04	10	1
5,8	8	8	6,77	8	8,14	6,95	8	2
7,7	8	13	12,7	13	8,74	7,58	13	3
8,8	8	9	7,11	9	8,77	8,81	9	4
8,5	8	11	7,81	11	9,26	8,33	11	5
7	8	14	8,84	14	8,1	9,96	14	6
5,3	8	6	6,08	6	6,13	7,24	6	7
13	19	4	5,39	4	3,1	4,26	4	8
5,6	8	12	8,15	12	9,13	10,84	12	9
7,9	8	7	6,42	7	7,26	4,82	7	10
6,9	8	5	5,73	5	4,74	5,68	5	11

المصدر: Anscombe(1973)

- فيما يلي التمثيل البياني للثنائيات (x_i, y_i) بحيث: i=1,2,3,4



<u>الشكل رقم 1.2</u>

<u>الشكل رقم 1.3</u>

<u>الشكل رقم 1.1</u>

الشكل رقم 1-4

المطلوب:

1. اعط قراءة لكل شكل من الأشكال الأربعة
2. بعد افتراض أن العلاقة خطية بين كل من X وY بافتراض أن كل الفرضيات اللازمة لتطبيق طريقة المربعات الصغرى العادية محققة قمنا بتقدير النماذج الثلاثة وحصلنا على نفس النتائج كما يلي:

Dependent Variable: Y2
Method: Least Squares
Included observations: 11
Y2=C(1)+C(2)*X2

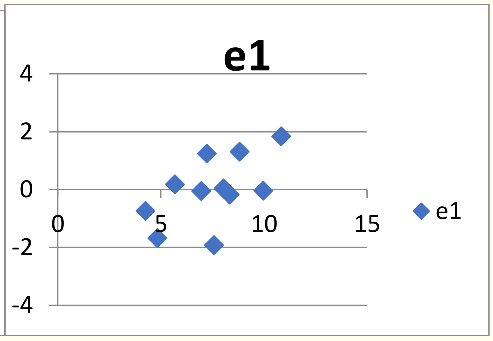
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	3.000909	1.125302	2.666758	0.0258
C(2)	0.500000	0.117964	4.238590	0.0022
R-squared	0.666242			

2.1. اعط قراءة لهذه النتائج، ثم اختبر صحة النماذج الأربعة.

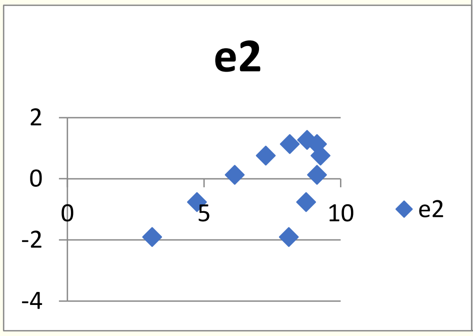
3. للتأكد من أن فرضيات تطبيق المربعات الصغرى محققة قمنا بحساب البواقي و البواقي المعيارية (الجدول2)ثم مثلناها بيانيا

e1	e1/s1	e2	e2/s2	e3	e3/s3	e4	e4/s4
0,039	0,0255	1,139	0,744	-0,541	-0,35	-0,4	-0,275
-0,051	-0,033	1,1392	0,745	-0,2308	-0,15	-1,2	-0,811
-1,921	-1,256	-0,7613	-0,498	3,23873	2,117	0,71	0,4635

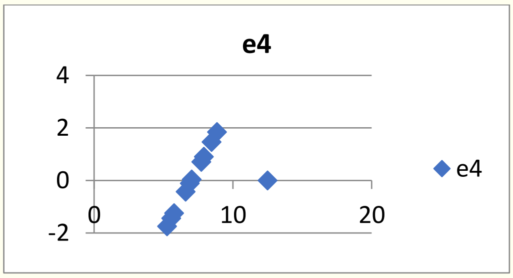
1,3091	0,8556	1,2691	0,829	-0,3909	-0,26	1,84	1,2021
-0,171	-0,112	0,7589	0,496	-0,6911	-0,45	1,47	0,9602
-0,041	-0,027	-1,9014	-1,243	-1,1614	-0,76	0,04	0,0256
1,2394	0,81	0,1294	0,085	0,07936	0,052	-1,8	-1,144
-0,74	-0,484	-1,9005	-1,242	0,38955	0,255	0	-0,001
1,8388	1,2018	0,1288	0,084	-0,8512	-0,56	-1,4	-0,942
-1,681	-1,099	0,7593	0,496	-0,0807	-0,05	0,91	0,5942
0,1795	0,1173	-0,7605	-0,497	0,22945	0,15	-0,1	-0,072



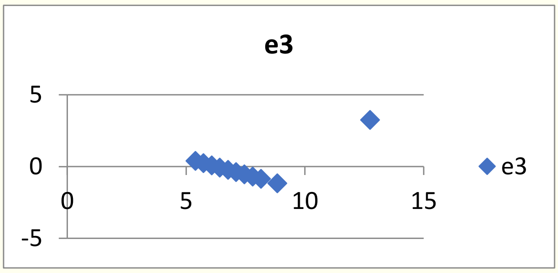
الشكل رقم 2.1



الشكل رقم 2.2



الشكل رقم 2.4



الشكل رقم 2.3

3.1. ماذا تلاحظ حول توزيع البواقي ؟ أعط قراءة لكل شكل على حدا .

تمرين تطبيقي:

تهتم مؤسسة لتسويق وإصلاح أجهزة الكمبيوتر بالتنبؤ بعدد المهندسين الواجب توفيرهم للسنوات القليلة المقبلة ، من أجل ذلك يجب تحليل المدة اللازمة لخدمة الطلبات و هذه الأخير مرتبطة بعدد الوحدات

الالكترونية الواجب استبدالها أو إصلاحها و لإنشاء هذه العلاقة أخذت عينة من تسجيلات خدمة الطلبات . المعطيات تتمثل في مدة خدمة الطلبات بالدقائق و عدد الوحدات المصلحة و هي كما يلي:

x	1	2	3	4	4	5	6	6	7	8	9	9	10	10
y	23	29	49	64	74	87	96	97	109	119	149	145	154	166

حيث عدد الدقائق ممثلة بـ Y وعدد الوحدات ممثلة بـ X.

تبين انه يمكن افتراض العلاقة الخطية التالية: بعد رسم لوحة الانتشار للثنائيات (x_i, y_i)

$$Minutes = \beta_0 + \beta_1\ Unités + \varepsilon$$

و بافتراض أن فرضيات تطبيق المربعات الصغرى محققة قدرنا معاملات النموذج المفترض و حصلنا على النتائج التالية:

$$\widehat{Minutes} = 4.162 + 15.509\ Unités$$

$$se \qquad 3.355 \qquad 0.505$$

$$n = 14\ ; \quad R^2 = 0.987\ ; \ S = 5.392$$

تمثل الخطأ المعياري للبواقي S تمثل الخطأ المعياري و se حيث

فسر قيم β_1 و β_0 المقدرة

ما هي النسبة المشروحة من خدمة الطلبات بواسطة عدد الوحدات المصلحة

اختبر الفرضية $_1 \neq 0$ ضد الفرضية $\beta_1 = 0$ ، ماذا تستنتج؟

أنشئ مجال الثقة حول القيمة الحقيقية لـβ_1 بدرجة ثقة 95%

ماهي المدة المتنبأ بها لتصليح 4 وحدات ، أوجد مجال التنبؤ للقيمة الحقيقية بدرجة ثقة 95%

6- اختبر $\beta_1 = 12$ ضد الفرضية $\beta_1 \neq 12$ بمستوى دلالة 5%، كيق يمكنك تفسير اختبار $\beta_1 = 12$

في مرحلة ثانية قامت المؤسسة بإضافة 10 ملاحظات للمعطيات السابقة:

I	1..........14	15	16	17	18	19	20	21	22	23	24
X	....	11	11	12	12	14	16	17	18	18	20
Y	....	162	174	180	176	179	193	193	195	198	205

بعد عملية التقدير كانت النتائج كما يلي:

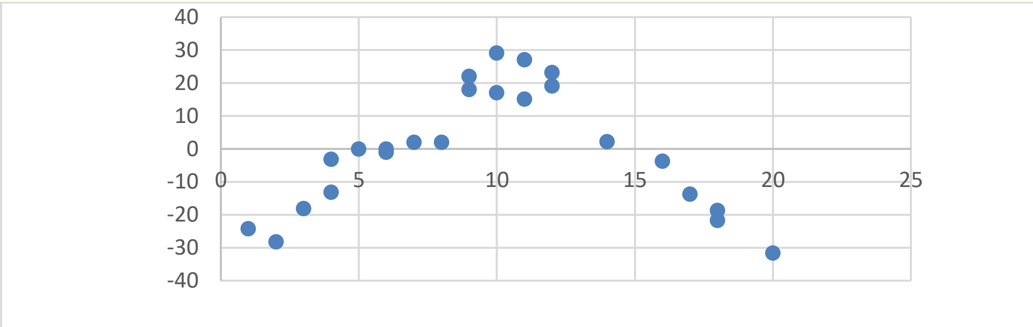
$$\widehat{Minutes} = 37.213 + 9.969\ Unités$$

$$se \qquad 7.985 \qquad 0.722$$

$$n = 24\ ; \quad R^2 = 0.900\ ; \quad S = 18.753$$

أ-قارن بين الخطأ المعياري للبواقي للنموذجين، ماذا يوحي لك الفرق بينهما

ب-نتيجة المقارنة بيت الخطأ المعياري للنموذجين تم رسم التمثيل البياني للبواقي بدلالة عدد الوحدات للنموذج الثاني



-ما هو تعليقك على هذا التمثيل البياني، ما هي الأسباب الممكنة التي أدت للحصول على هذا الشكل للبواقي.

-كيف يمكن تصحيح هذا الشكل كي يصبح في شكل توزيع متناظر.

10. تمارين غير محلولة

1. حدد فرضيات نموذج الانحدار الخطي البسيط واشرح بعناية أهمية كل فرضية السؤال 1.2 اكتب مقالاً عن خصائص عينة المقدرات.
2. أذكر خصائص المقدرات في العينة
3. وضح العلاقة بين التوزيع الطبيعي المعياري، وتوزيع كاي تربيع وتوزيع ستودنت وتوزيع فيشر
4. في سياق نموذج الانحدار الخطي، قارن وقارب بين طريقتي التقدير التاليتين: طريقة المربعات الصغرى العادية وطريقة الترجيح الاعظم
5. كتب ملاحظة موجزة عن العلاقة بين معامل التحديد ومعامل التحديد المصحح وإحصائية فيشر (موضحا أهميتهم في نمذجة نموذج الانحدار الخطي)
6. ناقش العبارة التالية : "الانحدار المتعدد أكثر فائدة من الانحدار البسيط في تحديد العلاقات الاقتصادية"
7. ماذا يعني الاقتصاد القياسي ب" مشكلة العينة الكبيرة جدا"؟
8. اشرح بالتفصيل معني الثابت في نموذج الانحدار الخطي
9. لنعتبر الانحدار التالي $Y = X\beta + u$ حيث Y هومتجه $n \times 1$ للملاحظات ، X هو مصفوفة $n \times k$ للمتغيرات التفسيرية بما في ذلك متجه العمود للواحد ، β هو $k \times 1$ متجه للمعلمات وu متجه $n \times 1$ من الاضطرابات

أ) أثبت أن معامل التحديد R2 هو مربع معامل الارتباط بين Y وقيمته المقدرة $\hat{}$

ب) بين أنه في حين أن مقدر المربعات الصغرى العادية β (OLS) يحقق Cramer-Rao MVB (الحد الأدنى للتباين)، فإن مقدر التباين σ2 لا يحقق ذلك .

ج) بين أن كل من مقدر OLS S^2 ومقدر MV لσ^2 كلاهما متسقان

10. نعتبر الانحدار $Y = X\beta + U$ حيث Y هو متجه $n \times 1$ للملاحظات ، X عبارة عن مصفوفة $n \times k$ للمتغيرات التفسيرية بما في ذلك متجه العمود الأحادي ، β هو متجه $k \times 1$ للمعلمات و U هو متجه $n \times 1$ للمتغير العشوائي . بين ما يلي :

أ) ليس بالضرورة أن يكون مجموع البواقي معدوما إذا لم يكن الحد الثابت موجودا،

ب) اذا كان $\bar{Y} \neq \hat{Y}$ أ محقق

11. نعتبر نموذج الانحدار التالي :

Y = Xβ + U

حيث Y هو متجه 1 × n للملاحظات على المتغير التابع ، X هو مصفوفة n × k للمتغيرات التفسيرية بما في ذلك متجه العمود الأحادي ، β هو 1 × k متجه للمعلمات و U هو متجه 1 × n للمتغير العشوائي .كما أن عدد الملاحظات ، n، يساوي عدد المعلمات ،.k

ما هو ثير n = k من حيثX ؟

ب (اشتقاق ، مقدر .OLS β ج (اشتقاق القيم المتوقعة من .Y التعليق على النتيجة .

د (اشتقاق .R2 التعليق على النتيجة .هـ (اشتقاق إحصائية .F التعليق على النتيجة .و (ما هي مسألة الاقتصاد القياسي التي يثيرها هذا الانحدار؟UWI) ،EC36C ، اختبار الفصل الدراسي2003

<u>تمرين تطبيقي:</u>

نعتبر نموذج الانحدار الخطي البسيط : $Y_i = \beta X_i + \varepsilon_i \quad i \in \{1,2,\dots.n\} \quad ;$

حيث : **أ)**β معلمة مجهولة يراد تقديرها ؛ **ب)**Y_iمتغيرات عشوائية مستقلة **ت)**X_iمتغيرات غير عشوائية

ث) ε_iمتغيرات عشوائية مستقلة لها نفس التوزيع الطبيعي بتوقع معدوم وتباين ثابت أي:

$$\varepsilon_i \sim N(0,\sigma^2) \text{ و } cov(\varepsilon_i,\varepsilon_j) = 0;\ \forall i \neq j$$

1. استعمل طريقة المربعات الصغرى لتقدير β وليكن $\widehat{\beta_0}$ مقدرها ،
2. بين أن$\widehat{\beta_0}$مقدر خطي و غير متحيز وأن تباينه هو$\frac{\sigma^2}{\sum_{i=1}^{n} {X_i}^2}$ ، ثم حدد توزيعه
3. نقترح مقدرين $\widehat{\beta_1}$ و $\widehat{\beta_2}$ لتقدير β بحيث $\widehat{\beta_1} = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i}{X_i}$: و $\widehat{\beta_2} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i^2}$

3.1بين أن $\widehat{\beta_1}$ مقدر غير متحيز لـ β **وأن**$\widehat{\beta_2}$ مقدر متحيز ثم حدد الشرط الذي من أجله $\widehat{\beta_2}$ يصبح غير متحيز،

قارن بين $\widehat{\beta_0}$ و $\widehat{\beta_1}$ من حيث التباين. (علما أنه حسب متباينة Cauchy- Schwarz لدينا : $n^2 \leq \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n}\frac{1}{X_i^2}$).

