الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة مصطفى اسطمبولي بمعسكر
Université Mustapha STAMBOULI de Mascara
كلية العلوم الاقتصادية، العلوم التجارية وعلوم التسيير
قسم العلوم الاقتصادية
Faculté des sciences économiques, s commerciales et sciences de gestion
Département des sciences économiques

مطبوعة محاضرات بعنوان:

الاقتصاد القياسي

مقدمة لطلبة السنة الثالثة

تخصص اقتصاد كمي

من اعداد:

د/قـــــادري نـــــورية

المـــوسم الجامعي 2020-2021

الفهــــرس

الفهــــرس

المقدمـــــة

إن التطور في النظرية الإحصائية، النظرية الاقتصادية و ثورة المعلومات و ما صاحبها من توفر للبيانات بالإضافة إلى التطور في مجال الإعلام الآلي، ساعدت كلها على حدوث تطور كبير في مجال الاقتصاد القياسي.

فالاقتصاد القياسي يحتوي على فرعين هما الاقتصاد القياسي النظري و الاقتصاد القياسي التطبيقي، حيث يهتم الاقتصاد القياسي النظري بتقديم الطرق الملائمة لقياس العلاقات الاقتصادية المختلفة كما يناقش الافتراضات التي تقوم عليها هذه الطرق، و خصائصها الإحصائية بالإضافة إلى وسائل علاج هذه المشاكل.

أما الاقتصاد القياسي التطبيقي فهو يختص بتطبيق الطرق القياسية النظرية في مجالات واقعية عديدة ترتبط بالاقتصاد، التجارة و التسيير و لكن لا يمكن فصل هذين الفرعين حيث في كثير من الحالات يترتب على عملية القياس التطبيقي الوصول إلى طرق قياس جديدة تتغلب على المشاكل التي تواجه الاقتصاد القياسي النظري.

على هذا الأساس، حاولنا من خلال هذه المطبوعة الإلمام بالعلاقات و القوانين الأساسية التي تقوم عليها الطرق النظرية للاقتصاد القياسي، و ذلك من خلال أربعة فصول : عموميات حول الاقتصاد القياسي، نماذج الانحدار الهطي البسيط، نماذج الانحدار الخطي المتعدد و مشاكل القياس الاقتصادي ، حيث تم شرح العلاقات و القوانين النظرة المختلفة من خلال برهنتها الرياضية و الإحصائية بطريقة بسيطة و مختصرة.

و باعتبار أن الاقتصاد القياسي يتضمن فرع تطبيقي، حرسنا على إضافة مجموعة أمثلة و تمارين تطبيقيه مع حلولها لتمكين الطلبة من فهم هذه العلاقات النظرية و تطبيقها في أبحاثة الميدانية مستقبلا.

تعتبر هذه المطبوعة نتاج لتجربة أكاديمية في مجال تدريس لعدة سنوات لمقياس الاقتصاد القياسي، لطلبة السنة الثالثة، علوم اقتصادية، وفق البرنامج العلمي المقترح من طرف وزارة التعليم العالي و البحث العلمي و يبقى هذا العمل المتواضع تحت تصرف الطلبة و الباحثين، متقبلين بذلك كل ملاحظاتهم و انتقاداتهم التي تزيد من إثراءه.

الفصل الأول: مقدمة حول الاقتصاد القياسي

1– تعريف الاقتصاد القياسيEconometric:

هو أحد فروع علوم الاقتصاد التي تهتم بالبحوث التجريبية ويتميز باستخدامه الأساليب الكمية في تقدير العلاقات بين المتغيرات الاقتصادية معتمدا في ذلك على النظرية الاقتصادية (Economic theory).

كانت البداية الحقيقية للاقتصاد القياسي في سنة 1930 في الولايات م.أ ومعها ظهرت دورية علمية في جانفي (Econometrica)1933

كلمة اقتصاد قياسي (Econometric) مقسمة إلى اقتصاد (Economy) وقياسي (metrics) أي القياس في الاقتصاد.

هذا يعني أن الاقتصاد القياسي يحاول الاستعانة بداية بالنظرية الاقتصادية لتحديد المشكلة المراد دراستها وأهم المتغيرات الاقتصادية التي تؤثر فيها ،ومن ثم يستعين بعلوم أخرى أولها ،الاقتصاد الرياضي لتوصيف العلاقات القائمة بين المتغيرات على شكل رموز ومعادلات ،ثم أيضا بعلم الإحصاء من أجل استنباط طرق القياس لتقدير معالم العلاقات المقترحة واختبار الفروض من أجل الوصول إلى نتائج دقيقة حتى يمكن الاعتماد عليها في شرح والتنبؤ للمشكلة المدروسة.

2– أهداف الاقتصاد القياسيThe Goals of Econometric:

يقوم الاقتصاد القياسي باقتباس وتحليل العلاقات الاقتصادية وتكميمها من خلال بيانات اقتصادية واقعية من أجل الوصول إلى أهداف مسطرة ،ويمكن ذكر أهمها :

أولا: تحليل واختبار النظريات الاقتصادية :أي من خلال اختبار العلاقات الاقتصادية بين المتغيرات على أساس ما تقدمه النظرية الاقتصادية (ولا يمكن عد النظرية الاقتصادية صحيحة ومقبولة مالم تجتاز اختبارا كميا بوضوح قوة تفسير النموذج للعلاقة بين المتغيرات) من خلال اختبار الفروض النظرية.

ثانيا: رسم السياسات واتخاذ القرارات : من خلال توفيره لقيم عددية لمعلمات العلاقات الاقتصادية بين المتغيرات، التي تساعد متخذي القرار على إجراء المقارنات واتخاذ لقرار المناسب سواءا على مستوى المؤسسات أو الحكومات (حيث يوفر مثلا: أساليب مختلفة لتقدير المرونات)

ثالثا: التنبؤ والتوقع بالسلوك الذي تأخذه السياسات الاقتصادية مستقبلا: من خلال توفير القيم العددية لمعلمات المتغيرات الاقتصادية والتنبؤ بما ستكون عليه الظاهرة الاقتصادية مستقبلا، مما يمكن واضعي السياسة ومتخذي القرار لتنظيم الوضع الاقتصادي واتخاذ إجراءات التأثير في متغيرات اقتصادية معينة.

3- **أدوات الاقتصاد القياسي**: يعتمد الاقتصاد القياسي في منهجه على ثلاث علوم أساسية وأخرى ثانوية تتكامل من أجل توفير قيم عددية لمعلمات المتغيرات الاقتصادية ويمكن توضيح ذلك كما يلي(حسين بخيت، سحر فتح الله، 2019):

أولا: النظرية الاقتصادية: من خلال الفرضيات والعلاقات التي تقدمها المدارس الاقتصادية الموجودة بين المتغيرات الاقتصادية سواءا كان ذلك على المستوى الجزئي أو الكلي (مثلا زيادة في سعر سلعه معينة له علاقة عكسية مع الطلب أو كمية الطلب عليها)، ولكن تبقى هذه العلاقات المقدمة من طرف النظرية والاقتصادية مجردة أي لم تقدم أي قياس عددي لهذه العلاقات، ما لم يتم تقدير معالمها في ضوء البيانات الإحصائية الواقعية

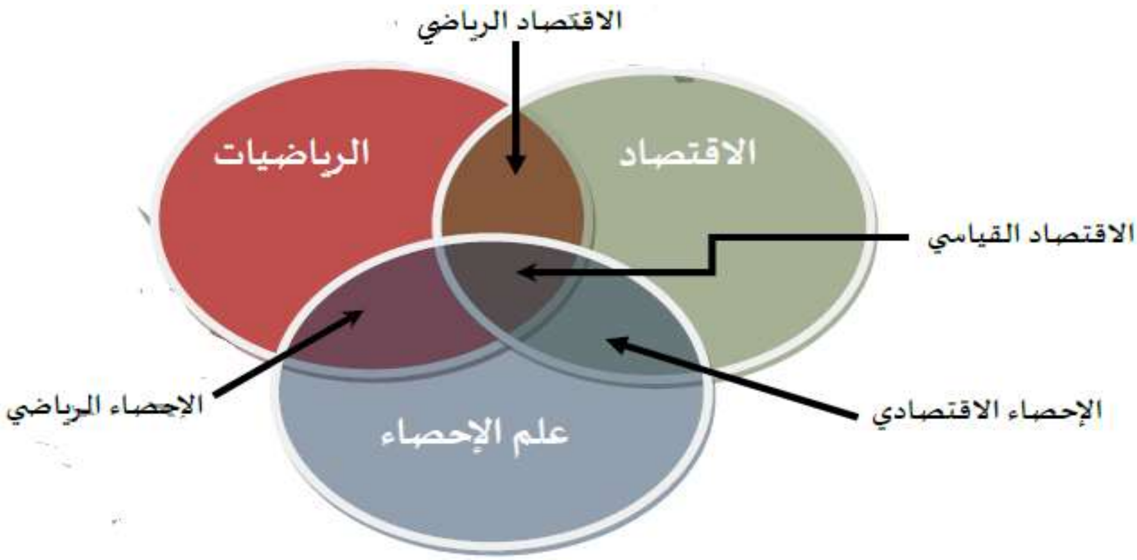
ثانيا: الاقتصاد الرياضي (الرياضيات): يهتم بإعادة صياغة العلاقات التي تم تحديدها بالاعتماد على النظرية الاقتصادية بالشكل الرياضي، أي على هيئة معادلات ورموز رياضية بدون قياسها.

ثالثا: الإحصاء الاقتصادي (علم الإحصاء): دوره هو تجميع البيانات الإحصائية الخاصة بالتغيرات الاقتصادية (طرق المعاينة)، تم تسجيلها ورسمها وجدولتها، أما تحليل واختبار العلاقات ومدى مطابقتها أو مطابقة النتائج مع النظرية الاقتصادية فهو مهمة الاقتصاد القياسي.

رابعا: الإحصاء الرياضي: يهتم بإعطاء الباحث أدوات التحليل من أجل دراسة العلاقات بين المتغيرات الاقتصادية وطرق معالجة أخطاء التقدير لاستخدامها في أهداف القياس الاقتصادي

من خلال ما سبق يمكن القول أن علم القياس الاقتصادي هو نقطة التقاء ثلاث علوم أساسية وهي الرياضيات الاقتصاد والإحصاء ويمكن توضيحه في الشكل الموالي:

شكل1-1: الاقتصاد القياسي وعلاقته بالعلوم الأخرى



و يمكن توضيح ما يلي:

تعطينا النظرية الاقتصادية شكل العلاقة بين المتغيرات الاقتصادية بطريقة مجردة، ثم يأتي دور الاقتصاد القياسي من اجل تحديد المقدار الكمي لتلك العلاقة.

اعتمادا على الاقتصاد الرياضي الذي يحاول تصوير أو توصيف العلاقة بشكل معادلة رياضية وترميز للمتغيرات ثم دور الإحصاء الاقتصادي الذي يمد القياس الاقتصادي بالمادة الأولية اللازمة للتحليل في صورة بيانات مجمعة و مبوبة، ودور الإحصاء الرياضي في تحليل هذه العلاقات ومعالجة أخطاء التقدير.

4. النموذج الرياضي الاقتصادي Economic Model:

و هو تمثيل بسيط خالي من التعقيدات من خلال العلاقات بين المتغيرات الاقتصادية لظاهرة معينة واقعية، بهدف تحليلها أو التنبؤ بها وأيضا السيطرة عليها وقد يتكون من معادلة واحدة منفردة أو من مجموعة معادلات الآنية (معادلة الطلب أو العرض معادلة منفردة)

نموذج السوق: مجموعة معادلات آتية

الهدف من النموذج الاقتصادي هو تقدير قيم عددية لمعلمات العلاقة بين متغيرات اقتصادية بغية تحليل هذه العلاقة والتنبؤ بها.

ويستخدم النموذج الاقتصادي الرموز الرياضية فمثال على ذلك.

نفترض النظرية الاقتصادية بأن الاستهلاك هو C وترمز لدالة الدخل بY أي أن، أي استهلاك هو دالة تابعة للدخل C :f(y) حيث يمثل C المتغير التابع أو المفسّر أو الاستجابة Dependent variable ،أما Y فهو المتغير المستقل أو المفسِر (Independent variable)

وباعتبار العلاقة بين المتغيرين في شكلها المبسط وهو العلاقة الخطية يصبح لدينا: $C = \beta_0 + \beta_1 Y$

حيث تمثل β_0 , β_1 معلمات النموذج (Coefficients)

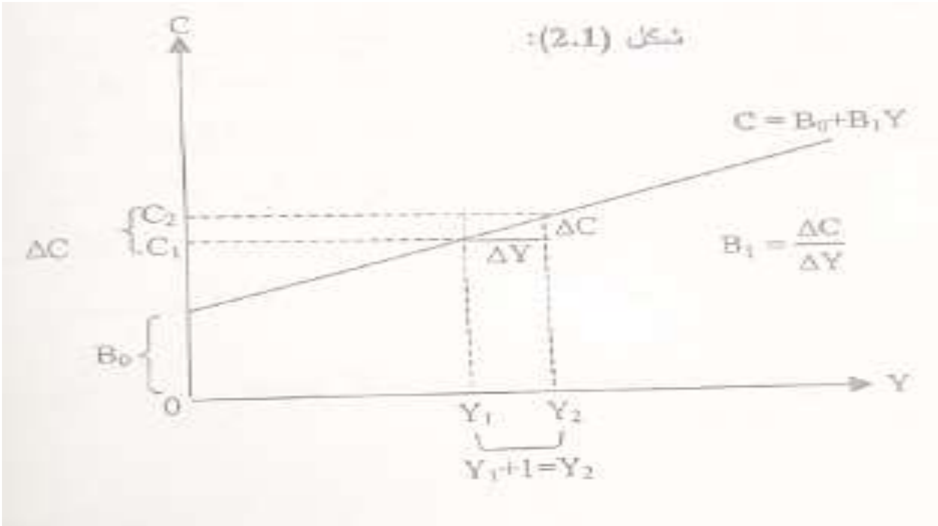
حيث β_0 هي قيمة C عندما يكون الدخل Y صفر أي :

$$Y = 0 \Longrightarrow C = \beta_0$$

ويمكن التعبير عنه $C = \beta_0 = C_0$ وهو الاستهلاك الذاتي

β_1 هي ميل المستقيم $\beta 1 = \frac{\Delta C}{\Delta Y} = b$ أي عندما يتغير الدخل بوحدة واحدة، يتغير الاستهلاك بقيمة $\beta_1 = b$ ميل الحدي لاستهلاك

ويمكن توضيحه أيضا بيانيا كما يلي:



1-4 ويجب توفر عدة خصائص في النموذج الاقتصادي أهمها:

- يجب أن يصف الظاهرة الاقتصادية بما يتطابق مع النظرية الاقتصادية
- أن يحاكي واقع الاقتصادي أي يتناسق مع السلوك الفعلي للمتغيرات الاقتصادية التي تحدد العلاقة بين هذه المتغيرات.
- الدقة في تقدير المعلمات أي أن تكون المقدرات أقرب مما يمكن للمعلمات الحقيقية (خاصية عدم التحيز وهذا ما تتطرق له فيما بعد)
- قدرة النموذج على التنبؤ.
- أن يكون النموذج بسيط وغير معقد

2-4 كما يمكن تصنيف النماذج كالتالي:

أولا: النماذج الاقتصادية الكلية والجزئية:

أ- النماذج الاقتصادية الكلية: تخص متغيرات الاقتصاد الكلي كالدخل الوطني أو الاستثمار العام...الخ

ب- النماذج الاقتصادية الجزئية: تخص المتغيرات الاقتصادية للوحدات الاقتصادية للوحدات الاقتصادية الجزئية (الاقتصاد الجزئي) : كنماذج العرض والطلب على سلعة معينة

ثانيا: النماذج الاقتصادية الساكنة والمتحركة:

أ- النماذج الاقتصادية الساكنة : وهي النماذج التي لا يكون الزمن أحد متغيرها أو ذو تأثير في قيم أحد متغيراتها أي لا وجود فترات ارتداد زمني.

ب- النماذج الاقتصادية المتحركة: وهي النماذج التي يكون فيها الزمن أحد متغيراتها أو مؤثرا في أحد متغيراتها وتوضح هذه النماذج كيفية تأثير الزمن على المتغيرات الاقتصادية وتعتبر أكثر واقعية

فمثلا: دالة العرض المتحركة $S_t = f(f_{t-1})$، حيث توضح أن العرض الحالي يعتمد على عر السلعة في الفترة السابقة ويسمى المتغير الحركي بالمتغير المرتد زمنيا.

3-4 يتكون النموذج الاقتصادي من ثلث أنواع من المعادلات والمتغيرات:

أولا: معادلات النموذج: تعتبر معادلات النموذج هيكل النموذج وهي تختلف على حسب نوع وهدف النموذج وهي تنقسم إلى :

أ- المعادلات السلوكية: تعبر على العلاقات الدالية بين المتغيرات الاقتصادية أي متغير تابع المتغير مستقل أو عدة متغيرات مستقلة $C = C_0 + ßY_d$

ب- المعادلات التعريفية: هي معادلات تعبر عن علاقات اقتصادية ناتجة عن تعاريف مثقف عليها أو تعريف متغير تابع في صورة علاقة مساواة $Y = C + I$

ج- المعادلات التوازنية: وهي تنتج عن علاقة تضاد بين المتغيرات الاقتصادية : العرض = الطلب $D = S$

ثانيا: متغيرات النموذج: يدخل في تكوين المعادلات الهيكلية للنموذج عدد من المتغيرات يمكن تصنيفها إلى عدة أنواع كما يلي :

أ- المتغيرات الداخلية (التابعة): هي المتغيرات التي تؤثر وتتأثر في النموذج، و تحدد قيمتها في النموذج عن طريق معاملات و قيم المتغيرات المستقلة.

ب- المتغيرات الخارجية (المستقلة): و هي المتغيرات التي تئثر في النموذج، ولا تتأثر به ،وتحدد قيمها بعوامل خارجية عن النموذج.

ت- المتغيرات المرتدة زمنيا: وهي القيم التي تؤخد قيمها من الفترات الزمنية السابقة.

5- منهجية الاقتصاد القياسي Methodology of Econometrics:

حتى يتم تقدير معاملات النموذج المستخدم في القياسي الاقتصادي من أجل تحليل العلاقة بين المتغيرات الاقتصادية والتنبؤ بها مستقبلا، يجب إتباع منهجية معينة والتي تتم حسب الخطوات التالية:

أولا: مرحلة صياغة النموذج: يتم في هذه المرحلة المهمة جدا تحديد المتغيرات التي يجب أن يشمل عليها النموذج ويتم في هذه المرحلة الاعتماد على النظرية الاقتصادية أي استنباط العلاقة بين المتغيرات بالشكل المجرد، ثم تدخل الاقتصاد الرياضي لتحويل هذه العلاقة إلى معادلات رياضية باستخدام الترميز، لتحديد نوع واتجاه العلاقة بين المتغيرات الاقتصادية تم صياغة هذه العلاقة في شكل معادلة رياضية.

مثلا: معادلة الطلب على سلعة ما X

$D_X = ß_0 + ß_1P_X + ß_2Y_d$

تم صياغة المعادلة الرياضية ومن النظرية الاقتصادية يتم توقع أن$ß_1$ هو سالب لكون وجود علاقة عكسية بين سعر السلعة X وكمية الطلب عليها، و$ß_2$ موجب لكون العلاقة طردية بين الكمية المطلوبة ودخل المستهلك.

وفي هذه الخطوة يتم جمع البيانات الخاصة بمتغيرات النموذج باستخدام طبعا الإحصاء الاقتصادي .

ثانيا: مرحلة التقدير والاختبارEstimation stage and Testing:

يتم تقدير معالم العلاقة التي تم وصفها وصياغتها رياضيا أي تقدير قيم معالم النموذج مثلا في المثال السابق تقدير $ß_0, ß_1 , ß_2$ في دالة الطلب، يصبح لدينا مقدرات النموذج $\hat{\beta}_2$، $\hat{\beta}_1$، $\hat{\beta}_0$ ، كما يتم في هذه المرحلة تقييم هذه المعالم المقدرة من النواحي القياسية (معنوية هذه المقدرات والاحصائية والاقتصادية).

ومحاولة إجراء المقارنة بين قيم وإشارات هذه المقدرات مع القيم والإشارات المتوقعة لهذه لها، على حسب ما تنص عليه النظرية الاقتصادية، بالإضافة إلى اختبار معنوية هذه المقدرات باستخدام اختبار ستودنت و تحليل جودة النموذج من الناحية الإحصائية

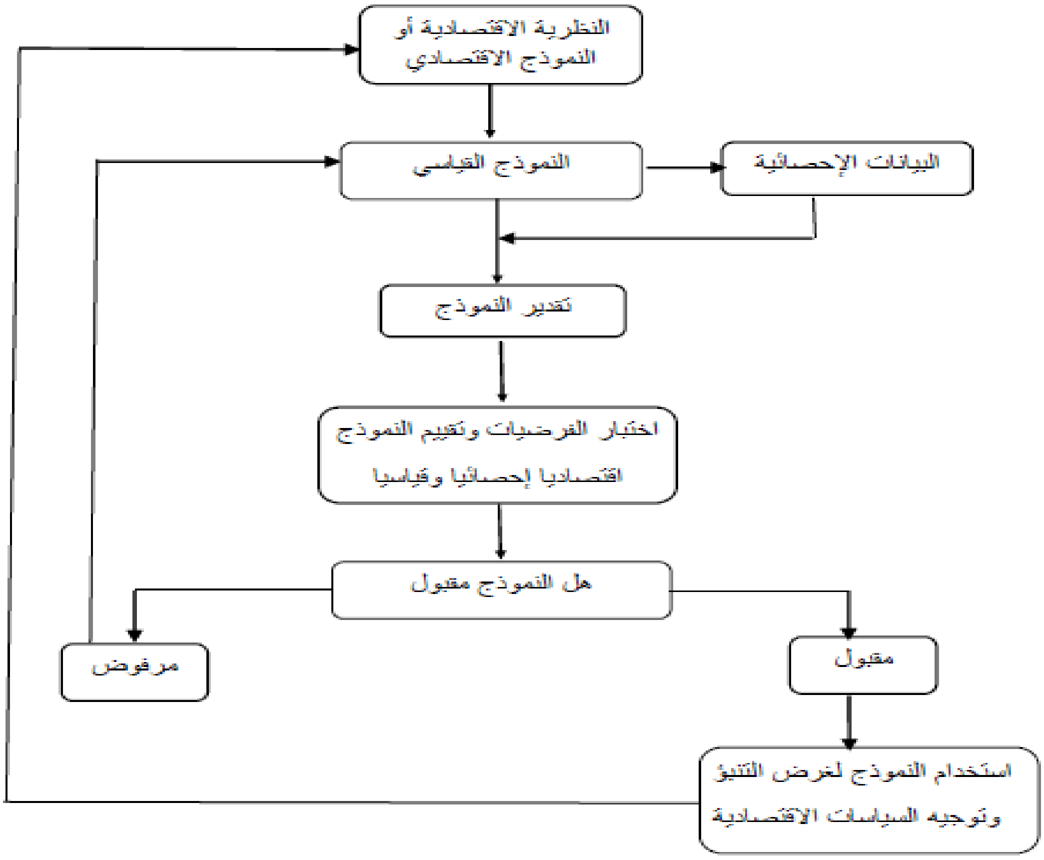
أما من الناحية القياسية التأكد من تحقق الفروض الخاصة بالنموذج القياسي أي وجود مشاكل القياس الاقتصادي مثل مشكلة الارتباط الذاتي، التعدد الخطي وعدم ثبات التباين تم معالجة هذه المشاكل.

كل هذا سنتطرق له في المحاضرات القادمة إن شاء الله

ثالثا: مرحلة التنبؤ Prediction Stage:

يتم في هذه المرحلة وبعد التأكد من صلاحية النموذج وجودته استخدامه في عملية التنبؤ بالمتغيرات الاقتصادية والعلاقة بينهم مستقبلا سواءا على المستوى القصير أو على المدى الطويل.

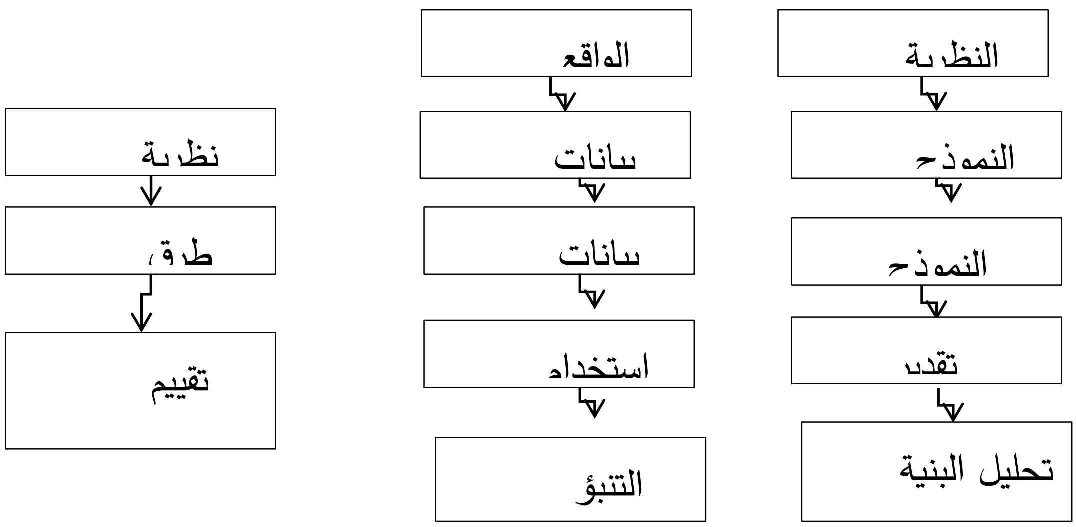
ويمكن تلخيص منهجية البحث في الاقتصاد القياسي كما يلي :



ويمكن تلخيص أيضا أسلوب الاقتصاد القياسي كما يلي:

حيث يعتمد القياس الاقتصادي على النظرية الاقتصادية التي تزودنا بالنماذج الاقتصادية التي يستخرج منها النماذج القياسية ،ويتم بناء هذه النماذج اعتمادا على البيانات الواقعية التي يتم تهيئتها للاستخدام.

باستخدام النظريات الاحصائية، يتم تقدير هذه النماذج وتقييمها وتصحيحها لاستخدامها في تحليل البنية الاقتصادية أو التنبؤ أو رسم السياسات الاقتصادية.



6-أنواع البيانات :وقف نجاح التحليل الاقتصادي على ملائمة واتاحة البيانات اللازمة .ولذلك من الأهمية تناول طبيعة وأنواع البيانات،يمكن تقسيم البيانات إلى ثلاثة أنواع هي:

1-6-بيانات السلاسل الزمنية Time series data:هي مجموعة من المشاهدات التي يأخذها المتغير في أوقات مختلفة، مثل البيانات اليومية والأسبوعية والشهرية...، تحتوي السلسلة الزمنية على عدد من القياسات لمتغير ما عند نقاط زمنية مختلفة، وهي تصف بذلك سلوك المتغير الإقتصادي عبر الزمن أي من خلال معطيات تاريخية للمتغير الاقتصادي يمكننا التنبؤ أو الاستشراف لسلوكه في المستقبل.

2-6 بيانات القطع العرضي أو العينات Cross-sectional data: توضح البيانات القطاعية القياسات التي يأخذها متغير ما بالنسبة لمفردات عينة ما عند نقطة زمنية معينة، مثال ذلك بيانات خاصة بدخول عينة من المستهلكين عند نقطة زمنية معينة .وتوضح البيانات القطاعية بذلك مدى تغير قيمة متغير ما من مفردة لأخرى عند نفس النقطة من الزمن.

3-6 البيانات المزدوجة:أو الهجينة Panel data: وهي تحتوي على مزيج من بيانات السلسلة الزمنية والبيانات المقطعية أي هي تدمج كلا النوعين السابقين .فهي تعطي بيانات عن مجموعة من المفردات عبر سلسلة زمنية، الامر الذي يساعد على تشكيل أكبر قدر ممكن من المعلومات التاريخية مع الاخد بعين الاعتبار الاختلاف بين مفردات العينة مما يعطي صورة اكثر تمثيلا لواقع الظاهرة موضوع البحث.

7-نظرية الارتباط

1-7 شكل الإنتشار:

غالبا ما تتاثر الظواهر الاقتصادية فيما بينها، نجد من بين هذه علاقات او التاثيرات علاقة الارتباط، من خلال معامل الإرتباط الذي يقيس درجة الإرتباط الموجودة بين ظاهرتين على الأقل، ولذلك يميز ما بين الارتباط البسيط والمتعدد وكذلك الخطي والغير خطي.

فإذا افترضنا أنه لدينا متغيرين فقد يكون لهما إحدى الحالات التالية:

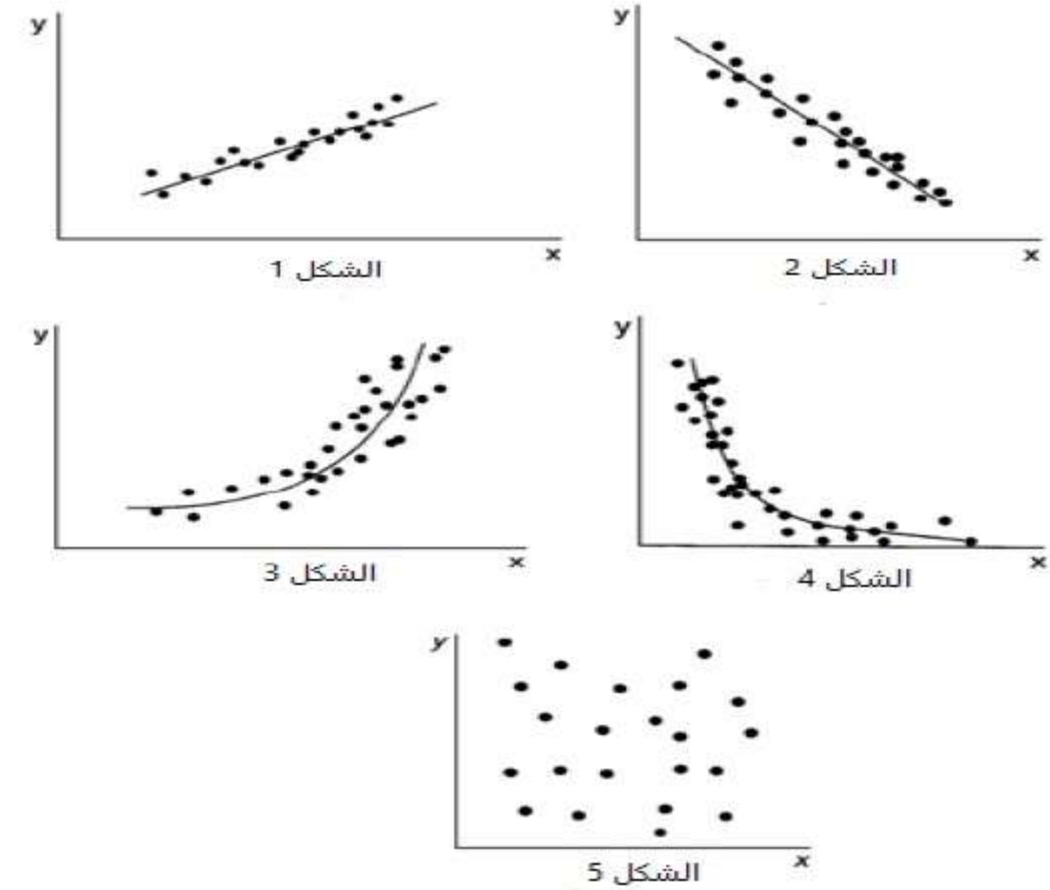
-ارتباط موجب.

-ارتباط سالب.

-عدم وجود ارتباط بينهما.

الأشكال التالية توضح الخطية والإرتباطية بين ظاهرتين

الشكل(2)- توضح الخطية والإرتباطية بين ظاهرتين



Régis Bourbonnais, Econométrie, 10e édition, Dunod, 2018 , p 18

حيث يمكن توضيح الاشكال التالية في الجدول الموالي كما يلي:

جدول رقم 1

لا وجود للارتباط	ارتباط موجب	ارتباط سالب	
الشكل 5	شكل 1	شكل 2	العلاقة خطية
الشكل 5	شكل 3	شكل4	العلاقة غير خطية

المصدر: من اعداد الباحثة

2-7 قياس معامل الارتباط الخطي البسيط:

لكي نحصل على قياس ذي معنى لقوة الاقتران بين متغيرين، فنحن في حاجة إلى معلمة تكون مستقلة القيمة عن وحدات القياس و يمكن توضيحها بمعامل الإرتباط :

$$r_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

حيث:

cov(x,y): هو التباين المشترك بين x و y.

σ_x و σ_y: هما الإنحرافان المعياريان لكل من x و y.

كما يمكن إستخدام العلاقة التالية لقياس معامل الارتباط الخطي البسيط:

$$r_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

حيث $\bar{y}$ و $\bar{x}$، هما متوسط كل من y، x، على الترتيب.

و يمكن التعبير عن العلاقة السابقة كما يلي:

$$\boxed{r_{x,y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}}$$

حيث أن:

$$-1 \leq r_{x,y} \leq +1$$

وباختصار، فإن معامل الارتباط يوضح كلا من إشارة العلاقة الخطية وقوتها بٌن متغيرين .والقيم الموجبة والسالبة ل $r_{x,y}$ توضح العلاقات الطردية والعكسية على التوالي، وكلما إقترب $r_{x,y}$ من زائد واحد أو ناقص واحد كامل زادت قوة العلاقة الخطية، أي كلما زادت قوة الإرتباط بين التغيرين.

7-3 اختبار معنوية معامل الارتباط الخطي البسيط:

-تشكيل الفرضيات:

$$\begin{cases} H_0 : r_{x,y} = 0 & (\text{لا يوجد إرتباط}) \\ H_1 : r_{x,y} \neq 0 & (\text{يوجد إرتباط}) \end{cases}$$

لدينا:

$$t^* = \frac{\gamma_{x,y} - r_{x,y}}{\sqrt{\frac{1 - \gamma_{x,y}^2}{n-2}}} \rightarrow T(n-2)$$

تحت شرط الفرضية الصفرية لدينا:

$$t^* = \frac{\gamma_{x,y} - 0}{\sqrt{\frac{1 - \gamma_{x,y}^2}{n-2}}} \rightarrow T(n-2)$$

- اتخاذ القرار:

إذا كان ستودنت المحسوب أكبر من ستودنت الجدولية عند مستوى ثقة $\alpha/2$ودرجة حرية (n-2)

نرفض الفرضية الصفرية وبالتالي يوجد إرتباط معنوي بنٌ المتغيرين x, y وذلك كالتالي:

$$t^{*}_{calculer} > t^{\alpha/2}_{n-2} \rightarrow on\ regette\ H_0$$

إذا كان حجم العينة n>30 فإنه يمكن استخدام القانون الطبيعي$\left(z\alpha_{/2}\right)$

مثال:

مهندس فلاحي يهتم بدراسة الارتباط بين مردودية الذكتار(x) متغنً مستقل وكمية السماد(y) ، أخذت عينة موضحة بالجدول الآتي:

تطور مردودية الهكتار وكمية السماد

مردودية الهكتار(X)	16	18	23	24	28	29	26	31	32	34
كمية السماد (Y)	20	24	28	22	32	28	32	36	41	41

1-أرسم سحابة النقاط وعلق عليها ؟

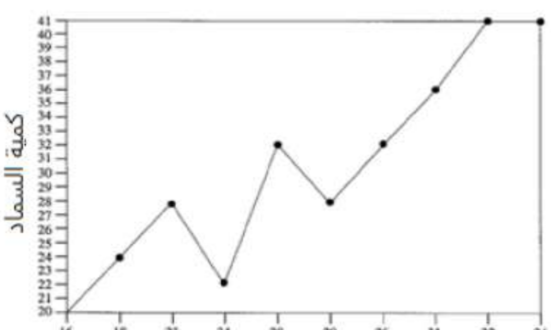
2-أحسب معامل الارتباط البسيط واختبر معنويتو مع العلم أن$\alpha = 5\%$

الحل:

1-رسم سحابة النقاط والتعليق على شكل الانتشار:

رسم شكل الانتشار من بيانات الجدول، يبين لنا أن سحابة النقاط ذات انتشار خطي موجب أي هناك ارتباط موجب مابين مردودية الهكتار وكمية السماد وبالتالي هناك ارتباط موجب ما بين المتغيرتين.

سحابة نقاط قيم مردودية الهكتار وكمية السماد



Régis Bourbonnais, Econométrie, 10e édition, Dunod, 2018, p 10

2-حساب معامل الارتباط البسيط وإختبار معنويته مع العلم أن$\alpha = 5\%$

2-1-حساب معامل الارتباط الخطي البسيط:

نحسب بالعلاقة التالية:

$$r_{x,y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

نقوم إجراء الحسابات المبينة في الجدول التالي:

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
16	20	256	400	320
18	24	324	576	432
23	28	529	784	644
24	22	576	484	528
28	32	784	1024	896
29	28	841	784	812
26	32	676	1024	832
31	36	961	1296	1116
32	41	1024	1681	1312
34	41	1156	1681	1394
$\sum x_i = 261$	$\sum y_i = 304$	$\sum x_i^2 = 7127$	$\sum y_i^2 = 9734$	$\sum x_i y_i = 8286$

$$\gamma_{x,y} = \frac{10(8286) - (261)(304)}{\sqrt{10(7127) - (261)^2}\sqrt{10(9734) - (304)^2}}$$

$$\gamma_{x,y} = 0.89 \rightarrow \gamma_{x,y}^2 = 0.79$$

2–2اختبار معنوية معامل الارتباط الخطي البسيط:

-تشكيل الفرضيات:

$$\begin{cases} H_0 : r_{x,y} = 0 & \text{(لا يوجد إرتباط)} \\ H_1 : r_{x,y} \neq 0 & \text{(يوجد إرتباط)} \end{cases}$$

لدينا:

$$t^* = \frac{\gamma_{x,y} - r_{x,y}}{\sqrt{\dfrac{1 - \gamma_{x,y}^2}{n-2}}} \rightarrow T(n-2)$$

تحت شرط الفرضية الصفرية لدينا:

$$t^* = \frac{\gamma_{x,y} - 0}{\sqrt{\dfrac{1 - \gamma_{x,y}^2}{n-2}}} \rightarrow T(n-2)$$

وبعد الحساب نجد قيمة ستودنت الاحصائي مساوية ل5.49

بالاستعانة بجدول ستودنت نجد أن قيمة ستودنت الجدولية عند مستوى ثقة2.5%و درجة حرية 8 تساوي2.316

قاعدة القرار:

بما أن ستودنت الإحصائي أكبر من ستودنت الجدولية، اذن نرفض الفرضية الصفرية وبالتالي يوجد إرتباط معنوي و موجب بين المتغيرين xوy

الفصل الثاني: الانحدار الخطي البسيط Simple Linear Regression

ي

يعتبر الانحدار الخطي أبسط أنواع نماذج الانحدار حيث يوجد العديد من العلاقات الاقتصادية التي يمكن قياسها باستخدام هذا الأسلوب، بالإضافة إلى أن دراسة العلاقة بين متغيرين أحدهما تابع والآخر مستقل باستخدام الانحدار الخطي تعتبر من أهم وأبسط أنواع العلاقات في التقدير والتحليل الإحصائي والاقتصادي، واذا رمزنا إلى المتغيرين ب Y (المتغير التابع)، X (المتغير المستقل) يمكن كتابة العلاقة كما يلي : $Y = F(x)$

أي y دالة تابعة ل x ويعتمد عليه، وكل قيمة من x لها قيمة مناظرة لها من قيم y

ولمعرفة نوعية العلاقة بين x وy إن كانت خطية أم غير خطية يجب الاستعانة بالنظرية الاقتصادية، ثم صياغتها. وأبسط هذه الصيغ هي الخطبة البسيطة

يبدأ الانحدار الخطي البسيط برسم لوحة الانتشار ثم محاولة تحديد إذا كان هناك احتمالية علاقة خطية تقريبية. وصيغة العلاقة الرياضية كما يلي :

$$Y = \beta_0 + \beta_1 X_i$$

حيث β_0،β_1 هما معلمات النموذج مجهولة القيم ويمكن شرحها رياضيا كما يلي:

β_0: تمثيل ثابت وهي تقاطع خط الانحدار مع محور التراتيب (Y)أما قيمتها فهي Y عندما تكون قيمة X مساوية للصفر

β_1: تمثل ميل خط الانحدار $\beta_1 = \frac{\Delta Y}{\Delta x}$ أو معامل الانحدار .

وبما أنه غير متوقع أن تقع النقاط تماما على خط الانحدار أي لا يمكن أن تشرح العلاقة بين المتغيرين x وy بشكل دقيق مما يجعل هذه المعادلة لا تعبر عن العلاقة بين المتغيرتين تعبيرا كاملا، مما يُكون انحراف بين العلاقة الحقيقية (الواقعية) والمعادلة الاحصائية، التي تنتج عنها أخطاء في القياس أو اختيار المتغير المستقل مما يتطلب تعديل ذلك بإضافة متغير آخر يسمى الحد العشوائي أو حد الخطأ العشوائي ونرمز له ب (ε) لامتصاص هذه الأخطاء نذكر منها:

- اهمال المتغيرات المستقلة التي تدخل في تقسيم المتغير التابع وعدم إدراجها وأخذها بعين الاعتبار.
- الصيغة الرياضية الغير سليمة للنموذج.
- حدوث خطأ في جمع البيانات وقياس المتغيرات الاقتصادية .

وتصبح صيغة المعادلة بعد اضافة أقل حد عشوائي كما يلي

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

ي

1- فرضيات النموذج: (الفرضيات الخاصة بالمتغير العشوائي ε_i

1-1: التوقع الرياضي لحد الخطأ العشوائي يساوي الصفر: أي متوسط القيم التي يأخذها ε_i تساوي الصفر

$$E(\varepsilon_i) = 0$$

يعني ذلك أن الأخطاء لا تدخل في تفسير المتغير التابع Y_i ، ويمكن توضيح ذلك كما يلي:

لدينا:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\Longrightarrow \quad \varepsilon_i = Y_i - \beta_0 - \beta_1 X_i$$

بإدخال Σل طرفي المعادلة تصبح لدينا

$$\sum \varepsilon_i = \sum (Y_i - \beta_0 - \beta_1 X_i)$$

$$\Rightarrow \sum \varepsilon_i = \sum Y_i - n\beta_0 - \beta_1 \sum X_i$$

نعلم أن :

$$\sum Y_i = n\bar{Y} \quad \Lambda \quad \sum X_i = n\bar{X} \quad \Lambda \ \beta_0 = \bar{Y} - \beta_1 \bar{X}$$

يصبح لدينا:

$$\sum \varepsilon_i = \sum Y_i - n(\bar{Y} - \beta_1 \bar{X}) - \beta_1 n\bar{X}$$
$$\sum \varepsilon_i = \sum Y_i - \sum Y_i + \beta_1 n\bar{X} - \beta_1 n\bar{X}$$

$$\Rightarrow \sum \varepsilon_i = 0 \quad \Rightarrow \quad E(\varepsilon_i) = 0$$

2-1 المتغيرالعشوائي (ε_i) يتوزع توزيعا طبيعيا: بمتوسط حسابي يساوي صفر عند كل قيمة من قيم المتغيرX وتباين ثابت. أي شكل جرس

$$\varepsilon_i \sim N(0, \sigma^2_{\varepsilon_i})$$

3-1: تجانس أو تباث تباين حد الخطأ Homoscedastisity: أي تشتت المتغير العشوائي (ε_i) حول الوسط الحسابي هو مقدار ثابت عند كل قيمة بين قيم X ويعبر عنها:

$$\mathrm{Var}(\varepsilon_i) = E(\varepsilon_i - E(\varepsilon_i))^2 \quad \Lambda \quad E(\varepsilon_i) = 0$$
$$\Rightarrow \mathrm{Var}(\varepsilon_i) = E(\varepsilon_i)^2 = \sigma^2_{\varepsilon_i}$$

ي

4-1: القيم المختلفة لحد الخطأ العشوائي (ε_i) تكون مستقلة عن بعضها البعض: أي التباينات المشتركة لأخطاء للمشاهدات المختلفة تكون معدومة أي:

$$\text{Cov}(\varepsilon_i, \varepsilon_j)=\text{E}(\varepsilon_i \varepsilon_j) = 0 \quad \forall i, j = 1{:}n \ \Lambda \ i \neq j$$

أي عدم وجود لارتباط ذاتي بين الأخطاء Autocorrelation

5-1: قيم (ε_i) غير مرتبطة بأي قيمة لمتغيرات المستقلة : أي:

$$\begin{aligned}\text{Cov}(\varepsilon_i, X_i) &= E(\varepsilon_i X_I) \\ &= X_I E(\varepsilon_i) \qquad \Lambda \ E(\varepsilon_i) = 0 \\ &= 0\end{aligned}$$

6-1: و هي فرضية تخص الانحدار الخطي المتعدد: عدم **وجود ارتباط أو علاقة بين المتغيرات المستقلة** وهذا سنشاهده فيما بعد.

2- **تقدير معالم النموذج**: للحصول على مقدرات النموذج هناك أهم طريقتين:

- طريقة المربعات الصغرى.
- طريقة معقولية العظمى

وفي محاضرتنا هذه سنتطرق فقط لطريقة المربعات الصغرى، سنرمز للقيمة المقدرة ل Y_{i} (الفعلية)، ب $\widehat{Y_1}$

1-2: طريقة المربعات الصغرىThe Ordinary Least Squares:

لمعرفة العلاقة بين متغيرين اقتصاديين X و Y في مجتمع ما، نكتفي بسحب عينة مكونة من n ثنائية (X_i, Y_i) يكون البحث عن خط الانحدار الذي يعبر بشكل دقيق عن العلاقة:

$$Y_{\text{i}} = \beta_0 + \beta_1 \text{X}_{\text{i}} + \varepsilon_{\text{i}}$$

لذلك نقوم بتقدير المعلمينβ_0 و β_1 اللذان يجعلان هذا الخط أفضل خط يمثل هذه البيانات، أي اتخاذ أحسن تصحيح خطي بتدنية مربعات الانحراف بين المشاهدات الفعلية والمقدرة $\sum_{i=1}^{\text{n}} \hat{\varepsilon}_{\text{i}}{}^2$ حيث لدينا:

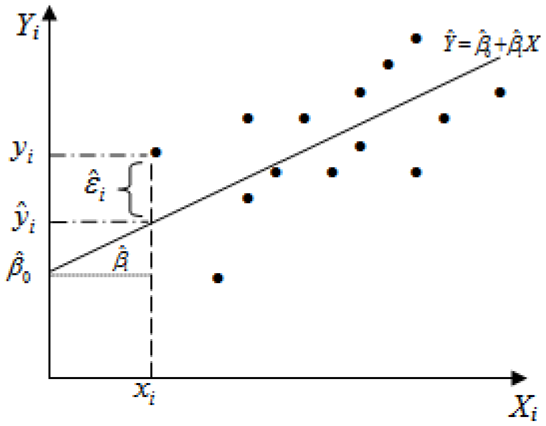
$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i$$

$$\hat{\text{Y}}_i = \hat{\beta}_0 + \hat{\beta}_1 \text{X}_{\text{i}} \quad \text{مع}$$

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

ويمكن توضيح ذلك من الشكل الموالي

ي



ويمكن صياغة ذلك رياضيا :

$$\mathrm{Min}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = Min\sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)^2$$

والشرط الرياضي لتدنية $\sum\hat{\varepsilon}_i^2$ هو أن تكون المشتقات الجزئية لهذه العبارة بالنسبة ل $\hat{\beta}_0$ و $\hat{\beta}_1$ مساوية للصفر أي:

$$\begin{cases}\dfrac{\delta\sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)^2}{\delta\hat{\beta}_0}=0\ \ldots\ldots(1)\\ \dfrac{\delta\sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)^2}{\delta\hat{\beta}_1}=0\ \ldots\ldots\ldots(2)\end{cases}$$

$$(1)\Leftrightarrow 2\sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)(-1)=0$$
$$\Leftrightarrow\sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)=0$$
$$\Leftrightarrow\sum Y_i=n\hat{\beta}_0+\hat{\beta}_1\sum X_i\ \ldots\ldots(1-1)$$
$$(2)\Leftrightarrow 2\sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)(-X_i)=0$$
$$\Leftrightarrow\sum X_i\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)=0$$
$$\Leftrightarrow\sum X_iY_i=\hat{\beta}_0\sum X_i+\hat{\beta}_1\sum X_i^2.(2-2)$$

ولاستخراج قيم كل من $\hat{\beta}_0$ و $\hat{\beta}_1$ سنستخدم طريقة المحددات كما يلي:

لدينا:

$$(1\text{-}1)\ldots\ldots\sum Y_i=n\hat{\beta}_0+\hat{\beta}_1\sum X_i$$
$$(2\text{-}2)\ldots\sum X_iY_i=\hat{\beta}_0\sum X_i+\hat{\beta}_1\sum X_i^2$$

أي من الشكل:

$$\begin{pmatrix}\sum Y_i\\ \sum X_iY_i\end{pmatrix}=\underbrace{\begin{bmatrix}n & \sum X_i\\ \sum X_i & \sum X_i^2\end{bmatrix}}_{(X)}\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_1\end{pmatrix}$$

المصفوفة

لدينا محدد المصفوفة X كالتالي:

$$|D|=\begin{bmatrix}n & \sum X_i\\ \sum X_i & \sum X_i^2\end{bmatrix}=\mathrm{n}(\sum X_i^2)-(\sum X_i)^2$$

ي

لدينا ايضا:

$$|A_0| = \begin{bmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{bmatrix} = \left(\sum Y_i \sum X_i^2\right) - \left(\sum X_i Y_i\right)\left(\sum X_i\right)$$

:

$$|A_1| = \begin{bmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{bmatrix} = \mathrm{n}\left(\sum X_i Y_i\right) - \left(\sum \mathrm{X_i}\right)\left(\sum \mathrm{Y_i}\right)$$

$$\hat{\beta}_1 = \frac{|A_1|}{|D|} = \frac{\mathrm{n}(\sum X_i Y_i) - (\sum \mathrm{X_i})(\sum \mathrm{Y_i})}{\mathrm{n}(\sum X_i^2) - (\sum X_i)^2}$$

$\hat{\beta}_1$وهي أيضا بالصيغة التالية:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

كما نتحصل على قيمة $\hat{\beta}_0$ كما يلي:

$$\hat{\beta}_0 = \frac{|A_0|}{|D|} = \frac{(\sum Y_i \sum X_i^2) - (\sum X_i Y_i)(\sum X_i)}{\mathrm{n}(\sum X_i^2) - (\sum X_i)^2}$$

وهي تعطى بالصيغة النهائية التالية بدلالة $\hat{\beta}_1$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

مثال1: الجدول الموالي يوضح بيانات حول عدد سنوات الخدمة(X_i) ومعدل الأجر السنوي(Y_i) بآلاف الدينار لعينة من الموظفين. نريد دراسة وجود العلاقة بين معدل الأجر السنوي وسنوات لخدمة.

X_i	4	8	12	16	20	24	28	32
Y_i	25.6	32.7	45.4	53.9	59	62.6	65	65.8

إذن نقوم بتقدير خط الانحدار التالي :

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 \mathrm{X_i}$$

من خلال المعادلتين التاليتين(طريقة 1)

$$\hat{\beta}_1 = \frac{\mathrm{n}(\sum X_i Y_i) - (\sum \mathrm{X_i})(\sum \mathrm{Y_i})}{\mathrm{n}(\sum X_i^2) - (\sum X_i)^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \frac{(\sum Y_i \sum X_i^2) - (\sum X_i Y_i)(\sum X_i)}{\mathrm{n}(\sum X_i^2) - (\sum X_i)^2}$$

نقوم بحساب المجاميع كالتالي:

ي

X_i	Y_i	X_iY_i	$(X_i)^2$
4	25.6	102.4	16
8	32.7	261.6	64
12	45.4	544.8	144
16	53.9	862.4	256
20	59	1180	400
24	62.6	1502.4	576
28	65	1820	784
32	65.8	2105.6	1024
$\sum X_i = 144$	$\sum Y_i = 410$	$\sum X_iY_i = 8379.2$	$\sum (X_i)^2$

بالإضافة لدينا:

$$\overline{\mathrm{X}} = \frac{\sum_{i=1}^{n} \mathrm{X_i}}{\mathrm{n}} = \frac{144}{8} = 18$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{410}{8} = 51,25$$

اذن نقوم بالتعويض في معادلات $\hat{\beta}_0$ و $\hat{\beta}_1$ كالتالي:

$$\hat{\beta}_1 = \frac{8(8379.2) - (144)(410)}{8(3264) - (144)^2}$$

$$= \frac{7993.6}{5376} \Rightarrow \hat{\beta}_1 = 1.487$$

$$\hat{\beta}_0 = \frac{(410)(3264) - (144)(8379.2)}{8(3264) - (144)^2}$$

$$= \frac{131635.2}{5376} \Rightarrow \hat{\beta}_0 = 24.48$$

و يمكن ايضا تطبيق العلاقة الثانية:

$$\hat{\beta}_0 = 51.25 - (1.487)18 = 24.48$$

اذن المعادلة المقدرة هي من الشكل:

$$\hat{Y}_i = 24.48 + 1.487\mathrm{X_i}$$

كما يمكن حساب $\hat{\beta}_1$ التالية:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

X_i	Y_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})(Y_i - \bar{Y})$
4	25.6	-14	196	-25.65	359.1

ي

8	32.7	-10	100	-18.55	185.5
12	45.4	-6	36	-5.85	35.1
16	53.9	-2	4	2.65	5.3
20	59	2	4	7.75	15.5
24	62.6	6	36	11.35	68.1
28	65	10	100	13.75	137.5
32	65.8	14	196	14.55	203.7
$\sum X_i = 144$	$\sum Y_i = 410$	Σ=0	Σ = 672		Σ =1009.8

إذن نعوض في المعادلة السابقة

$$\hat{\beta}_1 = \frac{1009.8}{672} = 1.50$$

هناك طريقة أخرى في تقدير معالم النموذج وهي طريقة المصفوفات سنتعرض لها بشكلها المفصل في المحور الثالث (نموذج الانحدار الخطي المتعدد)

2-2 الخصائص الأساسية لمقدرات المربعات الصغرى

بعد تقدير معالم النموذج هناك خصائص يجب أن تتوفر في هذه المقدرات التي يتم استخراجها بطريقة المربعات الصغرى، لهذا نحاول أن نبين مدى توفر هذه الخصائص في مقدرات المربعات الصغرى كما يلي:

خاصية عدم التحيز و الخطية: ويقصد بها مقدرتا المربعات الصغرى العادية $\hat{\beta}_1$ و $\hat{\beta}_0$ أفضل مقدرتين خطيتين لβ_0 وβ_1 ونعني بها أن يكون الفرق بين القيمة المتوقعة للمقدرتين$E(\hat{\beta}_0)$ و $E(\hat{\beta}_1)$ وقيمة المعلمتين الحقيقتين للمجتمع الإحصائي تساوβ_0 وβ_1 ي الصغرى أي:

$$E(\hat{\beta}_0) - \beta_0 = 0 \Rightarrow E(\hat{\beta}_0) = \beta_0$$

$$E(\hat{\beta}_1) - \beta_1 = 0 \Rightarrow E(\hat{\beta}_1) = \beta_1$$

و لإثبات ذلك نقوم بالبرهان على ذلك:

- بالنسبة ل $\hat{\beta}_1$

لدينا:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

بوضع:

$$y_i = Y_i - \bar{Y} \quad \wedge \quad x_i = X_i - \bar{X}$$

يصبح لدينا:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

نضع $K_i = \frac{x_i}{\sum(x_i)^2}$ اذن نكتب $\hat{\beta}_1 = \sum_{i=1}^{n} K_i y_i$

ي

و هذا يدل على انه خطي أي يكتب بالشكل: $\hat{\beta}_1 = K_1 y_1 + K_2 y_2 + \cdots \ldots + K_n y_n$

و من خصائص K_i أنه:

- حيث: $\sum_{i=1}^{n} K_i = 0$: $K_i = \frac{x_i}{\sum (x_i)^2} \Rightarrow \sum_{i=1}^{n} K_i = \frac{\sum x_i}{\sum (x_i)^2} = \frac{\sum (X_i - \bar{X})}{\sum (X_i - \bar{X})^2} = 0$
- حيث: $\sum_{i=1}^{n} K_i x_i = 1$: $\sum_{i=1}^{n} K_i x_i = \sum K_i (X_i - \bar{X})$

$$= \sum K_i X_i - \bar{X} \sum K_i$$

$$\sum_{i=1}^{n} K_i x_i = \sum K_i X_i = \frac{\sum X_i}{\sum {X_i}^2} X_i = \frac{\sum {X_i}^2}{\sum {X_i}^2} = 1$$

- حيث: $\sum {K_i}^2 = \frac{1}{\sum {x_i}^2}$: $\sum_{i=1}^{n} K_i = \frac{\sum x_i}{\sum (x_i)^2} \Rightarrow \sum {K_i}^2 = \frac{\sum (X_i - \bar{X})^2}{(\sum (X_i - \bar{X})^2)^2}$

$$\sum {K_i}^2 = \frac{1}{\sum (X_i - \bar{X})^2} = \frac{1}{\sum {x_i}^2}$$

و نعلم أنه:

$$\hat{\beta}_1 = \sum_{i=1}^{n} K_i y_i$$

$$\hat{\beta}_1 = \sum_{i=1}^{n} K_i (Y_i - \bar{Y})$$

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{\varepsilon}$$

نعوض في معادلة $\hat{\beta}_1$ يصبح:

$$\hat{\beta}_1 = \sum_{i=1}^{n} K_i \left[\beta_0 + \beta_1 X_i + \varepsilon_i - \beta_0 - \beta_1 \bar{X} - \bar{\varepsilon}\right]$$

$$\hat{\beta}_1 = \sum_{i=1}^{n} K_i \left[\beta_1 X_i - \beta_1 \bar{X} + \varepsilon_i - \bar{\varepsilon}\right]$$

$$\hat{\beta}_1 = \sum_{i=1}^{n} K_i \left[\beta_1 (X_i - \bar{X}) + (\varepsilon_i - \bar{\varepsilon})\right]$$

$$\hat{\beta}_1 = \beta_1 \sum_{i=1}^{n} K_i X_i + \sum_{i=1}^{n} K_i \varepsilon_i$$

لأن $\bar{\varepsilon}$ يؤول إلى الصفر إذن :

$$\hat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} K_i \varepsilon_i$$

بإدخال التوقع الرياضي على طرفي المعادلة يصبح لدينا :

$$E(\hat{\beta}_1) = E(\beta_1) + \sum_{i=1}^{n} K_i E(\varepsilon_i)$$

ي

β_1 معلم نظري اذن $E(\varepsilon_i) = 0 \quad \Lambda \quad E(\beta_1) = \beta_1$

اذن نستنتج أن : $E(\hat{\beta}_1) = \beta_1$

اذن المقدر $\hat{\beta}_1$ مقدار غير متحيز أي أحسن مقدر ل β_1 هو $\hat{\beta}_1$

- بالنسبة ل $\hat{\beta}_0$

لدينا:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$$

$$\hat{\beta}_0 = \beta_0 + \beta_1\bar{X} + \bar{\varepsilon} - \hat{\beta}_1\bar{X}$$

ندخل الوقع الرياضي على طرفي المعادلة:

$$E(\hat{\beta}_0) = \beta_0 + \beta_1\bar{X} + \frac{1}{n}\sum E(\varepsilon_i) - E(\hat{\beta}_1)\bar{X}$$

لدينا: $E(\hat{\beta}_1) = \beta_1$

اذن:

$$E(\hat{\beta}_0) = \beta_0 + \beta_1\bar{X} + 0 - \beta_1\bar{X} \Rightarrow E(\hat{\beta}_0) = \beta_0$$

و منه $\hat{\beta}_0$ مقدر غير متحيز أي أحسن مفدر ل β_0

كما يمكننا أن نستنتج أن المقدر $\hat{\beta}_0$ هو خطي كما يلي:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$$
$$= \frac{1}{n}\sum Y_i - \sum K_i Y_i \bar{X}$$
$$= \sum\left(\frac{1}{n} - K_i\bar{X}\right)Y_i$$

بوضع: $W_i = \left(\frac{1}{n} - K_i\bar{X}\right)$

نكتب:

$$\hat{\beta}_0 = \sum_{i=1}^{n} W_i Y_i$$

أي:

$$\hat{\beta}_0 = W_1Y_1 + W_2Y_2 + \cdots W_nY_n$$

و هو المطلوب

إذا وجدنا تحيز لأحد المقدرتين، فإننا ننظر إلى الخاصية التقاربية لذلك المقدر، ويحدث ذلك لما يكون المتغير المستقل X_i عبارة عن متغير تابع ومبطأ بفترة زمنية ما، ونقول عن $\hat{\beta}_1$ بأنه مقدر متسق (Consistent Estimator) فإذا كان: $n \to \infty$ أي حجم العينة كبير، فإن توزيع المعاينة لـ $\hat{\beta}_1$ يقترب من القيمة الحقيقية β_1، ونقول أن النهاية الاحتمالية للمقدر $\hat{\beta}_1$ هي β_1 ونكتب : $p\lim_{n\to\infty}(\hat{\beta}_1) = \beta_1$

ي

بالإضافة إلى ذلك، يجب أن تكون قيمتا التحيز والتباين تقتربان أو تساويان الصفر كلما اقترب n من ما لا نهاية

أي:

$$1.\lim_{n\to\infty} E(\hat{\beta}_1) = p\lim_{n\to\infty}(\hat{\beta}_1) = \beta_1$$

$$2.\lim_{n\to\infty} \mathrm{var}(\hat{\beta}_1) = p\lim_{n\to\infty} \mathrm{var}(\hat{\beta}_1) = 0$$

في هذه الحالة، نقول عن المقدر $\hat{\beta}_1$ بأنه مقدر متسق للمعلمة الحقيقية. إن المقدرات المتحصل عليها لكل من β_0 ، β_1 و σ^2 سواء بطريقة المربعات الصغرى أو غيرها هي تقديرات نقطية، ولذلك يجب أن نبني مجالا لهذه المقدرات وذلك بقبول مستوى ثقة معين وهو ما نسميه بالتقدير المجالي للمعالم.