الحل:

- استعمال طريقة المربعات الصغرى لتقدير β وليكن $\widehat{\beta_0}$ مقدرها ،

$$/ \quad \text{SCR}= \sum e_i^2 \quad = \min \widehat{\beta_0}(MCO)$$

$$= \sum(y_i - \hat{y}_i)^2 = \sum(y_i - \beta x_i)^2 \sum e_i^2$$

$$\begin{cases} \dfrac{\partial \sum e_i^2}{\partial \beta} = 0 \\ \dfrac{\partial^2 \sum e_i^2}{\partial \beta^2} > 0 \end{cases} \begin{cases} \widehat{\beta_0} = \dfrac{\sum x_i y_i}{\sum {x_i}^2} \\ +2 \sum {x_i}^2 > 0 \end{cases} \rightarrow \begin{cases} -2 \sum x_i \left((y_i - \beta x_i)^2 = 0\right. \\ +2 \sum {x_i}^2 > 0 \end{cases} \rightarrow$$

- حدد توزيع المتغير العشوائي $\widehat{\beta_0}$، ثم بين أن هذا المقدر خطي و غير متحيز وأن تباينه هو $\frac{\sigma^2}{\sum_{i=1}^{n} {X_i}^2}$

$$\widehat{\beta_0} = \frac{\sum x_i y_i}{\sum {x_i}^2} = \frac{\sum x_i(\beta x_i + \varepsilon_i)}{\sum {x_i}^2} = \beta + \frac{\sum x_i \varepsilon_i}{\sum {x_i}^2} \rightarrow$$ خطي $\widehat{\beta_0}$

$$E(\widehat{\beta_0}) = E\left(\beta + \frac{\sum x_i \varepsilon_i}{\sum {x_i}^2}\right) = \beta + \frac{\sum x_i E(\varepsilon_i)}{\sum {x_i}^2} = \beta \rightarrow$$ غير متحيز $\widehat{\beta_0}$

$$V(\widehat{\beta_0}) = V\left(\beta + \frac{\sum x_i \varepsilon_i}{\sum {x_i}^2}\right) = V\left(\frac{\sum x_i \varepsilon_i}{\sum {x_i}^2}\right) = \frac{1}{\left(\sum {x_i}^2\right)^2} \sum {x_i}^2 V(\varepsilon_i) = \frac{\sigma^2}{\sum {x_i}^2}$$

$\rightarrow$ و هو المطلوب

$$\begin{cases} \widehat{\beta_0} = \beta + \dfrac{\sum x_i}{\sum {x_i}^2} \varepsilon_i \\ \varepsilon_i \sim N(O, \sigma^2) \\ \text{التوزيع الطبيعي مستقر بالنسبة للجمع} \end{cases} \rightarrow \widehat{\beta_0} \sim N\left(\beta, \frac{\sigma}{\sqrt{\sum {x_i}^2}}\right)$$

نقترح مقدرين $\widehat{\beta_1}$ و $\widehat{\beta_2}$ لتقدير β بحيث : $\widehat{\beta_1} = \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i}{X_i}$ و $\widehat{\beta_2} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i^2}$

- تبيان أن $\widehat{\beta_1}$ مقدر غير متحيز لـ β **وأن** $\widehat{\beta_2}$ مقدر متحيز ثم تحديد الشرط الذي من أجله $\widehat{\beta_2}$ يصبح غير متحيز،

$$\widehat{\beta_1} = \frac{1}{n}\sum_{i=1}^{n} \frac{\beta X_i + \varepsilon_i}{X_i} = \widehat{\beta_1} = \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i}{X_i} = \beta + \frac{1}{n} \sum \frac{\varepsilon_i}{X_i} = \beta + \frac{1}{n}\sum_{i=1}^{n}\left(\beta + \frac{\varepsilon_i}{X_i}\right) =$$

$$E(\widehat{\beta_1}) = E\left(\beta + \frac{1}{n}\sum\frac{\varepsilon_i}{X_i}\right) = \beta + \frac{1}{n}\sum\frac{E(\varepsilon_i)}{X_i} = \beta \quad \rightarrow$$ غير متحيز $\widehat{\beta_1}$

$$\widehat{\beta_2} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i^2} = \widehat{\beta_2} = \frac{\sum_{i=1}^{n}(\beta X_i + \varepsilon_i)}{\sum_{i=1}^{n} X_i^2} = \beta\frac{\sum X_i}{\sum {X_i}^2} + \frac{\sum \varepsilon_i}{\sum {X_i}^2}$$

$$E(\widehat{\beta_2}) = E\left(\beta\frac{\sum X_i}{\sum {X_i}^2} + \frac{\sum \varepsilon_i}{\sum {X_i}^2}\right) = \beta\frac{\sum X_i}{\sum {X_i}^2} + \frac{\sum E(\varepsilon_i)}{\sum {X_i}^2} = \beta\frac{\sum X_i}{\sum {X_i}^2} \neq \beta$$

$\rightarrow$ متحيز $\widehat{\beta_2}$

يصبح $\widehat{\beta_2}$ غير متحيز اذا كان $\sum X_i = \sum {X_i}^2$

- المقارنة بين $\widehat{\beta_0}$ و $\widehat{\beta_1}$ من حيث التباين. (علما أنه حسب متباينة Cauchy- Schwarz لدينا : $(n^2 \leq \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n}\frac{1}{X_i^2}$.

$$V(\widehat{\beta_1}) = V\left(\beta + \frac{1}{n}\sum\frac{\varepsilon_i}{X_i}\right) = \frac{1}{n^2}\sum\frac{V(\varepsilon_i)}{X_i} = \frac{\sigma^2}{n^2}\sum\frac{1}{X_i^2}$$

$$\frac{V(\widehat{\beta_0})}{V(\widehat{\beta_1})} = \frac{\frac{\sigma^2}{\sum {x_i}^2}}{\frac{\sigma^2}{n^2}\sum\frac{1}{X_i^2}} = \frac{n^2}{\sum {x_i}^2 \sum\frac{1}{X_i^2}} < 1 \quad (n^2 \leq \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n}\frac{1}{X_i^2})$$

الفصل الثالث: نموذج الانحدار الخطي المتعدد

إن نموذج الانحدار الخطي المتعدد هو تعميم لنموذج الانحدار الخطي البسيط ، أو هو توسيع نموذج الانحدار البسيط ذو المتغير المستقل الواحد إلى نموذج ذو K متغير مستقل .حيث نهتم بشرح قيم المتغير التابع y بواسطة K متغير مستقل $_j$ (j= 1,2,.......k) .

1. النموذج

يعرف نموذج الانحدار الخطي المتعدد بالمعادلة التالية:

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots\dots\dots\dots b_k x_{ik} + \varepsilon_i = \sum_{j=1}^{k} b x_i + \varepsilon_i$$

و علينا تقدير ال (k+1) معلمة $(b_0, b_1, \dots\dots\dots\dots, b_k)$ من خلال عينة ذات n ملاحظة .حيث :

- i : رقم الملاحظة i = 1 ,2,.........,n
- y_i : الملاحظة رقم i للمتغير y .
- x_{ij} : الملاحظة رقم للمتغير المستقل $_j$.
- ε_i : خطأ النموذج ، المتغير العشوائي و يشمل كل المعلومات التي لم يأخذها بعين الاعتبار النموذج من أجل شرح التغيرات في قيم y.
- b_j : معالم النموذج

كما يمكن كتابة النموذج في شكله المصفوفي كالتالي :

Y = X * b + ε

$$\begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & \dots\dots\dots & x_{1,k} \\ 1 & x_{2,1} & \dots\dots\dots & x_{2,k} \\ . & . & \dots\dots\dots & . \\ . & . & \dots\dots\dots & . \\ . & . & \dots\dots\dots & . \\ . & . & \dots\dots\dots & . \\ 1 & x_{n,1} & \dots\dots\dots & x_{n,k} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ . \\ . \\ . \\ b_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}$$

$$(n \times 1) \qquad \big(n \times (k+1)\big) \qquad (k+1 \times 1) \qquad (n \times 1)$$

نلاحظ أن العمود الأول للمصفوفة X عناصره = 1 . كونه يرافق المعلمة b_0 (الثابت في النموذج)

2. فرضيات النموذج

هي نفس فرضيات النموذج البسيط (الفرضيات العشوائية) بالإضافة الى الفرضيات الهيكلية التالية:

الفرضية الخامسة: غياب التعدد الخطي بين المتغيرات المستقلة. أي أن المصفوفة $(x'x)$ نظامية = (rang x = rang x' k+1) و (det $(x'x) \neq 0$) و منه $(x'x)^{-1}$ موجودة .

الفرضية السادسة: $\left(\frac{x'x}{n}\right)$تؤول إلى مصفوفة محددة وغير شاذة fini et non singulière

الفرضية السابعة: n $> k+1$: عندما n $> k+1$ نحصل على نظام ل n معادلة و n مجهول و منه يمكن تحديد كل b_i (الخط المستقيم يمر من كل النقاط) .

3. تقدير معالم النموذج باستعمال طريقة المربعات الصغرى العادية

إن مبدأ المربعات الصغرى العادية (MCO) هو تصغير مجموع مربعات البواقي أي :

$$\sum e_i^2 = Min \;....(1)$$

$$\sum e_i^2 = e'e = (y - xb)'(y - xb) = (\, y'y - y'xb - b'x'y + b'x'xb\,)$$

$$= (y'y - 2b'x'y + b'x'xb\,)$$

ملاحظة :

$b'x'y = y'xb$ حيث الاثنان سلميان $(1\times k+1)\times 1)$ (n,1) (k+1, k) (1 × k+1)

ان حل المعادلة(1) هو قيم التي تعدم المشتق الأول:

$$\frac{\partial \sum e_i^2}{\partial b} = -2x'y + 2x'xb = 0 \;\Rightarrow\; \hat{b} = (x'x)^{-1}x'y$$

و يكون هذا الحل مقبولا إذا كانت : $(x'x)$ تقبل معكوس .

ملاحظة : في حالة وجود ارتباط خطي بين متغيرين مستقلين ، المصفوفة $(x'x)$ تصبح شاذة و لا تقبل معكوس المعادلات الطبيعية للنموذج نحصل عليها من المعادلة التالية : $x'y = x'xb$ و صيغتها التحليلية هي كما يلي:

$$\begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{1i}y_i \\ \sum_{i=1}^{n} x_{2i}y_i \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=1}^{n} x_{ki}y_i \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} & \dots\dots\dots \sum_{i=1}^{n} x_{ki} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{1i}^2 & \sum_{i=1}^{n} x_{1i}\,x_{2i} & \dots\dots\dots \sum_{i=1}^{n} x_{1i}x_{ki} \\ \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{1i}\,x_{2i} & \sum_{i=1}^{n} x_{2i}^2 & \dots\dots\dots \sum_{i=1}^{n} x_{2i}x_{ki} \\ \dots \quad \dots & \dots & \dots & \dots\dots\dots \quad \dots \\ \dots & \dots & \dots & \dots\dots\dots \quad \dots \\ \sum_{i=1}^{n} x_{ki} & \sum_{i=1}^{n} x_{1i}\,x_{ki} & \sum_{i=1}^{n} x_{2i}\,x_{ki} & \dots\dots\dots \sum_{i=1}^{n} x_{ki}^2 \end{pmatrix} \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{b}_k \end{pmatrix}$$

4. خواص المقدرات

1.4 خاصية عدم التحيز

$y = x\,b + \varepsilon$

$\hat{y} = x\,\hat{b} + e \rightarrow e = y - \hat{y}$

$$\hat{b} = (x'x)^{-1}x'y = (x'x)^{-1}x'^{(xb+\varepsilon)} \underset{\hat{y}\,=\,x\,\hat{b}}{=} (x'x)^{-1}(x'x)b + (x'x)^{-1}x'\varepsilon$$

$\boxed{\hat{b} = b + (x'x)^{-1}x'\varepsilon \;\Rightarrow\; \text{خطي } \hat{b}}$

$\Rightarrow \hat{b} - b = (x'x)^{-1}(x'\varepsilon)$

$E(\hat{b}) = b + (x'x)^{-1}x'E(\varepsilon) \quad ; \qquad E(\varepsilon) = 0$

$\boxed{E(\hat{b}) = b \Rightarrow \text{غير متحيز } \hat{b}}$

2.4 مصفوفة التباينات للشعاع $\hat{b}$

لتكن $\Sigma_{\hat{b}}$مصفوفة التباينات (أو مصفوفة التباينات البيانات المشتركة) لـ $\hat{b}$ والمعرفة بـ:

$$\Sigma_{\hat{b}} = E(\hat{b}-b)(\hat{b}-b)' \quad ; \; \hat{b}-b = (x'x)^{-1}x'\varepsilon \; ;$$

$$(\hat{b}-b)' = \varepsilon' x\,(x'x)^{-1}$$

$$\Rightarrow \; \Sigma_{\hat{b}} = \mathrm{E}\,[(x'x)^{-1}x'\varepsilon\varepsilon' x\,(x'x)^{-1}] \; = \; (x'x)^{-1}x'E(\varepsilon\varepsilon')x(x'x)^{-1}$$

لنحسب $E(\varepsilon\varepsilon')$:

$$E(\varepsilon\varepsilon') = \begin{bmatrix} E(\varepsilon_1\varepsilon_1) & E(\varepsilon_1\varepsilon_2) & \dots & E(\varepsilon_1\varepsilon_n) \\ E(\varepsilon_2\varepsilon_1) & E(\varepsilon_2\varepsilon_2) & \cdots & E(\varepsilon_2\varepsilon_n) \\ \vdots & & \ddots & \vdots \\ E(\varepsilon_n\varepsilon_1) & & \cdots & E(\varepsilon_n\varepsilon_n) \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \dots & 0 \\ 0 & \sigma_\varepsilon^2 & \cdots & \cdots \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \sigma_\varepsilon^2 \end{bmatrix} = \; \sigma^2 I$$

$$\Sigma_{\hat{b}} = \sigma^2(x'x)^{-1}x'x(x'x)^{-1} \boxed{\Rightarrow \Sigma_{\hat{b}} = \sigma^2(x'x)^{-1}}$$

3.4 نظرية قوس ماركوف *Gauss-Markov*

إن مقدر MCO لديه خاصية Blue حيث أنه أحسن مقدر خطي غير متحيز(المقدرات لها أصغر تباين)

ملاحظة : إن طريقة MVالترجيح الأعظم تعطي نفس مقدرات المربعات الصغرى العادية MCO .

4.4 مقدر تباين الأخطاء

نظرية : إن مقدر تباين الأخطاء $\hat{\sigma}^2$ معطى بالعبارة التالية:

$$\hat{\sigma}^2 = \frac{SCR}{n-(\,k+1)} = \frac{SCR}{n-k-1} = \frac{\sum e_i^2}{n-k-1} = \frac{e'e}{n-k-1}$$

وعليه يمكن إيجاد مقدر مصفوفة التباينات:

بتعويض $\hat{\sigma}_\varepsilon^2$ تباين الأخطاء بمقدرها في عبارة مصفوفة التباينات-التباينات المشتركة بمعاملات النموذج $\hat{b}_{k+1}$ نحصل على مقدر مصفوفة التباينات –التباينات المشتركة ل $\hat{b}$.

$$\hat{\Sigma}_{\hat{b}} = \hat{\sigma}^2_{\hat{\varepsilon}}(x'x)^{-1}$$

حيث أن عناصر القطر الرئيسي تمثل مقدرات تباينات المعاملات $\hat{b}_j$.

مثال : لتكن المعطيات التالية حول المتغيرات y ، x_1 ، x_2 :

x_1	5	4	6	4	6
x_2	3	1	5	2	4
Y	3	1	8	3	5

نعتبر العلاقة بين y ، x_1 ، x_2 ، من الشكل التالي :

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \varepsilon_i$$

1.أكتب العلاقة على الشكل المصفوفي.

2.أعط المعادلات الطبيعية للنموذج، ثم قدر b_0، b_1، b_2 .

$$y = x \quad b + \varepsilon$$

$$\begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix}$$

المعادلات الطبيعية للنموذج: بفرض أن الفرضيات الكلاسيكية محققة و بتطبيق MCO يمكن أن نجد :

$$(x'x)\hat{b} \;=\; x'y$$

$$\begin{pmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{pmatrix}\begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \end{pmatrix} = \begin{pmatrix} 20 \\ 76 \\ 109 \end{pmatrix}$$

تقدير $\hat{b}$:

$$\hat{b} = (x'x)^{-1}x'y$$

$$(x'x)^{-1} = \frac{1}{|\Delta|}\,(adj(x'x)^{-1})'$$

$\Delta = 20$;

$$(x'x)^{-1} = \frac{1}{20}\begin{pmatrix} 534 & 90 & -160 \\ 90 & 20 & -30 \\ -160 & -30 & 50 \end{pmatrix} = \begin{pmatrix} 26.5 & 4.5 & -8 \\ 4.5 & 1 & -1.5 \\ -8 & -1.5 & 2.5 \end{pmatrix}$$

$$\hat{b} = \begin{pmatrix} 26.5 & 4.5 & -8 \\ 4.5 & 1 & -1.5 \\ -8 & -1.5 & 2.5 \end{pmatrix}\begin{pmatrix} 20 \\ 76 \\ 109 \end{pmatrix} = \begin{pmatrix} 4 \\ 2.5 \\ -1.5 \end{pmatrix}$$

$$\Rightarrow \hat{y} = 4 + 2.5x_1 - 1.5x_2$$

كما يمكن استعمال طريقة حل نظام معادلات لإيجاد قيم b_j من خلال المعادلات الطبيعية.

5. توزيع المعاينة للمقدرات

لدينا :

$$\hat{b} = b + (x'x)^{-1}x'\varepsilon$$

$$\Rightarrow \hat{b} \;\sim N(\,b\,;\sigma_\varepsilon^2(x'x)^{-1}) \Rightarrow \hat{b}_j \;\sim N\left(\,b_j\,;\,\delta_\varepsilon^2 a_{jj}\right)$$

$$\frac{\hat{b}_j - b_j}{\sigma_{\hat{b}}} = \frac{\hat{b} - b}{\sigma_\varepsilon\sqrt{a_{jj}}} \sim N\,(\,0\,;1\,)$$

وعند تعويض σ_ε^2 بمقدره $\hat{\sigma}_\varepsilon^2$

$$\frac{\hat{b}-b}{\sigma_{\varepsilon}\sqrt{a_{jj}}}=\frac{\hat{b}_j-b_j}{\sigma_{\hat{b}}}=\frac{\hat{b}_j-b_j}{Se\,(\hat{b})}\backsim t_{n-k-1}$$

$$(n-k-1)\frac{\hat{\sigma}_{\varepsilon}^2}{\sigma^2}\backsim \chi^2_{n-k-1}$$

$$\Longrightarrow I.C_{(1-\alpha)}(b_j)=\left[\hat{b}_j\pm\ t_{n-k-1}\ Se(\hat{b})\right]$$

6. الاختبارات الإحصائية للدلالة الإحصائية للنموذج المقدر

1.6 اختبار تحليل التباين

إن تحليل تغيرات y (SCT) أي $\sum(y-\bar{y})^2=(y-\bar{y})'(y-\bar{y})$:

إلى تغيرات مشروحة بواسطة النموذج (SCE) و تغيرات متبقية (SCR) ، يسمح لنا بتشكيل جدول لتحليل التباين، انظر الجدول رقم() .

الجدول رقم() : جدول تحليل التباين

مصدر التغير	مجموع المربعات	ddl	المربع المتوسط
المتغيرات المستقلة (النموذج)	SCE= $\sum(\hat{y}_i-\bar{y})^2$	k	SCE/k
البواقي	SCR= $\sum(y_i-\hat{y})^2$	$n-k-1$	$\mathrm{SCR}/n-k-1$
المجموع	SCT= $\sum(y-\bar{y})^2$	$n-1$	$\mathrm{SCT}/n-1$

كلما كانت عبارة مجموع المربعات للانحرافات المقدرة كبيرة و عبارة مجموع المربعات للبواقي صغيرة، كلما كان النموذج أفضل، أي أن المتغيرات المستقلة تشرح جيد المتغير التابع ولقياس ذلك يمكن استعمال اختبار فيشر .

2.6 اختبار جودة التوفيق (معامل التحديد):

إن الجزء المشروح من y بواسطة النموذج محدد بواسطة معامل التحديد .

$$R^2=\frac{\mathrm{SCE}}{\mathrm{SCT}}=1-\frac{\mathrm{SCR}}{\mathrm{SCT}}$$

ويمكن استنتاج العبارة التالية لمعامل التحديد في حالة غياب الثابت من النموذج الذي نريد تقدير معالمه:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y - \bar{y})^2} = \frac{\hat{b}'x'x\hat{b}}{y'y}$$

و نعلم أنه $(0 \leq R^2 \leq 1)$ فكلما كانت قيمة R^2 تقترب من "1" كلما كان النموذج جيد، و كلما إقتربت قيمته من "0" دل ذلك على أن المتغيرات المستقلة x_j لا تشرح إلا جزء قليل من y .

إلا أنه قيمة R^2 في النموذج المتعدد تتأثر بزيادة عدد المتغيرات المستقلة ولنزع هذا التأثير نحسب ما يسمى **بمعامل التحديد المصحح:**

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1}$$

ويتم تفسير معامل التحديد المصحح مثل معامل التحديد، أي أنه يمثل النسبة المشروحة من المتغير التابع بواسطة المتغيرات المستقلة منزوعا منه أثر زيادة عدد المتغيرات.

3.6 اختبار فيشر(اختبار الدلالة الكلية)

يسمح هذا الاختبار بقياس مدى قبول النموذج المقترح لتقدير العلاقة بين المتغير التابع والمتغيرات المستقلة. ويمكن بناء الفرضيات التالية لإجراء الاختبار عند مستوى الدلالة α:

$$\begin{cases} H_0: b_1 = b_2 \dots = b_k = 0 & \text{النموذج غير صحيح(غير مقبول)} \\ H_1: \quad \exists b_j \ / \ b_j \neq 0 & j = \overline{1,k} \quad \text{النموذج مقبول} \end{cases}$$

ومن أجل ذلك نستعمل متغير القرار التالي:

$$F = \frac{SCE/_k}{SCR/_{(n-k-1)}} = \frac{R^2/_k}{(1-R^2)/_{(n-k-1)}}$$

والذي نعلم أن توزيعه هو فيشر بدرجات الحرية $\boldsymbol{k}$ و $(n-k-1)$:

$$F \sim f_{k,(n-k-1)}$$

القيمة الحرجة التي من أجلها نرفض H_0 (أي قبول الفرضية البديلة) هي تلك القيم التي من أجلها تكون قيمة متغير القرار المحسوبة بواسطة معلومات العينة أكبر من قيمة متغير القرار النظرية المجدولة F_t والتي نحصل عليها من جدول الاحتمالات للتوزيع فيشر.

$$w = \{(\text{نموذج الانحدار}) / F_c > F_t\}$$

4.6 اختبار ستودنت

يمكن القيام باختبار الدلالة الاحصائية لكل معلمة على حدى بإجراء الاختبار التالي :

$$\begin{cases} H_0: b_j = 0 & \Longrightarrow \text{المتغير } x_j \text{ لا يشرح y أو أن } b_j \text{ ليس لها دلالة إحصائية} \\ H_1: b_j \neq 0 & \Longrightarrow \text{المتغير } x_j \text{ يشرح y و } b_j \text{ لها دلالة إحصائية} \end{cases}$$

لإجراء هذا الاختبار نستعمل الإحصائية :

$$T = \frac{\hat{b}_j - b_j}{Se\,(\hat{b})} \sim t_{n-k-1}$$

تحت H_0 محققة نحسب :

$$S \setminus H_0: \quad T_C = \frac{\hat{b}_j}{Se\,(\hat{b}_j)}$$

ثم نقارنها بقيمة t (الجزئي بالترتيب $1 - \frac{\alpha}{2}$ لتوزيع ستودنت) المجدولة و بدرجة حرية $n - k - 1$ بحيث تكون المنطقة الحرجة معرفة بـ:

$$w = \left\{|T_C| > t_{n-k-1}\left(1 - \frac{\alpha}{2}\right)\right\}$$

نرفض H_0 أي أن x_j يشرح y و b_j لها دلالة احصائية

ملاحظة : نقوم باختبار معالم المتغيرات المستقلة فقط دون المعلمة الثابتة .

7. الاختبارات الخطية حول شعاع معالم النموذج: b .

نهتم بكيفية استعمال مقدرb من أجل إجراء مختلف الاختبارات حول b .

لنعتبر الفرضيات الممكنة التالية :

الفرضية(1): $\boldsymbol{b_j = 0}$ ، أي نختبر إذا كان المتغير المستقل x_j يؤثر على y أي : هل الصفة x_j لها معنوية أم لا.

الفرضية(2) $\boldsymbol{b_j = b_0}$ ، حيث b_0 هي قيمة معطاة ، فمثلا إذا كانت b_j تمثل مرونة السعر ، نهتم باختبار ما إذا كانت $b_j = -1$.

الفرضية(3): $\boldsymbol{b_2 + b_3 = 1}$ إذا كانت تمثل............

الفرضية(4): $b_3 = b_4$ أو $b_3 - b_4 = 0$

الفرضية(5): $[b_1 \quad b_2 \ldots\ldots\ldots b_k] = [0 \quad 0 \ldots\ldots\ldots 0]$

الفرضية(6): $b_2 = 0$ حيث الشعاع b مقسم من شعاعين جزئيين b_1 و b_2 يحتوي كل منهما على K_1 و K_2 عنصر

بصفة عامة يمكن إستعمال الإحصائية التالية لاختبار كل الفرضيات السابقة :

$$\boldsymbol{F = \frac{(SCR_R - SCR_{NR})/q}{SCR_{NR}/n-k-1} \sim f_{q;(n-k-1)}}$$

بحيث:

$\boldsymbol{SCR_R}$: مجموع مربعات البواقي تحت H_O محققة (النموذج المختزل) او (النموذج المقيد)

$\boldsymbol{SCR_{NR}}$: مجموع مربعات البوقي على النموذج بدون قيود (نحدر y على كل المتغيرات المستقلة أو (النموذج الكامل).

$\boldsymbol{q}$: عدد القيود

8. التنبؤ

لنفترض أنه لدينا نموذج متعدد و قمنا بتقديره . لدينا :

$$\hat{b} = (x'x)^{-1}x'y \quad و \quad \hat{y} = x\hat{b} \quad و \quad y = xb + \varepsilon$$

ثم تحصلنا على القيم الخاصة بالمتغيرات المستقلة للفرد $(n+1)$ نكتبها في شكل شعاع كما يلي:

$$C' = [1 \quad x_{1n+1} \quad x_{2n+1} \quad \dots \quad x_{kn+1}]$$

وعند تعويضها في المعادلة المقدرة نحصل على:

$$\hat{y}_{n+1} = \hat{b}_0 + \hat{b}_1 x_{1n+1} + \dots + b_k x_{kn+1}$$

$$\hat{y}_{n+1} = C'^{\hat{b}}$$

حيث أن $\hat{y}_{n+1}$ هو التنبؤ النقطي لـ y_{n+1}

إلا أننا نهتم بإيجاد التنبؤ في شكل مجال ثقة و عليه يجب تحديد توزيع $\hat{y}_{n+1}$

$$\hat{b} \backsim N\left(E(\hat{b}), V(\hat{b})\right) \Rightarrow C'\hat{b} \backsim N\left(E(C'\hat{b}), V(C'\hat{b})\right)$$

$$\Rightarrow \hat{y}_{n+1} \backsim N\left(E(C'\hat{b}), V(C'\hat{b})\right)$$

$$E(C'\hat{b}) = C'E(\hat{b}) = C'b$$

$$V(C'\hat{b}) = C'V(\hat{b})C$$

$$\hat{y}_{n+1} \backsim N(c'^{\hat{b}}, c'V(\hat{b})c)$$

$$\Rightarrow \frac{\hat{y}_{n+1} - \grave{c}'b}{c'v(\hat{b})c} = \frac{c'^{\hat{b}} - c'b}{c'v(\hat{b})c} \backsim N(0,1)$$

بما أن $V(\widehat{\ }) = \sigma^2(x'x)^{-1}$ و σ^2 غير معلوم نعوضه ب $\hat{\sigma}^2$.

$$\widehat{V(\hat{b})} = \hat{\sigma}^2(x'x)^{-1}$$

و منه :

$$\frac{\hat{y}_{n+1} - c'b}{c'\widehat{V(\hat{b})}c} \sim t_{n-k-1}$$

وعليه يمكن استنتاج مجال الثقة حول القيمة الحقيقية المتنبئ بها كما يلي:

$$y_{n+1}\epsilon\left[\hat{y}_{n+1} \pm t_{1-\alpha/_2}\hat{\sigma}\sqrt{c'^{(x'x)^{-1}}c}\right]$$

9. استعمال المتغيرات المؤشرة في النمذجة

إن المتغيرات المستعملة في النمذجة كانت كلها كمية أي أن المتغيرات المستقلة كمية و المتغير التابع أيضا كمي ، و عندما نقول متغير كمي نقصد بذلك أنه قابل للقياس إلا أنه في الواقع يمكن أن نصادف متغيرات غير كمية أي غير قابلة للقياس (كيفية) و لها أثر في شرح الظاهرة المدروسة أو تكون هي الظاهرة التي نهتم بدراستها ، فمثلا عند دراسة الإستهلاك ، نفترض أن هذه الدالة تختلف بإختلاف مكان الأفراد أو ظروف البلاد من حيث الرخاء

- **استعمال المتغيرات المؤشرة كمتغيرات مستقلة**

قد تكون هذه المتغيرات تشرح الظاهرة بالإضافة إلى متغيرات مستقلة كمية و قد تكون هي المتغيرات المستقلة الوحيدة .

ليكن y المتغير التابع و $_1$ و المتغير المستقل الكمي.

$$y = b_o + b_1\, x_{1i} + \ \varepsilon_i$$

و لنفترض أن الأفراد i ينتمون إلى حالتين مختلفتين. :

مكانين (ريف1، مدينة2) ؛

زمنين (فترة1، فترة 2)،

صنفين (مجال 1 للأجور ، مجال 2 للأجور)

مرحلتين (مرحلة 1 الحرب ،مرحلة 2 السلم)

فترتين (فترة 1 صيف ، فترة2 شتاء) ....

ولنعرف المتغير المؤشر :

$$D = \begin{cases} 1 & \text{إذا كان } i \text{ ينتمي للحالة 1} \\ 0 & \text{العكس} \end{cases}$$

ونريد معرفة ما إذا كان تغير الحالة يؤثر على y. فنكتب النموذج على النحو التالي :

- **الحالة الأولى الاثر يكون على الثابت فقط:**

$$y = b_0 + b_1 x_{1i} + b_2 D_i + \varepsilon_i$$

$$y = \begin{cases} b_0 + b_1 x_{1i} + b_2 + \varepsilon_i = (b_0 + b_2) + b_1 x_{1i} + \varepsilon_i & si\ D = 1 \\ b_0 + b_1 x_{1i} + \varepsilon_i & si\ \ D = 0 \end{cases}$$

ونلاحظ أن الاختلاف بين النموذجين يكون بالنسبة للثابت.

بحيث:

- **الحالة الثانية الاثر يكون على الميل فقط:**

$$y = b_0 + b_1 x_{1i} + b_2 x_{1i} D_i + \varepsilon_i$$

بحيث:

$$y = \begin{cases} b_0 + b_1 x_{1i} + b_2 x_{1i} + \varepsilon_i = b_0 + (b_1 + b_2) x_{1i} + \varepsilon_i & si\ D = 1 \\ b_0 + b_1 x_{1i} + \varepsilon_i & si\ D = 0 \end{cases}$$

و نلاحظ أن الاختلاف يكون في الميل بين النموذجين.

- **الحالة الثالثة الاثر يكون على الثابت و الميل معا:**

$$y = b_0 + b_1 x_{1i} + b_2 D_i + b_3 D_i x_{1i} + \varepsilon_i$$

بحيث :

$$y = \begin{cases} b_0 + b_1 x_{1i} + b_2 + b_3 x_{1i} + \varepsilon_i = (b_0 + b_2) + (b_1 + b_3) x_{1i} + \varepsilon_i & si\ D = 1 \\ b_0 + b_1 x_{1i} \quad + \varepsilon_i & si\ D = 0 \end{cases}$$

و هنا يكون النموذجين مختلفين في الثابت و الميل.

و سواءا كان لدينا الشكل الأول أو الثاني أو الثالث فإننا نقوم بتقدير النموذجين ثم نقارن بينهما من حيث المعنوية

ملاحظة : يمكن أن نعرف عدة متغيرات مؤشرة في حالة متغير لديه أكثر من حالتين تعريفيتين .

مثلا: ليكن المتغير: التخصص لطلبة السنة الثانية ماستر بالمدرسة العليا للتجارة

$$x = \begin{cases} 1 & \text{التخصص مالية} \\ 2 & \text{التخصص مناجمنت} \\ 3 & \text{التخصص محاسبة} \\ 4 & \text{التخصص تسويق} \end{cases}$$

في هذه الحالة يمكننا تعريف اربع متغيرات مؤشرة:

$$D_4 = \begin{cases} 1 \\ 0 \end{cases} \qquad D_3 = \begin{cases} 1 \\ 0 \end{cases} \quad D_2 = \begin{cases} 1 \\ 0 \end{cases} \qquad D_1 = \begin{cases} 1 \\ 0 \end{cases}$$

الا أنه لا يمكن استعمال الا (4-1) متغير مؤشر على الأكثر كي لا نقع في مشكل التعدد الخطي

تمرين تطبيقي:

نعتبر النموذج الخطي المتعدد التالي: $Y_i = a_0 + a_1X_{1i} + a_2X_{2i} + \varepsilon_i$; $i \in \{1,2, \dots .25\}$

ونفترض أن ε_i متغيرات عشوائية مستقلة لها نفس التوزيع الطبيعي بتوقع معدوم وتباين ثابت σ^2

1. أعط عبارة مقدر المربعات الصغرى $\hat{a} = \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}$ للشعاع $a = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ ،

2. تقدير النموذج أعلاه أعطى النتائج التالية:

$$\hat{\sigma}^2 = 0{,}02 ؛ (X'X)^{-1} = \begin{pmatrix} 10 & -5{,}5 & -2{,}25 \\ -5{,}5 & 3{,}9 & 1{,}45 \\ -2{,}25 & 1{,}45 & 1{,}55 \end{pmatrix} ؛ \hat{a} = \begin{pmatrix} 0{,}69 \\ -1{,}51 \\ 0{,}82 \end{pmatrix}$$

$R^2 = 0{,}87$ ؛

أ- فسر قيمة معامل التحديد،

ب- أوجد عبارة وقيمة مقدر مصفوفة التباينات- التباينات المشتركة لـ $\hat{a}$: $\widehat{V(\hat{a})}$،

ت- ما هو توزيع ومعالم $\hat{a}$ ،

ث- اعط تفسيرا للمعلمة a_1، **كيف يتغير هذا التفسير اذا كان:** $X_1 = lnx_1$ و$Y = lny$،

ج- هل لمعامل المتغير a_1 دلالة إحصائية، قم بالاختبار المناسب عند مستوى الدلالة $\alpha = 5\%$ ،

ح- اكتب جدول تحليل التباين ثم اختبر صلاحية النموذج ككل عند مستوى الدلالة $\alpha = 5\%$ ،

خ- اذا افترضنا أنه نظريا قيمة a_3 تساوي الواحد، اختبر مدى صحة هذه النظرية بالنسبة للنموذج المقدر : اكتب صيغة فرضيات الاختبار ثم قم بالاختبار عند مستوى الدلالة $\alpha = 5\%$.