3-توزيع المعاينة للمقدرات وتقدير مجال ثقة المعالم:

نقوم أولا بحساب تباين $\hat{\beta}_0$ و $\hat{\beta}_1$ حيث أنهما يستخدما في تحديد مجال المعلمتين β_0 و β_1 وكذا في إجراء الاختبارات المعنوية للمقدرات ولكن يجب قبل ذلك تقدير تباين الأخطاء، باعتباره تباين الأخطاء النظري غير معرف ويرمز له ب $\sigma^2{}_{\varepsilon_i}$ والمقدر $\hat{\sigma}^2{}_{\hat{\varepsilon}_i}$

3-1: تقدير تباين حد الخطأ العشوائي:

لدينا:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$
$$\bar{Y} = \beta_0 + \beta_1\bar{X} + \bar{\varepsilon}$$
$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

من جملة المعادلات نحصل على:

$$\hat{\varepsilon}_i = Y_I - \hat{Y}_i$$
$$= (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)X_i + \varepsilon_i \quad \ldots.(^*)$$

لدينا: $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$

اذن يصبح لدينا:

$$(\beta_0 - \hat{\beta}_0) = \beta_0 - \bar{Y} + \hat{\beta}_1\bar{X}$$
$$= \beta_0 - (\beta_0 + \beta_1\bar{X} + \bar{\varepsilon}) + \hat{\beta}_1\bar{X}$$
$$= (\hat{\beta}_1 - \beta_1)\bar{X} - \bar{\varepsilon}$$

نعوض في المعادلة (*) يصبح لدينا:

$$\hat{\varepsilon}_i = (\beta_1 - \hat{\beta}_1)X_i - (\hat{\beta}_1 - \beta_1)\bar{X} + (\varepsilon_i - \bar{\varepsilon})$$
$$= (\beta_1 - \hat{\beta}_1)(X_i - \bar{X}) + (\varepsilon_i - \bar{\varepsilon})$$

نربع طرفي المعادلة السابقة

$$\hat{\varepsilon}_i{}^2 = (\beta_1 - \hat{\beta}_1)^2(X_i - \bar{X})^2 - (\varepsilon_i - \bar{\varepsilon})^2 + 2(\beta_1 - \hat{\beta}_1)(X_i - \bar{X})(\varepsilon_i - \bar{\varepsilon})$$

نضع: $(X_i - \bar{X})^2 = x_i{}^2$

ي

$$\left(\beta_1-\hat{\beta}_1\right)=\sum_{i=1}^{n}K_i\varepsilon_i \Rightarrow \left(\beta_1-\hat{\beta}_1\right)^2=(\sum_{i=1}^{n}K_i\varepsilon_i)^2$$

اذن ينتج لدينا المعادلة التالية:

$$\hat{\varepsilon}_i{}^2=(\sum_{i=1}^{n}K_i\varepsilon_i)^2x_i{}^2-(\varepsilon_i-\bar{\varepsilon})^2+2(\sum_{i=1}^{n}K_i\varepsilon_i)x_i(\varepsilon_i-\bar{\varepsilon})$$

نقوم بإدخال التوقع الرياضي كما يلي:

$$\mathrm{E}\left(\hat{\varepsilon}_i{}^2\right)=\mathrm{E}[(\sum_{i=1}^{n}K_i\varepsilon_i)^2x_i{}^2-(\varepsilon_i-\bar{\varepsilon})^2+2(\sum_{i=1}^{n}K_i\varepsilon_i)x_i(\varepsilon_i-\bar{\varepsilon})]$$

إذن لدينا مايلي:

أ-

$$\left(\sum_{i=1}^{n}K_i\varepsilon_i\right)^2=\sum_{i=1}^{n}K_i{}^2\varepsilon_i{}^2+2\sum\sum K_iK_j\varepsilon_i\varepsilon_j$$

$$E\left[\left(\sum_{i=1}^{n}K_i\varepsilon_i\right)^2\right]=\sum_{i=1}^{n}K_i{}^2E(\varepsilon_i{}^2)+2\sum\sum K_iK_jE\left(\varepsilon_i\varepsilon_j\right)$$

لدينا: $\mathrm{E}\left(\varepsilon_i\varepsilon_j\right)=0$

إذن: $E[(\sum_{i=1}^{n}K_i\varepsilon_i)^2]=\sigma^2{}_{\varepsilon_i}\sum_{i=1}^{n}K_i{}^2$

لدينا من قبل: $\sum_{i=1}^{n}K_i{}^2=\frac{1}{\sum_{i=1}^{n}x_i{}^2}$

يصبح لدينا : $E[(\sum_{i=1}^{n}K_i\varepsilon_i)^2]=\frac{\sigma^2{}_{\varepsilon_i}}{\sum_{i=1}^{n}(x_i)^2}$

ب- $(\varepsilon_i-\bar{\varepsilon})=\varepsilon_i-\frac{1}{n}\sum\varepsilon_i=\varepsilon_i-\frac{1}{n}(\varepsilon_1+\varepsilon_2+\cdots\varepsilon_n)$

نربع طرفي المعادلة:

$$(\varepsilon_i-\bar{\varepsilon})^2=\varepsilon_i{}^2+\frac{1}{n^2}(\varepsilon_1{}^2+\varepsilon_2{}^2+\cdots\varepsilon_n{}^2)+2\sum\sum\varepsilon_i\varepsilon_j-\frac{2}{n}\varepsilon_i(\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_n)$$

بإدخال التوقع الرياضي كما يلي:

$$\mathrm{E}((\varepsilon_i-\bar{\varepsilon})^2)=E(\varepsilon_i{}^2)+\frac{1}{n^2}E(\sum\varepsilon_i{}^2)+2\sum\sum E\left(\varepsilon_i\varepsilon_j\right)-\frac{2}{n}\sum E\left(\varepsilon_i\varepsilon_j\right)\quad i\neq j$$

لدينا: $2\sum\sum E\left(\varepsilon_i\varepsilon_j\right)=0$

$$\frac{1}{n^2}E\left(\sum\varepsilon_i{}^2\right)=\frac{\sigma_{\varepsilon}{}^2}{n}$$

$$E(\varepsilon_i{}^2)=\sigma_{\varepsilon}{}^2$$

$$\frac{1}{n}\sum E\left(\varepsilon_i\varepsilon_j\right)=\frac{\sigma_{\varepsilon}{}^2}{n}$$

بالتعويض في المعادلة السابقة يصبح لدينا :

$$\mathrm{E}((\varepsilon_i-\bar{\varepsilon})^2)=\sigma_{\varepsilon}{}^2+\frac{\sigma_{\varepsilon}{}^2}{n}+2\,\frac{\sigma_{\varepsilon}{}^2}{n}$$

ي

ث- $(\sum_{i=1}^{n} K_i\varepsilon_i)(\varepsilon_i-\bar{\varepsilon}) = (\sum_{i=1}^{n} K_i\varepsilon_i)(\varepsilon_i) - (\sum_{i=1}^{n} K_i\varepsilon_i)(\bar{\varepsilon})$

ندخل التوقع الرياضي كما يلي:

$$E\left[\left(\sum_{i=1}^{n} K_i\varepsilon_i\right)(\varepsilon_i-\bar{\varepsilon})\right] = E\left[\left(\sum_{i=1}^{n} K_i\varepsilon_i\right)(\varepsilon_i)\right] - E\left[\left(\sum_{i=1}^{n} K_i\varepsilon_i\right)(\bar{\varepsilon})\right]$$

$$= \sum_{i=1}^{n} K_i E({\varepsilon_i}^2) - E\left[\sum K_i E(\,\varepsilon_i)\bar{\varepsilon}\right]$$

و عليه بعد الحساب نجد:

$$E\left[\left(\sum_{i=1}^{n} K_i\varepsilon_i\right)(\varepsilon_i-\bar{\varepsilon})\right] = \sum x_i K_i \sigma^2$$

وفي الأخير نعوض في المعادلة كالتالي:

$$\mathrm{E}({\hat{\varepsilon}_i}^2) = \frac{\sigma^2}{\sum (x_i)^2}\sum (x_i)^2 + n\sigma^2 + \sigma^2 - (2\sigma^2) - 2\sigma^2 \sum x_i K_i$$

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

$$\mathrm{E}({\hat{\varepsilon}_i}^2) = (n-2)\sigma^2$$

$$\Rightarrow E\left(\frac{\sum {\hat{\varepsilon}_i}^2}{n-2}\right) = {\sigma_{\varepsilon_i}}^2$$

إذن أحسن تقدير لتباين الأخطاء النظري ${\sigma_{\varepsilon_i}}^2$ هو $\boldsymbol{{\hat{\sigma}_{\hat{\varepsilon}_i}}^2 = \frac{\sum {\hat{\varepsilon}_i}^2}{n-2}}$

مع $\sum {\hat{\varepsilon}_i}^2 = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)$

(n-2) هي درجة الحرية و تعني حجم العينة n مطروحا منه2 لوجود معلمتين في النموذج β_0، β_1

2-3: تقدير تباين المقدرة β_1

لدينا ما يلي: $\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} K_i\varepsilon_i$

إذن $\left(\hat{\beta}_1 - \beta_1\right)^2 = (\sum_{i=1}^{n} K_i\varepsilon_i)^2 = \sum {K_i}^2{\varepsilon_i}^2 + 2\sum\sum K_i K_j \varepsilon_i \varepsilon_j$

ندخل التوقع الرياضي للمعادلة:

$$E\left[\left(\hat{\beta}_1 - \beta_1\right)^2\right] = \sum {K_i}^2 E({\varepsilon_i}^2) + 2\sum\sum K_i K_j E(\varepsilon_i\varepsilon_j)$$

لدينا: $E(\varepsilon_i\varepsilon_j) = 0 \quad \Lambda \quad \mathrm{E}({\hat{\varepsilon}_i}^2) = {\sigma_{\varepsilon_i}}^2$

إذن ينتج عن ذلك:

$$var(\,\hat{\beta}_1) = {\sigma_{\varepsilon_i}}^2 \sum {K_i}^2 = \frac{{\sigma_{\varepsilon_i}}^2}{\sum (X_i - \bar{X})^2}$$

ويمكن تعويض ${\sigma_{\varepsilon_i}}^2$ و ${\hat{\sigma}_{\hat{\varepsilon}_i}}^2$

إذن: $\boldsymbol{var(\,\hat{\beta}_1) = {\hat{\sigma}_{\hat{\beta}_1}}^2 = \frac{{\hat{\sigma}_{\varepsilon_i}}^2}{\sum (X_i - \bar{X})^2}}$

ي

3-3: تقدير تباين المقدرة $\hat{\beta}_0$

لدينا من قبل: $\hat{\beta}_1 - \beta_1 = \sum K_i \varepsilon_i \quad \Lambda \quad \hat{\beta}_0 - \beta_0 = \bar{\varepsilon} - (\hat{\beta}_1 - \beta_1)\bar{X}$

$$\hat{\beta}_0 - \beta_0 = \sum_{i=1}^{n} \left(\frac{1}{n} - K_i \bar{X}\right) \varepsilon_i$$

نضع: $\dot{K}_i = \frac{1}{n} - K_i \bar{X}$ إذن: $\hat{\beta}_0 - \beta_0 = \sum \dot{K}_i \varepsilon_i$

نربع الطرفين كالعادة يصبح لدينا :

$$(\hat{\beta}_0 - \beta_0)^2 = \left(\sum_{i=1}^{n} \dot{K}_i \varepsilon_i\right)^2 = \sum \dot{K}_i^{\,2} \varepsilon_i^{\,2} + 2 \sum \sum \dot{K}_i \dot{K}_j \varepsilon_i \varepsilon_j$$

ندخل التوقع الرياضي على الطرفين :

$$E\left[(\hat{\beta}_0 - \beta_0)^2\right] = E\left[\left(\sum_{i=1}^{n} \dot{K}_i \varepsilon_i\right)^2\right] = \sum \dot{K}_i^{\,2} E(\varepsilon_i^{\,2}) + 2 \sum \sum \dot{K}_i \dot{K}_j E(\varepsilon_i \varepsilon_j)$$

لدينا $E(\varepsilon_i \varepsilon_j) = 0 \quad : \quad \Lambda \quad E(\hat{\varepsilon}_i^{\,2}) = \sigma_{\varepsilon_i}^{\,2} \quad \Lambda \quad \sum K_i = 0$

وأيضا $\sum \dot{K}_i^{\,2} = \sum \left(\frac{1}{n^2} + K_i^{\,2} \bar{X}^2 - \frac{2}{n} K_i \bar{X}\right) = \frac{1}{n} + \sum K_i^{\,2} \bar{X}^2$

إذن نستنتج أن:

$$\boldsymbol{var(\hat{\beta}_0) = \hat{\sigma}_{\hat{\beta}_0}^{\,2} = \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2}\right) \hat{\sigma}_{\varepsilon_i}^{\,2}}$$

و ذلك بنعويض $\sigma_{\varepsilon_i}^{\,2}$ بمقدرة $\hat{\sigma}_{\varepsilon_i}^{\,2}$

يمكن استنتاج علاقة بين تبايني مقدرتين $\hat{\sigma}_{\hat{\beta}_0}^{\,2}$، $\hat{\sigma}_{\hat{\beta}_1}^{\,2}$ تعطى كما يلي:

$$\boldsymbol{\hat{\sigma}_{\hat{\beta}_0}^{\,2} = \frac{\hat{\sigma}_{\hat{\varepsilon}_i}^{\,2}}{n} + \bar{X}^2 \hat{\sigma}_{\hat{\beta}_1}^{\,2}}$$

4-3: تحديد مجال الثقة للمعلمتين $\hat{\beta}_0$ و $\hat{\beta}_1$

يمكن تشكيل أو تكوين مجال للقيم التي بالامكان أن تأخذها معالم الانحدار الفعلية أو الحقيقية وذلك بمعرفة توزيع المقدرتين $\hat{\beta}_0$ و $\hat{\beta}_1$. فإذا كان حجم العينة صغير حيث $n \leq 30$ يتم تكوين مجال الثقة من التوزيع ستودنت Student (t) بالنسبة للمعلمتين β_0 و β_1 أما إذا كانت العينة $n > 30$ فيتم تكوين مجال الثقة من التوزيع الطبيعي كما يلي:

- $n \leq 30$ و $\sigma_{\varepsilon_i}^{\,2}$ غير معروف

لدينا: $\frac{\hat{\beta}_0 - E(\hat{\beta}_0)}{\hat{\sigma}_{\hat{\beta}_0}} \sim t(\alpha/2, n-2)$

ي

$$\frac{\hat{\beta}_1 - E(\hat{\beta}_1)}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(\alpha/_2, n-2)$$

أي

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \rightsquigarrow t_{(n-2)}$$

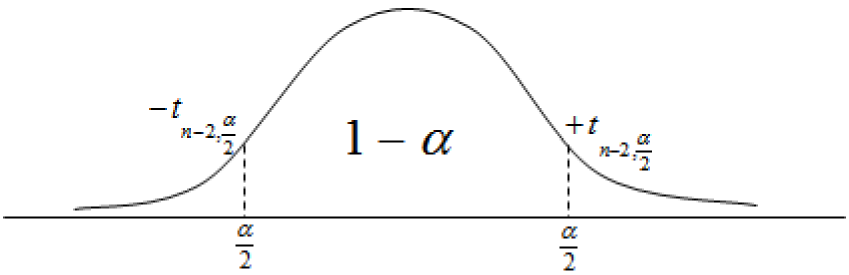
$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \rightsquigarrow t_{(n-2)}$$

ومنه احتمال احتواء المجال المذكور (مجال الثقة) على معلمة الانحدار الفعلية عند معنوية $(\%\alpha)$ هو واحد مطروح منه مستوى المعنوية $(1-(\%\alpha))$ كما يلي:

$$\Pr\left[-t_{n-2,\frac{\alpha}{2}} \le \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \le +t_{n-2,\frac{\alpha}{2}}\right] = 1-\alpha$$

$$\Pr\left[-t_{n-2,\frac{\alpha}{2}} \le \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \le +t_{n-2,\frac{\alpha}{2}}\right] = 1-\alpha$$

الشكل رقم (3) : توزيع المعاينة لـ $\hat{\beta}_1$ و $\hat{\beta}_0$ ثنائي الطرف



ويتم استخراج مجال الثقة للمعلمتين من خلال ضرب كل الأطراف بواسطة $\hat{\sigma}_{\hat{\beta}_0}{}^2$ و $\hat{\sigma}_{\hat{\beta}_1}{}^2$ على التوالي و اضافة معلمتي الانحدار لأطراف المتراجحة وينتج لدينا مجالي الثقة كما يلي :

$$\beta_0 \in \left[\hat{\beta}_0 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}\right]$$

ي

$t(\alpha/_2, n-2)$ هي قيم مستخرجة من جدول توزيع استودنت وهي تعبير القيمة الحرجة لتوزيع ستودنت عند درجة حرية(n-2) وبنسبة معنوية (α%)

- n> 30 و $\sigma_{\varepsilon_i}{}^2$ معروف

لدينا"

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\hat{\sigma}_\varepsilon^2}{\sum_i x_i^2}\right)$$

$$\hat{\beta}_0 \sim N\left(\beta_0, \hat{\sigma}_\varepsilon^2 \frac{\sum_i X_i^2}{n\sum_i x_i^2}\right)$$

أي

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim N(0,1)$$

$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim N(0,1)$$

بنفس المراحل السابقة نستخرج احتمال احتواء مجال الثقة على معلمتي الانحدار الفعلية عند معنوية α% كما يلي:

$$\Pr\left[-z_{\frac{\alpha}{2}} \le \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \le +z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

$$\Pr\left[-z_{\frac{\alpha}{2}} \le \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \le +z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

ثم نستخرج مجال الثقة للمعلمتين كما يلي:

$$\beta_0 \in \left[\hat{\beta}_0 - z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}\right]$$

حيث تمثل $Z_{\alpha/2}$: القيم الحرجة للتوزيع الطبيعي عند معنوية α% تستخرج من جدول التوزيع الطبيعي وكلما كان مجال الثقة ضيق كلما كانت المقدرتين أدق وذلك لصغر الخطأ العشوائي.

ي

مثال2: نفس معطيات المثال السابق ونقوم باستخراج مجالي الثقة لمعلمتي النموذج السابق عند 5%= α%

قمنا بتقدير نموذج الانحدار وتحصلنا على:

$$\hat{Y}_i = 24.48 + 1.48X_i$$

نقوم بحساب قيم $\hat{Y}$ثم نستخرج قيم حد الخطأ العشوائي$\hat{\varepsilon}_i$ في المراحل التالية :

X_{i}	Y_i	$\hat{Y}$	$\hat{\varepsilon}_i = Y_i - \hat{Y}_i$	$(\hat{\varepsilon}_i)^2$
4	25.6	30.4	-4.8	23.04
8	32.7	36.3	-3.6	12.96
12	45.4	42.2	3.2	10.24
16	53.9	48.2	5.7	32.5
20	59	54.1	4.9	24.01
24	62.6	60.0	2.6	6.76
28	65	65.9	-0.9	0.81
32	65.8	71.8	-6	36
				Σ= 146.32

نقوم بحساب تباين الخطأ العشوائي

$$\hat{\sigma}_{\hat{\varepsilon}_i}{}^2 = \frac{\sum_{i=1}^{8} \hat{\varepsilon}_i{}^2}{n-2} = \frac{146.32}{6}$$

$$\hat{\sigma}_{\hat{\varepsilon}_i}{}^2 = 24.38$$

- حساب تباين مقدرات النموذج

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{\hat{\sigma}^2_{\hat{\varepsilon}_i}}{\sum_{i=1}^{8}(X_i - \bar{X})^2} = \frac{24{,}38}{672} = 0.036$$

$$\Rightarrow \hat{\sigma}_{\hat{\beta}_1} = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{0.036} = 0.189$$

$$\hat{\sigma}^2_{\hat{\beta}_0} = \frac{\hat{\sigma}^2_{\hat{\varepsilon}_i}}{n} + \bar{X}^2\hat{\sigma}^2_{\hat{\beta}_1} = \frac{24.38}{8} + (18)^2(0.036) = 14.71$$

$$\Rightarrow \hat{\sigma}_{\hat{\beta}_0} = \sqrt{\hat{\sigma}^2_{\hat{\beta}_0}} = \sqrt{14.71} = 3.84$$

استخراج القيمة الحرجة لتوزيع ستودنت (n=8 <30)، عند معنوية5% و درجة حرية 6 $t_T(0.025,6)$=20447

اذن لدينا

$$\beta_0 \in \left[\hat{\beta}_0 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}\right]$$

$$\beta_0 \in [24.48 - 2.447(3.28); 24.48 + 2.447(3.28)]$$

$$\beta_0 \in [16.45; 32.50]$$

$$\beta_1 \in [1.48 - (2.447)(0.15); 148 + (2.447)1.48]$$

ي

$$\beta_1 \in [1.11; 1.84]$$

وهما مجالا ثقة للمعلمتين β_0 و β_1

4– تحليل القدرة التفسيرية للنموذج المقدر:

لدينا مما سبق $Y_i - \widehat{Y}_i = \hat{\varepsilon}_i \Rightarrow Y_i = \widehat{Y}_i + \hat{\varepsilon}_i$

وهذا يعني الخطأ العشوائي أو البواقي $\hat{\varepsilon}_i$ تستعمل كمقياس لمدى تمثيل المعادلة المقدرة (النموذج) للمشاهدات الفعلية وكلما كانت قيمة هذه البواقي كبية إذن يكون النموذج المقدر غير جيد والعكس.

على هذا الأساس نقوم بتحليل التباين كما يلي:

لدينا: $Y_i = \widehat{Y}_i + \hat{\varepsilon}_i$

نقوم بطرح $\overline{Y}$ من طرفي المعادلة :

$$Y_i - \overline{Y} = \widehat{Y}_i - \overline{Y} + \hat{\varepsilon}_i$$

نقوم بجمع وتربيع طرفي المعادلة كالتالي :

$$\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \sum_{i=1}^{n}(\widehat{Y}_i - \overline{Y})^2 + \sum_{i=1}^{n}(\hat{\varepsilon}_i)^2$$

ولتحليل القدرة التفسيرية للنموذج نقوم بتحليل هذه المجاميع

$\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ وهو مجموع مربعات الانحرافات الكلية للمتغير Y_i المراد تفسيرها أي هي التغيرات الكلية للمتغير وهو (Total Sum of Squares - TSS)

$\sum_{i=1}^{n}(\widehat{Y}_i - \overline{Y})^2$ وهو مجموع مربعات الانحرفات المفسرة من طرفي النموذج المقدر (Explained Sum Of Squares-ESS).

$\sum_{i=1}^{n}(\hat{\varepsilon}_i)^2$ هو مجموع مربعات البواقي (الجزء الغير المفسر من طرف النموذج المقدر لانحرافات Y_i وهي (Residual Sum of Squares-RSS)

أي لدينا TSS=ESS+RSS

وبتفسير طرفي المعادلة على الانحرافات الكلية أو مجموع مربعات الانحرافات الكلية أو مجموع مربعات الانحرافات الكلية ل Y_i و هي TSS، يصبح:

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

ومن خلال ذلك يمكن تعريف معامل التحديد R^2 والذي يقيس أو يشرح التغيرات التي تحدث في المتغير Y_i والتي تم شرحها بواسطة التغيرات في المتغير المستقل X_i، أو التي سببها هذا المتغير وهو نسبة الانحرافات الموضحة من قبل خط الانحدار (النموذج المقدر) للانحرافات الكلية.

ي

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 1 - \frac{\sum_{i=1}^{n}(\hat{\varepsilon}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

اذن: R^2 هو مقياس للقدرة التفسيرية للنموذج أي جودة النموذج في تفسير تغيرات Y_i

كما أن R^2 هو مقياس للارتباط أيضا حيث يعطينا أو يقيس العلاقة بين المتغيرين Y_i وX_i بالاضافة لجودة تفسير احد المتغيرين في احد المتغيرين للتغيرات التي تحدث في المتغير الآخر (السببية).

ولدينا في هذه الحالة: $R^2 = r^2$

فقط في حالة الانحدار الخطي البسيط معامل التحديد هو نفسه مربع معامل الارتباط بين المتغيرين التابع والمستقل

وتتراوح قيمة R^2 بين الصفر والواحد أي: $0 \leq R^2 \leq 1$

ويمكن البرهان عليها كالتالي :

لدينا:

$$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 \leq \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \Rightarrow \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \leq 1$$

$$\Rightarrow R^2 \leq 1$$

وأيضا لدينا :

$$\begin{cases} \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 \geq 0 \\ \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \geq 0 \end{cases} \Rightarrow \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \geq 0 \Rightarrow R^2 \geq 0$$

هو المطلوب,

$R^2 = 1$ أي جميع نقاط الانتشار على خط الانحدار المقدر أي $Y_i = \hat{Y}_i$

$R^2 = 0$، عندما يكون خط انحدار العينة أفقيا $Y_i = \bar{Y}_i$ ومعنى ذلك لا توجد علاقة بين المتغير التابع والمتغير المستقل وهاتين الحالتين نادرتي الحدوث.

مثال3: نفس المثال السابق نقوم بحساب معامل التحديد

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 1 - \frac{\sum_{i=1}^{n}(\hat{\varepsilon}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

لدينا: $\bar{Y} = 51.25$

ي

$(Y_i - \overline{Y})$	$(Y_i - \overline{Y})^2$	$(\hat{Y}_i - \overline{Y})$	$(\hat{Y}_i - \overline{Y})^2$
25.65-	657.92	20.83-	433.88
18.55-	344.10	14.95-	223.50
5.85-	34.22	9.05-	81.90
2.65	7.02	3.05-	9.30
7.75	60.06	2.85	8.12
11.35	128.82	8.75	76.56
13.75	189.06	14.65	214.62
14.55	211.70	20.55	422.30
	Σ= 1632.9		Σ= 1470.18

إذن لدينا:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} = \frac{1470.18}{1632.9} = 0.909$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{\varepsilon}_i)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} = 1 - \frac{146.32}{1632.9} = 1 - 0.09 = 0.91$$

إذن الجودة التفسيرية للنموذج المقدر تقدر بالتقريب91% هذا يعني أن التغير في المتغير Xيفسر 91% من التغيرات التي تحدث في Y و 9% المتبقية تمثل تأثير متغيرات أخرى، لم تدخل في النموذج أو أخطاء القياس.

91% تعني أن جودة النموذج عالية.

5-اختبار المعنوية الإحصائية للمعالم (اختبار الفرضيات)

قبل استخدام النموذج المقدر في تحليل العلاقة بين المتغيرتين X وY والتنبؤ بها مستقبلا، يجب إثبات صحته واختباره وذلك من خلال اختبار المعنوية الإحصائية لمعلمات النموذج، وذلك بفرض فرضية العدم H_0 والتي مفادها انعدام العلاقة بين المتغيرتين X_i و Y_iأي خط الانحدار هو أفقي أي أن $\beta_1 = 0$ وأيضا $\beta_0 = 0$، تم اختبارها وبالتالي يقابلها فرض البديل$H_1 : \beta_1 \neq 0$ أو$\beta_1 > 0$ أو $\beta_1 < 0$ (في حالة معرفة نوع العلاقة بين المتغيرتين من النظرية الاقتصادية) وأيضا اختبار الفرضية $\beta_0 \neq 0$

ويتم اختيار هذه الفرضية الصفرية كالتالي:

5-1: اختبار معنوية المعلمة β_1:

لدينا:

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

ي

- في حالة $n \leq 30$ نقوم بحساب احصائية ستودنت

كما يلي:

$$t_c = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$$

$$t_c = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$$

يتم قبول H_0 عند معنوية α% إذا كانت $\left|\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}\right| \leq t_{n-2,\frac{\alpha}{2}}$، أي المعلمة β_1 ليست ذات معنوية احصائية وتساوي معنويا الصفر أي المتغيرة X لا تدخل في تفسير تغيرات المتغير Y.

- في حالة $n < 30$ في هذه الحالة تقوم بمقارنة القيمة المحسوبة $\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$ مع القيمة الحرجة للتوزيع الطبيعي $Z_{\alpha/2}$

، ففي حالة $\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}} < Z\alpha/_2$ تقبل H_0.

2-5: اختيار معنوية المعلمة β_0: بنفس الطريقة التي تم بها اختبار معنوية المعلمة β_1.

2-5 اختبار المعنوية الكلية للنموذج – اختبار فيشر Fisher

إن اختيار معنوية النموذج المقدر ككل يكون في شكل توزيع Fisher، وهو يتمثل في اختبار أثر المتغير المستقل X_i ويتم ذلك من خلال طرح الفرضية الصفرية التالية :($H_0 : \beta_1 = 0$) ((أثر) المتغير المستقل X_i)

ثم نقوم بحساب إحصائية فيشر F_c حيث:

لدينا:

$$F = \frac{ESS/1}{RSS/(n-2)} \sim F_{1,n-2}$$

$$ESS = R^2 TSS$$

$$RSS = (1 - R^2)TSS$$

نحصل على:

$$F = \frac{R^2/1}{(1-R^2)/(n-2)} = \frac{R^2}{(1-R^2)}.(n-2) \sim F_{1,n-2}$$

ونقوم برفض $H_0 : \beta_1 = 0$، أي أن النموذج المقدر معنوي (وجود لأثرX_i)، إذا كان:

$$F_c > F_{T(\alpha\%,1,(n-2))}$$

ي

حيث تمثل $F_{T(\alpha\%,1,(n-2))}$ القيمة الجدولية لتوزيع فيشر عند معنوية α% و درجة حرية 1في أبسط (عدد المتغيرات المستقلة) و (n-2) في المقام (حجم العينة مطروحا منه عدد المعلمات).

مثال4: نفس المثال السابقة المطلوب اختبار المعنوية الجزئية لمعالم النموذج والمعنوية الكلية للنموذج المقدر عند معنوية $\alpha\% = 5\%$

- اختبار معنوية β_0

لدينا:
$$H_0: \beta_0 = 0$$
$$H_1: \beta_0 \neq 0$$

لدينا n=8، إذن القيمة $t_c = \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_0}$ تتبع توزيع t

إذن نقوم بحساب إحصائية student.

$$t_c = \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}0}} \Rightarrow t_c = \frac{\hat{\beta}_0}{\hat{\sigma}_{\hat{\beta}_0}} = \frac{24.48}{3.28}$$
$$t_c = 7.46$$

لدينا
$$t_{T(\alpha/_2,n-2)} = t_{T(0.025,6)} = 2.447$$
$$|t_c| > t_T$$

هذا يعني أن نرفض H_0 وبالتالي المعلمة β_0 ذات معنوية احصائية

- اختبار معنوية β_1

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

القيمة $t_c = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$ تتبع أيضا توزيع t، n< 30

إذن نحسب إحصائية ستودنت

لدينا: $t_c = \frac{1.48-0}{0.15} \Rightarrow t_c = 9.86$

$|t_c| > t_T \Leftrightarrow$ المعلمة β_1 ذات معنوية إحصائية

أي قبول المقدرتين $\hat{\beta}_1$،$\hat{\beta}_0$ للوصول إلى معلمي المجتمع.

- المعنوية الكلية للنموذج المقدرة: بالاعتماد على إحصائية فيشر

لدينا

$$H_0: \beta_0 = \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

ي

$$F_c = \frac{R^2/1}{(1-R^2)/(n-2)}$$

$$F_c = \frac{0.91}{(1-0.91)}(8-2)$$

$$F_c = 60.66$$

نقوم باستخراج القيمة الجدولية لتوزيع فيشر عند معنوية $\alpha\%=5\%$ ودرجة حرية 1 في البسط و6 في المقام وذلك لاتخاذ قرار اختبار فرضية العدم.

$$F_{T(5\%,1,6)} = 5.987$$

إذن نلاحظ أن $F_c > F_T$ يعني رفض فرضية العدم وبالتالي النموذج ذو معنوية إحصائية بنسبة معنوية $\alpha\%=5\%$

6- التنبؤ Forecasting:

من بين الأهداف الأساسية للنموذج المقدر هو استخدامه للتنبؤ بقيم المتغير التابع، مستقبلا استنادا لقيم المتغير المستقبل من اجل استشراف مسار الظاهرة موضوع البحث في المستقبل، إذن بعد تقييم نموذج الانحدار المقدر و التأكد من استفاءة للفرضيات الإحصائية، يمكن بالإمكان استخدامه في عملية التنبؤ.

ليكن لدينا نموذج الانحدار التالي:

$$Y_i = f(X_i)$$

و لدينا قيم X في فترة التنبؤ X_{T+h} و الشكل الهيكلي لا يتغير مستقبلا، إذن قيم المتغير Y في هذه الفترة T+h هي كما يلي:

$$Y_{T+h} = \beta_0 + \beta_1 X_{T+h} + \varepsilon_{T+h}$$

حيث يسمى h أفق التنبؤ و T هو حجم العينة T.......t=1,2,3

Y_{T+h} هو التنبؤ النظري

و لعدم معرفتنا للمعلمتين β_0 و β_1 ، نعتمد على مقدراتها $\hat{\beta}_0$ و $\hat{\beta}_1$ ، أي على النموذج المقدر كما يلي:

$$\hat{Y}_T(h) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+h}$$

و هذا المقدر يسمى بالتنبؤ التقديري و هو مقدر غير متحيز، أي $E(Y_{T+h}X_{T+h}) = \beta_0 + \beta_1 X_{T+h}$ و هو أحسن المقدرات الخطية له، أي له اصغر تباين.