الحل:

- اعطاء عبارة مقدر المربعات الصغرى $\hat{a} = \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}$ للشعاع $a = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ ،

$$\hat{a} = (X'X)^{-1}X'Y \quad \rightarrow \quad \text{مقدر } MCO$$

- تفسير قيمة معامل التحديد :

87% من تغيرات Y مشروحة بواسطة X_1 و X_2

- ايجاد عبارة وقيمة مقدر مصفوفة التباينات- التباينات المشتركة لـ $\hat{}$: $\widehat{V(\hat{a})}$،

$$\hat{a} = (X'X)^{-1}X'Y = \ \hat{a} = (X'X)^{-1}X'(Xa+\varepsilon) = a + \ (X'X)^{-1}X'\varepsilon$$

$$V(\hat{a}) = V(a + \ (X'X)^{-1}X'\varepsilon\ ) = V(\ (X'X)^{-1}X'\varepsilon) = \ (X'X)^{-1}X'X(X'X)^{-1}\ V(\varepsilon)$$
$$= (X'X)^{-1}\ \sigma^2 \quad \rightarrow \quad \widehat{V(\hat{a})} = \ (X'X)^{-1}\ \hat{\sigma}^2$$

$$\widehat{V(\hat{a})} = 0{,}02\begin{pmatrix} 10 & -5{,}5 & -2{,}25 \\ -5{,}5 & 3{,}9 & 1{,}45 \\ -2{,}25 & 1{,}45 & 1{,}55 \end{pmatrix} = \begin{pmatrix} 0{,}2 & -0{,}11 & -0{,}045 \\ & 0{,}078 & 0{,}029 \\ & & 0{,}031 \end{pmatrix}$$

- **توزيع ومعالم** $\hat{a}$:

$$\hat{a} = a + \ (X'X)^{-1}X'\varepsilon$$

$$\varepsilon \sim \ N(0,\sigma)$$

$$E(\ \hat{a}) = \boldsymbol{E}(a + \ (X'X)^{-1}X'\varepsilon) = a + (X'X)^{-1}X'E(\varepsilon) = a$$

$$\rightarrow \ \hat{a} \sim N(a\ , \qquad V(\hat{a}))$$

- تفسيرا المعلمة a_1، و**كيف يتغير هذا التفسير اذا كان :** $X_1 = lnx_1$ و $Y = lny$ ،

لدينا $a_1 = \frac{\Delta Y}{\Delta X}$: وهي تمثل مقدار التغير الذي يحدث في Y اذا تغير X بوحدة واحدة (بفرض ثبات باقي العوامل المؤثرة في Y)

في حالة $X_1 = lnx_1$ و $Y = lny$ يصبح $a_1 = \frac{\partial Y}{\partial X} = \frac{\Delta y}{\Delta x}\frac{x}{y}$ وهي تمثل المرونة أو حساسية y للتغير في x أي أنه اذا تغير x بـ 1% فان y سيتغير بـ $a_1\%$ مع افتراض ثبات باقي العوامل

- دراسة الدلالة الإحصائية لمعامل المتغير a_1 عند مستوى الدلالة $\alpha = 5\%$ ،

نختبر $H_0: a_1 = 0$ ضد $H_1: a_1 \neq 0$

لدينا متغير القرار: $T = \frac{\hat{a}_1 - a_1}{Se(\hat{a}_1)} \sim t_{n-k-1}$

و قاعدة القرار: $d(T) = \begin{cases} a_0: & |T| < t(^{\alpha}/_{2}) \\ a_1: & |T| \geq t(^{\alpha}/_{2}) \end{cases}$

$$S/H_0 \ : T = \frac{\hat{a}_1}{Se(\hat{a}_1)} = \frac{-1{,}51}{\sqrt{0{,}078}} = -5{,}39$$

نرفض H_0 أي a_1 دال احصائيا $\rightarrow$ $t_{22}(0{,}025) = 2{,}08 \ < \ |T|$

- انشاء جدول تحليل التباين و اختبار صلاحية النموذج ككل عند مستوى الدلالة $\alpha = 5\%$ ،

حساب SCRو SCE و SCT

لدينا : $SCR = \widehat{\sigma^2}(n-k-1) \rightarrow \widehat{\sigma^2} = \frac{66SCR}{n-k-1}$

$$\rightarrow SCR = 0{,}02(22) = 0{,}44$$

لدينا : $SCE = 2{,}58 \leftarrow SCT = 3{,}38 \leftarrow \frac{SCR}{SCT} = 0{,}13 \leftarrow R^2 = 1 - \frac{SCR}{SCT} = 0{,}87$

جدول ANOVA

مصدر التغير	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	
النموذج البواقي المجموع	SCE=2,58 SCR=0,44 SCT=3,38	k=2 n-k-1=22	SCE/k=1,29 SCR/n-k-1=0,02	F=$\frac{\mathrm{SCE/k}}{\mathrm{SCR/n-k-1}}$ $\sim F_{k,n-k-1}$

اختبار صلاحية النموذج ككل :

نختبر $H_0: a_0 = a_1 = 0$ ضد $a_1 \neq 0$ أو $H_1: a_0 \neq 0$

F=$\frac{\mathrm{SCE/k}}{\mathrm{SCR/n-k-1}} = 64{,}5 \gg F_{2,22} = 3{,}22$

ومنه نرفض H_0

- اذا افترضنا أنه نظريا قيمة a_2 أقل من الواحد، اختبار مدى صحة هذه النظرية بالنسبة للنموذج المقدر عند مستوى الدلالة $\alpha = 5\%$.

نختبر: $H_0: a_2 = 1$ ضد $H_1: a_2 < 1$

لدينا : $T = \frac{\hat{a}_2 - a_2}{Se(\hat{a}_2)} \sim t_{n-k-1}$

و قاعدة القرار : $d(T) = \begin{cases} a_0: & T < t(\alpha) \\ a_1: & T \geq t(\alpha) \end{cases}$

$$S/H_0 \ : T = \frac{\hat{a}_2 - 1}{Se(\hat{a}_2)} = \frac{-0{,}18}{\sqrt{0{,}032}} = -1{,}02 > -t_{22}(0{,}05) = 1{,}721$$

نقبل H_0 أي نتائج النموذج تتوافق مع النظرية النموذج

تمرين تطبيقي 2

لدينا عينة من n=50 شركة مختصة في النقل الجوي ، حيث لاحظنا عليها :

Q الإنتاج مقاسا بالمنتوج المُنقل (الوزن × المسافة)

L العمل مقاسا بعدد الموظفين

K رأس المال مقاسا بوزن المركبة (بالطن)

PP صفة الشركة ما اذا كانت عمومية أو خاصة

I. نعتبر النموذج الخطي البسيط :

$$logQ = \alpha + \beta.logK + \varepsilon \dots\dots(1)$$

تقدير النموذج (1) باستعمال طريقة المربعات الصغرى أعطى النتائج التالية:

$$\widehat{logQ} = 6{,}93 \quad + \quad 0{,}89logK$$

(0,11) (0,258)

$$\sum_{i=1}^{50}(\widehat{logQ}_i - \overline{logQ})^2 = 1{,}78\sum_{i=1}^{50}(logQ_i - \overline{logQ})^2 \quad ; \quad = 2{,}0382$$

القيم بين قوسين تمثل، هنا و في باقي التمرين، الأخطاء المعيارية لمقدرات المعالم .

1. ما هي نسبة تغيرات لوغاريتم الانتاج المشروحة بواسطة لوغاريتم رأس المال ؟
2. فسر اقتصاديا معامل متغير رأس المال.
3. تحت فرضية أن البواقي تتبع التوزيع الطبيعي ، اختبر الدلالة الاحصائية لـ βعند مستوى الدلالة 5%.
4. عين مجال الثقة حول المعلمة بدرجة ثقة 95 %.

II. في هذا الجزء سنأخذ بعين الاعتبار التغيرات في عامل العمل التي اعتبرت ثابتة في النموذج (1)، عندها يصبح النموذج كما يلي:

$$(2) logQ = \alpha + \beta.logK + \gamma.logL + \varepsilon$$

تقدير هذا النموذج أعطى النتائج التالية:

$$\widehat{logQ} = 5{,}922 + 0{,}899 logK + 0{,}123 logL$$

(0,239) (0,118) (0,123)

$$,771 \sum_{i=1}^{50}(\widehat{logQ}_i - \overline{logQ})^2 =$$

1. فسر معامل متغير العمل $\hat{\gamma}$
2. تحت فرضية أن البواقي تتبع التوزيع الطبيعي ، اختبر الدلالة الاحصائية لـ γعند مستوى الدلالة 5%، ثم علق على النتيجة.
3. أكتب جدول تحليل التباين لهذا النموذج، ثم اختبر الدلالة الاحصائية للنموذج ككل عند مستوى الدلالة 5%.

III. من أجل اختبار فرضية ما اذا كان هناك اختلاف في الانتاجية بين الشركات الخاصة و الشركات العمومية، أضفنا للنموذج(2) المتغير PPوحصلنا على النموذج التالي :

$$(3) logQ = \alpha + \beta.logK + \gamma.logL + \delta.PP + \varepsilon$$

حيث PP متغير مؤشر معرف كما يلي :

$$PP = \begin{cases} 1 & \text{اذا كانت الشركة خاصة} \\ 0 & \text{اذا كانت الشركة عمومية} \end{cases}$$

نتائج تقدير النموذج (3) كانت كما يلي

$$\widehat{logQ} = 5{,}98 + 0{,}86 logK + 0{,}14 logL + 0{,}16 PP$$

(0,056) (0,115) (0,11) (0,224)

1. اعط تفسيرا اقتصاديا لمعامل PP.

2. اختبر فرضية أن الانتاجية تختلف ما بين الشركات الخاصة العمومية.

الفصل الرابع: أثار، اكتشاف وتصحيح خروقات فرضيات النموذج

. تعتبر طريقة المربعات الصغرى من الأساليب المستخدمة على نطاق واسع في تحليل الانحدار، ولكنها تعتمد على عدة افتراضات للحصول على نتائج دقيقة وموثوقة. يمكن أن يؤدي خرق هذه الافتراضات إلى تقديرات متحيزة وغير كفؤة، بالإضافة إلى استنتاجات غير موثوقة. الخروقات الشائعة لهذه الافتراضات.

1. الخطية: تفترض طريقة المربعات الصغرى وجود علاقة خطية بين المتغيرات المستقلة والمتغير التابع. إذا تم انتهاك هذا الافتراض ووجود علاقة غير خطية، فإن التقديرات باستخدام طريقة المربعات الصغرى قد لا تعكس بدقة العلاقة الحقيقية الكامنة.

2. الاستقلالية: يجب أن تكون المشاهدات المستخدمة في طريقة المربعات الصغرى مستقلة بعضها عن بعض. يمكن أن يحدث انتهاك للاستقلالية بأشكال مختلفة، مثل التباين الذاتي (الاعتماد بين المشاهدات المتتالية)، أو أخذ العينات المجتمعية (الاعتماد داخل المجموعات أو التجمعات)، أو بيانات السلاسل الزمنية ذات الترابط التسلسلي. في مثل هذه الحالات، قد تكون الأخطاء القياسية واختبارات الفرضيات غير صحيحة.

3. تجانس التباين: يفترض تجانس التباين أن تكون تباينات الأخطاء ثابتة عبر جميع مستويات المتغيرات المستقلة. انتهاك هذا الافتراض يؤدي إلى تحولة التباين، حيث تختلف تباين الأخطاء بين مستويات مختلفة من المتغيرات المستقلة. يمكن أن يؤدي التباين غير المتجانس إلى تقديرات متحيزة للأخطاء القياسية وتقديرات غير كفوءة واختبارات فرضيات غير صحيحة.

4. التوزيع الطبيعي: تفترض طريقة المربعات الصغرى أن الأخطاء تتبع توزيعاً طبيعياً. قد يؤدي التباعد عن الطبيعة إلى تحويلات متحيزة للمعلمات، وأخطاء قياسية غير صحيحة، واختبارات فرضيات غير موثوقة. ومع ذلك، تعتبر طريقة المربعات الصغرى متينة تجاه التباعد عن الطبيعة، خاصةً مع حجم العينة الكبير.

5. عدم وجود التأثير الداخلي: يحدث التأثير الداخلي عندما يكون هناك ترابط بين المتغيرات المستقلة ومصطلح الخطأ. يمكن أن يحدث هذا الانتهاك بسبب وجود متغيرات محذوفة، أخطاء قياس، التزامن، أو انحياز الاختيار. يمكن أن يؤدي التأثير الداخلي إلى تقديرات متحيزة وغير ثابتة للمعلمات، مما يجعل طريقة المربعات الصغرى غير موثوقة للاستعمال في الاستدلال السببي.

6. عدم وجود تعدد الاندماج: تشير تعددية الاندماج إلى وجود درجة عالية من الترابط بين المتغيرات المستقلة في نموذج الانحدار. يمكن أن يتسبب التعدد في مشاكل في طريقة المربعات الصغرى عن طريق صعوبة تمييز التأثيرات الفردية للمتغيرات المترابطة. يمكن أن ينتج عن ذلك تقديرات غير دقيقة وأخطاء قياسية كبيرة وصعوبة في تفسير النموذج.

عند انتهاك هذه الافتراضات، يمكن استخدام تقنيات التقدير البديلة واختبارات التشخيص لمعالجة المشكلات. تشمل هذه الطرق تقنيات التحليل القوي، ومربعات الصغرى العامة، وطرق المتغير المعياري، وغيرها.

من المهم أن يقوم الباحثون بفحص الافتراضات لطريقة المربعات الصغرى بعناية وتقييم ما إذا كانت تتعرض لانتهاكات في بياناتهم. يمكن أن تساعد الاختبارات التشخيصية، مثل تحليل الباقي، واختبارات تعدد الاندماج، وتحولات المتغيرات، في تحديد احتمالات الانتهاك. من خلال الاعتراف بالانتهاكات ومعالجتها، يمكن للباحثين تحسين دقة وموثوقية نماذج الانحدار الخاصة به

1.عدم تجانس تباين الأخطاء Heteroscedasticity

9.4 وهو الفرض الثاني الذي تكلمنا علية في طريقة المربعات الصغرى، وهو أن الخطأ العشوائي له تباين ثابت.

1.1. طبيعة عدم ثبات تباين الأخطاء، أسبابه وآثاره:

إذا كانت فرضية تجانس التباين غير محققة، فإن مصفوفة التباين-التباين المشترك للأخطاء تعرف كما يلي:

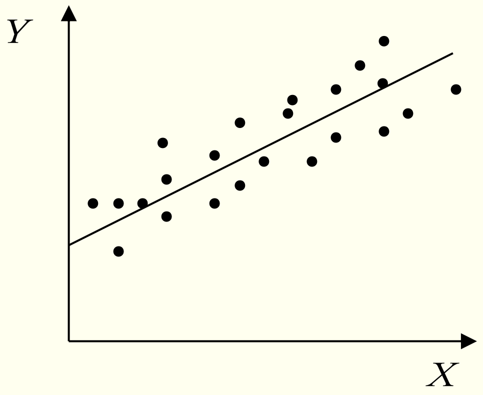
$$\Omega_{\varepsilon} = E(\varepsilon\varepsilon') = \begin{pmatrix} \sigma^2_{\varepsilon,1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\varepsilon,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{\varepsilon,n} \end{pmatrix} \neq \sigma^2_{\varepsilon} I_n$$

من الملاحظ أن تباينات الأخطاء ليست ثابتة على القطر الأول وبالتالي تباين الأخطاء مرتبط بقيم المتغير المستقل كما يظهر الشكل (3-3).

يوضح الشكل رقم (2) العلاقة المتوقعة بين المتغيرين التابع Y والمستقل X في حالة ثبات تباين الخطأ، ويلاحظ من خلال هذا الشكل أن تباين حد الخطأ لا يعتمد على قيم X

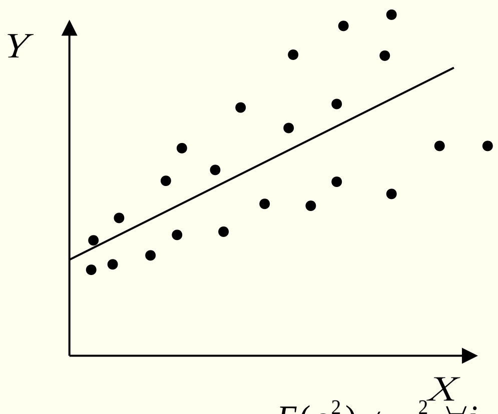
الشكل رقم (2)

ثبات تباين الخطأ في نموذج الانحدار البسيط



الشكل رقم (3)

عدم ثبات تباين الخطأ في نموذج الانحدار البسيط



ويوضح الشكل رقم (3) حالة عدم ثبات التباين لحد الخطأ $E(\varepsilon_i^2) \neq \sigma^2, \forall i$ ، حيث نلاحظ أن زيادة X سوف

تؤدي إلى زيادة تباين حد الخطأ، ويرتبط هذا المشكل ببيانات المقطع المستعرض Cross-section date أكثر من بيانات السلسلة الزمنية Cross-series date، حيث أن الأولى عبارة عن بيانات يتم تجميعها عن متغير ما في لحظة زمنية معينة (مثال: بيانات الإنفاق الاستهلاكي عند مستويات مختلفة لدخول الأفراد لسنة 2005)، أما بيانات السلسلة الزمنية فيتم تجميعها عن متغير ما عبر فترة زمنية معينة. وهناك عدة أسباب لعدم تجانس تباين حد الخطأ منها تحسن أساليب تجميع البيانات، وهذا يُقلل من الأخطاء المرتكبة في القياس، ومن ثم سوف يقل تباين حد الخطأ.

ويترتب على مشكلة عدم ثبات التباين عددا من الآثار تتمثل في[1]:

- تبقى المعالم المقدرة باستخدام المربعات الصغرى متصفة بعدم التحيز والاتساق، ولكنها تفقد صفة الكفاءة.
- تصبح التباينات المقدرة وكذلك التباينات المشتركة Covariances الخاصة بالمعالم المقدرة متحيزة وغير متسقة، ولذا فإن اختبارات الفرضيات لا تصبح دقيقة أو ملائمة.

[1] عبد القادر محمد عبد القادر عطية، ص 439.

- بالرغم من أن التنبؤات القائمة على أساس المعالم المقدرة باستخدام المربعات الصغرى العادية تظل غير متحيزة، إلا أنها تفقد صفة الكفاءة، وهو ما يعني أنها تكون أقل مصداقية من التنبؤات الأخرى.

في حالة عدم تجانس تباين الأخطاء، مقدر BLUE بطريقة المربعات الصغرى المعممة يكتب كما يلي:

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

$$\Omega_{\hat{\beta}} = (X'\Omega_\varepsilon^{-1}X)^{-1}$$

عكس تصحيح النموذج من الارتباط الذاتي، لا توجد منهجية موحدة للتصحيح من عدم ثبات تباين الأخطاء، فالطرق مختلفة مرتبطة بسبب وجود هذا المشكل.

7. 2.1. اختبارات اكتشاف عدم تباين الخطأ

يتم اكتشاف عدم ثبات تباين الأخطاء بواسطة عدة اختبارات منها ما يلي:

8. 1.2.1. اختبار Goldfeld-Quandt

بافتراض النموذج التالي $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, i = 1,..., n$ ، يمكن تبيان كيفية استخدام اختبار Goldfeld-Quandt في اكتشاف عدم ثبات تباين الخطأ من خلال الخطوات التالية:

- ترتيب مشاهدات X ترتيبا تصاعديا.

- استبعاد المشاهدات الوسطى لكل من X و Y،ثم تكوين مجموعتين من المشاهدات بحيث يكون لكل مجموعة على حدا معادلة خاصة بها كما يلي :

- المجموعة الأولى: وتتمثل في المشاهدات الخاصة بكل من X و Y الواردة قبل المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي : $Y_{1i} = a + bX_{1i} + \varepsilon_{1i}$

- المجموعة الثانية: وتتمثل في المشاهدات الخاصة بكل من X و Y الواردة بعد المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي: $Y_{2i} = c + dX_{2i} + \varepsilon_{2i}$

- تقدير معاملات المعادلتين السابقتين باستعمال المربعات الصغرى :

$$\hat{Y}_{1i} = \hat{a} + \hat{b}X_{1i}$$
$$Y_{2i} = \hat{c} + \hat{d}X_{2i}$$

❖ الحصول على القيم المقدرة لحد الخطأ:

$$\hat{\varepsilon}_{1i} = Y_{1i} - \hat{Y}_{1i}$$
$$\hat{\varepsilon}_{2i} = Y_{2i} - \hat{Y}_{2i}$$

❖ إيجاد القيمة المحسوبة لإحصائية F كما يلي:

$$F = \frac{\sum \hat{\varepsilon}_{2i}^2}{\sum \hat{\varepsilon}_{1i}^2}$$

❖ إيجاد درجات الحرية: $DF = \frac{n - m - 2(k+1)}{2}$

حيث k : عدد المتغيرات المستقلة، m : عدد المشاهدات المستبعدة.

❖ إيجاد القيمة المجدولة لإحصائية F عند درجات الحرية لكل من البسط والمقام، ومستوى معنوية معين.

❖ مقارنة بين القيم المحسوبة لإحصائية F والقيمة المجدولة لها :

- فإذا كانت F المحسوبة أكبر من F المجدولة، نقبل الفرضية البديلة أي فرضية عدم ثبات تباين الأخطاء.

- أما إذا كانت F المحسوبة أقل من F المجدولة، يتم قبول فرضية العدم.

لاحظ أن اختبار Goldfeld-Quandt لا يمكن تطبيقه إلا في حالة ما إذا كانت إحدى المتغيرات المستقلة هي المسببة في وجود مشكلة عدم ثبات تباين حد الخطأ.

9. 2.2.1اختبار White

اقترح (1980) White اختبارا يعتمد على العلاقة بين مربعات البواقي و جميع المتغيرات المستقلة و كذا مربعاتها. يمكن إبراز خطوات هذا الاختبار كما يلي:

- تقدير النموذج العام $Y = X\beta + \varepsilon$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_t^2$.

- تقدير المعادلة الوسطية التالية:

$$\hat{\varepsilon}_t^2 = \beta_0 + \beta_1 X_{t1} + \alpha_1 X_{t1}^2 + \ldots + \beta_k X_{tk} + \alpha_k X_{tk}^2 + u_t$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2.

- فرضية ثبات تباين الأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0 : \beta_0 = \alpha_1 = \beta_1 = \ldots = \alpha_k = \beta_k = 0$$

إحصائية مضاعف لاغرانج $LM = n \times R^2$ تتبع توزيع χ^2 بدرجة حرية $2k$. إذا كان $n \times R^2$ أكبر من $\chi^2(2k)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 أي إذا كان هناك على الأقل معامل واحد من معاملات المعادلة الوسيطية يختلف معنويا عن الصفر فإن تباين الأخطاء غير متجانس.

10. 3.2.1 اختبار ثبات التباين الشرطي للأخطاء ARCH-LM

تسمح نماذج ARCH[2] بنمذجة المتغيرات المالية التي تحتوي على تباين شرطي غير ثابت للأخطاء العشوائية حيث أن التطاير الشرطي الذي يعبر في الغالب عن المخاطرة غير ثابت. يعتمد إذن هذا الاختبار على مضاعف لاغرانج *LM*. خطوات الاختبار كالتالي:

- تقدير النموذج العام $Y = X\beta + \varepsilon$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_t^2$.

- تقدير المعادلة التالية:

$$\hat{\varepsilon}_t^2 = \theta_0 + \theta_1 \hat{\varepsilon}_{t-1}^2 + \ldots + \theta_q \hat{\varepsilon}_{t-q}^2 + u_t$$

مع حساب معامل التحديد الخاص بهذه المعادلة R^2. نفقد في هذه الحالة q مشاهدة.

[2] Engle (1982)

❖ فرضية ثبات التباين الشرطي للأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0 : \theta_0 = \theta_1 = ... = \theta_q = 0$$

إحصائية مضاعف لاغرانج $LM = (n-q) \times R^2$ تتبع توزيع χ^2 بدرجة حرية q. إذا كان $(n-q) \times R^2$ أكبر من $\chi^2(q)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 أي إذا كان هناك على الأقل معامل واحد من معاملات معادلة ARCH يختلف معنويا عن الصفر فإن التباين الشرطي للأخطاء غير متجانس.

3.1.11 معالجة عدم ثبات تباين حد الخطأ

من أبرز الطرق المستخدمة لتصحيح المشكلة هي طريقة المربعات الصغرى المرجحة، وتقوم هذه الفكرة على إعطاء القيم ذات الانحراف الأقل على خط الانحدار وزنا أكبر من القيم ذات الانحراف الأكبر في تقدير العلاقة محل الاعتبار[3]. ويتوقف شكل النموذج الأصلي المُحوَّل على نمط عدم ثبات التباين المكتشف في النموذج الأصلي المقدر.

وبفرض أن النموذج الأصلي كان كما يلي : $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, i = 1, ..., n$ فإن هناك عدة أنماط (افتراضات) لعدم ثبات تباين الأخطاء، ويختلف النموذج أو المعادلة المحولة من افتراض إلى آخر.

❖ الافتراض الأول: $E(\varepsilon_i^2) = \sigma^2 X_i^2$ وطبقا لهذا الافتراض يتم تحويل النموذج الأصلي إلى الشكل التالي:

$$\frac{Y_i}{X_i} = \frac{\beta_0}{X_i} + \beta_1 + \frac{\varepsilon_i}{X_i} = \beta_0 \frac{1}{X_i} + \beta_1 + \theta_i$$

حيث : θ_i عبارة عن حد الخطأ المحول $\frac{\varepsilon_i}{X_i}$

وبإجراء انحدار $\frac{Y_i}{X_i}$ على $\frac{1}{X_i}$ مستخدما طريقة المربعات الصغرى العادية نحصل على:

$$\left(\frac{\hat{Y}_i}{X_i}\right) = \hat{\beta}_0 \frac{1}{X_i} + \hat{\beta}_1$$

[3] عبد القادر محمد عبد القادر عطية، 1990، ص 452.

وبضرب المعادلة المحولة المقدرة السابقة في X_i يتم الحصول على النموذج الأصلي $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ بعد معالجة عدم ثبات التباين σ_ε^2، ويتضح مما سبق أن الحد الثابت في النموذج المحول (β_1) هو عبارة عن ميل معامل الانحدار للنموذج الأصلي، وميل معامل الانحدار للنموذج المحول هو عبارة عن الحد الثابت في النموذج الأصلي.

- الافتراض الثاني: $E(\varepsilon_i^2) = \sigma^2 X_i$ وطبقا لهذا الافتراض يتم تحويل النموذج الأصلي إلى المعادلة التالية:

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\beta_0}{\sqrt{X_i}} + \beta_1\sqrt{X_i} + \frac{\varepsilon_i}{\sqrt{X_i}} = \beta_0\frac{1}{\sqrt{X_i}} + \beta_1\sqrt{X_i} + \omega_i$$

حيث ω_i عبارة عن حد الخطأ المحول $\frac{\varepsilon_i}{\sqrt{X_i}}$، $X_i > 0$

وبنفس الحالة الأولى نجري انحدار $\frac{Y_i}{\sqrt{X_i}}$ على $\frac{1}{\sqrt{X_i}}$، $\sqrt{X_i}$ بواسطة المربعات الصغرى العادية.

- الافتراض الثالث: $E(\varepsilon_i^2) = \sigma^2 Y_i^2$، وطبقا لهذا الافتراض تكون المعادلة المحولة من الشكل:

$$\frac{Y_i}{Y_i} = \frac{\beta_0}{Y_i} + \beta_1\frac{X_i}{Y_i} + \frac{\varepsilon_i}{Y_i} = \beta_0\frac{1}{Y_i} + \beta_1\frac{X_i}{Y_i} + \varphi$$

- الافتراض الرابع: $E(\varepsilon_i^2) = \sigma^2 |\hat{\varepsilon}_i|$، ويتضمن هذا الافتراض أن تباين حد الخطأ دالة خطية لبواقي طريقة المربعات الصغرى العادية، وطبقا لهذا تكون المعادلة المقدرة كما يلي :

$$\frac{Y_i}{\sqrt{|\hat{\varepsilon}_i|}} = \beta_0\frac{1}{\sqrt{|\hat{\varepsilon}_i|}} + \beta_1\frac{X_i}{\sqrt{|\hat{\varepsilon}_i|}} + \frac{\varepsilon_i}{\sqrt{|\hat{\varepsilon}_i|}}$$

- الافتراض الخامس : التحويلات اللوغاريتمية، إن تحويل النموذج الأصلي إلى الصيغة اللوغاريتمية المزدوجة سوف يؤدي غالبا إلى تقليل درجة عدم ثبات تباين حد الخطأ، ومن ثم طبقا لهذا الافتراض تكون المعادلة المحولة المناسبة للنموذج الأصلي كما يلي :

❖ $LnY_i = \beta_0 + \beta_1 LnX_i + \varepsilon_i$.

2. الارتباط الذاتي للأخطاء

2.الارتباط الذاتي للأخطاء

12.1,2. الارتباط الذاتي من الدرجة الأولى AR(1):

في حالة أن جميع الفروض موجودة كل عشوائي يمثل عنصر مستقل من مجتمع موزع توزيع طبيعي ذا وسط صفري وتباين σ^2 . عندما يكون الارتباط الذاتي من الدرجة الأولى فأن العنصر غير مستقل بل يتبع النموذج التالي

$$u_t = \rho u_{t-1} + v_t$$

حيث ρ معلمه تقيس درجة الارتباط وقيمتها اقل من 1 و عشوائي موزعه توزيع طبيعي ومستقلة وذات وسط صفري وتباين ثابت = σ^2

$$u_t \sim N(0, \sigma_u^2) \quad \text{for all } t$$

$$E(u_t, u_s) = 0 \quad \text{for all } t$$

$$E(u_t, u_{t-1}) = 0 \quad \text{for all } t$$

إن العلاقة $u_t = \rho u_{t-1} + v_t$ تدل على إن التغير جزء من التغير السابق مضافا أليه تأثير يمثل بـ v_t. الارتباط هو درجة التغاير بين القيم الحالة والقيم التي قد تقع بعد s من الفترات، إذا كان الارتباط بين العناصر في الفترة s والفترة t-s يعني الارتباط بين العشوائي الحالي والعشوائي القادم مماثل للارتباط بين العشوائي الحالي والعشوائي السابق يكون Symmetric

2.2.13. خواص الارتباط الذاتي من الدرجة الأولى

الارتباط الذاتي من الدرجة الأولى وهو اكثر نموذج يستعمل ويأخذ الشكل التالي:

$$u_t = \rho u_{t-1}$$

حيث u تعتمد على u السابقة، ودرجة الاعتماد تعتمد على المعلمة ρ حيث تقيس درجة الارتباط الذاتي بين العناصر الحالية والعناصر السابقة. وتتراوح قيمته ρ بين $-1 \le \rho \le +1$ إذا كانت 1 = ρ معناه أن العشوائي الحالي يساوي العشوائي السابق وإذا كانت ρ = 1/2 معناه أن الحدث السابق يؤثر على الحدث الحالي ومقدار التأثير أن الحدث الحالي يساوي 1/2 الحدث السابق آي أن u_t تعتمد على قيمة u_{t-1} حسب قيمة ρ ولكن في العلاقة عادة تكون مصحوبة بمتغير عشوائي يمثل التغيرات العشوائية التي قد تصاحب الحدث. ولذلك نضيف العنصر العشوائي الذي هو v_t وهذا ما يسمى بمشروع الارتباط الذاتي من الدرجة الأولى ويسمى بالدرجة الأولى لأن تعتمد على u t-1 الدرجة الأولى آتية من 1 .

وتسمى Markov First - order Autoregressive scheme مشروع الانحدار الذاتي من الدرجة الأولى(AR1).

$$u_t = \rho u_{t-1} + v_t \qquad v_t \sim N(0, \sigma_u^2)$$

- E(u)=0

- تباين u^2 يساوي $Var(u_t) = \sigma_u^2$

إذا كانت البيانات ربع سنوية مثلا نقول $u_t = \rho u_{t-4} + u_t$ وفي هذه الحالة يكون ارتباط من الدرجة الرابعة.

إذا كانت ρ =صفر معناه إن u_t مستقلة عن u_{t-1}

لدينا عشوائيين u_t التي تؤثر على نموذج الانحدار الأصلي وثم u_t التي تؤثر على نفسها، فيما يختص بـ u_t التي تؤثر على النموذج الأصلي تعاني من مشكلة وجود الارتباط الذاتي فيها لان قيمتها تتحدد حسب القيم السابقة: $Y_i = \alpha + \beta X_i + u_i$ آي أن ترتبط ذاتيا مع القيم السابقة $u_t = \rho u_{t-1} + v_t$ و $u_t \sim N(0, \sigma_u^2)$ والتغاير بين u_t و u_{t-s} آي آن u_t يستوفي كل الفروض اللازمة لتطبيق م ص ع عليه ويمكن كتابة نموذج الانحدار الذاتي من الدرجة الأولى على النحو أعلاه. ويسمى في هذه الحالة Pure White Noise.