نقوم بحساب تباين هذا التنبؤ التقديري $Var\left(\hat{Y}_T(h)\right)$

لدينا:

$$Var\left(\hat{Y}_T(h)\right) = Var\left(\hat{\beta}_0 + \hat{\beta}_1 X_{T+h}\right)$$
$$= Var\left(\hat{\beta}_0\right) + {X_{T+h}}^2 Var\left(\hat{\beta}_1\right) + 2X_{T+h} cov\left(\hat{\beta}_0 \beta_1\right)$$

لدينا:

ي

$$cov(\hat{\beta}_0\beta_1) = \frac{-\bar{X}\sigma_\varepsilon{}^2}{\sum X_i{}^2}$$

$$Var(\hat{\beta}_1) = \frac{\sigma_\varepsilon{}^2}{\sum(X_i - \bar{X})^2}$$

$$Var(\hat{\beta}_0) = \frac{\sigma_i{}^2}{T} + \bar{X}^2 Var(\hat{\beta}_1)$$

نعوض قيم كل تباين، نجد في الأخير:

$$Var\left(\hat{Y}_T(h)\right) = \sigma_\varepsilon{}^2\left[\frac{1}{T} + \frac{(X_{T+h} - \bar{X})^2}{\sum_1^T (X_i - \bar{X})^2}\right]$$

و كلما انخفضت قيمة $(X_{T+h} - \bar{X})^2$وازداد حجم العينة، انخفضت قيمة تباين التنبؤ التقديرين أي تحسن قيمة التنبؤ.

و يمكن التعبير عنه بالعبارة التالية:

$$\hat{\varepsilon}_{T+h} = Y_{T+h} - \hat{Y}_T(h)$$

و يسمى هذا المقدر $\hat{\varepsilon}_{T+h}$ بمقدر خطأ التنبؤ Forecast error

لدينا:

$$E(\hat{\varepsilon}_{T+h}) = E\left(Y_{T+h} - \hat{Y}_T(h)\right) = 0$$

كما أن تباينه هو:

$$Var(\hat{\varepsilon}_{T+h}) = Var\left(Y_{T+h} - \hat{Y}_T(h)\right)$$

$$= Var(Y_{T+h}) + Var\left(\hat{Y}_T(h)\right) + 2cov\left(Y_{T+h}, \hat{Y}_T(h)\right)$$

لدينا $cov\left(Y_{T+h}, \hat{Y}_T(h)\right)$ مساويا للصفر، باعتبار أن Y_{T+h} تعتمد على ε_{T+h} و بينما $\hat{Y}_T(h)$ فهي تعتمد على ε_t مع T: t=1 بواسطة المقدرتين$\hat{\beta}_0$ و $\hat{\beta}_1$.

كما لدينا أيضا:

$$Var\left(\hat{Y}_T(h)\right) = \sigma_\varepsilon{}^2\left[\frac{1}{T} + \frac{(X_{T+h} - \bar{X})^2}{\sum_1^T (X_i - \bar{X})^2}\right]$$

$$Var(Y_{T+h}) = \sigma_{\varepsilon_{T+h}}{}^2$$

إذن يصبح:

$$Var(\hat{\varepsilon}_{T+h}) = \sigma_\varepsilon{}^2\left[1 + \frac{1}{T} + \frac{(X_{T+h} - \bar{X})^2}{\sum_1^T (X_i - \bar{X})^2}\right]$$

إذن المقدر الغير متحيز لتباين خطأ التنبؤ هو$\hat{\sigma}_{\varepsilon_{T=h}}{}^2$ حيث:

ي

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = \hat{\sigma}_{\varepsilon}{}^2\left[1+\frac{1}{T}+\frac{(X_{T+h}-\bar{X})^2}{\sum_1^T(X_i-\bar{X})^2}\right]$$

و كلما كانت العينة كبيرة كلما اقتربت قيمة $\hat{\sigma}_{\varepsilon_{T+h}}{}^2$ من $\hat{\sigma}_{\varepsilon}{}^2$

و بعد استخراجنا لمقدر خطأ التنبؤ، يمكننا بذلك استخراج أو تحديد مجال الثقة للتنبؤ، و الذي يعبر عن دقة التنبؤ، من خلال الخطوات التالية:

لتحديد دقة التنبؤ ل Y_{T+h} باعتبار ε_{T+h} موزعة طبيعيا و أيضا Y_{T+h} إذن $\hat{Y}_T(h)$هي موزعة أيضا طبيعيا، إذن $\hat{\varepsilon}_{T+h} = Y_{T+h} - \hat{Y}_T(h)$هي موزعة طبيعيا.

$$\frac{\hat{\varepsilon}_{T+h}}{\hat{\sigma}_{\varepsilon_{T+h}}} \sim N(0;1) \Rightarrow \frac{Y_{T+h}-\hat{Y}_T(h)}{\hat{\sigma}_{\varepsilon_{T+h}}} \sim N(0;1)$$

و يمكن استخراج قيمة$Z_{\alpha/2}$ عند معنوية$\alpha\%$ و هي القيمة الحرجة للتوزيع الطبيعي، إذن:

$$Pr\left[-Z_{\frac{\alpha}{2}} \le \frac{Y_{T+h}-\hat{Y}_T(h)}{\hat{\sigma}_{\varepsilon_{T+h}}} \le +Z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

إذن بنفس الخطوات التي استخدمن من قبل في تحديد مجال الثقة نستنتج أن:

مجال الثقة للتنبؤ هو:

$$Y_{T+h} \in \left[\hat{Y}_T(h) \mp Z_{\frac{\alpha}{2}}\hat{\sigma}_{\varepsilon_{T+h}}\right]$$

في حالة العينة الصغيرة$n \le 30$ نستخدم اختبار ستودنت$\left(t_{(\alpha/_2,n-2)}\right)$.

مثال5: نفس معطيات المثال السابق، نحاول التنبؤ للقيمة $\hat{\varepsilon}_{T+h}$ لما نكون قيمة X_{8+1}، كما يلي:

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = \hat{\sigma}_{\varepsilon}{}^2\left[1+\frac{1}{T}+\frac{(X_{T+h}-\bar{X})^2}{\sum_1^T(X_i-\bar{X})^2}\right]$$

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = 24.38\left[1+\frac{1}{8}+\frac{(36-18)^2}{672}\right]$$

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = 39.17 \Rightarrow \hat{\sigma}_{\varepsilon_{T+h}} = 6.25$$

لدينا: $\hat{Y}_8(1) = 24.48 + 1.487 * 36$

$$\hat{Y}_8(1) = 77.7$$

إذن نستخرج مجال الثقة عند معنوية $\alpha\%$ تساوي 5% ل Y_{T+1} كما يلي:

$$Y_{T+1} \in [77.7 \mp 2.447 * 6.25] \Rightarrow Y_{T+1} \in [62.4; 99.9]$$

تمرين شامل: نحاول دراسة أثر معدل النمو الاقتصادي X على معدلات البطالةY لبلد ما، خلال فترة زمنية قدرت ب 13 مشاهدة سنوية، ثم التنبؤ من خلالها للفترات الموالية و كانت البيانات كالتالي:

ي

X	1.2	0.9	1.7	2.1	2.5	3.1	4.1	4.0	4.7	5.0	5.1	3.5	4.1
Y	25	24.2	22	20	19	17	15	14	12.3	10	12.3	15.2	12.7

باعتبار العلاقة خطية بين المتغيرتين كما يلي:

$$\mathrm{Y_i} = \beta_0 + \beta_1 \mathrm{X_i} + \varepsilon_i$$

1- تقدير نموذج الانحدار:

$$\bar{X} = \frac{\sum_{i=1}^{13} X_i}{n} = \frac{42}{13} = 3.23 \quad \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{218.7}{13} = 16.82$$

X	Y	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(Y_i - \bar{Y})^2$
1.2	25	-2.03	4.12	8.17	-16.58	66.86
0.9	24.2	-3.33	5.43	7.37	-24.54	54.41
1.7	22	-1.53	234	5.17	-7.91	26.80
2.1	20	-1.13	1.27	3.17	-3.58	10.09
2.5	19	-.073	0.53	2.17	-1.58	4.73
3.1	17	-0.13	0.01	0.17	-0.02	0.03
4.1	15	0.86	0.75	-1.82	-1.56	3.32
4.0	14	0.76	0.59	-2.82	-2.14	7.96
4.7	12.3	1.46	2.15	-4.52	-6.59	20.45
5.0	10.0	1.76	3.13	-6.82	-12.00	46.55
5.1	12.3	1.86	3.49	-4.52	-8.40	20.45
3.5	15.2	0.26	0.072	-1.62	-0.42	2.63
4.1	12.7	0.86	0.75	-4.12	-3.45	16.99
			Σ= 24.68		Σ=-88.86	Σ=281.34

لدينا:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{-88.86}{24.68}$$

$$\hat{\beta}_1 = -3.60$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 16.82 - (3.6 -)3.23$$

$$\hat{\beta}_0 = 28.48$$

إذن النموذج المقدر هو:

$$\hat{\mathrm{Y}}_i = 28.48 - 3.6\mathrm{X_i}$$

بعد تقدير معالم نموذج البطالة، نستخرج **مجال الثقة لهذه المعالم** عند معنوية $\alpha\%$ تساوي 5%:

ي

$$\beta_0 \in \left[\hat{\beta}_0 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_0}\right]$$

$$\beta_1 \in \left[\hat{\beta}_1 - t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_1}\right]$$

كمرحلة أولى نقوم بتقدير تباين حد الخطأ العشوائي خيث لدينا:

$$\hat{\sigma}_{\hat{\varepsilon}_i}{}^2 = \frac{\sum_{i=1}^{8}\hat{\varepsilon}_i{}^2}{n-2} = \frac{\sum_{i=1}^{13}\left(Y_i - \hat{Y}_i\right)^2}{n-2}$$

Y	$\hat{Y}_i$	$\hat{\varepsilon}_i = Y_i - \hat{Y}_i$	$\hat{\varepsilon}_i{}^2$
25	24.16	0.84	0.70
24.2	25.24	-1.04	1.08
22	22.36	-0.36	0.13
20	20.25	-0.25	0.062
19	19.48	-0.48	0.23
17	17.32	-0.32	0.10
15	13.72	1.28	1.63
14	14.08	-0.08	0.006
12.3	11.56	0.74	0.54
10.0	10.48	-0.48	0.23
12.3	10.12	2.18	4.75
15.2	15.88	-0.68	0.46
12.7	13.72	-1.02	1.04
		\	
			Σ= 10.95

$$\hat{\sigma}_{\hat{\varepsilon}_i}{}^2 = \frac{10.95}{11} \Rightarrow \hat{\sigma}_{\hat{\varepsilon}_i}{}^2 = 0.99$$

$$\hat{\sigma}_{\hat{\beta}_1}{}^2 = \frac{\hat{\sigma}_{\hat{\varepsilon}_i}{}^2}{\sum_{i=1}^{8}(X_i - \bar{X})^2}$$

$$= \frac{0.99}{24.68}$$

$$\hat{\sigma}_{\hat{\beta}_1}{}^2 = 0.04 \Rightarrow \hat{\sigma}_{\hat{\beta}_1} = \sqrt{0.04} = 0.2$$

$$\hat{\sigma}_{\hat{\beta}_0}{}^2 = \frac{\hat{\sigma}_{\hat{\varepsilon}_i}{}^2}{n} + \bar{X}^2\hat{\sigma}_{\hat{\beta}_1}{}^2$$

$$= \frac{0.99}{11} + (3.23)^2 0.04$$

$$\hat{\sigma}_{\hat{\beta}_0}{}^2 = 0.49 \Rightarrow \hat{\sigma}_{\hat{\beta}_0} = \sqrt{0.49} = 0.7$$

كما لدينا القيمة الحرجة لتوزيع ستودنت: $t_{(0.025,11)} = 2.201$

ي

إذن يصبح لدينا:

$$\beta_1 \in [-3.6 \mp 2.201 * 0.2] \Rightarrow \beta_1 \in [-3.16; 0.44]$$
$$\beta_0 \in [28.48 \mp 2.201 * 0.7] \Rightarrow \beta_0 \in [26.94; 30.02]$$

-يتم **دراسة جودة تفسير النموذج** من خلال حساب معامل التحديد R^2 حيث:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{\varepsilon}_i)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} = 1 - \frac{10.95}{281.34} = 1 - 0.038 = 0.961$$

إذن الجودة التفسيرية للنموذج تقدر ب 96%، أي أن التغيرات في المتغير معدلات النمو الاقتصادي X يفسر 96% من تغيرات معدلات البطالة Y

-**دراسة المعنوية الكلية للنموذج**، باستخدام توزيع فيشر كما يلي:

$$H_0: \beta_0 = \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

$$F_c = \frac{R^2/1}{(1-R^2)/(n-2)}$$
$$F_c = \frac{0.961}{(1-0.961)}(13-2)$$
$$F_c = 271.05$$

لأخد قرار معنوية النموذج أي قبول أو رفض H_0، نقوم باستخراج القيمة الجدولية لتوزيع فيشر عند معنوية $\alpha\%=5\%$ ودرجة حرية 1 في البسط و11 في المقام.

$$F_T(0.05, 1, 11) = 4.84$$

لدينا $F_c > F_T$ أذن نرفض H_0 و بالتالي النموذج معنوي.

-**دراسة المعنوية الجزئية** لأثر معدلات النمو على معدلات البطالة أي معنوية β_1

لدينا:

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

القيمة $t_c = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$ تتبع أيضا توزيع t، n< 30

إذن نحسب إحصائية ستودنت t

لدينا: $t_c = \frac{-3.6-0}{0.2} \Rightarrow t_c = -18$

$|t_c| > t_T \Leftrightarrow$ المعلمة β_1 ذات معنوية إحصائية، أي قبول المقدرتين $\hat{\beta}_1$ للوصول إلى معلمي المجتمع.

ي

معامل أثر النمو الاقتصادي على البطالة معنوي.

النموذج معنوي و الأثر أيضا معنوي، إذن يمكن استخدام النموذج المقدر في التنبؤ.

-**نتنبأ** بقيمة معدل البطالة للفترة الموالية حيث يكون معدل النمو الاقتصادي مقدرا ب 4.7%

$$\hat{Y}_i = 28.48 - 3.6X_i$$

إذن:

$$\hat{Y}_{13}(1) = 28.48 - 3.6(4.7) = 11.56$$

و لاستخراج مجال الثقة للتنبؤ نقوم:

-تقدير تباين خطأ التنبؤ ${\hat{\sigma}_{\varepsilon_{T+h}}}^2$

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = \hat{\sigma}_{\varepsilon}{}^2 \left[1 + \frac{1}{T} + \frac{(X_{T+h} - \bar{X})^2}{\sum_1^T (X_i - \bar{X})^2}\right]$$

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = 0.99 \left[1 + \frac{1}{13} + \frac{(4.7 - 3.23)^2}{24.68}\right]$$

$$\hat{\sigma}_{\varepsilon_{T+h}}{}^2 = 1.15 \Rightarrow \hat{\sigma}_{\varepsilon_{T+h}} = 1.07$$

$$Y_{T+h} \in \left[\hat{Y}_T(h) \mp t_{(\alpha/_2, n-2)} \hat{\sigma}_{\varepsilon_{T+h}}\right]$$

$$Y_{T+h} \in [11.56 \mp 2.201 * 1.07]$$

$$Y_{T+h} \in [11.56 \mp 2.35]$$

$$Y_{T+h} \in [9.21; 14]$$

تمرين 1: **(للحل)**

نرغب في تقدير العلاقة الخطية بين الاستهلاك المحلي y بالانتاج x لمادة الاسمنت بالمليون برميل خلال عدة سنوات، و ذلك حسب نموذج الانحدار الخطي البسيط التالي:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i=1,2,\ldots\ldots n$$

لهذا الغرض تم اختيار عينة لعشر سنوات و كانت النتائج كما يلي:

i	1	2	3	4	5	6	7	8	9	10
Y_i	6	8	9	8	7	6	5	6	5	5
X_i	10	13	15	14	9	7	6	6	5	5

المطلوب:

ي

1- قم بتقدير نموذج يبين اثر انتاج الاسمنت على كمية استهلاكه ؟

2- اوجد مجال الثقة لمعلمتي نموذج الانحدار عند مجال ثقة 95%

3- قم بدراسة جودة النموذج، المعنوية الكلية للنموذج عند معنوية $\alpha = 5\%$

4- كم ستصيح Y_i إذا كانت X_{10+1} مساوية ل 8

التمرين الثاني:(للحل)

نفترض أن إحدى الشركات ترغب في تحديد العلاقة بين إنفاقها على الترويج (X) وعوائد المبيعات (Y) ، كلاهما بالمليون دولار ، وذلك باستخدام البيانات في الجدول الموالي، ولعمل ذلك يجب أن تقوم الشركة بتقدير نموذج يجعل Y تابعة ل X كمايلي:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i=1,2,\ldots\ldots n$$

t (سنة)	1	2	3	4	5	6	7	8	9	10
X_t (إعلانات)	4	5	6	6	7	8	7	9	8	10
Y_t (مبيعات)	44	42	52	48	50	60	58	62	64	70

المطلوب:

1- هل للترويج اثر على حجم المبيعات ؟

2- اوجد مجال الثقة لمعلمتي نموذج الانحدار عند مجال ثقة 95%

3- قم بدراسة جودة النموذج، المعنوية الكلية للنموذج و معنوية المقدرة β_1 للنموذج عند دلالة 5%
$\alpha =$

4- أوجد عائد المبيعات المتوقع عندما يتم الإنفاق على الإعلانات بما قيمته 13 مليون دولار، مع تحديد مجال الثقة للقيمة المطلوبة عند مستوى خطأ $\alpha = 5\%$

الفصل الثالث: تحليل الانحدار الخطي المتعدد -Multiple Linear Regression

من الفصل السابق اتضح ان الانحدار الخطي البسيط يركز على دراسة العلاقة بين متغيرين احدهما المتغير المستقل (X). لكن معظم الظواهر الاقتصادية لا تفسر فقط بمحدد مستقل واحد فقط، بل هناك العديد من المحددات أو العوامل المؤثرة التي تجعل الدراسة أكثر شمولية. على هذا الأساس، النموذج الخطي البسيط لا يكفي و لابد من توسيعه ليشتمل انحدار للمتغير التابع (Y) على العديد من المتغيرات المستقلة $X_1, X_2......., X_k$ لذا يتم الاستعانة بالنموذج الخطي المتعدد. حيث يعد الانحدار الخطي المتعدد من الأساليب الإحصائية المتقدمة والتي تضمن دقة الاستدلال ممن يريد تحسين نتائج البحث عن طريق الاستخدام الأمثل للبيانات في إيجاد علاقات سببية بين الظواهر موضوع البحث .

إن الانحدار الخطي المتعدد ليس مجرد أسلوب واحد وإنما مجموعة من الأساليب التي يمكن استخدامها لمعرفة العلاقة بين متغير تابع مستمر وعدد من المتغيرات المستقلة التي عادةً ما تكون مستمرة).

1- شكل نموذج الانحدار المتعدد:

يستند النموذج الخطي المتعدد على افتراض وجود علاقة خطية بين متغير تابع Y_i وعدد من المتغيرات المستقلة $X_1, X_2...,...., X_k$ وحد عشوائي ε_i، ويعبر عن هذه العلاقة بالنسبة لـn من المشاهدات وk من المتغيرات المستقلة بالشكل الأتي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \varepsilon_i, \quad i = 1,......., n.....................(1)$$

وفي واقع الأمر فإن هذه المعادلة هي واحدة من جملة معادلات يبلغ عددها (n) تكون نظام المعادلات الأتي:

$$\begin{aligned}
i = 1: Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + ... + \beta_k X_{1k} + \varepsilon_1 \\
i = 2: Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + ... + \beta_k X_{2k} + \varepsilon_2 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
i = n: Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + ... + \beta_k X_{nk} + \varepsilon_n
\end{aligned}$$

هذه المعادلة تتضمن $(k+1)$ من المعلومات المطلوب تقديرها علما بان الحد الأول منها β_0 يمثل الحد الثابت الآمر الذي يتطلب اللجوء إلى المصفوفات والمتجهات لتقدير تلك المعلمات، عليه يمكن صياغة هذه المعادلات في صورة مصفوفات كآلاتي:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & ... & X_{1K} \\ 1 & X_{21} & X_{22} & ... & X_{2K} \\ . & . & . & ... & . \\ 1 & X_{n1} & X_{n2} & ... & X_{nk} \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ . \\ B_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ \varepsilon_n \end{bmatrix} \quad .... \quad (2)$$

وباختصار يمكن كتابة العلاقة السابقة كالآتي:

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

Y : متجه عمودي أبعاده $(n\times 1)$ يحتوي مشاهدات المتغير التابع .

X : مصفوفة أبعادها $(n\times k+1)$ تحتوي مشاهدات المتغيرات المستقلة يحتوي عمودها الأول على قيم الواحد الصحيح ليمثل الحد الثابت .

β : متجه عمودي أبعاده $(k+1\times 1)$ يحتوي على المعالم المطلوب تقديرها .

ε : متجه عمودي أبعاده $(n\times 1)$ يحتوي على الأخطاء العشوائية .

وبما أن المعادلة (1) هي العلاقة الحقيقية المجهولة والمراد تقديرها باستخدام الإحصاءات المتوفرة عن المتغير التابع Y والمتغيرات المستقلة $X_1, X_2 ...,...., X_k$ ، فإنه يستوجب تحقق الفروض الأساسية الخاصة بـ ε_i التالية:

2- الفرضيات التي يقوم عليها نموذج الإنحدار الخطي المتعدد:

$H.1$ * تأخذ علاقة النموذج الخطي المتعدد الصيغة التالية:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_k X_{ik} + \varepsilon_i , \quad i = 1,......., n$$

$H.2$ * ε_{i} يتوزع توزيعا طبيعيا متعدد المتغيرات.

$H.3$ * القيمة المتوقعة لمتجه حد الخطأ تساوي صفرا أي أن: $E(\varepsilon_i) = 0$

$$E(\varepsilon_i) = E\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ . \\ E(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ . \\ 0 \end{bmatrix}$$

$H.4$ * تباين العناصر العشوائية ثابت، والتباين المشترك بينها يساوي صفرا أي أن:

$$\begin{cases} \operatorname{var}(\varepsilon_i) = \sigma^2 , & \forall i = 1.........n \\ \operatorname{cov}(\varepsilon_i , \varepsilon_j) = 0, & \forall i \neq j \end{cases}$$

$$E(\varepsilon\varepsilon') = E\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ \varepsilon_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 & \varepsilon 2 & ... & \varepsilon_n \end{bmatrix}$$

$$= E\begin{bmatrix} \varepsilon_1^2 & \varepsilon_1\varepsilon_2 & ... & \varepsilon_1\varepsilon_n \\ \varepsilon_2\varepsilon_1 & \varepsilon_2^2 & ... & \varepsilon_2\varepsilon_n \\ . & . & . & . \\ \varepsilon_n\varepsilon_1 & \varepsilon_n\varepsilon_2 & ... & \varepsilon_n^2 \end{bmatrix}$$

$$E(\varepsilon'\varepsilon)=\begin{bmatrix} E(\varepsilon_1^2) & E(\varepsilon_1\varepsilon_2) & .... & E(\varepsilon_1\varepsilon_n) \\ E(\varepsilon_2\varepsilon_1) & E(\varepsilon_2^2) & .... & E(\varepsilon_2\varepsilon_n) \\ . & . & ..... & . \\ E(\varepsilon_n\varepsilon_1) & E(\varepsilon_n\varepsilon_2) & .... & E(\varepsilon_n^2) \end{bmatrix}$$

$$E(\varepsilon\varepsilon') = \begin{bmatrix} \operatorname{var}(\varepsilon_1) & Cov(\varepsilon_1\varepsilon_2) & .... & Cov(\varepsilon_1\varepsilon_n) \\ Cov(\varepsilon_2\varepsilon_1) & Var(\varepsilon_2) & .... & Cov(\varepsilon_2\varepsilon_n) \\ . & . & .... & . \\ Cov(\varepsilon_n\varepsilon_1) & Cov(\varepsilon_n\varepsilon_2) & .... & Var(\varepsilon_n) \end{bmatrix}$$

$$\begin{cases} \operatorname{var}(\varepsilon_i) = E(\varepsilon^2{}_i) = \sigma^2 \\ \operatorname{cov}(\varepsilon_i, \varepsilon_j) = 0, \quad \forall i \neq j \end{cases}$$

$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \begin{pmatrix} \sigma_\varepsilon^2 & 0 & \cdots & 0 \\ 0 & \sigma_\varepsilon^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_\varepsilon^2 \end{pmatrix}$$

حيث أن:
$$\sigma_1{}^2 = \sigma^2{}_2 = \sigma^2 = ...\sigma^2{}_n$$

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

وتسمى المصفوفة العددية أعلاه بمصفوفة التباين والتباين المشترك "Variance -Covariance Matrix" لحد الخطأ ε_i، حيث تشكل العناصر القطرية في المصفوفة تباين قيم ε_i، بينما تبقى العناصر غير القطرية (أعلى وأسفل القطر) مساوية للصفر لانعدام التباين المشترك والترابط بين قيم ε_i .

ويمكن تلخيص الفرضيات السابقة رياضيا كما يلي: $\varepsilon_i \sim N\left(0, \sigma^2 I_n\right)$

$* H.5$ $\frac{1}{n}(X'X)$ تؤول إلى مصفوفة محدودة غير فردية.

$H.6$ * عدم وجود ارتباط خطي بين المتغيرات المستقلة كما أن عدد المشاهدات يجب أن يزيد على عدد المعلمات المطلوب تقديرها أي أن : $\rho(X) = K+1 \prec n$ حيث أن ρ رتبة مصفوفة البيانات و X عدد المتغيرات المستقلة $K+1$ ، وهي أصغر من عدد المشاهدات (n)، وهذه الفرضية ضرورية جدا لضمان أيجاد معكوس المصفوفة $(X'X)$ إذ أن عدم توفر هذا الفرض يجعل رتبة المصفوفة (X) اقل من $K+1$ وبالتالي فان رتبة $(X'X)$ التي

تستخدم في الحصول على مقدرات OLS بدورها اقل من $K+1$ ولا يمكن إيجاد معكوسا لها أي أن : $(X'X)^{-1}$ غير معرفة لن محدده يؤول الى الصفر وهذا ما يسبب ما يسمى **بمشكلة الارتباط الخطي المتعدد**، وبالتالي لا يمكن الحصول على المقدرات باستخدام طريقة المربعات الصغرى العادية (OLS).

$H.7^*$ (X) مصفوفة غير عشوائية، وتعني هذه الفرضية أنه إذا أخذنا عينة أخرى تتكون من (n) مشاهدة فإن المصفوفة (X) (مصفوفة المتغيرات المفسرة) تبقى دون تغيير، المصدر الوحيد للتغير هنا هو شعاع الخطأ العشوائي (ε) وهذا ما يؤثر على الشعاع (Y) أي $\mathrm{cov}(X,\varepsilon)=E(X'\varepsilon)=0$;

3– تقدير شعاع المعالم $\hat{\beta}$:

في النموذج $Y=X\beta+\varepsilon$، المجاهيل الوحيدة هي β و ε، المصفوفة X و الشعاع Y هي معطيات النموذج، ويجب الإشارة إلى أن شعاع الأخطاء غير مشاهد ولذلك حتى معرفة قيمة β لا تسمح للمتغيرات المستقلة بإعطاء القيمة الحقيقية لـ Y بالضبط.

وعلينا إذن تقدير β بشكل يجعل $\hat{Y}$ أقرب ما يمكن للمتغير التابع Y، ولهذا الغرض توجد عدة طرق، فيما نستعرض نحن طريقة المربعات الصغرى.

تهدف هذه الطريقة إلى إيجاد تقدير للشعاع β الذي يُصَغِّر مجموع مربعات الانحراف $\hat{\varepsilon}_i$ بين القيمة المقدرة $\hat{Y}$ والقيمة الحقيقية Y.

$$Min\sum_{i=1}^{n}\hat{\varepsilon}_i^2=Min\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2$$

$$\hat{\varepsilon}_i=Y_i-\hat{Y}_i \qquad i=1.......n \qquad\qquad \Rightarrow \quad \hat{\varepsilon}=Y-\hat{Y}=\begin{pmatrix}\hat{\varepsilon}_1\\ \hat{\varepsilon}_2\\ \vdots\\ \vdots\\ \hat{\varepsilon}_n\end{pmatrix}$$

$$Min\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2=Min\left(Y-\hat{Y}\right)'\left(Y-\hat{Y}\right)=Min\ \hat{\varepsilon}'\hat{\varepsilon}$$

نسمي :
$$\Gamma\left(Y,X,\hat{\beta}\right)=\left(Y-\hat{Y}\right)'\left(Y-\hat{Y}\right)=\hat{Y}'\hat{Y}-2\hat{Y}'Y+Y'Y=\hat{\beta}'X'X\hat{\beta}-2\hat{\beta}'X'Y+Y'Y$$

حيث: $\hat{Y}=X\hat{\beta}$. ومنه الهدف هو

$$\min_{\hat{\beta}}\ \Gamma\left(Y,X,\hat{\beta}\right)$$

وإذا كان $\hat{\beta}$ موجود فيجب أن يحقق الشرط الضروري:

$$\frac{\partial\Gamma\left(Y,X,\hat{\beta}\right)}{\partial\hat{\beta}'}=0\Leftrightarrow 2(X'X)\hat{\beta}-2X'Y=0$$

وبما أن رتبة X هي $k+1$ فإن : $\left(X'X\right)$ مصفوفة مربعة $\left((k+1)\times(k+1)\right)$ رتبتها $k+1$ وتقبل معكوس $\left(X'X\right)^{-1}$.

ومنه:
$$2(X'X)\hat{\beta}-2X'Y=0 \quad\Rightarrow (X'X)\hat{\beta}-X'Y=0$$

نضرب طرفي المعادلة بـ$(X'X)^{-1}$ لنحصل على : $\hat{\beta}=(X'X)^{-1}X'Y$ وهو تقدير لـ β .

وللتأكد من أن $\hat{\beta}$ المتحصل عليه هو قيمة دنيا لـ $\Gamma(Y,X,\hat{\beta})$، يجب تحقيق الشرط من الدرجة الثانية:

$$\frac{\partial^2\Gamma(Y,X,\hat{\beta})}{\partial\hat{\beta}'\partial\hat{\beta}}=(X'X)>0$$

وهي مصفوفة موجبة معرفة ومنه فإن $\hat{\beta}$ هو نهاية صغرى.

المصفوفة $(X'X)$ هي على الشكل التالي:

$$X'X=\begin{bmatrix} n & \sum_{i=1}^{n}X_{1i} & \sum_{i=1}^{n}X_{2i} & \cdots & \sum_{i=1}^{n}X_{ki} \\ \sum_{i=1}^{n}X_{1i} & \sum_{i=1}^{n}X_{1i}^2 & \sum_{i=1}^{n}X_{1i}X_{2i} & \cdots & \sum_{i=1}^{n}X_{1i}X_{ki} \\ \sum_{i=1}^{n}X_{2i} & \sum_{i=1}^{n}X_{2i}X_{1i} & \sum_{i=1}^{n}X_{2i}^2 & \cdots & \sum_{i=1}^{n}X_{2i}X_{ki} \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \sum_{i=1}^{n}X_{ki} & \sum_{i=1}^{n}X_{ki}X_{1i} & \sum_{i=1}^{n}X_{ki}X_{2i} & \cdots & \sum_{i=1}^{n}X_{ki}^2 \end{bmatrix}$$

المصفوفة $(X'Y)$ هي على الشكل التالي:

$$\hat{\beta}=\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_k\end{bmatrix} \qquad X'Y=\begin{bmatrix}\sum_{i=1}^{n}Y_i\\ \sum_{i=1}^{n}X_{1i}Y_i\\ \sum_{i=1}^{n}X_{2i}Y_i\\ \vdots\\ \sum_{i=1}^{n}X_{ki}Y_i\end{bmatrix}$$

مثال لدينا البيانات الخاصة بالاستيراد كمتغير تابع (Y) والدخل الوطني كمتغير مستقل أول (X_1) وأسعار الاستيراد كمتغير مستقل ثاني (X_2) لدولة ما للفترة من : 2002- 2010.