3.2.14. اختبارات الكشــف عن الارتباط الذاتي

من بين الاختبارات المخصصة عن مشكلة الارتباط الذاتي اختبار دير بن واتسون، اختبار h واختبار مضاعف لاجرانج :

1.3.2. اختبار دير بن واتسـون:

- اذن ينشأ الارتباط الذاتي من عدة أسباب منها :

- ❖ إهمال بعض المتغيرات التفسيرية في النموذج المراد تقديره.
- ❖ الصياغة الرياضية الخاطئة للنموذج.
- ❖ عدم دقة بيانات السلاسل الزمنية.

أما وجوده يؤثر سلبا على نتائج المربعات الصغرى العادية من حيث :

- ❖ القيم المقدرة للمعاملات سوف تكون غير متحيزة.
- ❖ تباين القيم المقدرة لمعاملات النموذج سوف لا يكون أقل ما يمكن.لذلك تستعمل عدة اختبارات للكشف على هذا الاختلال منها ما يلي :

2.3.2.15 اختبار دربين واتسون Durbin-Watson test (1950 et 1951)

يعتبر اختبار Durbin-Watson من أهم الاختبارات الشائعة المستخدمة في اكتشاف الارتباط الذاتي من الدرجة الأولى حسب الشكل:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t\ ,\quad u_t \sim N(0,\sigma_u^2)$$

ويهدف إلى اختبار الفرضيات التالية :

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

لاختبار فرضية العدم H_0 يجب حساب إحصائية دربين واتسون DW:

$$DW = \frac{\sum_{t=2}^{n}\left(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}\right)^2}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

يمكن كتابة الإحصائية أيضا بدلالة مقدر معامل الارتباط ρ ، لدينا:

$$DW = \frac{\sum_{t=1}^{n}\hat{\varepsilon}_t^2 + \sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2 - 2\sum_{t=1}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

نلاحظ أن: $\sum_{t=1}^{n}\hat{\varepsilon}_t^2 \cong \sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2$ إذن:

$$DW \cong \frac{2\sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2} - \frac{2\sum_{t=1}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

نعلم أن:

$$\hat{\rho} = \frac{\sum_{t=1}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2}$$

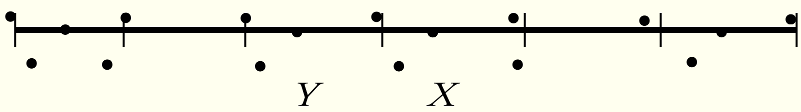
ومنه:

$$DW \cong 2(1-\hat{\rho})$$

حيث أن الإحصائية DW تمثل القيمة المحسوبة للاختبار وتأخذ قيمها بين 0 و4. ويتضح من المعادلة السابقة أنه إذا كانت $\rho = 0$ فإن $DW \cong 2$.

ويوضح الشكل التالي قيم *d* (القيم المجدولة للاختبار)، التي تشير إلى وجود أو عدم وجود ارتباط ذاتي من الدرجة الأولى موجب أو سالب، أو تجعل نتيجة الاختبار غير محددة، وتوجد قيم كل من الحدين الأعلى والأدنى لـ *d* (d_L,d_U) في الجدول الإحصائي لتوزيع دربين واتسون.

الشكل رقم (1) : مناطق القبول والرفض لاختبار Durbin-Watson



بالاعتماد على الشكل رقم (1) يمكن أن تُستخرج نتيجة اختبار DW كالتالي:

- ❖ إذا كانت $DW < d_L$ أو $DW > 4 - d_L$ يرفض H_0 .
- ❖ إذا كانت $4 - d_U > DW > d_U$ يقبل H_0 .
- ❖ إذا كانت $4 - d_U \le DW \le 4 - d_L$ أو $d_L \le DW \le d_U$ تكون نتيجة الاختبار غير محددة، ومن ثم يجب إضافة بيانات أكثر.

لا يمكن استعمال هذا الاختبار إلا في ظل الشروط التالية:

- ❖ يجب أن يكون النموذج متضمنا للمعلم الثابت β_0
- ❖ النموذج المقدر لا يتضمن متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة[4] لا يختبر دربين واتسون إلا الارتباط الذاتي من الدرجة الأولى.

3.3.2.16 اختبار Breusch-Godfrey

يرتكز هذا الاختبار على مضاعف لاغرانج و الذي يسمح باختبار وجود ارتباط ذاتي من درجة أكبر من الواحد. نموذج الانحدار الذاتي للأخطاء من الدرجة p يكتب على الشكل التالي:

$$\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \ldots + \rho_p\varepsilon_{t-p} + u_t$$

ليكن النموذج العام حيث أن الأخطاء مرتبطة ذاتيا:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \ldots + \beta_k X_{tk} + \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \ldots + \rho_p\varepsilon_{t-p} + u_t$$

هناك ثلاث خطوات لإجراء هذا الاختبار:

- ❖ تقدير النموذج العام بطريقة المربعات الصغرى ثم حساب البواقي $\hat{\varepsilon}_t$

[4] قد يتضمن هذا النوع من النماذج متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة، سيتم التطرق إلى هذا النوع من النماذج في الفصل الرابع.

- تقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}_t = \beta_0 + \beta_1 X_{t1} + ... + \beta_k X_{tk} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + ... + \rho_p \hat{\varepsilon}_{t-p} + u_t$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2. نذكر أن باستعمال هذه المعادلة، سنفقد p مشاهدة.

- فرضية استقلالية الأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0 : \rho_1 = \rho_2 = ... = \rho_p = 0$$

الإحصائية $LM = (n-p) \times R^2$ تتبع توزيع χ^2 بدرجة حرية p. إذا كان $(n-p) \times R^2$ أكبر من $\chi^2(p)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 فرضية استقلالية الأخطاء.

4.2.17 طرق التقدير في حالة وجود ارتباط ذاتي بين الأخطاء

إذا أخذنا بعين الاعتبار الارتباط الذاتي بين الأخطاء، النموذج الخطي العام يكتب على الشكل التالي:

$$Y = X\beta + \varepsilon$$

مع:

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t, \ |\rho| < 1$$

تعبر هذه المعادلة عن سيرورة الانحدار الذاتي من الدرجة الأولى $AR(1)$ الذي يحقق الفرضيات التالية:

$$\begin{aligned} &E(u_t) = 0 \\ &E(u_t^2) = \sigma_u^2, \forall t \\ &\text{cov}(u_t, u_{t'}) = 0, \forall t \neq t' \\ &\text{cov}(u_t, \varepsilon_{t-1}) = 0, \forall t \end{aligned}$$

من نموذج الانحدار الذاتي، نعوض ε_{t-1} بما يساويه، نحصل على:

$$\varepsilon_t = \rho(\rho\varepsilon_{t-2} + u_{t-1}) + u_t = \rho^2 \varepsilon_{t-2} + \rho u_{t-1} + u_t$$

نعوض ε_{t-2} بما يساويها، نحصل أيضا:

$$\varepsilon_t = \rho^2(\rho\varepsilon_{t-3} + u_{t-2}) + \rho u_{t-1} + u_t = \rho^3\varepsilon_{t-3} + \rho^2 u_{t-2} + \rho u_{t-1} + u_t$$

في الأخير، لدينا العادلة:

$$\varepsilon_t = u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + \ldots$$

تؤول هذه السيرورة إلى الصفر لأن: $|\rho| < 1$.

نقوم إذن بدراسة خصائص ε_t.

لدينا $E(\varepsilon_t) = 0$ و بتربيع الخطأ ε_t، نحصل على:

$$\varepsilon_t^2 = u_t^2 + \rho^2 u_{t-1}^2 + \rho^4 u_{t-2}^2 + \rho^6 u_{t-3}^2 + \ldots$$

ثم بإدخال التوقع الرياضي على الطرفين:

$$E(\varepsilon_t^2) = E(u_t^2) + \rho^2 E(u_{t-1}^2) + \rho^4 E(u_{t-2}^2) + \rho^6 E(u_{t-3}^2) + \ldots$$

نعلم أن: $$E(u_t^2) = E(u_{t-1}^2) = E(u_{t-2}^2) = E(u_{t-3}^2) = \ldots = \sigma_u^2, \forall t$$

ومنه: $$\mathrm{var}(\varepsilon_t) = E(\varepsilon_t^2) = (1 + \rho^2 + \rho^4 + \rho^6 + \ldots)\sigma_u^2$$

إذن: $$\mathrm{var}(\varepsilon_t) = E(\varepsilon_t^2) = \frac{\sigma_u^2}{1 - \rho^2}$$

وباستعمال خصائص u_t، نقوم بحساب التباين المشترك للخطأ ε_t، لدينا:

$$\begin{aligned}\varepsilon_t\varepsilon_{t-1} &= (u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + \ldots)(u_{t-1} + \rho u_{t-2} + \rho^2 u_{t-3} + \rho^3 u_{t-4} + \ldots) \\ &= \rho u_{t-1}^2 + \rho^3 u_{t-2}^2 + \rho^5 u_{t-3}^2 + \ldots\end{aligned}$$

بإدخال التوقع الرياضي على الطرفين، نحصل في الأخير على:

$$E(\varepsilon_t \varepsilon_{t-1}) = \frac{\rho \sigma_u^2}{1-\rho^2}$$

نفس الشيء مع $E(\varepsilon_t \varepsilon_{t-2})$:

$$E(\varepsilon_t \varepsilon_{t-2}) = \frac{\rho^2 \sigma_u^2}{1-\rho^2}$$

في الحالة العامة، سيكون لدينا:

$$E(\varepsilon_t \varepsilon_{t-i}) = \frac{\rho^i \sigma_u^2}{1-\rho^2}$$

مصفوفة التباين-التباين المشترك للأخطاء في حالة وجود ارتباط ذاتي من الدرجة الأولى تكتب على الشكل التالي:

$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \frac{\sigma_u^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}$$

مع $|\rho| < 1$

نذكر أن مقدر طريقة المربعات الصغرى يكتب على الصيغة التالية:

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

و هذا يعني أن معكوس مصفوفة التباين-التباين المشترك للأخطاء معرف كما يلي:

$$\Omega_\varepsilon^{-1} = \frac{1}{\sigma_u^2}\begin{pmatrix} 1 & -\rho & 0 & 0 & \dots & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \dots & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 1+\rho^2 & -\rho \\ 0 & \dots & \dots & \dots & 0 & -\rho & 1 \end{pmatrix}$$

من الملاحظ أن قيمتي ρ و σ_u^2 غير معروفتين وهذا يعني أنه ينبغي إيجاد مقدر لكل منهما وذلك بالبحث عن مصفوفة M حيث النموذج $MY = MX\beta + M\varepsilon$ يحقق جميع الفرضيات الأساسية.

لدينا: $E((M\varepsilon)(M\varepsilon)^{'})=E(M\varepsilon\varepsilon^{'}M^{'})=ME(\varepsilon\varepsilon^{'})M^{'}=M\Omega_{\varepsilon}M^{'}=\sigma_{\varepsilon}^{2}I$

إذن نحدد المقدر BLUE كما يلي: $\hat{\beta}=\left(X^{'}M^{'}MX\right)^{-1}X^{'}M^{'}MY$

و بالنتيجة التالية: $M^{'}M=\lambda\Omega_{\varepsilon}^{-1}=\sigma_{u}^{2}\Omega_{\varepsilon}^{-1}\quad\forall\lambda$

نحصل على المصفوفة:

$$\underset{(n-1,n)}{M}=\begin{pmatrix}-\rho & 1 & 0 & 0 & \dots & 0\\ 0 & -\rho & 1 & 0 & \dots & 0\\ 0 & 0 & -\rho & 1 & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & 0 & 0 & -\rho & 1\end{pmatrix}$$

التي تعتبر كاستنتاج للمصفوفة:

$$M^{'}M=\begin{pmatrix}\rho^{2} & -\rho & 0 & 0 & \dots & \dots & 0\\ -\rho & 1+\rho^{2} & -\rho & 0 & \dots & \dots & 0\\ 0 & -\rho & 1+\rho^{2} & -\rho & \dots & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ 0 & \dots & \dots & \dots & \dots & 1+\rho^{2} & -\rho\\ 0 & \dots & \dots & \dots & 0 & -\rho & 1\end{pmatrix}$$

و هي نفسها المصفوفة $\sigma_{u}^{2}\Omega_{\varepsilon}^{-1}$. وعليه تؤول طريقة المربعات الصغرى المعممة إلى طريقة المربعات الصغرى العادية، حيث يتم تقدير النموذج العام المصحح من الارتباط الذاتي عن طريق تحويل المتغيرات عن طريق شبه الفروقات من الدرجة الأولى، لدينا:

$$MY=\begin{pmatrix}Y_{2}-\rho Y_{1}\\ Y_{3}-\rho Y_{2}\\ \dots\\ Y_{n}-\rho Y_{n-1}\end{pmatrix}\quad \text{و} \quad MX_{j}=\begin{pmatrix}X_{2,j}-\rho X_{1,j}\\ X_{3,j}-\rho X_{2,j}\\ \dots\\ X_{n,j}-\rho X_{n-1,j}\end{pmatrix}$$

حيث: $j=1,2,\dots,k$

عند استخدام شبه الفروقات، نفقد المشاهدة الأولى لكل متغير ولتجنب ضياعها، نضع:

$$Y_1^* = Y_1\sqrt{1-\rho^2}$$
$$X_{1,j}^* = X_{1,j}\sqrt{1-\rho^2}$$

يكتب النموذج المصحح على النحو التالي:

$$Y_t - \rho Y_{t-1} = \beta_0(1-\rho) + \beta_1(X_{t1} - \rho X_{t-1,1}) + \beta_2(X_{t2} - \rho X_{t-1,2}) + ... + \beta_k(X_{tk} - \rho X_{t-1,k}) + \varepsilon_t - \rho\varepsilon_{t-1}$$

كما ذكرنا، يتم تقدير النموذج الأخير بطريقة المربعات الصغرى العادية. المشكل الأساسي الذي نواجهه يتمثل في تقدير معامل الارتباط الذاتي بين الأخطاء من الدرجة الأولى. هناك إذن عدة طرق للتقدير منها ما يلي :

1.4.2..18 تقدير ρ عن طريق إحصائية Durbin-Watson

<u>الخطوة الأولى</u>: تقدير ρ انطلاقا من إحصائية DW، حيث: $\hat{\rho} \cong 1 - DW/2$

<u>الخطوة الثانية</u>: تقدير النموذج التالي بعد إجراء التعديلات على المشاهدات بحساب شبه الفروقات:

$$Y_t - \hat{\rho} Y_{t-1} = \beta_0(1-\hat{\rho}) + \beta_1(X_{t1} - \hat{\rho} X_{t-1,1}) + \beta_2(X_{t2} - \hat{\rho} X_{t-1,2}) + ... + \beta_k(X_{tk} - \hat{\rho} X_{t-1,k}) + u_t$$

أي:
$$Y_t^* = \beta_0^* + \beta_1 X_{t1}^* + \beta_2 X_{t2}^* + ... + \beta_k X_{tk}^* + u_t$$

المعالم المقدرة بطريقة المربعات الصغرى هي: $\hat{\beta}_1, ..., \hat{\beta}_k$ و $\hat{\beta}_0^* = \hat{\beta}_0(1-\hat{\rho})$.

2.4.2.19 تقدير ρ بطريقة Theil-Nagar

اقترح Theil و Nagar تقديرا لـ ρ من خلال العلاقة التالية :

$$\hat{\rho} = \frac{n^2[1-(DW/2)] + (k+1)^2}{n^2 - (k+1)^2}$$

حيث k هي عدد المتغيرات المستقلة. نستخدم نفس الخطوات لتقدير معالم النموذج بطريقة المربعات الصغرى.

3.4.2.20 طريقة Cochrane-Orcutt

اقترح Cochrane و Orcutt تقديرا بإعطاء قيمة ابتدائية لـ ρ بواسطة القيم المقدرة لحد الخطأ.

الخطوة الأولى: إعطاء قيمة ابتدائية لمعامل الارتباط وذلك بتقنية تقدير مباشرة: $\hat{\rho}=\frac{\sum\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum\hat{\varepsilon}_{t-1}^2}$ ، فليكن:

$$\hat{\rho}_0=\hat{\rho}$$

الخطوة الثانية: تقدير النموذج التالي بطريقة المربعات الصغرى العادية:

$$Y_t-\hat{\rho}_0Y_{t-1}=\beta_0(1-\hat{\rho}_0)+\beta_1(X_{t1}-\hat{\rho}_0X_{t-1,1})+\beta_2(X_{t2}-\hat{\rho}_0X_{t-1,2})+...+\beta_k(X_{tk}-\hat{\rho}_0X_{t-1,k})+u_t$$

المعالم المقدرة هي: $\hat{\beta}_1,...,\hat{\beta}_k$ و $\hat{\beta}_0^*=\hat{\beta}_0(1-\hat{\rho})$.