Y_i الاستيراد	X_1 الدخل الوطني	X_2 السعر
20	**3**	**02**
22	**5**	**02**
23	**6**	**02.5**
24	**7**	**3.5**
26	**8**	**4.5**
29	**9**	**06**
30	**11**	**7**
34	**12**	**10**
37	**13**	**10.5**
40	**16**	**12**
$\sum Y_i=285$	$\sum X_{1i}=90$	$\sum X_{2i}=60$

المطلوب: إيجاد $\hat{\beta}=(X'X)^{-1}X'Y$

$$(X'X)=\begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} \\ \sum X_{1i} & \sum X_{1i}^{2} & \sum X_{1i}X_{2i} \\ \sum X_{2i} & \sum X_{1i}X_{2i} & \sum X_{2i}^{2} \end{bmatrix}$$

بالتطبيق العددي نتحصل على :

$$(X'X)=\begin{bmatrix} 10 & 90 & 60 \\ 90 & 954 & 671 \\ 60 & 67 & 486 \end{bmatrix}$$

$$(X'Y)=\begin{bmatrix} \sum Y_i \\ \sum X_{1i}Y_i \\ \sum X_{2i}Y_i \end{bmatrix}=\begin{bmatrix} 285 \\ 2804 \\ 1935 \end{bmatrix}$$

$$\hat{\beta}=(X'X)^{-1}(X'Y)=\begin{bmatrix} 10 & 90 & 60 \\ 90 & 954 & 671 \\ 60 & 671 & 486 \end{bmatrix}^{-1}\begin{bmatrix} 285 \\ 2804 \\ 1935 \end{bmatrix}=\begin{bmatrix} 15.99 \\ 0.65 \\ 1.11 \end{bmatrix}$$

$$\boxed{\hat{Y}_i=15.99+0.65X_{1i}+1.11X_{2i}}$$

4- الخصائص الإحصائية للمعالم المقدرة

لدينا: $\hat{\beta}=(X'X)^{-1}X'Y$

و أيضا : $Y=X\beta+\varepsilon$

لنرمز بـ A للمصفوفة $(X'X)^{-1}X'$، حيث:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix}\Rightarrow \hat{\beta}=A.Y \quad \therefore \hat{\beta}_i=\sum_{j=1}^{n}a_{ij}Y_j \quad ,i=1......k$$

ومنه نرى أن مختلف المقدرات $(\hat{\beta}_k........\hat{\beta}_2,\hat{\beta}_1)$ هي على شكل خطي مع المتغير Y.

كذلك:

$$\hat{\beta}=(X'X)^{-1}X'[X\beta+\varepsilon]=(X'X)^{-1}X'X\beta+(X'X)^{-1}X'\varepsilon\Rightarrow\hat{\beta}=\beta+(X'X)^{-1}X'\varepsilon$$

بإدخال التوقع الرياضي :

$$E(\hat{\beta})=\beta+(X'X)^{-1}X'E(\varepsilon) \quad / \quad E(\varepsilon)=0$$

نحصل في الأخير : $E(\hat{\beta})=\beta$

نستنتج أن التقدير $\hat{\beta}$ لـ β المحصل عليه بطريقة المربعات الصغرى غير متحيز. بالإضافة إلى ذلك فإن $\hat{\beta}$ هو التقدير الأفضل من ضمن كل التقديرات الخطية غير المتحيزة لـ β $(BLUE)$.

5– تقدير تباين الأخطاء σ^2 و مصفوفة التباين–التباين المشترك للمقدرات $\Omega_{\hat{\beta}}$:

لدينا $E(\varepsilon\varepsilon') = \Omega_\varepsilon = \sigma^2 I_n$ وبما أن σ^2 غير معروف، فيجب تقديره:

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

نضع : $M_X = (I_n - X(X'X)^{-1}X')$، حيث M_X *تسمى المصفوفة الدورية* أي :

$$M_X = M'_X M_X = M_X^2 = M'$$

و أيضا:
$$M_X X = 0$$

ومنه: $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon' M'_X M_X \varepsilon$ أي : $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon' M_X \varepsilon$

ندخل التوقع الرياضي على الطرفين : $E(\hat{\varepsilon}'\hat{\varepsilon}) = E(\varepsilon' M_X \varepsilon)$

نعلم أن أثر $\hat{\varepsilon}'\hat{\varepsilon}$ يساوي أثر $\varepsilon' M\varepsilon$، ونعلم أيضا أن أثر (AB)=أثر (BA).

يصبح لدينا إذن : أثر $(\varepsilon' M\varepsilon)$=أثر $(\varepsilon\varepsilon' M)$

$$E(\hat{\varepsilon}'\hat{\varepsilon}) = E(\varepsilon'\varepsilon)Tr(M_X)$$

نعلم أن: $E(\varepsilon'\varepsilon) = \sigma^2$ وعليه: $E(\hat{\varepsilon}'\hat{\varepsilon}) = \sigma^2\{Tr(I_n) - Tr(X(X'X)^{-1}X')\}$

ومنه :
$$E(\hat{\varepsilon}'\hat{\varepsilon}) = \sigma^2(n-k-1)$$

حيث : $Tr(I_n) = n \quad Tr(X(X'X)^{-1}X') = k+1$;

لكي نحصل على تقدير غير متحيز لـ σ^2 نقوم بقسمة العبارة على : $(n-k-1)$

$$E\left(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k-1}\right) = \sigma^2$$

في الانحدار المتعدد هناك $k+1$ معلم للتقدير و n عدد المشاهدات، اذن عدد درجات الحرية $n-k-1$، و بالتالي :

$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k-1} = \frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-k-1}$$

لدينا من قبل: $\hat{\beta} - \beta = (X'X)^{-1}X'\varepsilon$ وإحدى فرضيات النموذج هي $E(\varepsilon\varepsilon') = \Omega_\varepsilon = \sigma^2 I_n$.

نقوم بحساب $(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$، حيث :

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

بإدخال التوقع الرياضي، نحصل على مصفوفة التباينات–التباينات المشتركة للمقدرات :

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

إذن :

$$\Omega_{\hat{\beta}} = \sigma_{\varepsilon}^2 (X'X)^{-1}$$

بما أن σ_{ε}^2 غير معروف، حيث يمكن استبداله بمقدر تباين الأخطاء $\hat{\sigma}_{\varepsilon}^2$ وعليه نحصل على :

$$\hat{\Omega}_{\hat{\beta}} = \hat{\sigma}_{\varepsilon}^2 (X'X)^{-1}$$

مثال : في المثال التطبيقي السابق أوجد :

2*اوجد الانحراف المعياري المقدر للمعلمات المقدرة باستخدام المعطيات الأصلية والانحرافات:

* **حساب تباين المقدر للمعلمات المقدرة** .

$$\boxed{Var(\hat{\beta}) = diag\Omega_{\hat{\beta}}}$$

$$\Omega_{\hat{\beta}} = \begin{bmatrix} \mathrm{var}(\hat{\beta}_0) & Cov(\hat{\beta}_0\hat{\beta}_1) & Cov(\hat{\beta}_0\hat{\beta}_2) \\ Cov(\hat{\beta}_1\hat{\beta}_0) & Var(\hat{\beta}_1) & Cov(\hat{\beta}_1\hat{\beta}_2) \\ Cov(\hat{\beta}_2\hat{\beta}_0) & Cov(\hat{\beta}_2\hat{\beta}_1) & Var(\hat{\beta}_2) \end{bmatrix}$$

لدينا : $\boxed{\Omega_{\hat{\beta}} = \sigma_{\varepsilon}^2 (X'X)^{-1}}$ و $\hat{\sigma}_{\varepsilon}^2 = \dfrac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k-1} = \dfrac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-k-1}$

إيجاد $\sum_{i=1}^{n}\hat{\varepsilon}_i^2$:

لدينا : $\sum_i (Y_i-\bar{Y})^2 = \sum_i (\hat{Y}_i-\bar{Y})^2 + \sum_i \hat{\varepsilon}_i^2$

$$TSS = \sum_i (Y_i-\bar{Y})^2 = 408.50$$

$$TSS = \sum Y_i^2 - n\bar{Y}^2 = 408.5$$

$$ESS = \sum_I \left(\hat{Y}-\bar{\hat{Y}}\right)^2 = \hat{\beta}'X'Y - n\bar{Y}^2 = 405.1$$

$$\mathrm{ESS} = \begin{pmatrix} 15.99 & 0.65 & 1.11 \end{pmatrix}\begin{pmatrix} 285 \\ 2804 \\ 1935 \end{pmatrix} - 10\left(\frac{285}{10}\right)^2 = 405.1$$

$$RSS = \sum_{i=1}^{n}\hat{\varepsilon}_i^2 = TSS - ESS = 3.4$$

$$\hat{\sigma}_{\varepsilon}^2 = \frac{3.4}{10-2-1} = 0.48$$

$$\Omega_{\hat{\beta}} = \sigma_e^2 (X'X)^{-1} = 0.48\begin{bmatrix} 10 & 90 & 60 \\ 90 & 954 & 671 \\ 60 & 671 & 486 \end{bmatrix}^{-1}$$

$$\Omega_{\hat{\beta}} = 0.48\begin{bmatrix} 1.36 & -0.35 & 0.32 \\ -0.35 & 0.12 & -0.13 \\ 0.32 & -0.13 & 0.14 \end{bmatrix}$$

- **حساب الانحراف المعياري للمقدرات :**

$$Var(\hat{\beta}) = diag\Omega_{\hat{\beta}}$$

$$Var(\hat{\beta}_0) = 0.48*1.36 = 0.65 \quad \Rightarrow \quad \sigma_{\hat{\beta}0} == \sqrt{0.65} = 0.80$$

$$Var(\hat{\beta}_1) = 0.48*0.12 = 0.06 \quad \Rightarrow \quad \sigma_{\hat{\beta}1} == \sqrt{0.06} = 0.24$$

$$Var(\hat{\beta}_1) = 0.48*0.15 = 0.07 \quad \Rightarrow \quad \sigma_{\hat{\beta}2} = \sqrt{0.07} = 0.26$$

6- قياس حدود الثقة لمعالم النموذج:

عند مستوى معنوية $(\alpha\%)$ يكون مجال الثقة للمعالم :

$$\Pr\left[-t_{\left(n-k-1,\frac{\alpha}{2}\right)} \le \frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}_{\hat{\beta}_i}} \le +t_{\left(n-k-1,\frac{\alpha}{2}\right)}\right] = 1-\alpha \qquad i = 0,1,\dots,k$$

والصيغة الرياضية لتقدير حدود الثقة هي:

الانحراف المعياري المعلمة المقدرة$(t\alpha/2)$ $\pm$ المعلمة المقدرة $=$ معلمة المجتمع أي:

$$i = 0,1,\dots,k \qquad \beta_i \in \left[\hat{\beta}_i - t_{\left(n-k-1,\frac{\alpha}{2}\right)}\hat{\sigma}_{\hat{\beta}_i}, \hat{\beta}_i + t_{\left(n-k-1,\frac{\alpha}{2}\right)}\hat{\sigma}_{\hat{\beta}_i}\right]$$

$t_{\left(n-k-1,\frac{\alpha}{2}\right)}$: القيمة الحرجة لتوزيع *Student* بدرجة حرية $n-k-1$ و نسبة معنوية $\left(\%\frac{\alpha}{2}\right)$ ونجدها من جدول لتوزيع القيمة المحسوبة.

ملاحظة: عندما يكون حجم العينة كبيرا ($n \ge 30$) فينبغي استعمال التوزيع الطبيعي و يمكن أخذ القيمة الحرجة $z_{\alpha/2}$ و ذلك بحساب المساحة المظلة للتوزيع الطبيعي.

وتصبح العلاقة السابقة كالأتي:

$$\beta_i \in \left[\hat{\beta}_i - z_{\left(\frac{\alpha}{2}\right)} \hat{\sigma}_{\hat{\beta}_i} , \hat{\beta}_i + z_{\left(\frac{\alpha}{2}\right)} \hat{\sigma}_{\hat{\beta}_i} \right] \qquad i = 0,1,...,k$$

$Z_{\left(n-k-1, \frac{\alpha}{2}\right)}$: القيمة الحرجة للتوزيع *الطبيعي* بدرجة حرية $n-K-1$ و نسبة معنوية $\left(\%\frac{\alpha}{2}\right)$ ونجد من جدول التوزيع الطبيعي القيمة المجدولة.

مثال: نفس المثال السابق المطلوب استخراج مجال الثقة لمعالم النموذج المقدر.

$$\beta_0 \in \left[\hat{\beta}_0 \pm t\left(\alpha/_2, n-k-1\right).\hat{\sigma}_{\hat{\beta}_0}\right] \Rightarrow \beta_0 \in [15.99 \pm 2.365.0.80] \Rightarrow \beta_0 \in [14.09; 17.88]$$

$$\beta_1 \in \left[\hat{\beta}_1 \pm t\left(\alpha/_2, n-k-1\right).\hat{\sigma}_{\hat{\beta}_1}\right] \Rightarrow \beta_1 \in [0.65 \pm 2.365.0.24] \Rightarrow \beta_0 \in [0.08; 1.21]$$

$$\beta_2 \in \left[\hat{\beta}_2 \pm t\left(\alpha/_2, n-k-1\right).\hat{\sigma}_{\hat{\beta}_2}\right] \Rightarrow \beta_2 \in [1.11 \pm 2.365.0.26] \Rightarrow \beta_0 \in [0.49; 1.72]$$

7- اختبار جودة التوفيق والارتباط:

7-1 معامل التحديد R^2 : **Multiple Coefficient of determination**

ويعد مؤشر أساس في تقييم مدى معنوية العلاقة بين المتغير التابع (Y) والمتغيرات المستقلة (X_k) حيث : ($k = 1,2,3.......k$) ، بعبارة أخرى هو مقياس يوضح نسبة مساهمة المتغيرات المستقلة في تفسير التغير الحاصل في المتغير التابع . ويمكن حسابه كالأتي:

$$R^2 = \frac{ESS}{TSS} = \frac{TSS - RSS}{TSS} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

مع:

TSS : تمثل الانحرافات الكلية .

ESS : تمثل الانحرافات الموضحة من قبل خط الانحدار.

$\varepsilon\varepsilon'$: تمثلا الانحرافات غير الموضحة.

$$ESS = \sum_i (\hat{Y}_i - \bar{Y})^2 = \sum \hat{Y}_i - n\bar{Y}^2$$

$$\sum \hat{Y}_i - n\bar{Y}^2 = \hat{Y}'\,\mathrm{\hat{Y}} - \mathrm{n}\,\bar{\mathrm{Y}}^2$$

$$\hat{Y} = \mathrm{X}\,\hat{\beta}$$

$$\hat{Y}' = \hat{\beta}'\,\mathrm{X}'$$

$$\hat{Y}'\,\mathrm{Y} = \left(\hat{\beta}'\,\mathrm{X}'\right)\mathrm{X}\,\hat{\beta}$$

لدينا :

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

$$\hat{Y}\,'\mathrm{Y} = \beta\,'\mathrm{X}\,'\ \mathrm{X}\left(X\,'\mathrm{X}\right)^{-1}\left(X\,'\mathrm{Y}\right)$$

$$\hat{Y}\,'\mathrm{Y} = \hat{\beta}\,'\left(X\,'\mathrm{Y}\right)$$

أي أن :

$$ESS \;\; = \sum\left(\hat{Y}_i - \overline{Y}\right)^2 = \hat{Y}\,'\mathrm{Y}\ -\ \mathrm{n}\,\overline{\mathrm{Y}}^{\,2} = \hat{\beta}\,'\left(X\,'\mathrm{Y}\right) - n\,\overline{\mathrm{Y}}^{\,2}$$

$$TSS \;\; = \sum_i \ (Y_i - \overline{Y})^2 = \sum\ Y_i^{\ 2} - \ \mathrm{n}\overline{\mathrm{Y}}^{\,2}$$

$$\frac{ESS \;\; = \sum\left(\hat{Y}_i - \overline{Y}\right)^2 = \hat{Y}\,'\mathrm{Y}\ -\ \mathrm{n}\,\overline{\mathrm{Y}}^{\,2} = \hat{\beta}\,'\left(X\,'\mathrm{Y}\right) - n\,\overline{\mathrm{Y}}^{\,2}}{TSS\ = \sum\ \left(Y_i - \overline{\mathrm{Y}}\right)^2\ =\ \sum\ Y_{\mathrm{i}}^{\ 2} - \mathrm{n}\,\overline{\mathrm{Y}}^{\,2}}$$

$$R^2 = \frac{\hat{\beta}'X'Y - n\overline{Y}^2}{Y'Y - n\overline{Y}^2}$$

$$R^2 = 1 - \frac{\varepsilon'\varepsilon}{Y'Y - n\overline{Y}^2}$$

SCT : تمثل الانحرافات الكلية .

SCE : تمثل الانحرافات الموضحة من قبل خط الانحدار.

$\varepsilon\varepsilon'$:: تمثلا الانحرافات غير الموضحة.

وبما أن معامل التحديد R^2 عبارة عن نسبة الانحرافات الموضحة من قيل خط الانحدار إلى الانحرافات الكلية " Total variation " ، فانه يمثل نسبة مجموع مربعات التغير في المتغيرات المستقلة إلى مجموع المربعات الكلية :

$$\boldsymbol{R^2 = \frac{\hat{\beta}'X'Y - n\overline{Y}^2}{Y'Y - n\overline{Y}^2}}$$

$$\boldsymbol{R^2 = 1 - \frac{\varepsilon'\varepsilon}{Y'Y - n\overline{Y}^2}}$$

ملاحظة : $0 \le R^2 \le 1$

-عندما يكون R^2 قريب من الواحد فهذا يدل على جودة التوفيق وقوة القدرة التفسيرية للنموذج والعكس صحيح.

-إذاكان: "RSS=0" فهذا يعني أن النموذج هو عبارة عن خط مستقيم ولا توجد أخطاء في النموذج وهي حالة نادرة الحدوث.

-أما إذاكان: " ESS=0" فهذا يعني ان المتغيرات المستقلة لا تفسر إطلاقا المتغير التابع وهي حالة نادرة الحدوث أيضا.

يمكن إيجاد معامل التحديد بالانحرافات كما يلي:

حيث : $y_i = Y_i - \bar{Y} \quad , \quad x_i = X_i - \bar{X} \qquad \forall i = 1.......n$

$$R^2 = \frac{\hat{\beta}'x'y}{y'y} = \frac{\hat{\beta}'x'y}{\sum y^2}$$

7-2 ايجاد معامل التحديد المصحح $\bar{R}^2$:

إن إضافة متغيرات مستقلة جديدة إلى المعادلة يؤدي إلى رفع قيمة R^2 ، وذلك لثبات قيمة المقام وتغير قيمة البسط بمقدار $(\hat{\beta}'X'Y)$ غير أن الاستمرار بإضافة المتغيرات المستقلة سيؤدي إلى انخفاض درجات الحرية $(n-k-1)$ ، مما يتطلب استخراج معامل التحديد المعدل أو المصحح $\bar{R}^2$ على النحو الآتي:

$$\bar{R}^2 = \left[1 - \frac{RSS \; / \; n-k-1}{TSS \; / \; n-1}\right]$$

$$\bar{R}^2 = \left[1 - \left(\frac{RSS}{TSS}\right)\frac{n-1}{n-k-1}\right]$$

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

نلاحظ أن: $R^2 \geq \bar{R}^2$ إذا كانت $k > 1$

يمكن أن نلاحظ أنه عندما يكون عدد المشاهدات n كبير نسبيا فإن R^2 يؤول إلى $\bar{R}^2$

$$\lim_{n\to\infty} \bar{R}^2 = 1 - \lim_{n\to\infty} \frac{n-1}{n-k-1}\left(1 - R^2\right) = R^2$$

نتيجة : إذا كان حجم العينة n كبيرا، فإن R^2 و $\bar{R}^2$ يقتربان قي قيمتهما، لكن في العينات الصغيرة، إذا كان عدد المتغيرات المستقلة كبيرا بالمقارنة مع حجم العينة، فإن $\bar{R}^2$ يقل بكثير على R^2 ، ويمكن أن يأخذ قيما سالبة، في هذه الحالة يجب شرحه على أساس أن قيمته تساوي الصفر.

مثال:نفس المثال السابق المطلوب إيجاد معامل التحديد مع التفسير R^2:

$$R^2 = \frac{ESS}{TSS} = \frac{TSS - RSS}{TSS} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{405.1}{408.5} = 0.99$$

$$\text{ESS} = \hat{\beta}'\left(X'\text{Y}\right) - n\bar{\text{Y}}^2$$

$$ESS = (15.99 \quad 0.65 \quad 1.11)\begin{pmatrix} 285 \\ 2804 \\ 1935 \end{pmatrix} - 10\left(\frac{285}{10}\right)^2 = 405.1$$

$$TSS = \sum Y_i^2 - n\overline{Y}^2 = 408.5$$

أو :

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{3.4}{408.5} = 0.99$$

التفسير : الاستيراد (Y) مفسر ب : 99% عن طريق الدخل الوطني (X_1) و السعر (X_2) وتبقى 01% تدخل ضمن هامش الخطأ وهي متغيرات أخرى لم تدرج في النموذج أو أخطاء إرتكبناها أثناء القياس ، على العموم هو هامش قليل جدا دلالة على قوة النموذج التفسيرية.

إيجاد معامل التحديد المصحح $\overline{R}^2$:

$$\overline{R}^2 = \left[1 - (1 - R^2)\frac{n-1}{n-k-1}\right]$$

$$\overline{R}^2 = \left[1 - (1 - 0.99)\frac{10-1}{10-2-1}\right] = 0.98$$

3-7 معامل الارتباط الجزئي: Partial Correlation

في بعض الظواهر والدراسات يوجد هناك عدد من المتغيرات (ثلاثة فأكثر) مرتبطة بعلاقة رياضية فيما بينها مثل: إنفاق أسرة يكون مرتبط بدخلها الشهري و عدد افرداها وكذلك حجم مبيعات سلعة معينة يرتبط بسعرها وحجم الدعاية لها وكذلك الفترة الزمنية للبيع ففي هذه الحالة، ولغرض حساب معامل الارتباط بين متغيرين اثنين في دراسة معينة مع وجود متغيرات أخرى نلجأ إلى حساب ما يسمى بالارتباط الجزئي .

الارتباط الجزئي هو: العلاقة الرياضية الصافية بين متغيرين اثنين فقط مع وجود متغيرات أخرى قيد الدراسة ويمكن حساب هذه العلاقة الرياضية من خلال معامل الارتباط الجزئي.

إن الفرق بينه وبين معامل الارتباط البسيط هو أن معامل بيرسون يستخرج العلاقة بين متغيرين اثنين لأي ظاهرة بدون يأخذ بنظر الاعتبار وجود متغيرات أخرى تؤثر في الظاهرة أو لا ، بينما معامل الارتباط الجزئي لا يأخذ بنظر الاعتبار وجود متغيرات أخرى تؤثر في الظاهرة فحسب وإنما يقوم باستبعاد أثرها لكي يستخرج الارتباط الصافي بين أي متغيرين.

أ- حساب معامل الارتباط الجزئي

ليكن لدينا نموذج انحدار متكون من متغيرتين مستقلتين كالاتي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

- **معامل الارتباط الجزئي بين** (X_1, Y):

$$r_{(Y,x_1)}(X_2) = \frac{r_{YX_1} - r_{YX_2}\, r_{X_1X_2}}{\sqrt{\left(1-r^2_{\;Yx_2}\right)\left(1-r^2_{\;x_1x_2}\right)}}$$

حيث تعني العبارة بين القوسين X_2 تثبيت المتغيرة عند مستواها المتوسط.

r_{YX1} : معامل الارتباط البسيط بين المتغيرتين (Y, X_1).

r_{YX2} : معامل الارتباط البسيط بين المتغيرتين (Y, X_2).

r_{X1X2} : معامل الارتباط البسيط بين المتغيرتين (X_1, X_2).

- **معامل الارتباط الجزئي بين** (X_2, Y):

$$r_{(Y,x_2)}(X_1) = \frac{r_{YX_2} - r_{YX_1}\, r_{X_1X_2}}{\sqrt{\left(1-r^2_{\;YX_1}\right)\left(1-r^2_{\;X_1X_2}\right)}}$$

حيث تعني العبارة بين القوسين X_1 تثبيت المتغيرة عند مستواها المتوسط.

ب–خصائص معامل الارتباط الجزئي:

* إن قيمة معامل الارتباط الجزئي تتراوح بين $(1,-1)$

* تفسر قيمته كما تفسر قيمة معامل الارتباط البسيط.

* إن معامل الارتباط الجزئي لأي متغيرين تكون إشارته مماثلة لإشارة معامل الارتباط البسيط بينهما.

مثال تطبيقي :

إنطلاقا من المعطيات السابقة أوجد معامل الارتباط الجزئي (X_1,Y) مع تثبيت X_2 ثم أوجد معامل الارتباط الجزئي (X_2,Y) مع تثبيت X_1.

- **معامل الارتباط الجزئي بين** (X_1,Y) :

$r_{(Y,x_1)}(X_2)=\dfrac{r_{YX_1}-r_{YX_2}\,r_{X_1X_2}}{\sqrt{\left(1-r^2_{Yx_2}\right)\left(1-r^2_{x_1x_2}\right)}}$ حيث تعني العبارة بين القوسين X_2 تثبيت المتغيرة عند مستواها المتوسط.

r_{YX1} : معامل الارتباط البسيط بين المتغيرتين (Y, X_1).

r_{YX2} : معامل الارتباط البسيط بين المتغيرتين (Y, X_2).

r_{X1X2} : معامل الارتباط البسيط بين المتغيرتين (X_1, X_2).

$$r_{(Y,x_1)}=\frac{COV(Y\,,\mathrm{X}_1)}{\sigma_Y\ \ \sigma_{X1}}=\frac{23.9}{\sqrt{40.85}\sqrt{14.4}}=0.98$$

$$r_{(Y,x_2)}=\frac{COV(Y\,,\mathrm{X}_2)}{\sigma_Y\ \ \sigma_{X2}}=\frac{22.5}{\sqrt{40.85}\sqrt{12.6}}=0.99$$

$$r_{(X_1,X_2)}=\frac{COV(X_1\,,\mathrm{X}_2)}{\sigma_{X1}\ \ \sigma_{X2}}=\frac{13.1}{\sqrt{14.4}\sqrt{12.6}}=0.97$$

$$r_{(Y,X_1)}(X_2)=\frac{r_{YX_1}-r_{YX_2}\,r_{X_1X_2}}{\sqrt{\left(1-r^2_{Yx_2}\right)\left(1-r^2_{x_1x_2}\right)}}=\frac{0.98-0.99*0.97}{\sqrt{\left(1-0.99^2\right)\left(1-0.97^2\right)}}=0.57$$

- **معامل الارتباط الجزئي بين** (X_2, Y):

$$r_{(Y,X_2)}(X_2)=\frac{r_{YX_2}-r_{YX_1}\,r_{X_1X_2}}{\sqrt{\left(1-r^2_{YX_1}\right)\left(1-r^2_{x_1x_2}\right)}}=\frac{0.99-0.98*0.97}{\sqrt{\left(1-0.98^2\right)\left(1-0.97^2\right)}}=0.81$$

نلاحظ من النتائج المتحصل عليها أن المتغيرة X_2 اكثر مساهمة في تفسير Y من المتغيرة X_1

8- اختبار الفرضيات لنموذج الخطي المتعدد :

يهدف هذا العنصر إلى توسيع معارفنا الأساسية لنموذج الانحدار وذلك بإجراء اختبار معنوية الانحدار المتعدد والمقدر باستخدام توزيع اختبار إحصاءه F ومقارنته باختبار t ومن ثم تقييم كفاءة الأداء العام لنموذج الانحدار المتعدد R^2 ومقارنته بمعامل التحديد المقدر المعدل $\bar{R}^2$ ، وكذلك اختبار العلاقة بين F و R^2 من خلال جدول تحليل التباين $ANOVA$, ثم علاقة R^2 بقيمة المتغير العشوائي $\sum\varepsilon_i^2$.

8-1- اختبار (t) لمعنوية المعالم :

يستخدم اختبار t لتقييم معنوية تأثير المتغيرات المستقلة $X_1, X_2 ..., ..., X_k$ في المتغير التابع y في نموذج الانحدار المتعدد يعتمد على نوعين من الفروض:

المعلمة ليس لها دلالة معنوية $H_0 : \beta_i = 0$ (فرضية العدم)

$i = 0,1,..., k$

المعلمة لها دلالة معنوية $H_1 : \beta_i \neq 0$ (الفرضية البديلة)

وبعد احتساب قيمة (t) تقارن مع قيمتها الجدولية لتحديد قبول او رفض فرضية العدم ومن ثم تقييم معنوية معلمات النموذج المقدر، والصيغة الرياضية لهذا الاختبار يمكن بيانها كما يلي :

$t_c = \dfrac{\hat{\beta}_i - \beta_i}{\hat{\sigma}_{\hat{\beta}_i}}$: القيمة المحسوبة.

وما دمنا نختبر فرضية العدم و تنص على انعدام β_i فإن قيمة t_c تصبح على الشكل التالي: $t_c = \dfrac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}}$، حيث

يتم قبول أو رفض الفرضية H_0 بمستوى معنوية $(\alpha\%)$ على أساس مقارنة $\lfloor t_c \rfloor$ مع القيمة المجدولة t_t حيث أن

: t_t يتم قراءتها من جدول ستودينت كالتالي: $t_{\left(n-k-1, \frac{\alpha}{2}\right)}$ حيث أن:

$\dfrac{\alpha}{2}$: مستوى المعنوية و k : عدد المتغيرات المستقلة في نموذج الانحدار الخطي المتعدد المراد دراسته.

إذا كانت $\left|\dfrac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}}\right| \leq t_{n-k-1, \frac{\alpha}{2}}$ ففي هذه الحالة المعلمة ليس لها معنوية إحصائية أي يساوي معنويا الصفر .

إذا كانت $\left|\dfrac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}}\right| > t_{n-k-1, \frac{\alpha}{2}}$ أي المعلمة لها معنوية إحصائية فهو يختلف معنويا عن الصفر.

ملاحظة: عندما يكون حجم العينة كبيرا ($n \geq 30$) فينبغي استعمال التوزيع الطبيعي و يمكن أخذ القيمة الحرجة $z_{\alpha/2}$ و ذلك بحساب المساحة المظللة للتوزيع الطبيعي.

8-2 اختبار فيشر F – Statistics لمعنوية النموذج المقدر:

يستهدف هذا الاختبار معرفة مدى معنوية العلاقة الخطية بين المتغيرات المستقلة $X_1, X_2, ...X_K$ على المتغير التابع Y، وكما هو الحال في الانحدار البسيط فإنه يعتمد على نوعين من الفروض:

(فرضية العدم) $H_0 : \beta_i = \beta_2 = \beta_3 = ...\beta_k = 0$ الانحدار ككل ليس له دلالة معنوي

(الفرضية البديلة) $H_1 : \exists$ معامل $\neq 0$ الانحدار ككل له دلالة معنوية

والصيغة الرياضية لهذا الاختبار هي :

$$F_c = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 \Big/ k}{\sum_{i=1}^{n}\varepsilon_i^2 \Big/ (n-k-1)} = \frac{\hat{\beta}'X'Y - n\bar{Y}^2 \Big/ k}{\sum_{i=1}^{n}\varepsilon_i^2 \Big/ (n-k-1)} \sim F_\alpha(k, n-k-1)$$

$$F_c == \frac{R^2 / k}{(1-R^2)/(n-k-1)} \sim F_\alpha(k, n-k-1)$$

وبعد احتساب قيمة F تقارن مع قيمتها الجدولية بدرجة حرية (k) و $(n-k-1)$ للبسط والمقام على التوالي ولمستوى معنوية معين .

- عندما تكون $F_{Tab} \langle |F_{Cal}|$ نرفض فرضية العدم (H_0) ونقبل الفرضية البديلة (H_1) مما يدل على أنه من بين معلمات النموذج يوجد واحد على الأقل يختلف عن الصفر أي أن هناك متغيرا مفسرا له تأثير جوهري على المتغير التابع بمعنى أن معادلة الانحدار المقدرة لها معنوية إحصائيا.
- عندما تكون $F_{Tab} \le |F_{Cal}|$ نقبل فرضية العدم (H_0)أي جميع المتغيرات التفسيرية لا تمارس أي تأثير على المتغير التابع وتكون معادلة الانحدار المقدرة غير معنوية إحصائيا.

3-8 جدول تحليل التباين **ANOVA**:

لغرض الوقوف على تأثير كل من (X_1)، (X_2) في المتغير التابع Y ، لابد من عمل جدول تحليل التباين لبيان أثر المتغيرين المستقلين (X_1)و (X_2)في النموذج.

جدول رفم 2: تحليل التباين

مصدر التباين	مجموع مربعات الخطأ	درجات الحرية	متوسط مربعات الخطأ	اختبار F
التغير الموضح من قبل (X_1) و (X_2) ESS	$\sum\left(\hat{Y}_i-\bar{Y}\right)^2$	k	$\sum\left(\hat{Y}_i-\bar{Y}\right)^2 / k$	$F_c=\dfrac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2 / k}{\sum_{i=1}^{n}\varepsilon_i^2 / (n-k-1)} \sim F_\alpha(k,n-k-1)$
التغير غير الموضح RSS	$\varepsilon'\varepsilon$	$n-k-1$	$\varepsilon'\varepsilon / n-k-1$	
الانحراف الكلي TSS	$\sum\left(Y_i-\bar{Y}\right)^2$	$n-1$		

مثال تطبيقي : اختبر معنوية المعلمات باستخدام اختبار t عند مستوى معنوية $\alpha=5\%$.

أ–اختبار معنوية ميل الدخل الوطني عند مستوى معنوية $\alpha=5\%$.

فرضيات هذا الاختبار هي:

(فرضية العدم) $H_0:\beta_1=0$ المعلمة ليس لها دلالة معنوية

(الفرضية البديلة) $H_1:\beta_1\neq 0$ المعلمة لها دلالة معنوية

يتم هذا لاختبار بإيجاد القيمة المحسوبة t_C وتساوي :

حيث أن: $$t_C=\frac{\left|\hat{\beta}_1-\beta_1\right|}{\hat{\sigma}_{\hat{\beta}_1}}$$

وما دمنا نختبر فرضية العدم وتنص على انعدام β_1 فإن قيمة t_C تصبح على الشكل التالي:

$$t_c=\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}=\frac{0.65}{0.24}=2.70$$

حيث أن : $t_t = t_{\left(n-k-1,\frac{\alpha}{2}\right)} = t_{\left(10-2-1,\ \frac{0.05}{2}\right)} = t_{(7\,,\,0.025)} = 2.365$

نلاحظ أن : $|t_C| \succ t_{(7\,,\,0.025)}$ نقبل الفرضية H_1 ونرفض الفرضية H_0 أي أن المعلمة لها معنوية إحصائية

أي أن الدخل الوطني يؤثر على الاستيراد.

ب–اختبار معنوية ميل السعر عند مستوى معنوية $\alpha = 5\%$.

فرضيات هذا الاختبار هي:

(فرضية العدم) $H_0 : \beta_2 = 0$ المعلمة ليس لها دلالة معنوية

(الفرضية البديلة) $H_1 : \beta_2 \neq 0$ المعلمة لها دلالة معنوية

يتم هذا لاختبار بإيجاد القيمة المحسوبة t_C وتساوي :

حيث أن: $$t_C = \frac{\left|\hat{\beta}_2 - \beta_2\right|}{\hat{\sigma}_{\hat{\beta}_2}}$$

وما دمنا نختبر فرضية العدم وتنص على انعدام β_2 فإن قيمة t_C تصبح على الشكل التالي:

$$t_c = \frac{\hat{\beta}_2}{\hat{\sigma}_{\hat{\beta}2}} = \frac{1.11}{0.26} = 4.26$$

حيث أن : $t_t = t_{\left(n-k-1,\frac{\alpha}{2}\right)} = t_{\left(10-2-1,\ \frac{0.05}{2}\right)} = t_{(7\,,\,0.025)} = 2.365$

نلاحظ أن : $|t_C| \succ t_{(7\,,\,0.025)}$ نرفض الفرضية H_0 أي أن المعلمة لها معنوية إحصائية أي أن السعر يؤثر

على الاستيراد.

ج–اختبر معنوية الحد الثابت عند مستوى معنوية $\alpha = 5\%$.

فرضيات هذا الاختبار هي:

(فرضية العدم) $H_0 : \beta_0 = 0$ المعلمة ليس لها دلالة معنوية

(الفرضية البديلة) $H_1 : \beta_0 \neq 0$ المعلمة لها دلالة معنوية

يتم هذا لاختبار بإيجاد القيمة المحسوبة t وتساوي :

حيث أن: $t_C = \dfrac{\left|\hat{\beta}_0 - \beta_0\right|}{\hat{\sigma}_{\hat{\beta}_0}}$

وما دمنا نختبر فرضية العدم وتنص على انعدام β_0 فإن قيمة t_C تصبح على الشكل التالي:

$$t_c = \frac{\hat{\beta}_0}{\hat{\sigma}_{\hat{\beta}_0}} = \frac{15.99}{0.80} = 19.89$$

حيث أن : $t_t = t_{\left(n-k-1, \frac{\alpha}{2}\right)} = t_{\left(10-2-1,\, \frac{0.05}{2}\right)} = t_{(7\,,\,0.025)} = 2.365$

نلاحظ أن : $|t_C| \succ t_{(7\,,\,0.025)}$ نرفض الفرضية H_0 و نأخد بديلتها H_1 أي أن المعلمة لها معنوية إحصائية.

د- اختبار معنوية الانحدار ككل باستخدام اختبار F عند مستوى معنوية $\alpha = 5\%$.

الصيغة الرياضية للفرضية المراد اختبارها كالأتي:

(فرضية العدم) $H_0 : \beta_1 = \beta_2 = 0$ الانحدار ككل ليس له دلالة معنوية

(الفرضية البديلة) $H_1 : \exists \beta_j \ / \ \beta_j \neq 0$ الانحدار ككل له دلالة معنوية

يتم أولا تحديد قيمة F المحسوبة كالتالي :

$$F_c = \frac{ESS/k}{RSS/(n-k-1)} \sim F_{(k,n-k-1)}$$

$$F_c = \frac{405.1/2}{3.4/(10-3)} = 346.5$$

أو : $$F_c = \frac{R^2/k}{\left(1-R^2\right)/(n-k-1)} = \frac{0.99/2}{(1-0.99)/7} = 346.5$$

في توزيع F القيمة المجدولة لإحصائية *Fischer* في هده الحالة تعتمد على درجتي حرية 2 (في البسط) و $n-3$ (في المقام).

$$F_t = F_{k,n-k-1} = F_{(2,7)} = 4.74$$

إذا كان $F_C \geq F_t$ نرفض H_0 و بالتالي نأخد بديلتها H_1 والانحدار ككل له دلالة معنوية إحصائية أي أن: الدخل الوطني (X_1) والسعر (X_2) يؤثران معا في الاستيراد (Y_i) .

" جدول رقم 3: تحليل التباين

مصدر التباين	مجموع مربعات الخطأ	درجات الحرية	متوسط مربعات الخطأ	اختبار F
الانحراف الموضح من قبل (X_1) و (X_2) **ESS**	$\sum(\hat{Y}_i-\bar{Y})^2=405.1$	$k=2$	$\sum(\hat{Y}_i-\bar{Y})^2 / k$	$F_c=\dfrac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2 / k}{\sum_{i=1}^{n} e_i^2 / (n-k-1)} \sim F(k, n-k-1)$
الانحراف غير الموضح **RSS**	$e'e=RSS=3.4$	$n-k-1=10-2-1=7$	$e'e / n-k-1$	
الانحراف الكلي **TSS**	$\sum(Y_i-\bar{Y})^2=408.5$	$n-1=10-1=9$		

9-اختبار صلاحية النموذج لكل فترة (اختبار CHOW)

أي اختبار ما إذا كان النموذج صالحاً لكل الفترة الزمنية المستخدمة ، خاصة إذا تعلق الأمر بالبيانات في شكل سلاسل زمنية، حيث غالباً ما تحدث تغيرات جوهرية اقتصادية كانت أم سياسية من شأنها أن تؤثر على معلمات النموذج، ومن ثم يصبح النموذج متكون من نقطة انعطاف التي حدثت فيها التغيرات، وبالتالي فالنموذج يصبح غير قادر على تفسير كل المدة الزمنية المستخدمة ونتائجه تكون مظللة، فيصبح من الغير الممكن الاعتماد على نموذج واحد لتمثيل لكل فترة، وبالتالي فلابد من تقسيم السلسلة الزمنية المستخدمة إلى قسمين ويتم تقدير نموذج الانحدار في كل فترة .

نقدر النموذج انطلاقا من عينتين جزئيتين n_1 و n_2 مع $n=n_1+n_2$، حيث :

$$\hat{Y}_i=\hat{\beta}_0^{(1)}+\hat{\beta}_1^{(1)}X_{i1}+\hat{\beta}_2^{(1)}X_{i2}+\ldots+\hat{\beta}_k^{(1)}X_{ik}$$
$$\hat{Y}_i=\hat{\beta}_0^{(2)}+\hat{\beta}_1^{(2)}X_{i1}+\hat{\beta}_2^{(2)}X_{i2}+\ldots+\hat{\beta}_k^{(2)}X_{ik}$$

نختبر الفرضيات التالية :

(فرضية العدم): H_0 النموذج يحافظ على نفس بنائه الهيكلي

(الفرضية البديلة): H_1 الانحدارين مختلفين

إن هذا النوع من الاختبارات يسمى باختبار المساواة ما بين مجموعات و بين معالم الانحدار أو اختبارات التغير الهيكلي أو اختبار CHOW وهو إحدى التطبيقات المهمة لتحليل التباين.

تعرف إحصائية فيشر كما يلي :

$$F_c = \frac{\left[RSS - \left(RSS^1 + RSS^2\right)\right]/df_1}{\left(RSS^1 + RSS^2\right)/df_2}$$

مع :

$$df_1 = (n-k-1) - \left[(n_1 - k - 1) + (n_2 - k - 1)\right] = k+1$$
$$df_2 = (n_1 - k - 1) + (n_2 - k - 1) = n - 2(k+1)$$

إذا كانت $F_c \leq F_\alpha\left(k+1, n-2(k+1)\right)$ ، ففي هذه الحالة نقبل الفرضية H_0 ، أي أن النموذج يحافظ على نفس بنائه الهيكلي.

10- استخدام المتغيرات الصورية

هناك بعض المتغيرات مهمة في تحليل الظواهر الاقتصادية و لكن من طبيعة نوعية أي لا تأخذ إلا القيمتين 0 أو 1. ويستعمل هذا النوع من المتغيرات عندما نريد دمج عامل مستقل ثنائي: "الظاهرة حدثت أو لم تحدث" أو أيضا عندما يكون العامل المستقل ذا طابع نوعي: "ذكر أو أنثى"...

الظاهرة تحدث:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$

الظاهرة لا تحدث:

$$Y_i = \alpha_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$

يمكن كتابة هاتين المعادلتين في معادلة واحدة:

$$Y_i = \beta_0 + (\alpha_0 - \beta_0)D_i + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$

مع $D_i = 1$ عندما لا تحدث الظاهرة

و $D_i = 0$ عندما تحدث الظاهرة.

في هذه الحالة نقوم بإدماج متغير مفسر إضافي للنموذج المحدد الأصلي و تطبيق الطرق الكلاسيكية للتقدير ويطلق على المتغير D_i الذي يظهر في المعادلة المتغير الصوري Dummy variable، وعلى هذا الأساس إن استخدام المتغيرات الصورية تعميم قوي لتحليل الانحدار، فإنه يسمح لنا تعميم مجال تحليلنا للانحدار ليشمل متغيرات مهمة لا يمكننا وصفها في وحدات كمية. فباستخدام هذا النوع من المتغيرات يمكن الأخذ في الحسبان تأثير العوامل النوعية المهمة التي تؤثر على المتغير التابع.

11–التنبؤ Forecasting

الفصل الثالث: تحليل الانحدار الخطي المتعدد -Multiple Linear Regression

يمكن التنبؤ بالملاحظات المستقبلية، (أو خارج العينة) لشعاع الملاحظات الخاصة بالمتغير التابع $Y_t : t=1,2,...,T$ ، وذلك بمعرفتنا لمصفوفة الملاحظات المستقبلية للمتغيرات المستقلة، حيث ليكن النموذج الخطي العام خلال العينة T والمقدر على الشكل: $\hat{Y} = X\hat{\beta}$

ومقدر المربعات الصغرى العادية $\hat{\beta} = (X'X)^{-1}X'Y = AY$ ، فيكون المقدر بملاحظة واحدة في المستقبل هو:

$$\hat{Y}_T(1) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+1,1} + \hat{\beta}_2 X_{T+1,2} + \hat{\beta}_3 X_{T+1,3} + ... + \hat{\beta}_k X_{T+1,k}$$

التنبؤ بعد فترتين في المستقبل:

$$\hat{Y}_T(2) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+2,1} + \hat{\beta}_2 X_{T+2,2} + \hat{\beta}_3 X_{T+2,3} + ... + \hat{\beta}_k X_{T+2,k}$$

التنبؤ بعد الفترة h في المستقبل:

$$\hat{Y}_T(h) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+h,1} + \hat{\beta}_2 X_{T+h,2} + \hat{\beta}_3 X_{T+h,3} + ... + \hat{\beta}_k X_{T+h,k}$$

حيث $h=1,2,....,H$ يسمى بأفق التنبؤ.

وعليه نصل إلى التنبؤ بالفترة H في المستقبل :

$$\hat{Y}_T(H) = \hat{\beta}_0 + \hat{\beta}_1 X_{T+H,1} + \hat{\beta}_2 X_{T+H,2} + \hat{\beta}_3 X_{T+H,3} + ... + \hat{\beta}_k X_{T+H,k}$$

إذن يكون شعاع القيم التقديرية إذا أردنا التنبؤ بمجموعة من الملاحظات المستقبلية بفترة تساوي H ملاحظة مرة واحدة :

$$\hat{Y}_T(H) = \underset{(H\times 1)}{\begin{pmatrix} \hat{Y}_T(1) \\ \hat{Y}_T(2) \\ \vdots \\ \vdots \\ \hat{Y}_T(H) \end{pmatrix}}$$

أما مصفوفة ملاحظة المشاهدات المستقبلية للمتغيرات المستقلة هي:

$$X_{T+H} = \underset{(H\times(k+1))}{\begin{pmatrix} 1\ X_{T+1,1} & X_{T+1,2} & \cdots & X_{T+1,k} \\ 1\ X_{T+2,1} & X_{T+2,2} & \cdots & X_{T+2,k} \\ \vdots\ \ \vdots & \vdots & \vdots & \vdots \\ \vdots\ \ \vdots & \vdots & \ddots & \vdots \\ 1 X_{T+H,1} & X_{T+H,2} & \cdots & X_{T+H,k} \end{pmatrix}}$$

ومنه يمكن كتابة النموذج الخطي العام المتنبأ به على الشكل: $Y_{T+H} = X_{T+H}\beta + \varepsilon_{T+H}$ ، كما يكون النموذج المقدر من الشكل: $\hat{Y}_T(H) = X_{T+H}\hat{\beta}$ ، ويكون متوسط مقدر التنبؤ هو:

$$E\left(\hat{Y}_T(H)\right) = X_{T+H}E\left(\hat{\beta}\right) = X_{T+H}\beta = E\left(Y_{T+H}\right)$$

ومنه نقول أن $\hat{Y}_T(H)$ هو تنبؤ خطي غير متحيز للعبارة:

$$X_{T+H}\beta = E(Y_{T+H})$$

ليكون التباين :

$$\begin{aligned}\operatorname{var}\left(\hat{Y}_T(H)\right) &= \left[\left(\hat{Y}_T(H) - X_{T+H}\beta\right)\left(\hat{Y}_T(H) - X_{T+H}\beta\right)'\right] \\ &= \sigma_\varepsilon^2 X_{T+H}\left(X'X\right)^{-1} X'_{T+H}\end{aligned}$$

لنعرف شعاع أخطاء التنبؤ:

$$\begin{aligned}\hat{\varepsilon}_{T+H} &= Y_{T+H} - \hat{Y}_T(H) \\ E\left(\hat{\varepsilon}_{T+H}\right) &= E\left(Y_{T+H} - \hat{Y}_T(H)\right) = 0\end{aligned}$$

أما تباين شعاع أخطاء التنبؤ فهو:

$$\operatorname{var}\left(\hat{\varepsilon}_{T+H}\right) = \operatorname{var}\left(Y_{T+H} - \hat{Y}_T(H)\right) = E\left[\left(-X_{T+H}\left(\hat{\beta} - \beta\right) + \varepsilon_{T+H}\right)\left(-X_{T+H}\left(\hat{\beta} - \beta\right) + \varepsilon_{T+H}\right)'\right]$$

و يعطى في الأخير بالعبارة:

$$\operatorname{var}\left(\hat{\varepsilon}_{T+H}\right) = \sigma_\varepsilon^2 X'_{T+H}\left(X'X\right)^{-1} X_{T+H} + \sigma_\varepsilon^2 I_H$$

هذا التنبؤ هو أحسن تنبؤ خطي غير متحيز (BLUP) يمكن الحصول عليه، حيث إذا عرفنا $\tilde{Y}_T(H)$ تنبؤ آخر خطي لعينة ملاحظات المتغير التابع مع متوسط خطا التنبؤ مساو للصفر $E\left(\tilde{\varepsilon}_{T+H}\right) = E\left(Y_{T+H} - \tilde{Y}_T(H)\right) = 0$ تكون لدينا المتراجحة:

$$\operatorname{var}\left(Y_{T+H} - \tilde{Y}_T(H)\right) - \operatorname{var}\left(Y_{T+H} - \hat{Y}_T(H)\right) \geq 0$$

ومنه يمكن القول أن $\hat{Y}_T(H) = X_{T+H}\hat{\beta}$ هو أحسن تنبؤ خطي غير متحيز.

وتكون اختبارات التنبؤ عن طريق إيجاد التوزيع الذي يعتبر فرضية العدم، والقائلة بأن النموذج الخطي العام يبقى محافظا على شكله من الملاحظة الأولى إلى الملاحظة $T+H$ في المستقبل، أي نفترض عدم تغير البناء الهيكلي للنموذج،

$$H_0: \quad \hat{Y} = X\hat{\beta} \quad t = 1,2,3,\ldots\ldots T, T+1,\ldots, T+h,\ldots., T+H$$

وذلك ضد الفرضية البديلة، والتي تقول أن نموذج العينة الأولى T يختلف عن نموذج التنبؤ للفترة H.

$$F = \frac{\left(Y_{T+H} - \hat{Y}_T(H)\right)'\left[X_{T+H}\left(X'X\right)^{-1} X'_{T+H} + I_H\right]^{-1}\left(Y_{T+H} - \hat{Y}_T(H)\right)/H}{\hat{\sigma}_\varepsilon^2} \quad \rightsquigarrow \quad F_{H,T-k-1}$$

وإذا كان $H = 1$ يصبح التوزيع أعلاه على الشكل:

$$F = \frac{\left(Y_{T+1} - \hat{Y}_T(1)\right)'\left[X_{T+1}\left(X'X\right)^{-1} X'_{T+1} + 1\right]^{-1}\left(Y_{T+1} - \hat{Y}_T(1)\right)}{\hat{\sigma}_\varepsilon^2} \quad \rightsquigarrow \quad F_{1,T-k-1}$$

مثال:

الفصل الثالث: تحليل الانحدار الخطي المتعدد -Multiple Linear Regression

لدينا 10 مشاهدات لمتغير تابع رفقة متغيرين مفسرين كالتالي:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \; \ldots i = 1, 2, \ldots, 10.$$

i	Y_i	X_{1i}	X_{2i}
1	10	5	22
2	11	8	19
3	9	4	17
4	7	2	24
5	13	8	22
6	15	7	20
7	12	7	19
8	10	5	18
9	6	5	25
10	14	11	23

المطلوب:

1. أحسب مقدرات المعاملات بطريقة المربعات الصغرى مع التفسير الاقتصادي.
2. أحسب التباينات و الانحرافات المقدرة للمعلمات.
3. اختبر معنوية الميول الحدية مع تحديد مجالات الثقة عند مستوى احتمال 95%
4. اختبر جودة النموذج المقدر ككل من خلال جدول تحليل التباين مع حساب معامل التحديد و معامل التحديد المرجح.
5. تنبأ للقيمة i=11 علما أن $x_{1,11} = 10$ و $X_{2,11} = 19$ مع تحديد مجال الثقة للقيمة المتنبأ بها.

الحل:

1-حساب المقدرات:

بداية علينا باستخراج المصفوفتين $(X'X)^{-1}$ و$X'Y$ على النحو التالي:

i	Y	X_1	X_2	X_1^2	X_2^2	X_1x_2	X_1y	X_2y
1	10	5	22	25	484	110	50	220
2	11	8	19	64	361	152	88	209
3	9	4	17	16	289	68	36	153
4	7	2	24	4	576	48	14	168
5	13	8	22	64	484	176	104	286
6	15	7	20	49	400	140	105	300
7	12	7	19	49	361	133	84	228
8	10	5	18	25	324	90	50	180
9	6	5	25	25	625	125	30	150
10	14	11	23	121	529	253	154	322

Σ	107	62	209	442	4433	1295	715	2216

منه لدينا المصفوفتين كالتالي:

$$(X'X) = \begin{pmatrix} 10 & 62 & 209 \\ 62 & 442 & 1295 \\ 209 & 1295 & 4433 \end{pmatrix} ; X'Y = \begin{pmatrix} 107 \\ 715 \\ 2216 \end{pmatrix}$$

و بنفس الطريقة في التمرين السابق نستخرج مرافق و معكوس المصفوفة (X'X) حيث يكون شعاع المقدرات على النحو التالي:

$$\hat{\beta} = (X'X)^{-1} X'Y = \frac{1}{37376}\begin{pmatrix} 282361 & -4191 & -12088 \\ -4191 & 649 & 8 \\ -12088 & 8 & 576 \end{pmatrix}\begin{pmatrix} 107 \\ 715 \\ 2216 \end{pmatrix} = \begin{pmatrix} 11.4794 \\ 0.8916 \\ -0.3018 \end{pmatrix}$$

بالتالي تكون معادلة الانحدار المقدرة على النحو التالي:

$$Y_i = 11.4794 + 0.8916\, x_{1i} - 0.3018\, x_{2i} + \varepsilon_i$$

التفسير الاقتصادي: من خلال المعادلة أعلاه نلاحظ أن المتغيرة Y لها علاقة طردية مع X_1 حيث كلما ارتفع X_1ب 10% سيؤدي ذلك إلى ارتفاع Y ب 8.916 %، كما نلاحظ وجود علاقة عكسية بين Y و X_2 حيث إذا ارتفع X_2 ب 10 % أدى إلى انخفاض Y ب 3.018%.

2-حساب التباينات و الانحرافات المقدرة:

بداية لابد من حساب التباين المقدر للأخطاء الذي بدوره يتطلب حساب مجموع مربعات الأخطاء الذي نستخلصه من الجدول التالي:

i	Y	$\hat{Y}$	$\hat{\varepsilon}_i$	$\hat{\varepsilon}_i^2$
1	10	9.2978	0.7022	0.49308484
2	11	12.878	-1.878	3.526884
3	9	9.9152	-0.9152	0.83759104
4	7	6.0194	0.9806	0.96157636
5	13	11.9726	1.0274	1.05555076
6	15	11.6846	3.3154	10.99187716
7	12	11.986	0.0136	0.00018496
8	10	10.5054	-0.505	0.255025
9	6	8.3924	-2.392	5.72357776
10	14	14.3456	-0.3456	0.11943936
Σ	/	/	/	23.96479124

$$\hat{\sigma}^2_{\hat{\varepsilon}} = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-k-1} = \frac{23.96479124}{10-2-1} = 3.4235$$

$$\hat{\Omega}_{\hat{\beta}} = \hat{\sigma}^2_{\hat{\varepsilon}} (X'X)^{-1} = 3.4235 \begin{pmatrix} 282361 & . & . \\ . & 649 & . \\ . & . & 576 \end{pmatrix} \frac{1}{37376}$$

حيث تساوي التباينات و الانحرافات المقدرة للمعالم مايلي:

$$\hat{\sigma}^2_{\hat{\beta}0} = 3.4235 * \frac{282361}{37376} = 25.8632 \rightarrow \hat{\sigma}_{\hat{\beta}0} = 5.0855$$

$$\hat{\sigma}^2_{\hat{\beta}1} = 3.4235 * \frac{649}{37376} = 0.0594 \rightarrow \hat{\sigma}_{\hat{\beta}1} = 0.2437$$

$$\hat{\sigma}^2_{\hat{\beta}2} = 3.4235 * \frac{576}{37376} = 0.0527 \rightarrow \hat{\sigma}_{\hat{\beta}2} = 0.2295$$

3-اختبار معنوية الميول الحدية:

بداية نقوم بحساب قيم المحسوبة لاختبار الفرضيات التالية:

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases} \begin{cases} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{cases}$$

حيث لغرض اختبار الفرضيتين العديمتين تكون إحصائية ستيودنت المحسوبة على النحو التالي:

$$t^*_{\hat{\beta}1} = \frac{|\hat{\beta}1|}{\hat{\sigma}_{\hat{\beta}1}} = \frac{0.8916}{0.2437} = 3.65$$

$$t^*_{\hat{\beta}2} = \frac{|\hat{\beta}2|}{\hat{\sigma}_{\hat{\beta}2}} = \frac{0.3018}{0.2295} = 1.31$$

أما القيمة الجدولية فهي $t^{0.05}_{n-k-1} = t^{0.05}_{7} = 2.365$ ففي ما يخص β_1 فإن قيمة ستيودنت المحسوبة أكبر من القيمة الجدولية مما يعني قبول الفرضية العدمية و نقول أن a_1 معنوية عند مستوى احتمال5% أي أن المعلمة تختلف معنويا عن الصفر و أن المتغيرة X_1 تِثر معنويا على المتغير التابع Y، أما فيما يخص المعلمة β_2 فإن قيمة ستيودنت المحسوبة أصغر من القيمة الجدولية مما يعني قبول الفرضية العدمية و القول أن a_2 غير معنوية عند مستوى احتمال 5% أي أن المعلمة لا تختلف معنويا عن الصفر و أن المتغيرة X_2 لا تؤثر معنويا التابعY.

في حين يتم تحديد مجالات الثقة كالتالي:

$$\beta_1 \in \left[\hat{\beta}_1 \pm t^{0.05}_{n-k-1} * \hat{\sigma}_{\hat{\beta}\,1}\right] \Rightarrow \beta_1 \in [0.8916 \pm 2.365 * 0.2437] \rightarrow \beta_1 \in [0.3152,1.4679]$$

$$\beta_2 \in \left[\hat{\beta}_2 \pm t^{0.05}_{n-k-1} * \hat{\sigma}_{\hat{\beta}\,2}\right] \Rightarrow \beta_2 \in [-0.3018 \pm 2.365 * 0.2295] \rightarrow \beta_2 \in [-0.8445,0.2409]$$

أهم ملاحظة من مجالي الثقة أعلاه هو أن مجال الثقة الول لا يتضمن القيمة الصفر مما يؤكد أن المعلمة a_1 لا يمكن ان تساوي الصفر عند مستوى الثقة 95 %،في حين نلاحظ العكس فيما يخص المعلمة a_2 التي يتضمن مجال ثقتها القيمة الصفر الأمر الذي يعني وجود احتمال كبير أن تساوي هذه المعلمة الصفر عند مستوى ثقة 95%،و هذه النتائج تؤكد نتائج اختبار المعنوية حيث المعلمة a_1 معنوية عنددرجة معنوية 5 % في حين المعلمة a_2 غير معنوية عند درجة 5%.

4-اختبار جودة النموذج ككل:

من خلال نتائج الأسئلة الماضية لدينا مجموع مربعات الأخطاء $\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = SSR = 23.9648$ و بالتالي

ننتقل إلى حساب مجموع مربعات الانحرافات الكلية SST من خلال الجدول التالي:

i	Y	$\bar{Y}$	$Y_i - \bar{Y}$	$(Y_i - \bar{Y})^2$
1	10	10.7	-0.7	0.49
2	11	10.7	0.3	0.09
3	9	10.7	-1.7	2.89
4	7	10.7	-3.7	13.69
5	13	10.7	2.3	5.29
6	15	10.7	4.3	18.49
7	12	10.7	1,3	1.69
8	10	10.7	-0,7	0.49
9	6	10.7	-4,7	22.09
10	14	10.7	3,3	10.89
Σ	/		/	76.1

هذا يعني أن SST=$(Y_i - \bar{Y})^2 = 76.1$ حيث من خلاله نستنتج:

SSE = SST – SSR = 76.1 – 23.9648 = 52.1352

فيكون جدول تحليل التباين كالتالي:

مصدر التغير	مجموع مربعات الانحرافات	عدد درجات الحرية	متوسط مربعات الانحرافات
X_1 , X_2	SSE = 52.1252	2	26.0676
البواقي	SSR = 23.9648	7	3.4235
المجموع	SST = 76.1	9	

حيث لدينا الفرضيتين التاليتين:

$$\begin{Bmatrix} H_0: \beta_1 = \beta_2 = 0 \\ H_1: \beta_1 \neq \beta_2 \neq 0 \end{Bmatrix}$$

منه يمكن حساب إحصائية المحسوبة على النحو التالي:

$$F_c = \frac{26.0676}{3.4235} = 7.61$$

أما القيمة الجدولية فهي $F_{2.7}^{0.05} = 4.74$ ، و بما أن القيمة المحسوبة أكبر من الجدولية فنقبل الفرضية البديلة و نرفض الفرضية العدمية و نقول أن النموذج المقدر معنوي ككل عند مستوى احتمال 5%.

أما لحساب معامل التحديد فنستعمل العلاقة التالية:

$$R^2 = \frac{SSE}{SST} = \frac{25.1252}{76.1} = 0.6849$$

في حين يحسب معامل التحديد المرجح كالتالي:

$$\overline{R}^2 = 1 - \left[\frac{n-1}{n-k-1}\left(1-R^2\right)\right] = 1 - \frac{SSR/_{(n-k-1)}}{SST/_{(n-1)}}$$

$$\overline{R}^2 = 1 - \left[\frac{n-1}{n-k-1}\left(1-R^2\right)\right] = 1 - \left[\frac{10-1}{10-2-1}(1-0.6849)\right] = 0.59511$$

$$\overline{R}^2 = 1 - \frac{SSR/_{(n-k-1)}}{SST/_{(n-1)}} = 1 - \frac{23.9648/_{(10-2-1)}}{76.1/_{(10-1)}} = 0.59511$$

5-التنبؤ:

لهذا الغرض نقوم يتطبيق العلاقة التالية لمجال الثقة:

$$Y_{T+h} \in \left[\hat{Y}_T(h) \pm t_{n-k-1}^{\alpha/2}.\hat{\delta}_\varepsilon\sqrt{1 + X'_{t+h}(X'X)^{-1}X_{t+h}}\right]$$

حيث من خلال المعطيات لدينا:

$$X_{t+h} = \begin{pmatrix} 1 \\ 10 \\ 19 \end{pmatrix} \rightarrow X'_{t+h} = (1 \quad 10 \quad 19)$$

و لدينا التباين لخطأ القيمة المتنبأ بها يحسب على النحو التالي:

$$\hat{\sigma}^2_{\hat{\varepsilon}11} = \hat{\sigma}^2_{\hat{\varepsilon}}[1 + X'_{t+h}(X'X)^{-1}X_{t+h}]$$

$$\hat{\sigma}^2_{\hat{\varepsilon}11} = 3.4235\left[1 + (1 \quad 10 \quad 19)\begin{pmatrix} \frac{282361}{37376} & \frac{-4191}{37376} & \frac{-12088}{37376} \\ \frac{-4191}{37376} & \frac{649}{37376} & \frac{8}{37376} \\ \frac{-12088}{37376} & \frac{8}{37376} & \frac{576}{37376} \end{pmatrix}\begin{pmatrix} 1 \\ 10 \\ 19 \end{pmatrix}\right]$$

$$\hat{\sigma}^2_{e11} = 3.4235 * 1.3988 = 4,7888 \rightarrow \hat{\sigma}_{\hat{\varepsilon}11} = 2.188$$

نعوض الآن في معادلة مجال فنتحصل على:

$$Y_{T+h} \in \left[\hat{Y}_T(h) \pm t_{n-k-1}^{\alpha/2} * 2.1883\right]$$

كما أن قيمة ستيودنت الجدولية $t_{n-k-1}^{\alpha/2} = t_7^{0.025} = 2.365$.