الخطوة الثالثة: إعادة تقدير ρ ببواقي تقدير جديدة $\hat{\varepsilon}_t^1$ حيث:

$$\hat{\varepsilon}_t^{(1)}=Y_t-\hat{\beta}_0-\hat{\beta}_1X_{t1}-...-\hat{\beta}_kX_{tk}$$

$$\hat{\rho}_1=\frac{\sum\hat{\varepsilon}_t^{(1)}\hat{\varepsilon}_{t-1}^{(1)}}{\sum(\hat{\varepsilon}_t^{(1)})^2}$$

الخطوة الرابعة: تقدير النموذج التالي على المتغيرات ذات شبه الفروقات:

$$Y_t-\hat{\rho}_1Y_{t-1}=\beta_0(1-\hat{\rho}_1)+\beta_1(X_{t1}-\hat{\rho}_1X_{t-1,1})+\beta_2(X_{t2}-\hat{\rho}_1X_{t-1,2})+...+\beta_k(X_{tk}-\hat{\rho}_1X_{t-1,k})+u_t$$

ثم نعيد تقدير ρ مرة أخرى ببواقي تقدير جديدة $\hat{\varepsilon}_t^{(2)}$ فنحصل على تقدير لـ $\hat{\rho}_2$. نكرر العملية مرات أخرى إلى غاية سكون مقدرات النموذج (عادة نكرر العملية ثلاث أو أربع مرات).

4.4.2.21 طريقة Hildreth-Lu

الخطوة الأولى: تحديد نوع الارتباط (موجب أو سالب) بواسطة إحصائية Durbin-Watson.

الخطوة الثانية: نحدد مجالا للقيم الممكنة لمعامل الارتباط ρ . نختار قيما تنتمي إلى المجال $[0,1]$ إذا كان المعامل موجبا و قيما تنتمي إلى هذا المجال $[-1,0]$ إذا كان سالبا.على سبيل المثال يكون

لدينا $\rho=\{0.1;0.2;\ldots;0.9;1\}$ إذا اعتبرنا أن معامل الارتباط موجب فيتم اختيار درجة سلم 0.1 أو 0.01 ومع كل قيمة يتم تقدير النموذج:

$$Y_t-\hat{\rho}_iY_{t-1}=\beta_0(1-\hat{\rho}_i)+\beta_1(X_{t1}-\hat{\rho}_iX_{t-1,1})+\beta_2(X_{t2}-\hat{\rho}_iX_{t-1,2})+\ldots+\beta_k(X_{tk}-\hat{\rho}_iX_{t-1,k})+u_t$$

ونأخذ قيمة ρ القيمة التي عندها يكون مجموع مربعات البواقي $\sum_t \hat{u}_t^2$ أصغر ما يمكن.

مثال:

u^2_{t-1}	$(u_t-u_{t-1})^2$	u_t^2	u_t-u_{t-1}	$(u_t.u_{t-1})$	ut-1	
–	–	–	–	–	–	4
16	1	25	1	20	4	5
25	1	36	1	30	5	6
36	4	16	-2	24	6	4
16	1	9	-1	12	4	3
9	0	9	0	9	3	3
9	25	4	-5	-6	3	2-
4	16	4	-4	-4	2	2-
4	25	9	-5	-6	2	3-
9	64	25	-8	-15	3	5-
25	4	49	-2	35	5-	7-
49	0	49	0	49	7-	7-
49	1	64	-1	56	7-	8-
64	1	81	-1	72	8-	9-
81	1	100	-1	90	9-	10-
396	144	480	-28	366	10-	
	0.363636			0.924242		المجموع

$$d=\frac{\sum(u_t-u_{t-1})^2}{\sum u^2}=\frac{144}{396}=0{,}36$$

الإحصاء المحسوب: أي اختبار معنوية يتطلب استعمال إحصاء محسوب ويقارن بالقيمة الجد وليه وبعد ذلك يتخذ القرار بقبول أو رفض فرضية العدم.

$$d = \frac{\sum (u_t - u_{t-1})^2}{\sum u^2}$$

ويحصل على الإحصاء الجد ولي من جدول دير بن واتسون. حيث يعطي قيمة d_L وهي القيمة الدنيا.Lower . حيث يعطى قيمة d_U القيمة القصوى Upper. نحصل على القيمتين من الجداول، ولكن نحصل على القيمتين نحتاج إلى عدد المشاهدات n وعدد المتغيرات المستقلة k-(1).

مقارنه بين القيم الحسوبه والقيم الجد وليه، جداول دير بن واتسون تكون في مستوى معين يكون عند جدول 5% وآخر 1% والذي يستخدم دائما 5%. و n تبدأ من الجدول 15 وهذه نقيضه من نقائض جدول واتسون لأنه يتطلب عدد من المشاهدات لا تقل عن 15 ولكن تم تفادي هذا الآمر بتطوير جداول تبدأ من 6 مشاهدات فقط تقوم على نفس الفكرة.

N	K=1	
	d_L	d_u
15	1.08	1.36
16	1.10	1.37
17	1.13	1.38
18	1.16	1.39
19	1.18	1.40
20	1.20	1.41

إذا كانت n=20 عند مستوى الثقة 5% d_L =1.20 d_u=1.41 هذا يعني أننا سنرفض Ho إذا كانت d<1.20 ولن نرفض Ho إذا كانت d>1.41. وإذا كانت $d_L \leq d \leq d_u$ لا نستطيع إن نتخذ قرار بشأن قبول فرض العدم آو رفضه. ارفض Ho إذا كانت $d< d_u$، أو $d > 4 - d_L$ وإذا كانت $4- d_L \leq d \leq d_u$ لا نستطيع أن نتخذ قرار بشأن قبول فرض العدم آو رفضه.

القيم من 2 تؤدي إلى قبول Ho والقيم القريبة من الصفر آو قيمة 4 تؤدي إلى رفض Ho.

5.2.22.اختبار h

في حالة وجود متغير متباطئ للمتغير التابع ونتيجة لأسباب إحصاءيه لوحظ إن إحصاء دير بن واتسون d يتجه نحو قيمة 2 وإذا استندنا على إحصاء دير بن واتسون سنتوصل إلى نتيجة خاطئة ونقول انه لا يوجد مشكلة ارتباط ذاتي بحكم انه قريب من 2 . إذا كان الاختبار لوجود الارتباط الذاتي من الدرجة الأولى والذي يتخذ شكل مشروع ماركوف: $u_t = \rho u_{t-1} + v_t$

$$H_0 : \rho = 0$$
$$H_A : \rho \neq 0$$

النماذج الاقتصادية تحكمها قوه معينه تحتم ظهور المتغيرات المتباطئة كمتغيرات مفسره لآي متغير تابع.

$$Y = \gamma_0 + \gamma_1 Y_{t-1} + \gamma_2 X_t + u_t$$

الإحصاء المحسوب $h = \hat{\rho}\sqrt{\dfrac{n}{1 - nV(\hat{\gamma}_1)}}$

تحسب بالطرق المقترحة سابقا ويفضل استخدام d . n عدد المشاهدات و $V(\gamma)$ تباين مقدرة معلمة Y_{t-1}.

بأخذ إحصاء دير بن واتسون

$$d = 2(1 - \hat{\rho})$$
$$\hat{\rho} = 1 - \frac{d}{2}$$

$$h = (1 - \frac{d}{2})\sqrt{\frac{n}{1 - nV(\hat{\gamma})}}$$

القيمة الجد وليه h يمكن إيجادها من جدول التوزيع الطبيعي Z .

ملاحظات على اختبار_h :

- اختبار عينه كبيره يفضل استعماله للعينات التي اكبر من 30 مشاهده.2

- تقل قوة الاختبار عند العينات التي اقل من 30-ينهار الإحصاء إذا كانت $nV(\gamma) > 1$ آي لا يستخدم عندما تكون $nV(\gamma) > 1$ يكون $\sqrt{-}$ بالسالب أي يكون الجذر التربيعي رقم تخيلي. - بحم هذه الملاحظة الأخيرة وبحكم ا اعتبارات أخرى يطبق هذا على الارتباط الذاتي من الدرجة الأولى ولا يصلح للارتباط الذاتي من الدرجة الثا

3.مفهوم التعدد (الازدواج) الخطي Multicollinearity

إحدى فرضيات النموذج الكلاسيكي للانحدار المتعدد هي أن لمصفوفة المشاهدات عن المتغيرات المستقلة رتبة تامة k، هذه الفرضية سمحت لنا باستنتاج مقدر $\hat{\beta}$ لـ β ، خطي وغير متحيز وذي تشتت أصغر، وذلك انطلاقا من المعادلة $(X'X)\hat{\beta} = X'Y$. فإذا رفعت هذه الفرضية، فإن $(X'X)$ لن تكون ذات رتبة تامة، أي تكون أقل من رتبة(X) (أو(X')) أي أقل من k. ومع أن $(X'X)$هي مصفوفة ذات حجم ($k \times k$) بالتالي تكون مصفوفة شاذة (محددها معدوم)، ومنه فإن $(X'X)^{-1}$تكون غير موجودة وبالتالي المعادلة $(X'X)\hat{\beta} = X'Y$ لا تقبل إذن حلا وحيدا (عدد لا نهائي من الحلول). يضع النموذج الكلاسيكي للانحدار المتعدد $Y = X\beta + \varepsilon$ المتغير التابع $Y_i : i = 1....n$ في علاقة خطية مع المتغيرات المستقلة $X_{i1}, X_{i2},......X_{ik} : i = 1....n$ ، وكذلك مع الأخطاء العشوائية$\varepsilon_i : i = 1....n$ ، فإذا كانت بالإضافة إلى ذلك رتبة X أقل من أو تساوي k فإن هذا يترجم بارتباط خطي بين أعمدة المصفوفة X.

وبعبارة أخرى يشير مشكل التعدد الخطي إلى وجود ارتباط خطي بين عدد من المتغيرات المفسِّرة، ومن تم فإن هذا المشكل لا يوجد في حالة الانحدار البسيط.[5] نسمي X_j العمود رقم j لـ X حيث $X = [X_1, X_2,....X_j.....X_k]$.

- قولنا أن رتبة X أقل من k يعني أنه يوجد شعاع C حيث : $C' = [C_1, C_2,....C_j.....C_k] \neq 0$

حيث :
$$C_1X_1 + C_2X_2 + ...+ C_kX_k = 0$$

- العلاقة الأخيرة تعبر عن وجود علاقة خطية بين المتغيرات المستقلة.

1.3.23أسباب التعدد الخطي وآثاره

ينشأ التعدد الخطي من عدة أسباب منها ما يلي :

- اتجاه المتغيرات الاقتصادية معا للتغير مع مرور الزمن: فبمرور الزمن سوف تتزايد المتغيرات الاقتصادية التالية معا: الدخل، الاستهلاك، الادخار، الاستثمار، المستوى العام للأسعار والعمالة، وحيث أن هناك ارتباط بين هذه المتغيرات فإن التعدد الخطي سوف يتحقق.

- استخدام متغيرات مستقلة ذات فترة إبطاء في المعادلة المراد تقديرها: فالدخل في الفترة الزمنية الحالية يتحدد جزئيا بواسطة قيمته في الفترة الزمنية السابقة، وحيث أن هناك ارتباط بين القيم المتتالية لمتغير ما فإن التعدد الخطي سوف يتحقق.

وفي وجود التعدد الخطي فإنه سوف يترتب عنه:

- زيادة التباين والتباين المشترك للمقدرات بدرجة كبيرة دون التأثير على التنبؤات المستمدة من الانحدار[6].
- القيم المقدرة لمعاملات الانحدار سوف تكون غير محددة وغير دقيقة.
- الأخطاء المعيارية للقيم المقدرة لمعاملات الانحدار سوف تكون كبيرة جدا.

2.3.24اختبارات اكتشاف التعدد الخطي

تعتمد درجة الخطورة لأثر التعدد الخطي على درجة الارتباط الجزئي، ومعامل الارتباط الكلي (أو معامل التحديد المضاعف)، ومنه يمكن القول بأن كلا من الأخطاء المعيارية ومعاملات الارتباط الجزئية $r_{xi,xj}$، معامل التحديد المضاعف R^2، يمكنها أن تستعمل لاختبار التعدد الخطي، لكن كل معيار من هذه المعايير الثلاثة المذكورة ليس بمؤشر على وجود التعدد الخطي بمفرده، وذلك لأن القيم العالية للأخطاء المعيارية لا تظهر دائما، بسبب التعدد الخطي، وإنما يمكن أن تظهر لأسباب أخرى، كما أن الارتباطات العالية فيما بين المتغيرات المستقلة لا تؤثر بالضرورة على قيم المقدرات $\hat{\beta}_j$، ومنه ليست هذه الأخيرة بمعيار مناسب لقياس واكتشاف التعدد الخطي بمفردها، وبالمقابل يمكن لقيمة معامل التحديد المضاعف R^2 أن تكون عالية بالمقارنة مع $r_{xi,xj}$.

[6] امتثال محمد حسن، محمد علي محمد أحمد،، 2000، ص 354.

ورغم ذلك، من المحتمل أن تحتوي نتائجنا على إشارات خاطئة أو على أخطاء معيارية كبيرة، ومع كل هذا يمكن القول بأن توفيق المعايير الثلاثة، أعلاه يساعدنا على اكتشاف التعدد الخطي.

1.2.3.25طريقة التحليل الترافدي لـ Frisch

تكمن هذه الطريقة في انحدار المتغير التابع على كل متغير مستقل على حدا، ومنه نحصل على كل الانحدارات الأولية، ثم نختار الانحدار الأولي الذي يعطي النتائج الأكثر مصداقية، ثم نضيف تدريجيا متغيرات أخرى ونختبر آثارها على كل من المعالم الفردية (أخطائها المعيارية، قيمة R^2) ويكون المتغير المضاف للانحدار ذا معنوية إذا تحققت فيه الشروط التالية :

- إذا حَسَّن المتغير المستقل الجديد من R^2 بدون أن يجعل المعالم الفردية مرفوضة بطريقة خاطئة، نحتفظ بهذا المتغير ونعتبره كمتغير مستقل.
- إذا لم يُحَسِّن المتغير الجديد من العلاقة ويؤثر على قيم المعالم الفردية، نعتبره مرفوضا ونحذفه من الانحدار.
- إذا أثر المتغير الجديد بشكل واضح على إشارات وقيم المعالم المقدرة، نعتبره متغيرا مُفَسِّرا، فإذا تأثرت المعالم الفردية بالطريقة التي تصبح فيها غير مقبولة على أساس الاعتبارات النظرية المعروفة مسبقا، فإنه يمكننا القول بأن هذا مؤشر على وجود التعدد الخطي بشكل معقد، يكون هذا المتغير مُهما، لكن بسبب الارتباطات الخطية مع المتغيرات المستقلة الأخرى، يكون أثره غير مقدر وغير معروف إحصائيا بواسطة المربعات الصغرى العادية.

إن التحليل الترافدي لـ Frisch ينص على تقدير كل الانحدارات الممكنة ما بين المتغيرات الموجودة بالعلاقة المدروسة، آخذين كل متغير، بالترتيب، كمتغير تابع واعتبار كل الانحدارات الممكنة لكل متغير في بقية المتغيرات، والتي ندخلها تدريجيا في التحليل، ومن الواضح أن التحليل الترافدي يتطلب منا حسابات كثيرة، ومنه تكون المقارنات ما بين النتائج معقدة أكثر.

2.2.3.26قياس التعدد الخطي أو شرط الأعداد Condition numbers

من خلال النموذج التالي : $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i \;:\quad i = 1.........n$

يكون لدينا :

$$\begin{cases} Var(\hat{\beta}_1) = \dfrac{\sigma_\varepsilon^2}{\sum X_{i1}^2(1-R_1^2)} \\ Var(\hat{\beta}_2) = \dfrac{\sigma_\varepsilon^2}{\sum X_{i2}^2(1-R_2^2)} \\ Cov(\hat{\beta}_1,\hat{\beta}_2) = \dfrac{\sigma_\varepsilon^2 R_1^2}{\sum X_{i1}X_{i2}(1-R_1^2)} \end{cases}$$

حيث أن R_1^2 هو مربع معامل الارتباط المتعدد ما بين المتغيرين المستقلين X_{i1} و X_{i2} ، بينما R_2^2 هو ما بين $X_{i1}X_{i2}$ ، وهما في الأخير متساويان، أما عند توسيع النموذج إلى *k* متغير مستقل ($k>2$) يصبح R_j^2 على أنه مربع معامل الارتباط المتعدد ما بين المتغير المستقل X_{ij} وبقية المغيرات المستقلة الأخرى، ومنه يمكننا استنتاج قانون عام لتباين المقدرات الفردية لشعاع معالم النموذج كما يلي :

$$Var(\hat{\beta}_j) = \frac{\sigma_\varepsilon^2}{\sum X_{ij}^2(1-R_j^2)}, \quad j=1.......k$$

وتكون قيمة $Var(\hat{\beta}_j)$ كبيرة كلما كانت: σ_ε^2 كبيرة؛ $\sum X_{ij}^2$ صغيرة، R_j^2 كبيرة.