أما القيمة المتنبأ بها $\hat{y}_{t+h}$ فتحسب على النحو التالي:

$$\hat{y}_{t+h} = \hat{y}_{11} = 11.4794 + 0.8916 * (10) - 0.3018 * (19) = 14.6612$$

و أخيرا يكون مجال الثقة كالتالي:

$$Y_{T+h} \in [14.6612 \pm\ 2.365 * 2.1883]\ \rightarrow Y_{T+h} \in [9.485\ ,19.836]$$

الفصل الرابع: مشاكل القياس الاقتصادي-Economic Measurement Problems

نماذج الانحدار بأشكالها المختلفة تعاني عادة من مشاكل قياسية متعددة ويؤدي وجود هذه المشاكل إلى اختلال أحد (أو كل) افتراضات طريقة المربعات الصغرى العادية، وتصبح هذه الطريقة غير ملائمة لتقدير معلمات العلاقات الاقتصادية، لذلك يتعين في هذه الحالة البحث عن طرق قياسية أخرى أكثر ملائمة وحتى نختبر مدى توفر هذه الافتراضات يتعين إجراء بعض الاختبارات مستخدمين بعض المعايير القياسية، وسوف نتعرض في هذا الفصل لأربعة مشاكل قياسية وهي: مشكلة التعدد الخطي، مشكلة عدم ثبات التباين و مشكلة الارتباط الذاتي للأخطاء إضافة إلى مشكلة ارتباط المتغير المستقل بالأخطاء.

1–التعدد الخطيMulticollinearity :

يرجع مصطلح التعدد الخطي الى المقال الذي نشره Frisch في سنة 1934 حيث وضح فيه مفهوم التعدد الخطي بأن المتغيرات التي تتعامل معها قد تكون واقعة تحت تأثير علاقتين أو أكثر، بعبارة أخرى يشير مصطلح الانحدار الخطي المتعدد إلى وجود ارتباط خطي بين عدد من المتغيرات التفسيرية في نموذج الانحدار، وبذلك يتم خرق أحد فرضيات نموذج الانحدار الخطي المتعدد "أي أن لا يكون هناك ارتباطاً خطياً متعدداً بين المتغيرات المستقلة"، ومن ثم فإن مشكلة التعدد الخطي المتعدد لا توجد في حالة الانحدار البسيط و إنما يوجد فقط في حالة الانحدار المتعدد.

عند وجود علاقة ارتباط تامة بين متغيرين تفسيريين أو بين جميع المتغيرات التفسيرية المكونة للنموذج يصبح محدد المصفوفة $(X'X)$ يساوي صفرا، حيث يستحيل إيجاد معكوس المصفوفة $(X'X)$، وبالتالي عدم إمكانية استخدام طريقة المربعات الصغرى الاعتيادية (OLS)، أو أن يكون محدد المصفوفة $(X'X)$ قريبا من الصفر في حالة الارتباط غير التام الذي معه تضمحل قدرة (OLS) على عكس الخصائص الحقيقية لمعلمات النموذج ويكون النموذج ذا قدرة تنبؤية ضعيفة.

على سبيل المثال إذا كان هناك نموذج ارتباط خطي متعدد يحتوي على متغيرين مستقلين : X_1 و X_2 التالي:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_t$$

إذا حدث ارتباط بين X_1 و X_2 كما يلي :

$$X1 = AX_2$$

نقول أن هناك مشكلة ارتباط خطي متعدد، ومعامل الارتباط البسيط بين المتغيرين المستقلين يساوي الواحد الصحيح:

$$r_{X1X2}=\frac{\sum\left(X_{1i}-\overline{X}_1\right)\left(X_{2i}-\overline{X}_2\right)}{\sqrt{\sum\left(X_{1i}-\overline{X}_1\right)}\sqrt{\sum\left(X_{2i}-\overline{X}_2.\right)}.}$$

باستخدام معامل الارتباط r_{X1X2} أو مربعها معامل التحديد r^2_{X1X2} نحدد إذا كان هناك ارتباط خطي متعدد.

- إذا كان معامل الارتباط $r_{X1X2}=0$ يساوي الصفر فانه ليس هناك مشكلة ارتباط خطي متعدد.
- إذا كانت هناك ارتباط كامل آي أن $r_{X1X2}=1$ فنول أن هناك ارتباط خطي متعدد تام، وعند حدوث الارتباط الخطي التام لا نستطيع إجراء التقدير باستخدام المربعات الصغرى العادية ، ولكن هذه المشكلة لا تحدث كثيرا في الدراسات العملية إلا في ظروف استثنائية ويمكن معالجتها بحذف أحد المتغيرات لان الآخر يقوم مقامه..

1-1 أسباب التعدد الخطي :

يرجع طهور مشكلة التعدد الخطي لعدة عوامل أهمها:

* تميل المتغيرات الاقتصادية لان تتغير معا عبر الزمن نظرا لأنها تتأثر جميعها بنفس العوامل، على سبيل المثال تزداد معظم المتغيرات الاقتصادية في أوقات الرواج أو النمو الاقتصادي السريع. فزيادة الطلب الكلي على السلع والخدمات يصاحبها زيادة في الإنتاج وزيادة في العمالة وزيادة في الدخل وزيادة في الاستثمار والاستهلاك والادخار وارتفاع الأسعار. والعكس يحدث في فترات الكساد.

* استخدام المتغيرات ذات الفجوة الزمنية كمتغيرات تفسيرية بنموذج الانحدار. فعلى سبيل المثال يظهر الدخل الجاري للفترة الحالية ودخل الفترة السابقة في دالة الاستهلاك كمتغيرات مستقلة تؤثر في استهلاك الفترة الحالية، فتأخذ دالة الاستهلاك الصيغة التالية:

$$C_t=\alpha+\beta_1Y_t+\beta_2Y_{t-1}+\varepsilon_t$$

* بالرغم من أن مشكلة التعدد الخطي عادة ما تظهر في حالة استخدام بيانات سلسلة زمنية إلا أنها قد تظهر في بعض الحالات عند استخدام بيانات قطاعية. فعلى سبيل المثال يلاحظ انه في حالة استخدام بيانات قطاعية لمجموعة مؤسسات صناعية لتقدير دالة إنتاج، فان الكميات المستخدمة من العمل ورأس المال كمتغيرات مستقلة قد ترتبط بشدة. ويرجع هذه إلى أن المؤسسات الكبيرة عادة ما تستخدم كميات كبيرة من كل من العمل ورأس المال، في حين أن المؤسسات الصغيرة عادة ما تستخدم كميات قليلة من كل من العمل ورأس المال.

* يؤدي صغر حجم العينة بحيث يصبح عدد المشاهدات قريبا من عدد المتغيرات التفسيرية إلى ظهور مشكلة التعدد الخطي. وتسمى هذه المشكلة بمشكلة صغر حجم العينة.

1-2 نتائج التعدد الخطي:

أولا: حالة تعدد خطي تام:

* تصبح القيم المقدرة للمعلمات غير محددة.

* الأخطاء المعيارية للمعلمات المقدرة تصبح كبيرة كبرا لا نهائيا.

ثانيا : التعدد الخطي غير التام :

* تصبح المعلمات المقدرة غير دقيقة وان كان من الممكن في هذه الحالة تقدير قيم منفصلة لكل منها.

*كبر الأخطاء المعيارية للمعلمات المقدرة كبرا محددا.

* قد يكون هناك بعض المتغيرات ذات الأهمية الكبيرة في تفسير الظاهرة محل البحث، أي أن معامل التحديد لدالة الانحدار المقدرة باستخدام بيانات عنها يكون مرتفعا، إلا أن وجود ارتباط بينها قد يؤدي لتضخم الأخطاء المعيارية للمعلمات المقدرة مما قد يدفع الباحث لحذف بعض هذه المتغيرات مؤديا بذلك إلى انخفاض معامل التحديد وإضعاف المقدرة التفسيرية للنموذج، بالإضافة إلى سوء تعيين النموذج وما يترتب عليه من خطأ في التقدير يسمى بخطأ الحذف.

* يؤدي وجود التعدد الخطي إلى كبر معامل التحديد مع عدم معنوية المعلمات المقدرة.

* تصبح مقدرات طريقة المربعات الصغرى حساسة للتغيرات الطفيفة في البيانات.

1-3 اختبار وجود مشكلة التعدد الخطي:

1-3-1 اختبار Klein:

يبدو أن كلاين يقبل بان الارتباط الخطي المتعدد ليس بالضرورة مشكلة ما لم يكن ذلك الارتباط الخطي المتعدد اكبر نسبياً من درجة الارتباط المتعدد الكلي بين كل المتغيرات آنياً، ويعتقد كلاين أن الارتباط $r^2_{X1X2} \succ R^2_{YX1X1....Xm}$ الخطي يكون مؤذياً إذا كان الارتباط الداخلي يكون اكبر من الارتباط الكلي، ولكن يعاب على هذا الاختبار أن درجة الارتباط الداخلي بين المتغيرات المستقلة لا تعتبر معيارا دقيقا لمدى التأثير الذي يحدثه وجود التعدد الخطي على قيم المعلمات المقدرة وقيم الأخطاء المعيارية. فقد تكون معاملات الارتباط البسيطة بين المتغيرات المستقلة منخفضة بالرغم من وجود مشكلة تعدد خطي خطير.

1-3-2 طريقة التحليل الترافدي ل Frisch:

تكمن هذه الطريقة في إجراء انحدار للمتغير التابع على كل متغير مستقل على حدى، ومنه نحصل على كل الانحدارات الأولية، ثم نختار الانحدار الأولي الذي يعطي أفضل النتائج في ضوء المعايير المستخدمة، ثم نضيف تدريجياً متغيرات مفسرة أخرى و نختبر أثاره على الأخطاء المعيارية و على 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 أثر المتغير المستقل المضاف بشكل واضح على إشارات و قيم معلمات الانحدار لتكون قيم غير مقبولة اقتصادياً فإنه يمكننا القول بأن هذا مؤشر على وجود التعدد الخطي بشكل معقد.

إن التحليل الترافدي لFrisch ينص على تقدير كل الانحدارات الممكنة ما بين المتغيرات الموجودة بالعلاقة المدروسة آخذين كل متغير بالترتيب كمتغير تابع و اعتبار كل الانحدارات الممكنة لكل متغير في بقية المتغيرات و التي ندخلها تدريجياً في التحليل، و من الواضح أن التحليل الترافدي يتطلب منا حسابات كثيرة، و منه تكون المقارنات ما بين النتائج معقدة أكثر.

مثال تطبيقي حول اختبار Firsche:

حاولنا من خلال هذا المثال بناء نموذج اقتصادي قياسي للاستثمار الأجنبي المباشر في الجزائر خلال الفترة : 1985-2013.

أعطت نتائج التقدير الأولية النتائج التالية:

$$IDE = 112.62 + 0.22GDP - 6.53_3\, INF - 1.84IBS - 3.83TCH + e_i$$

$$\left(\overline{R}^2 = 0.88\right) \qquad (F_c = 43.62) \qquad \left(R^2 = 0.90\right) \qquad (DW = 1.84)$$

اختبار FIRSH : وفقا لهذا الاختبار يتم إتباع الخطوات التالية :

سوف نقوم باختبار كل متغير مفسر على حدى من أجل معرفة مدى تأثير كل من هذه المتغيرات.

1* انحدار الاستثمار الأجنبي المباشر على GDP خلال الفترة : (1985– 2013).

تحصلنا على النموذج التالي:

$$IDE = 162.94 + 0.22GDP + e_i$$

$$\text{"-2.53 "} \qquad \text{"13.14 "}$$

$(\bar{R}^2 = 0.88)$ $(F_c = 43.62)$ $(R^2 = 0.89)$ $(DW = 1.53)$

من خلال النموذج نلاحظ أن الاستثمار الأجنبي المباشر يرتبط بعلاقة طردية بالناتج المحلي الإجمالي، وهذا يتوافق مع توقعاتنا إذن معامل$(\hat{\beta}_1)$ له معنوية اقتصادية ، أما من الناحية الإحصائية نقبل معلمة الناتج المحلي الإجمالي، وهذا لأن إحصائية ستودنت $(t = 13.14)$ أكبر من القيمة المجدولة ل(t_t) ، أما فيما يخص معامل التحديد $(R^2 = 0.89)$ مما يؤكد على قدرة النموذج التفسيرية بين الناتج المحلي الإجمالي والاستثمار الأجنبي المباشر، كما أن إحصائية فيشر(F_c) تساوي 170.78 وهي أكبر من(F_t) بمستوى معنوية5 ٪ . إذن معامل الناتج المحلي GDP $(\hat{\beta}_1)$له معنوية إحصائية واقتصادية.

2* انحدار الاستثمار الأجنبي المباشر على INF خلال الفترة : (1985– 2013).

نحصل على النموذج الموالي:

$$IDE = 856.55 - 35.27INF + e_i$$

"5.51 " "3.37- "

$(\bar{R}^2 = 0.34)$ $(F_c = 172.78)$ $(R^2 = 0.35)$ $(DW = 1.53)$

من خلال النموذج نلاحظ أن الاستثمار الأجنبي المباشر يرتبط بعلاقة عكسية بالتضخم، وهذا يتوافق مع توقعاتنا إذن معامل$(\hat{\beta}_1)$ له معنوية اقتصادية، أما من الناحية الإحصائية نقبل معلمة التضخم، وهذا لأن إحصائية ستودنت $(t = 3.37)$ أكبر من القيمة المجدولة ل(t_t) ، أما فيما يخص معامل التحديد$(R^2 = 0.35)$ مما يؤكد على ضعف قدرة النموذج التفسيرية بين الناتج التضخم والاستثمار الأجنبي المباشر ،كما أن إحصائية فيشر(F_c) تساوي 172.78 وهي أكبر من(F_t) بمستوى معنوية5 ٪ . إذن معامل التضخم $(\hat{\beta}_1)$ له معنوية إحصائية واقتصادية.

3* انحدار الاستثمار الأجنبي المباشر على IBS "معدل الضريبة على أرباح الشركات " خلال الفترة : (1985– 2013).

تحصلنا على النموذج التالي:

$$IDE = 2218 - 45.13IBS + e_i$$

"6.98 " "-5.69 "

$(\bar{R}^2 = 0.58)$ $(F_c = 32.40)$ $(R^2 = 0.60)$ $(DW = 0.66)$

من خلال النموذج نلاحظ أن الاستثمار الأجنبي المباشر يرتبط بعلاقة طردية مع معدل الضريبة على أرباح الشركات ، وهذا يتوافق مع توقعاتنا إذن معامل$(\hat{\beta}_1)$ له معنوية اقتصادية، أما من الناحية الإحصائية نقبل معلمة معدل

الضريبة على أرباح الشركات، وهذا لأن إحصائية ستودنت $(t = 5.69)$ أكبر من القيمة المجدولة ل(t_t) ، أما فيما يخص معامل التحديد $(R^2 = 0.60)$ مما يؤكد على متوسط قدرة النموذج التفسيرية بين معدل الضريبة على أرباح الشركات والاستثمار الأجنبي المباشر ،كما أن إحصائية فيشر(F_c) تساوي 32.4 وهي أكبر من(F_t) بمستوى معنوية5 ٪ . إذن معامل معدل الضريبة على أرباح الشركات$(\hat{\beta}_1)$ IBSS له معنوية إحصائية واقتصادية.

4* انحدار الاستثمار الأجنبي المباشر على سعر الصرف خلال الفترة : (1985– 2013).

تحصلنا على النموذج التالي:

$$IDE = -235.02 + 15.61TCH + e_i$$
$$"1.40-" \quad "4.95"$$

$$(\bar{R}^2 = 0.51) \qquad (F_c = 24.58) \qquad (R^2 = 0.53) \qquad (DW = 0.47)$$

من خلال النموذج نلاحظ أن الاستثمار الأجنبي المباشر يرتبط بعلاقة طردية مع سعر الصرف، حيث أن سعر الصرف الحقيقي له تأثير موجب على تدفق الاستثمار الأجنبي المباشر في الجزائر كما يمكن تفسير هذه العلاقة الموجبة بين سعر الصرف والاستثمار الأجنبي المباشر في الجزائر بأن نسبة كبيرة من مشروعات الاستثمار الأجنبي المباشر تتجه لهدف تصدير(أغلبية الاستثمارات في قطاع المحروقات) وهذا يعني أن العائد يكون بالعملة الأجنبية ومنه انخفاض العملة المحلية سيكون له أثر ايجابي على المستثمر الأجنبي حين يقيم هذا العائد بالعملة المحلية إذن معامل سعر الصرف لها معنوية اقتصادية.

أما من الناحية الإحصائية نقبل معلمة سعر الصرف، وهذا لأن إحصائية ستودنت $(t = 4.95)$ أكبر من القيمة المجدولة ل(t_t) ، أما فيما يخص معامل التحديد $(R^2 = 0.53)$ مما يؤكد على متوسط قدرة النموذج التفسيرية بين سعر الصرف والاستثمار الأجنبي المباشر ،كما أن إحصائية فيشر(F_c) تساوي 24.58 وهي أكبر من(F_t) بمستوى معنوية5 ٪ . إذن معامل سعر الصرف $(\hat{\beta}_1)$ له معنوية إحصائية واقتصادية.

بعد القيام بإجراء الانحدار للمتغير التابع على كل متغير مفسر على حدى سوف نختار معادلة الانحدار التي تعطي أفضل النتائج على أساس عدة اختبارات و التي سوف نلاحظها في الجدول التالي:

جدول رقم4: اختبار معادلة الانحدار التي تعطي أفضل النتائج:

	إحصائيات	إحصائيات	Schwarz	Akaike	S.E.of	المعنوية	(R^2)

	الإحصائية	regression	info criterion	criterion	(F)	(t)	
0.89	**نعم**	**204.9**	**13.6**	**13.6**	**172.78**	**13.14**	النموذج 01
0.35	**نعم**	**501.14**	**15.35**	**15.45**	**172.78**	**–3.37**	النموذج 02
0.60	**نعم**	**390.10**	**14.85**	**14.95**	**32.40**	**–5.69**	النموذج 03
0.53	**نعم**	**422**	**15.01**	**15.11**	**24.58**	**4.95**	النموذج 04

المصدر: بالاعتماد على برنامج Eviews9.0

نلاحظ من خلال الجدول أن أفضل معادلة انحدار التي تعطي أفضل النتائج في ضوء المعايير المستعملة هو النموذج الثاني فهي الأفضل سواء من حيث إحصائية ستودنت أو إحصائية فيشر أو معامل التحديد أو الخطأ المعياري ، ومنه نختار معادلة الانحدار في النموذج الثاني و إدخال المتغيرات المفسرة بالتدريج و فحص أثرها على الأخطاء المعيارية و معامل التحديد حيث يمكن تلخيص هذه النماذج في الجدول التالي:

جدول رقم(5): حساب أهمية المتغير المفسر الإضافي:

SERE	F	$\bar{R}^2$	R^2	TCH	INF	IBS	GDP	ثابت	النموذج

2	IDE=f(GDP)	162.94 (−2.53)	0.22 (13.14	–	–	–	0,891	0,886	172.78	204.79
6	IDE=f(GDP,IBS)	−808,6 (−1,59)	0.25	13.609 (1.28)	–	–	0.899	0.889	89.87	201.715
7	IDE=f(GDP,IBS)	–	0.16	−3.168 (−2.307)	–	–	0.877	0.881	165.09	208.98
8	IDE=f(GDP,INF)	60.105 –	0.196	–	−5.978	–	0.898	0.888	88.64	202.96
9	IDE=f(GDP,INF)	–	0.187	–	−8.26	–	0.897	0.892	183.01	199.59
10	IDE=f(GDP,TCH)	101.68 –	0.236	–	–	−3.26	0.899	0.889	89.13	202.46
11	IDE=f(GDP,TCH)	–	0.243	–	–	−5.345	0.861	0.886	172	205.20
12	IDE=f(GDP,IBS,INF)	619.52 –	0.233	11.344 (1.036)	−4.74	–	0.904	0.888	59.67	202.58
13	IDE=f(GDP,IBS,INF)	–	0.190	−0.733 (−0.339)	−6.965	–	0.897	0.887	87.70	203.93
14	IDE=f(GDP,IBS,TCH)	579.52 –	0.248	9.28 (0.451)	–	−1.27	0.9	0.884	57.123	206.62
15	IDE=f(GDP,IBS,TCH)	–	0.34	−1.9 (−1.21)	–	−3.74	0.898	0.888	88.65	202.961
16	IDE=f(GDP,INF,TCH)	4.43 (0.03)	0.224	–	−6.10	−0.332	0.906	0.891	61.33	200.09
17	IDE=f(GDP,INF,TCH)	–	0.224	–	−5.96	−3.28	0.906	0.897	96.83	195.0

المصدر: بالاعتماد على برنامج Eviews 9.0 .

نلاحظ من خلال الجدول أن جميع النماذج ليس لها معنوية إحصائية (t) إلا في معاملGDP باستثناء ثلاث نماذج وهي :

IDE= F(GDP, TCH) IDE= F (GDP,INF) IDE=F(GDP, IBS)

كما نلاحظ في النموذج IDE= F(GDP, TCH) أن إشارة معلمة TCH لا تتوافق مع توقعاتنا ومنطق النظرية الاقتصادية، ومنه هذا المتغير مرفوض من الناحية الاقتصادية ومنه سوف نختار النموذج الذي يعطينا أفضل النتائج من بين النموذجين الباقيين على أساس عدة اختبارات.

جدول رقم(6): مقارنة النتائج.

	إحصائيات (F)	(R^2)	$(\bar{R}^2)$	S.E.of regression	Schwarz	Akaike
النموذج الثاني IDE=f(GDP)	172.78	0.891	0.886	204.79	10.82	10.72
النموذج التاسع IDE=f(GDP,INF)	183.01	0.897	0.892	199.59	10.77	10.67
النموذج السابع IDE=f(GDP,IBS)	165.09	0.887	0.881	208.985	1.86	10.76

المصدر: بالاعتماد على برنامج Eviews.

* من خلال النموذج التاسع عند إضافة INF للنموذج الثاني أنه قد حسن من قيمة معامل التحديد وكذلك معامل التحديد المعدل، حيث أصبح يساوي 0.892 بعدما كان يساوي 0.886 في النموذج الثاني.

*من خلال النموذج السابع عند إضافة IBS للنموذج الثاني أنه قد خفض من قيمة معامل التحديد وكذلك معامل التحديد المعدل حيث أصبح يساوي 0.881 بعد ما كان يساوي 0.886 في النموذج الثاني.

ومنه نستنتج أن النموذج التاسع هو الأفضل من بين النماذج، ضف إلى ذلك أن النموذج:

IDE= F(GDP, INF) نحصل فيه على توفيق أفضل سواء من ناحية معامل التحديد أو معامل التحديد المعدل أو إحصائية فيشر أو الخطأ المعياري.

3-2-1 طريقة Farrar-Glauber:

إن أسلوب كلاين تم تحجيمه من قبل كل من (Farrar-Glauber) في بحثهما المعنون Multicollinerity in regression analysis مشكلة الارتباط الخطي المتعدد في تحليل الانحدار، والمنشور في مجلة Review of economics and statistic، سنة 1967 ويستند اختبار (Farrar-Glauber) الى إحصائية (χ^{2*}) حيث يتم اختبار الفرضية التالية:

(استقلال خطي) $H_0: \mathrm{D}=1$

$H_1 : \mathrm{D} < 1$ (ارتباط خطي)

ويمكن التعبير عن إحصائية **Farrar-Glauber** (القيمة المحسوبة) كما يلي:

$$\chi^2 * = -\left[n - 1 - \frac{1}{6}(2k+5)\right].\ln D$$

حيث n هو حجم العينة، k هو عدد المتغيرات المفسرة في النموذج و $\ln$ هو اللوغاريتم النبيري لمحدد مصفوفة معاملات الارتباط التالية :

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

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

إذا 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.

بعد ثبوت مشكلة التعدد الخطي بموجب الاختبار أعلاه، يستوجب تحديد أي متغير من المتغيرات المستقلة مرتبطة خطياً والتي أدى إلى حدوث مشكلة التعدد الخطي، ويتم التشخيص من خلال اختبار F وحسب الصيغة التالية:

$$F_c == \frac{R^2{}_{xj.x1.x2........XK} / k}{\left(1 - R^2{}_{xj.x1.x2......XK}\right) / n - k - 1} \sim F_\alpha(k, n-k-1)$$

وذلك لاختبار الفرضيات التالية:

$$H_0 : \mathrm{R}^2{}_{XJ.X1.X2.....XK} = 0$$

$$H_1 : \mathrm{R}^2{}_{XJ.X1.X2.....XK} \neq 0$$

نقارن قيمة F_c المحسوبة مع قيمة F_t النظرية بدرجتي حرية $(k, n-k-1)$ ومستوى معنوية α فإذا كانت $|F_{Cal}| \rangle\ F_{Tab}$ ترفض H_0 أي أن المتغيرات التوضيحية مرتبطة مع بعضها وبعكسه ترفض الفرضية البديلة H_1 اي ان المتغيرات التوضيحية لا ترتبط مع بعضها ولا يشكل مصدر قلق لمشكلة التعدد الخطي، وبذلك يتم تشخيص كافة المتغيرات المرتبطة مع بقية المتغيرات التوضيحية.

ولغرض تحديد العوامل المسببة لحصول مثل هذه المشكلة للمتغيرات التوضيحية لذلك يجب إجراء اختبار ثالث وهو اختبار t الذي يعتمد على قيم معاملات الارتباط الجزئية ما بين كل اثنين من المتغيرات التوضيحية وبموجب الصيغة التالية:

$$t_{ij} == \frac{r_{xixj.x1x2......xk} \Big/ \sqrt{n-k}}{\sqrt{\left(1-r^2{}_{xij.x1x2......xk}\right)}}$$

حيث أن: $r^2{}_{ij.}$ يمثل مربع معامل الارتباط الجزئي ما بين المتغيرين المستقلين $(X_i$ و $X_j)$ باعتبار أن بقية المتغيرات المستقلة ثابتة .

حيث أن :

$$H_0 : \mathrm{r}_{Kij.X1X2......XK} = 0$$

$$H_1 : \mathrm{r}_{Kij.X1X2......XK} \neq 0$$

نقارن القيمة المحسوبة t_c والقيمة الجدولية t_t بدرجة حرية $(n-k)$ ومستوى معنوية معين فإذا كانت $|t_{Cal}| \rangle\ t_{Tab}$ ترفض H_0 ، أي أن الارتباط الجزئي بين $(X_i$ و $X_j)$ معنوي، وبذلك يتم تشخيص بشكل نهائي المتغيرات التوضيحية التي تكون سبباً في حصول مشكلة التعدد الخطي.

مثال تطبيقي على اختبار Farrar-Glauber :

إذا كانت لدينا مصفوفة الارتباطات R التالية:

$$(R) = \begin{bmatrix} 1 & 0.8811 & 0.9366 \\ 0.8811 & 1 & 0.9866 \\ 0.9366 & 0.9866 & 1 \end{bmatrix}$$

$$(n=15) \qquad (k=4)$$

محدد الارتباط يصبح :

$$|R| = \left\| \begin{matrix} 1 & 0.8811 & 0.9366 \\ 0.8811 & 1 & 0.9866 \\ 0.9366 & 0.9866 & 1 \end{matrix} \right\| = 0.000969$$

نلاحظ أن قيمة المحدد تقترب من الصفر، فإن هناك دليل على وجود تعدد خطي.

$$\chi^2 * = -\left[n-1-\frac{1}{6}(2k+5)\right].\ln D$$

$$\chi^2 * = -\left[15-1-\frac{1}{6}(24+5)\right].\ln 0.000969$$

$$\chi^{2}* = 52$$

نلاحظ أن قيمة χ^2 الجدولية بدرجات حرية 6=2/3*4 هي 16.8 وهي أصغر من χ^2 المحسوبة وعليه نستنتج أن هناك مشكلة التعدد الخطي.

3-1 علاج مشكلة التعدد الخطي:

يعتمد العلاج الملائم لمشكلة التعدد الخطي على طبيعة المشكلة نفسها، ونفرق في هذا الصد بين حالات عديدة:

* إذا كانت المتغيرات التفسيرية المرتبطة متغيرات قليلة الأهمية في التأثير على الظاهرة محل البحث فقد يكون الحل هو إسقاط هذه المتغيرات، ولكن يلاحظ أن هذا الحل قد يؤدي لوجود مشكلة ارتباط ذاتي من ناحية أخرى.

* محاولة توسيع حجم العينة من خلال إضافة بيانات كافية عن متغيرات الظاهرة المدروسة، لأنه يساعد على تخفيض حجم التباينات نظرا لوجود علاقة عكسية بين حجم العينة و قيمة التباين:

* استخدام معلومات قبلية في حالة توافرها.

* ومن الأساليب المقترحة الأخرى لعلاج مشكلة التعدد الخطي تحويل المتغيرات، فمن أسباب التعدد الخطي المتعدد أن المتغيرات الاقتصادية تميل للتغير في نفس الاتجاه عبر الزمن. ولتفادي هذا الأثر نقوم باستخدام النسب و الفروقات عوضاً عن المتغيرات الأصلية ، ولكن ليس من المؤكد أن تؤدي هذه الطريقة بالضرورة للتخلص من مشكلة التعدد الخطي، أو تحويل شكل الدالة باستعمال

* خلط بيانات قطاعية وبيانات سلسلة زمنية في النموذج المدروس.

4-1 تمرين:

لتكن لديك المعطيات المتعلقة بالواردات (Y) والناتج الوطني الإجمالي (X_1) والرقم القياسي للأسعار (X_2) (بالألف دج) التالية:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \varepsilon_t$$

$$Y_t = 5.47 + \underset{(0.211)}{0.086 X_{1t}} + \underset{(0.231)}{0.414 X_{2t}} + \underset{(0.761)}{0.759 X_{3t}} + e_t$$

$R^2 = 0.95$ $\quad n = 20 \quad$ $R^2_1 = 0.89 \quad R^2_2 = 0.75 \quad R^2_3 = 0.90$

$SCR = 905.1 \quad \sum (Y_t - \bar{Y})^2 = 18102$

$\sum (X_{1t} - \bar{X}_1) = 5362 \quad \sum (X_{2t} - \bar{X}_2) = 4482 \quad \sum (X_{3t} - \bar{X}_3) = 5050$

$COV(X_1, X_2) = 216.5 \quad COV(X_1, X_3) = 254.91 \quad COV(X_1, X_3) = 232.96$

ملاحظة: الأرقام ما بين قوسين هو الانحراف المعياري المقدر

* حدد جدول تحليل التباين للنموذج؟

* أحسب معامل الارتباط البسيط بين المتغيرات المستقلة؟

* اختبر معنوية المعالم باستخدام اختبار t عند مستوى الدلالة $\alpha = 5\%$ ؟

* ما لمعني التحليلي الإحصائي لهذا النموذج؟ تأكد من ذلك باستخدام كل الاختبارات الإحصائية التي درستها؟

* تم حذف المتغير X_3 وتحصلنا على النتائج التالية:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

$$Y = 5.31 + 0.458 X_{1t} + 0.319 X_{2t} + e_t$$

$$(0.641) \qquad (0.590)$$

$R^2 = 0.90 \qquad n = 20$

$SCR = 850.025 \qquad F = 132.43$

*هل تحسن النموذج وهل استفدنا من حذف المتغير X_3 ؟

* قم بتقدير النموذج الأصلي باستخدام طريقة انحدار الحذف علما أن : $C = 0.08$ ؟

2- الارتباط الذاتي بين الأخطاء

من بين الفرضيات الأساسية لطريقة المربعات الصغرى هي استقلالية القيمة المقدرة لحد الخطأ في فترة ما عن القيمة المقدرة لحد الخطأ في فترات زمنية سابقة لها أي انعدام الارتباط الذاتي بين أخطاء الفترات المختلفة و يعني ذلك:

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0 \qquad \forall\, i \neq j$$

و عند افتراق هذه الفرضية هذا يغني وجود ارتباط ذاتي و تصبح مصفوفة التباين و التباين المشترك Ω_ε لا تحتوي على الصفر خارج القطر الأول أي:

$$E(\varepsilon\acute{\varepsilon}) = \Omega_\varepsilon \neq \sigma_\varepsilon^2 I$$

و بالتالي:

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

$$= (\acute{X}X)^{-1}(\acute{X}\Omega_\varepsilon X)(\acute{X}X)^{-1}$$

$$\neq \sigma_\varepsilon^2 (X'X)^{-1}$$

أسباب الارتباط الدايي بين أخطاء النموذج:

-الصياغة الرياضية الخاطئة للنموذج أو عدم دقة توصيف العلاقات بين المتغيرات.