ومنه نُعرِّف مقياسا جديد يسمى "معامل تضخم التباين"(Variance Inflation Factor (V.I.F " ، ومقياسا آخر يسمى "شرط العدد Condition number"، وهما مقياسان يحددان درجة التعدد الخطي.

- ويعرف معامل تضخم التباين كما يلي :

$$V.I.F(\hat{\beta}_j) = \frac{1}{1-R_j^2}$$

- وبناء على هذا التعريف نستطيع كتابة:

$$Var(\hat{\beta}_j) = \frac{\sigma_\varepsilon^2}{\sum X_{ij}^2} \times V.I.F(\hat{\beta}_j), \quad j=1.......k$$

أي أن :

$$V.I.F(\hat{\beta}_j) = \frac{\sum X_{ij}^2}{\sigma_\varepsilon^2} \times Var(\hat{\beta}_j) \qquad j=1.......k$$

انطلاقا من الانتقادات الموجهة لمعامل الارتباط، يكون مقياس *VIF* غير كافٍ لتحديد التعدد الخطي، ومنه نذكر مقياس شرط الأعداد المذكور من طرف Welsch, 1980 ، والذي يقيس حساسية مقدرات الانحدار للتغيرات الصغيرة في التباينات، ويعرف شرط الأعداد على أنه الجذر التربيعي لأكبر قيمة مقسمة على أصغر قيمة للقيم المميزة للمصفوفة $(X'X)$ وهو على الشكل :

$$K(X) = \frac{\sqrt{\lambda_{\max}}}{\sqrt{\lambda_{\min}}}$$

فكلما كانت القيمة أعلاه أقرب إلى الواحد، كلما كان الشرط أفضل لعدم جدية التعدد الخطي، ومع هذا، فإن المقياسين المذكورين أعلاه ليسا كاملين، حيث القانون الخاص بـ *VIF* ينظر إلى الارتباطات من خلال المتغيرات المستقلة فقط، وهذا ليس بالعامل الوحيد، كما أن شرط العدد يمكن أن يتغير بإعادة تحويل المتغيرات المستقلة، والتي ليست دائمة صحيحة، ويصلح المقياسان للاستعمال عند حذف بعض المتغيرات وفرض قيود على المعالم فقط في الحالات التي يكون فيها $R_j^2 \approx 1$ ، أو لما تكون القيمة المميزة الصغيرة $\lambda_{\min}$ أقرب من الصفر. نقدر النموذج في هذه الحالة تبعا لبعض القيود المفروضة على معالمه، ويقترح Theil مقياسا أخر لقياس درجة الارتباط فيما بين المتغيرات ومنه درجة التعدد الخطي على الشكل :

$$m = R^2 - \sum_{j=1}^{k}\left(R^2 - R_{-j}^2\right)$$

حيث أن R^2 هو معامل التحديد المضاعف المعروف من قبل، أما R_{-j}^2 فهو مربع معامل الارتباط المتعدد من انحدار y (المركزة) في $x_1, x_2, \ldots\ldots x_k$ مع حذف x_j ، لكن إحدى عيوب هذه الطريقة هي أن *m* يمكن أن تكون سالبة مما يجعل التحليل أصعب، وهناك من يقترح طرقا معينة لحل مشكلة التعدد الخطي كإضافة حد ثابت لتباينات مقدرات المعالم قبل حل المعادلات الطبيعية للمربعات الصغرى.

3.2.3.27طريقة Farrar-Glauber

لاكتشاف ظاهرة التعدد الخطي يتبع Farrar-Glauber الخطوات التالية:

- أولا : حساب محدد مصفوفة معاملات الارتباط بين المتغيرات المستقلة:

$$D = \begin{vmatrix} 1 & r_{X_1X_2} & r_{X_1X_3} & \cdots & r_{X_1X_k} \\ r_{X_2X_1} & 1 & r_{X_2X_3} & \cdots & r_{X_2X_k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{X_kX_k} & r_{X_kX_2} & r_{X_kX_3} & \cdots & 1 \end{vmatrix}$$

عندما تكون قيمة المحدد تقترب من الصفر، فإن هناك دليل على وجود تعدد خطي.

- ثانيا : نستعمل اختبار χ^2 وذلك بوضع الفرضيات التالية :

$H_0 : \mathrm{D} = 1$ (استقلال خطي)

$H_1: D<1$ (ارتباط خطي)

إحصائية Farrar-Glauber (القيمة المحسوبة) تعرف كما يلي:

$$\chi^2 * = -\left[n-1-\frac{1}{6}(2k+7)\right].\ln D$$

حيث n هو حجم العينة، k هو عدد المتغيرات المفسرة في النموذج و ln هو اللوغاريتم النبيري.

فإذا كانت قيمة $\chi^2 *$ أكبر تماما من القيمة المجدولة لتوزيع χ^2 بدرجة حرية $\frac{1}{2}k(k+1)$ و نسبة معنوية α، نقبل H_1 أي هناك تعدد خطي

3.3.28الحلول المقترحة للتعدد الخطي

عند وجود التعدد الخطي، فإن الحلول تكون مُعتمِدة على إمكانية إيجاد مصادر أخرى للبيانات وعلى أهمية العوامل التي تسببت في ظهورها، ثم على الهدف الذي من أجله نقوم بتقدير الدالة تحت الدراسة، فإذا لم يؤثر التعدد الخطي بشكل فعلي على مقدرات النموذج، يقترح بعض باحثي القياس الاقتصادي إهمال وجوده في النموذج، حيث يمكن تحاشي التعدد الخطي بتوسيع حجم العينة، فمثلا يمكن تحويل البيانات السنوية إلى بيانات موسمية أو شهرية إن أمكن ذلك، كما يمكن التخلص من التعدد الخطي بإسقاط (حذف) المتغير المسبب لهذا المشكل لكن هذه العملية يمكن أن تخلق مشاكل أخرى، وهناك من يقترح إدخال معلومات إضافية للنموذج.

إن وجود التعدد الخطي يجعل من الصعب فصل آثار المتغيرات المختلفة، ومنه نحتاج إلى معلومات خاصة تساعدنا على فصل أثر كل متغير لوحده، ويكون ذلك عن طريق فرض قيود على بعض المعالم بناءا على المعلومات المسبقة للنظرية الاقتصادية.

الأخطاء غير موزعة طبيعيا

عندما تكون الأخطاء غير موزعة طبيعيًا في نموذج الانحدار، يمكن استخدام عدة أساليب لتصحيح هذه الحالة. هنا بعض الطرق الشائعة لتصحيح الأخطاء غير الموزعة طبيعياً:

1. استخدام نماذج عددية غير معلمة :(Nonparametric Models) في بعض الحالات، يمكن استخدام نماذج عددية غير معلمة التي لا تفترض توزيعًا محددًا للأخطاء. تشمل هذه النماذج تقنيات مثل الانحدار الشجري (Regression Trees) والأشجار المحسنة (Random Forests) والأجهزة العصبية .(Neural Networks) تعتبر هذه الطرق أكثر مرونة في التعامل مع الأخطاء غير الموزعة طبيعياً.

2. استخدام تحويلات :(Transformations) مثلما تم ذكره في الإجابة السابقة، يمكن استخدام تحويلات على المتغيرات أو الأخطاء لجعلها أقرب إلى التوزيع الطبيعي. يمكن استخدام التحويلات الرياضية مثل تحويل لوغاريتمي (Logarithmic transformation) أو تحويل جيبية مربعة (Square-root transformation)أو تحويل بوكس-كوكس(Box-Cox transformation) ، وغيرها. يجب مراعاة أن بعض التحويلات قد يكون لها تأثير على التفسير الاقتصادي للمتغيرات.

3. استخدام نماذج مقاومة :(Robust Models) تعتبر النماذج المقاومة وسيلة أخرى للتعامل مع الأخطاء غير الموزعة طبيعياً. تعتمد هذه النماذج على تقنيات إحصائية تقلل من تأثير القيم الشاذة أو الأخطاء غير الموزعة بشكل طبيعي. على سبيل المثال، تشمل النماذج المقاومة تقنيات الانحدار المقاوم (Robust Regression) والانحدار بالكوارتيلات.(Quantile Regression)

4. استخدام نماذج مخصصة: في حالة عدم القدرة على تحقيق التوزيع الطبيعي للأخطاء باستخدام الطرق السابقة، يمكن استخدام نماذج مخصصة تم تطويرها لمعالجة حالات محددة. على سبيل المثال، إذا كان لديك أخطاء ذات توزيع غير موزع طبيعي وذات قيم قريبة من الصفر، يمكن استخدام نموذج تحويل بوكس-تيدويل.(Box-Tidwell transformation)

يجب ملاحظة أنه قد يكون من المفيد استشارة المزيد من المراجع الأكاديمية والكتب المتخصصة في الاقتصاد القياسي وتحليل البيانات للحصول على توجيهات أكثر تفصيلاً حول تصحيح الأخطاء غير الموزعة طبيعياً في سياقات محددة.

مشكل المتغيرات المفقودة

كيفية التعامل مع مشكلة المتغير المفقود:

1. تحديد المتغير المفقود: من المهم إجراء تحليل شامل للبيانات ومراجعة الأدبيات لتحديد المتغيرات المحتملة التي قد تكون مفقودة. يمكن أن تكون هذه المتغيرات ذات صلة نظرية بالمتغير المعتمد

أو المتغيرات المستقلة المدرجة في النموذج. يجب إيلاء اهتمام خاص للمتغيرات التي قد تكون متناظرة أو ذات علاقة بالأخطاء.

2. جمع بيانات إضافية: إذا كنت تحدد متغيرًا مفقودًا ذا صلة، يمكنك النظر في جمع بيانات إضافية لتضمين هذا المتغير في النموذج. يمكن ذلك من خلال إجراء استطلاعات، أو الحصول على بيانات من مصادر ثانوية، أو استخدام تقنيات نمذجة محددة لتقدير هذا المتغير المفقود.
3. تقنيات تقدير بديلة: إذا لم يكن من الممكن جمع بيانات إضافية، يمكنك استخدام تقنيات تقدير بديلة لمعالجة المتغير المفقود. على سبيل المثال، يمكنك استخدام متغيرات أداة لتصحيح التحيز الناجم عن الفقدان. المتغيرات الأداة هي متغيرات ذات علاقة مع المتغير المفقود ولكنها غير مرتبطة بالأخطاء. هذا يسمح بتقدير المعاملات بدون تحيز.
4. تحليل الحساسية: يمكن أيضًا إجراء تحليل الحساسية لتقييم الأثر المحتمل للمتغير المفقود على نتائج النموذج. يمكنك تقدير النموذج بتوصيفات مختلفة تشمل وتستبعد المتغير المفقود. هذا يساعد في تحديد ما إذا كان الفقدان المتغير له تأثير كبير على نتائج النموذج.
5. تفسير النتائج بحذر: عند مواجهة متغير مفقود، من المهم تفسير نتائج النموذج بحذر. يجب أن تدرك قيود النموذج وأن تؤكد على عدم اليقين المتعلق بالمتغير المفقود. التواصل الواضح بشأن قيود الدراسة ضروري لتفسير النتائج بشكل سليم.

من المهم أن تأخذ في الاعتبار السياق الخاص لبحثك والاستعانة بمصادر متخصصة في الاقتصاد القياسي لاختيار النهج الأفضل للتعامل مع مشكلة المتغير المفقود في دراستك.

29. تمارين تطبيقية

التمرين الأول: اليك المعطيات التالية المتعلقة بالنموذج المتعدد التالي:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

$\sum_{t=1}^{20} X_{2t} = 25$; $\sum_{t=1}^{20} Y_t^2 = 91.25$;و $\sum_{t=1}^{20} X_{1t} = 80$; $\sum_{t=1}^{20} Y_t = 37$

$\sum_{t=1}^{20} X_{1t} Y_t = 166$; $\sum_{t=1}^{20} X_{2t} Y_t = 92$;

$\sum_{t=1}^{20} X_{1t} X_{2t} = 120$; $\sum_{t=1}^{20} X_{1t}^2 = 360$; $\sum_{t=1}^{20} X_{2t}^2 = 165$

المطلوب

-قدر معالم النموذج الخطي المتعدد

- حددي معامل التحديد المصحح ثم قيم النموذج عند مستوى معنوية5%.

- تنبا بقيمة Y_{T+1} اذا علمت انه خلال هذه الفترة تكون $X_1 = 10$ و $X_2 = 8$.

التمرين الثاني:

اذا توفرت لدينا البيانات عن المتغيرات بالانحرافات في نموذج الانحدار التاي:

$$Y = B_1 + B_2 X_2 + B_3 X_3 + \mu$$

$$\sum y^2 = 67 \ , \ \sum x_2{}^2 = 38 \ , \ \sum x_3{}^2 = 17 \ , \ \sum x_2 y = 43.5 \ ,$$
$$\sum x_3 y = -5.1 \ , \ \sum x_2 x_3 = -8.5 \ , \ \bar{X} = 3{,}1 \ , \ \bar{X} = 2.7 \ ,$$
$$Y = 3.7$$

المطلوب:

1. أوجد مقدرات المربعات الصغرى العادية لمعالم النموذج،
2. أوجد قيمة معامل التحديد،
3. اختبر الفرضيات التالية عند مستوى الدلالة 5%:

$$i)\quad H_{\circ}:\quad \beta_2 = 0$$
$$H_1:\quad \beta_2 \neq 0$$

$$ii)\quad H_{\circ}:\quad \beta_3 = 0.5$$
$$H_1:\quad \beta_3 \neq 0.5$$

$$ii)\quad H_{\circ}:\quad \beta_2 = \beta_3 = 0$$
$$H_1:\quad \beta_2, \beta_3 \neq 0$$

4. أوجد معامل التحديد المعدل $\bar{R}^2$ الخاص بهذا النموذج.
5. أوجد جدول تحليل التباين ANOVA

تمرين الثالث:

لتكن النتائج التالية للانحدار بين متغيرين:

المشاهدة	1	2	3	4	5	6	7	8	9	10
البواقي u	3.9	4.8	5.8	3.7	2.5	2.0	1.2	2.1	1.7	0.3

1. قم باختبار دارين واتسون للكشف عن الارتباط الذاتي من الدرجة الأولى:
2. فرضية العدم: الفرضية البديلة :الإحصاء المحسوب: القيم الجد وليه: النتيجة:
3. احسب مقدر معلمة الارتباط الذاتي $\hat{\rho}$ من قيمة d

المراجع

1. Bourbonnais, R., Econométrie. Manuel et exercices corrigés, Dunod, 2è édition, 1998.
2. Bressoux P., Modélisation statistique appliquées aux sciences sociales, De Boeck, 2008.
3. Confais J., Le Guen M., Premier pas en régression linéaire avec SAS◯R , Revue Modulad *n*°35, pages 220 à 363, 2006.
4. Dagnelie P., Statistique théorique et appliquées - Inférence Statistique à une et deux dimensions, vol.2, de Boeck, 2006.
5. Dodge, Y, Rousson, V., Analyse de régression appliquée, Dunod, 2è édition, 2004.
6. Giraud, R., Chaix, N., Econométrie, Presses Universitaires de France (PUF), 1989.
7. Hardy M., Regression with Dummy Variables, Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-093, Newbury Park, CA : Sage, 1993.
8. Jacquard J., Turrisi R., Interaction e ects in multiple regression, (2*nd* ed). Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-072, Thousands Oaks, CA : Sage, 2003.
9. Johnston, J., DiNardo, J., Méthodes Econométriques, Economica, 4è édition, 1999.
10. Labrousse, C., Introduction à l'économétrie. Maîtrise d'économétrie, Dunod, 1983.
11. Rakotomalala R., Analyse de corrélation - Étude des dépendances - Variables quantitatives, http://eric. univ-lyon2.fr/~ricco/cours/cours/Analyse_de_Correlation.pdf.
12. Rakotomalala R., Pratique de la régression linéaire multiple - Diagnostic et sélection de variables, http:
13. //eric.univ-lyon2.fr/~ricco/cours/cours/La_regression_dans_la_pratique.pdf.
14. Rakotomalala, R., Pratique de la régression logistique - Régression Logistique Binaire et Polytomique, http:
15. //eric.univ-lyon2.fr/~ricco/cours/cours/pratique_regression_logistique.pdf.
16. Saporta, G., Probabilités, Analyse des données et Statistique, Technip, 2ème édition, 2006.
17. Scherrer B., Biostatistique, Volume 1, Gaëtan Morin Editeur, 2007.
18. Tenenhaus, M., Statistique - Méthodes pour décrire, expliquer et prévoir, Dunod, 2007.