-إهمال لبعض المتغيرات التفسيرية المهمة في النموذج.

-عدم دقة البيانات وعدم استقرارية السلاسل الزمنية.

-وجود هذه المشكلة في النموذج سوف يؤثر سلبا على نتائج التقدير بطريقة المربعات الصغرى حيث تكون تميز مقدرات النموذج غير متحيزة كما أن تباينها لا يكون أصغر ما يمكن.

2-1 الكشف على الارتباط الذاتي:

يستعمل للكشف على هذا الاختلال عدة اختبارات أهمها:

2-1-1 اختبار دربين واتسون Durbin-Watson test:

يعتبر من أهم الاختبارات المستخدمة في اكتشاف الارتباط الذاتي من الدرجة الأولى.

شروط استخدام هذا الاختبار هي:

. احتواء النموذج على المعلم الثابت β_0.

. لا يتضمن النموذج متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة.

. يجب أن يكون حجم العينة أكثر من 14 مفردة حتى نتمكن من إجراء الاختبار باعتبار جدول دربين واتسون Durbin-Watson تبدأ من n=15.

. لا يحتوي النموذج على المتغير التابع ذو الفجوة الزمنية المتأخرة.

عند تطبيق هذا الاختبار يكون لدينا:

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

يهدف إلى اختبار الفرضيات التالية:

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

نقوم بحساب إحصائية DW (دربين واتسون) كالتالي:

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

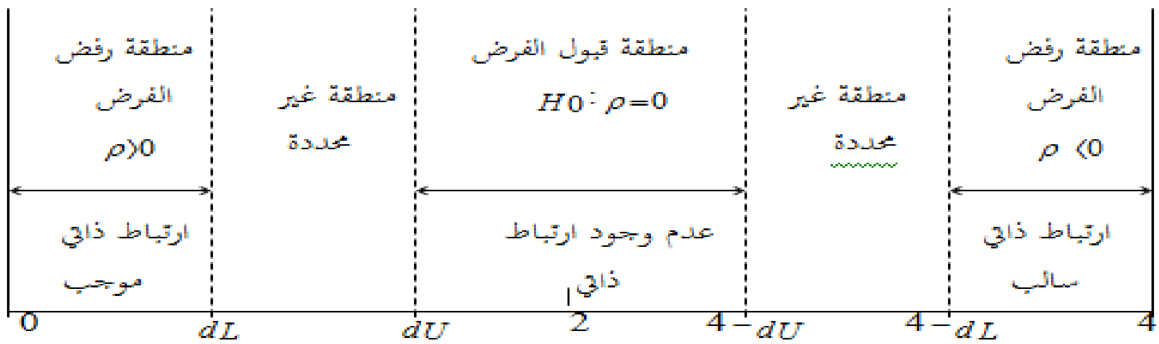
$$DW = 2(1-\hat{\rho}) \qquad \text{مع}$$

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

تأخذ قيمة DW المحسوب بين 0 و 4 إذا كانت $\rho = 0$ فإن DW=2 أي انعدام الارتباط الذاتي.

نقارن DWالمحسوب إلى قيم d المجدولة للاختبار و التي تشير إلى وجود أو عدم وجود ارتباط ذاتي من الدرجة الأولى موجب أو سالب أو تكون النتيجة غير محدودة حسب قيم كل من الحد الأعلى d_uو الحد الأدنى d_L و الجدول الاحصائي لتوزيع دربين واتسون DW.

نستخرج هذه القيم (d_L, d_u) حسب عدد المشاهدات n عدد المتغيرات المستقلة في النموذج عند مسرى دالة αو تكون مناطق القبول أو الرفض في اختبار DW كما يلي:



Source : BOURBONNIAIS. R, (2004), p 123.

و بالتالي نستنتج ما يلي:

- إذا كان $DW > d_L \wedge DW > 4-d_L$ نرفض H_0
- إذا كان $d_u < DW < 4-d_L$ نقبل H_0
- إذا كان $d_L < DW < d_u$[1] / $4-d_u < DW < 4-d_L$ تكون النتيجة غير محددة، و من ثمة يجب معالجة الأمر بإضافة بيانات أكثر مثلا.

إذن كما نعلم أن اختبار DW هو اختبار للارتباط الذاتي من الدرجة الأولى اذن هو غير كافي باعتبار وجود ارتباط ذاتي من درجة أكبر من الواحد لذى نستخدم اختبار آخر لذلك.

2-1-2 اختبار **Breusch-Godfrey (LM)**:

اختبار Breusch-Godfreyيركز على مضاعف لاغرانج للكشف على وجود ارتباط ذاتي من الدرجة الأكبر من الواحد (LM test) و في هذه الحالة يكون:

$$\varepsilon_t = \rho_1\varepsilon_{t-1} + \varepsilon_2\rho_{t-2} + \dots + \rho_p\varepsilon_{t-p} + \mu_t$$

حسب هذه الحالة يرتبط حد الخطأ بالفترة الحالية وبالحدود العشوائية (الأخطار) للفترات السابقة حتى الفترة P يكون لدينا النموذج العام باعتبار الأخطاء مرتبطة ذاتيا:

$y_t = \beta_0 + \beta_1 X_{t1} + \ldots + \beta_k X_{tK} + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \ldots \rho_p \varepsilon_{t-p} + \mu_t$

نود في هذه الحال اختبار فرضية استقلالية الأخطاء H_0 كما يلي:

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

لاجراء اختبار LM لدينا الخطوات الثلاث التالية:

– نقوم

بتقدير النموذج العام بطريقة المربعات الصغرى ثم نحسب البواقي $\hat{\varepsilon}_t$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t$$

-تقدير المعادلة الوسيطية (او الانحدار المساعد) التالية:

$$\hat{\varepsilon}_t = \hat{\beta}_0 + \hat{\beta}_1 X_{t_1} + \cdots + \hat{\beta}_k X_{t_k} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \ldots \rho_p \hat{\varepsilon}_{t-p} + \mu_t$$

ثم حساب معامل التحديد الخاص لهذا النموذج R^2 (سنفقد في هذه الحالة P مشاهدة)

لدينا $LM = (n-p).R^2$

n: حجم العينة Λ p: رتبة الارتباط الذاتي

$$LM = (n-p)\, R^2 \sim X^2_{(p)}$$

لمقارنة إحصائية $LM((n-p).R^2)$ عند مستوى معنوية معين و درجة حرية(p) ل X^2_p نستنتج ما يلي:

-إذا كان $(n-p)\, R^2 > X^2_p$ نرفض H_0 أي هناك ارتباط ذاتي على الأقل من الرتبة الأولى.

2-2 تصحيح النموذج في حالة الارتباط الذاتي بين الأخطاء

في حالة الارتباط الذاتي بين الأخطاء يصبح لدينا

$$_1\varepsilon = \mu_t + \rho \mu_{t-1} + \rho^2 \mu_{t-2} + \rho^3 \mu_{t-3} + \ldots$$

انطلاقا من :

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

مع

$$Var(\varepsilon_t) = E(\varepsilon_t^2) = \frac{\sigma_\mu^2}{1-\rho^2}$$

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

تصبح مصفوفة التباين و التباين المشترك للأخطاء:

$$\Omega_{\varepsilon} = E(\varepsilon\varepsilon') = \frac{\sigma\mu^2}{1-\rho^2} = \begin{bmatrix} 1 & \rho & \rho^2 \dots & \rho^{n-1} \\ \rho & 1 & \rho \dots. & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots \ \rho^{n-3} \\ & & \vdots & \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} \dots & 1 \end{bmatrix}$$

كما أن مقدر طريقة المربعات الصغرى في هذه الحالة يكتب من الشكل:

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

و لتقدير ذلك نستخدم طريقة المربعات الصغرى المعممة MCG

و حتى يتم تقدير النموذج العام المصحح من الارتباط الذاتي بطريقة المربعات الصغرى العادية يجب تحويل المتغيرات عن طريق شبه الفروقات من الدرجة الأولى كما يلي:

$$y^* = \begin{bmatrix} y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ \dots. \\ y_n - \rho y_{n-1} \end{bmatrix}$$

حيث j=1,2…k

$$X' = \begin{bmatrix} x_{2j} - \rho x_{1j} \\ x_{3,j} - \rho x_{2j} \\ \dots \\ x_{n,j} - \rho x_{n-1,j} \end{bmatrix}$$

عند استخدامنا لشبه الفروقات نفقد المتغير الاول و لتجنب ذلك نضع:

$$y_1^* = y_1\sqrt{1-\rho^2}$$
$$x_{1,j}^* = x_{1,j}\sqrt{1-\rho^2}$$

و يكتب النموذج المصحح على النحو التالي:

$$Y_t - \rho y_{t-1} = \beta_0\,(1-\rho) + \beta_1\,(x_{t,1} - \rho\, x_{t-1,1}) + \beta_2\,(x_{t,2} - \rho\, x_{t-1,2}) + \dots + \beta_k\,(x_{t,k} - \rho x_{t-1,k}) + \varepsilon_t - \rho\varepsilon_{t-1}$$

و ليتم تقدير النموذج السابق بطريقة المربعات الصغرى يجب أولا تقدير معامل الارتباط الذاتي بين الأخطاء من الدرجة الأولى (ρ) ، و هناك عدة طرق نذكر أهمها:

1–2–2 تقدير عن طريق إحصائية Durbin-Watson

- نقوم بتقدير انطلاقا من إحصائية DW حيث

$$\hat{\rho} \cong 1 - \frac{DW}{2}$$

- تقدير النموذج العام بعد إجراء التعديل في المشاهدات من خلال حساب الفروق من الدرجة الأولى:

$$Y_t - \hat{\rho}y_{t-1} = \beta_0(1-\hat{\rho}) + \beta_1(x_{t,1} - \hat{\rho}t_{t-1.1}) + \beta_2(x_{t,2} - \hat{\rho}x_{t-1.2}) + \ldots\ldots.. + \beta_K(x_{t,K} - \hat{\rho}X_{t-1,K}) + \mu_t$$

أي

$$Y_t^x = \beta_0^* + \beta_1 X_{t,1}^* + \beta_2 X_{t,2}^* + \cdots \ldots\ldots. + \beta_K X_{t,K}^* + \mu_t$$

إذن المعاملات المقدرة هي

$$\hat{\beta}_0^* = \hat{\beta}_0\ (1\text{-}\hat{\rho})\ ,\ (\hat{\beta}_1)\ ,\ \hat{\beta}_2\ ,\ldots,\ \hat{\beta}_k$$

2-2-2 تقدير ρ بطريقة Theil-Nagar

تم اقتراح تقديرا ل ρ من خلال العلاقة التالية:

$$\hat{\rho} = \frac{n^2\ \left[\left[1-\left(\frac{DW}{2}\right)\right]\ + (K+1)^2\right.}{n^2 - (K+1)^2}$$

حيث K هي عدد المتغيرات المستقلة ، ثم نستخدم الخطوات المتبعة في طريقة DW أي تقدير معالم النموذج بطريقة المربعات الصغرى.

3-2-2 طريقة Cochrane _ Orcutt (طريقة التكرار)

- إعطاء قيمة ابتدائية لمعامل الارتباط ρ من خلال العلاقة :

$$\hat{\rho} = \hat{\rho}_0 \qquad \text{أي} \qquad \hat{\rho} = \frac{\sum_{t=1}^{n}(\hat{\varepsilon}_t\hat{\varepsilon}_{t-1})}{\sum_{t=1}^{1}(\hat{\varepsilon}_t)^2}$$

تقدير النموذج العام المصحح باستخدام$\hat{\rho}_0$ و طريقة المربعات الصغرى العادية أي

$$Y_t - \hat{\rho}_0 Y_{t-1} = \beta_0(1-\hat{\rho}_0) + \beta_1(X_{t,1} - \hat{\rho}_0 X_{t-1,1}) + \beta_0\ (X_{t,2} - \hat{\rho}_0 X_{t-1,2}) + \cdots \ldots\ldots + \beta_K(X_{t,K} - \hat{\rho}_0 X_{t-1,K}) + \mu_t$$

- إعادة تقدير ρ باستخدام البواقي المقدرة من النموذج العام المصحح المقدر كما يلي:

$$\hat{\varepsilon}_t^{(1)} = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{t1} - \ \ldots\ldots\ldots..\hat{\beta}_k X_{tK}$$

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

- نقدر النموذج ذو المتغيرات ذات شبه الفروقات :

$$Y_t - \hat{\rho}_1 Y_{t-1} = \beta_0(1-\hat{\rho}_1) + \beta_1(X_{t,1} - \hat{\rho}_1 X_{t-1,1}) + \beta_2(X_{t,2} - \hat{\rho}_1 X_{t-1,2}) + \ \ldots\ldots\ldots + \beta_K(X_{t,K} - \hat{\rho}_1 X_{t-1,K}) + \mu_t$$

ثم نعيد تقدير ρ مرة أخرى ببواقي جديدة $\hat{\varepsilon}_t^{(2)}$ ونحصل على متغيرة أخرى $\hat{\rho}_2$، تكرر هذه العملية إلى غاية الحصول على متغيرات نموذج ساكنة.

مثال: لدينا معطيات الواردات كمتغير تابع للناتج الوطني بالمليون جنيه لبلد ما والمطلوب: اختبار الارتباط الذاتي ثم معالجته إن وجد:

السنوات	1	2	3	4	5	6	7	8
الواردات Yt	3748	4010	3711	4004	4469	4569	4582	4697
الناتج	21777	22418	22308	23319	24180	24893	25310	25799
السنوات	9	10	11	12	13	14	15	16
Yt	4753	5062	5669	5628	5736	5946	6501	6549
Xt	25886	28868	28134	29091	29450	30705	32372	33152
السنوات	17	18	19	20				
Yt	6705	7104	7609	8100				
Xt	33764	34411	35429	36200				

$$\hat{y}_t = -2489{,}25 + 0{,}28X_t$$

الحل: أولا نقوم لحساب ديربين واطسون DW

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

نستخرج كل من $\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2 \quad \Lambda \quad \sum_{t=1}^{1}(\hat{\varepsilon}_t)^2$ كما يلي:

لدينا :

السنة	Yt	Xi	Y_t^1	$\hat{\varepsilon}_t = y_t - \hat{y}_t$	$(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})$
1	3748	21777	3615.5	132.42	-
2	4010	22418	3795.2	214.73	82.3
3	3711	22308	3764.4	-53.42	-268.16

4	4004	23319	4047.8	-43.84	9.58
5	4151	24180	4289.2	-138.2	-94.73
6	4469	24893	4489.08	-20.08	118.12
7	4582	25310	4605.9	-23.98	-3.90
8	4697	25799	4743.07	-46.07	-22.08
9	4753	25886	4767.7	-14.46	31.61
10	5062	26868	5042.7	19.25	33.71
11	5669	28134	5397.6	271.35	252.10
12	5628	29091	5665.9	-37.92	-309.28
13	5736	29450	5766.5	-30.56	7.36
14	5946	30075	6118.3	-172.38	-141.82
15	6501	32372	6585.7	-84.70	-87.68
16	6549	33.52	6804.9	-255.36	-170.66
17	6705	33764	6975.9	-270.92	-15.56
18	7104	34411	7157.3	-53.30	217.62
19	7609	35429	7442.6	166.31	219.62
20	8100	36200	7658.8	441.18	274.86

$\sum \hat{\varepsilon}_t^2 = 567861$

$\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2 = 491847$

نحصل على إحصائية DW

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

$$DW = \frac{491847}{567861}$$

$$DW = 0.86$$

$$\sum_{t=2}^{n} \hat{\varepsilon}_t \, \hat{\varepsilon}_{t-1} = 321977$$

نرجع لجدول Durbin Watson عند مستوى معنوية $\alpha=0.05$ و متغير واحد K=1و حجم العتبة $n = 20$

نجد: $d_L=1.21 \quad \Lambda \quad d_u=1,41$

لدينا $DW<d_L$ إذن ترفض H_0 و منه وجود ارتباط ذاتي موجب و لمعالجة هذا الاشكال نقوم بحساب قيمة معامل الارتباط $\hat{\rho}$

لدينا :

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

$$\Rightarrow \hat{\rho} = 1- DW/2 = 1-0.866/2$$

و أيضا:

$$\hat{\rho}=\frac{\sum_{t=2}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{1}(\hat{\varepsilon}_t)^2} = \frac{321977}{567861}$$

$$\hat{\rho} = 0.567$$

بعد حسابنا ل $\hat{\rho}$نقوم باستخراج قيم المتغيرات بعد تحويلها باستخدام طريقة شبه الفروقات كما يلي:

أول المشاهدات:

$$y_1^*=y_1\sqrt{1-\hat{\rho}^2}$$

$$X_1^*=X_1\sqrt{1-\hat{\rho}^2}$$

و منه:

$$y_t^* = y_t - 0.567y_{t-1}$$

$$X_t^* = X_t - 0.567X_{t-1}$$

ولدينا أيضا:

$$\hat{\beta}_1^*=\hat{\beta}_1 \quad \Lambda \quad \hat{\beta}_0^*= \hat{\beta}_0 \ (1 - \hat{\rho})$$

باستخدامنا لطريقة المربعات الصغرى نقوم بتقدير النموذج العام باستعمال المتغيرات الجديدة Y_t^X و X_t^X و تحصلنا على النموذج التالي:

$$Y_t^*= -1727.4 +0.29X_t^* \qquad \text{مع} \qquad DW = 1.515$$

نلاحظ أن $d_U< DW< 2$ أي موجود في منطقة قبول H_0 أي عدم وجود ارتباط ذاتي.

مثال 2:لدراسة العلاقة بين معدلات الفائدة طويلة الأجل Y_t بدلالة معدلات الفائدة قصيرة الأجل X_t قمنا بجمع معطيات شهرية لمعدلات الفائدة لبلد ما كما يلي:

السنة	Y_t	X_t	السنة	Y_t	X_t
1 2004	6.56	7.77	5	7.01	5.23
2	6.54	7.12	6	6.86	5.34

6.13	6.89	7	7.96	6.81	3
6.44	7.11	8	8.33	7.04	4
6.42	7.28	9	8.23	7.1	5
5.96	7.29	10	7.90	7.02	6
5.48	7.21	11	7.55	7.18	7
5.44	7.17	12	8.96	7.33	8
4.87	6.93	2006 1	8.08	7.30	9
4.88	6.92	2	7.46	7.22	10
5.00	6.88	3	7.47	6.93	11
4.86	6.73	4	7.15	6.77	12
5.20	7.01	5	6.26	6.68	2005 1
5.41	6.92	6	5.50	6.66	2
			5.49	6.77	3
			5.61	7.05	4

المطلوب اختيار استقلالية الأخطاء لنموذج الانحدار ل Y_t على X_t باستخدام الطرق السالفة الذكر، تصحيح النموذج من الارتباط الذاتي ان وجد.

الحل:

الخطوة الأولى: هي تقدير نموذج الانحدار ل Y_t على X_t باستخدام طريقة المربعات الصغرى نجد النموذج التالي:

$$\hat{y}_t = 6.728 + 0.037X_t$$

نقوم بحساب إحصائيةDW حيث:

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

لدينا $\hat{\varepsilon}_t = Y_t - \hat{y}_t$ ⟸ $\sum \hat{\varepsilon}_t^2 = 1.360$

$$\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2 = 0.606$$

$$DW = \frac{0.606}{1.360} = 0.445$$

نستخرج من جدول ديربين واتسون قيم كل من d_L و d_U عند دلالة احصائية α=0.05 و $n = 30$ مع K=1 (متغير مستقل واحد) نجد $d_U = 1.49$ $d_L = 1.35$ إذن:

الخطوة الثانية : هي اختيار استقلالية الأخطاء من خلال مقارنة إحصائية DW مع القيم الجدولية d_U و d_L

نجد أن $0 < DW < d_L$ و هي منطقة ارتباط ذاتي موجب بين الأخطاء.

كما يمكننا استعمال إحصائية Breusch Godfrey كما يلي:

$$\hat{\varepsilon}_t = \hat{\beta}_0 + \hat{\beta}_1 X_{t_1} + \hat{\rho}_1 \hat{\varepsilon}_{t-1} + \hat{\rho}_2 \hat{\varepsilon}_{t-2} + \mu_t$$

يمكننا استخراجها مباشرة باستخدام البرنامج الإحصائي الإحصائي مثلا علىeviews (اختيارات الأخطاء ونختار serial correction LM test)

تحصلنا على المعادلة:

$$\hat{\varepsilon}_t = -0.07 + 0.014X_t + 0.93\hat{\varepsilon}_{t-1}\ \varepsilon^1_{t-1} - 0.35\hat{\varepsilon}_{t-2}$$
$$R^2 = 0.607$$

إذن نختبر فرضية استقلالية الأخطاء

$$H_0 = \rho_1 = \rho_2 = 0$$

فحسب إحصائية مضاعف لاغرانج

$$LM = (n - 2) \times R^2 = (30-2) \times 0.607$$
$$LM = 17.016$$

نستخرج إحصائية χ^2 بدرجة حرية 2 و $\alpha = 0.05$

$$\chi^2\ (2 \times 0.05) = 5.99$$

إذن لدينا: $LM > \chi^2\ (2 \times 0.05)$ أي $17.016 > 5.99$

أي رفض فرضي العدم وبالتالي وجود الارتباط الذاتي بين الأخطاء.

الخطوة التالية: تصحيح النموذج من وجود الارتباط الذاتي

أولا نستخرج معامل الارتباط الذاتي كما قمنا به في المثال الأول:

$$\hat{\rho} = 1 - DW/2$$

$$\hat{\rho} = 1 - 0.446/2$$

$$\hat{\rho} = 0.776$$

نحول جدول المعطيات الأصلية إلى متغيرات ذات شبه الفروقات

المشاهدات الأولى:

$$y_1^* = y_1\sqrt{1-\hat{\rho}^2} \qquad X_1^* = X_1\sqrt{1-\hat{\rho}^2}$$
$$y_t^* = y_t - \hat{\rho} y_{t-1} \qquad X_t^* = X_t - \hat{\rho} X_{t-1}$$

ثم نحسب القيم الجديدة كما هي موضحة في الجدول أدناه ثم نقوم بتقدير النموذج الجديد المحول:

$$Y_t^* = \beta_0^* = \beta_1^* X_t^* + \mu_t$$

بطريقة المربعات الصغرى حيث $\beta_0^* = (1-\hat{\rho})\beta_0$

$\mu_t = \varepsilon_t - \hat{\rho}\varepsilon_{t-1}$ Λ $\beta_1^* = \beta_1$

جدول المتغيرات ذات شبه الفروقات

السنة		X_t	Y_t	$X_t^* = X_t - \hat{\rho}X_{t-1}$	$y_t^* = y_t - \hat{\rho}y_{t-1}$
1	2004	7.77	6.65	4.89	2.60
2		7.12	6.54	1.08	1.44
3		7.96	6.81	2.42	1.72
4		8.33	7.04	2.14	1.74
5		8.23	7.10	1.75	1.63
6		7.90	7.02	1.50	1.50
7		7.55	7.18	1.41	1.72
8		8.96	7.33	3.09	1.75
9		8.08	7.30	1.09	1.60
10		7.46	7.22	1.19	1.54
11		7.47	6.93	1.67	1.32
12		7.15	6.77	1.34	1.38
1	2005	6.26	6.68	0.70	1.42
2		5.50	6.66	0.63	1.47
3		5.49	6.77	1.21	1.59
4		5.61	7.05	1.34	1.79
5		5.23	7.01	0.87	1.53
6		5.34	6.86	1.27	1.41
7		6.13	6.89	1.98	1.56
8		6.44	7.11	1.67	1.75
9		6.42	7.28	1.41	1.75
10		5.96	7.29	0.97	1.63
11		5.48	7.21	0.84	1.54
12		5.44	7.17	1.18	1.56
1	2006	4.87	6.93	0.64	1.35
2		4.88	6.92	1.09	1.53
3		5.00	6.88	1.20	1.50
4		4.86	6.73	0.97	1.38
5		5.20	7.01	1.42	1.78
6		5.41	6.92	1.37	1.47

بعد تقدير النموذج بطريقة المربعات الصغرى

$$\hat{y}_t^* = 1.2586 + 0.2298X_t^*$$

ويمكننا التأكد من أن : $\hat{\beta}_0^* = (1-\hat{\rho})\hat{\beta}_0$

$\sum_{t=2}^{30}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}) = 0.921$ $\qquad\qquad$ $\sum_{t}^{30} \hat{\varepsilon}_t^{\,2} = 0.5260$

نقوم بحساب إحصائية DW

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

$$DW = \frac{0.923}{0.526} = 1.755$$

نقوم بمقارنة إحصائية DW مع القيم المستخرجة من جدول ديربين واتسون d_u و d_L فنجد أن $d_U<DW<2$

, و هي منطقة قبول H_0 أي عدم وجود ارتباط ذاتي بين الأخطاء و هو المطلوب.

ويمكن استخدام طريقة التكرارCochrane-Orcuh باستخدام البرنامج الإحصائي R أوEviews و نحصل تقريبا على نفس النتيجة.

3-عدم ثبات تباين الأخطاء (عدم التجانس) Hétéroscedodtinty

عند تشخيصنا لنموذج الانحدار يجب التأكد من عدم خرق لإحدى فرضيات النموذج و هي تجانس تباينات الأخطاء (أي التباين غير متساو) في حد الخطأ ε_t هو σ_{ε}^2 .

هناك عدة أسباب لظهور هذه المشكلة مثل وجود قيم متطرفة في البيانات، شكل النموذج غير صحيح أو التحويل الغير صحيح للبيانات، ...الخ.

إن عدم ثبات تباين الأخطاء يخلف العواقب التالية:

مصفوفة التباين والتباين المشترك للأخطاء:

$$\Omega_{\varepsilon} = E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_{\varepsilon\,1}^2 & 0 \ldots\ldots 0 \\ 0\ \sigma_{\varepsilon,2}^2 0_2 \ldots .0 \\ 0\ \ 0\ \ \sigma_{\varepsilon,3}^2 \ldots 0 \\ \ldots\ldots\ldots.\ldots..\ \sigma_{\varepsilon,n}^2 \end{bmatrix}$$

تصبح تباينات الأخطاء غير ثابتة على القطر الأول، وهذا قد يرتبط بقيم المتغير المستقلK مما يترتب مايلي:

- لا يغير عدم ثبات التباين في خاصية عدم تحيز واتساق مقدرات النموذج بطريقة المربعات الصغرى ولكن تصبح غير ذات كفاءة أي أنها ليست أفضل مقاييس خطية غير متحيزة (Blue).

- تصبح اختبارات ستودنت وفيشر ((F, t ليست ذات ثقة أو ليست موثوق بها بشأن المعنوية، الاحصائية للنموذج ومقدراته.

3-1 اكتشاف عدم ثبات التباين

هناك العديد من الاختبارات للكشف على عدم ثبات التباين، سوف نتطرق لأهمها:

1-1-3 **اختبار Goldfeld–Quandt(1965)**: يستخدم في العينات الكبيرة و موزعة طبيعيا (اختبار معياري) وبافتراض النموذج التالي

$$Y_i = \beta_0 + \beta_i X_i + \varepsilon_i \quad , i= 1 :n$$

ثم نتبع الخطوات التالية للكشف على عدم تباين الخطأ:

- ترتيب مشاهدات X تصاعديا.
- نستبعد المشاهدات الوسطى لكل من X و Y ثم تكوين مجموعتين من المشاهدات بحيث يكون مجموعة على حدا معادلة خاصة كمايلي:

المجموعة الأولى: تتمثل في المشاهدات من X و Y التي تكون قبل المشاهدات المحذوفة و المعادلة الخاصة بهذه المجموعة هي:

$$Y_{1i} = a + bX_{1i} + \varepsilon_{1i} \quad \ldots(1)$$

المجموعة الثانية: تتمثل في المشاهدات الخاصة من X وY التي تأتي المشاهدات المحذوفة والمعادلة الخاصة بهم:

$$Y_{2i} = c + dx_{2i} + \varepsilon_{2i} \quad \ldots(2)$$

-نقوم بتقدير المعادلتين السابقتين (1) و (2) كما يلي:

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

-نستخرج حد الخطأ للمعادلتين المقدرتين

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

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

-نحسب إحصائية فيشر F بالمعادلة التالية

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

ثم إيجاد درجة الحرية $DF = \frac{n-m-2(k+1)}{2}$

K: عدد المتغيرات المستقلة

m: عدد المشاهدات المستبعدة

- المقارنة بين إحصائية فيشر المحسوبة و الجدولية عند معنوية $\alpha\%$ بعد طرح الفرضية العدم اللازمة و التي مفادها ثبات تباين الأخطاء أي $\sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2}$ حيث اذا كانت $F_c > F_T$ إذن نرفضH_0 أي عدم ثبات تباين الأخطاء.

أما اذا كانت $F_c < F_T$ اذن قبول H_0و بالتالي ثبات تباين الأخطاء.

ملاحظة هامة: يتم تطبيق هذا الاختبار فقط إذا كانت إحدى المتغيرات المستقلة هي سبب وجود مشكلة عدم ثبات التباين.

3-1-2 اختبار Breusch-Pagan(1979):

يعتمد هذا الاختبار على استخدام البواقي ومضاعف لاغرانج وتزداد قوة هذا الاختبار بزيادة حجم العينة، ويتم إتباع الخطوات التالية في استخدام هذا الاختبار للكشف عن عدم ثبات التباين:

-تقدير معادلة الانحدار باستخدام طريقة المربعات الصغرى ثم نستخرج البواقي كمايلي:

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i$$

ثم الحصول على مربعات البواقي ε_i^2

-نجري انحدار لـ ε_i^2على المتغيرات المستقلة التي عددها K المتضمنة في النموذج.

(الهدف من ذلك هو معرفة إذا كانت مربعات البواقي مرتبطة بواحد أو أكثر من المتغيرات) وهو ما يشير إلى احتمال وجود عدم ثبات التباين في البيانات.

أو يمكن ربط تباين الأخطاء $\sigma^2_{\varepsilon_i} = \frac{\Sigma\, \varepsilon_i^2}{n - K}$ بالمتغيرات المستقلة أو متغيرات أخرى لها أكثر تفسير في النموذج.

$$i : 1 : n \quad \sigma^2_{\varepsilon} = \alpha_0 + \alpha_1 z_{1\,i} + \alpha_2 z_{i\,2} + \ldots + \alpha_p z_{i\,p} \quad \ldots(2)$$

فرض العدم هو ثبات تباين الأخطاء من خلال:

H_0: $0 = \alpha_0 = \alpha_1 = \alpha_2 = \ldots = \alpha_p$

H_1: $0 \neq \alpha_0 \neq \alpha_1 \neq \alpha_2 \neq \ldots \neq \alpha_p$

لاختبار فرضية العدم يمكننا استخدام إحصائية فيشر للانحدار (2) الإضافي بين تباين الأخطاء و المتغيرات المستقلة أو الإضافية فإذا كانت إحصائية فيشر ذات معنوية أي $F_c > F_T$إذن نفرض فرضية العدم أي عدم ثبات تباين الأخطاء أما إذا كان العكس أي إحصائية فيشر للانحدار (2) ليست ذات معنوية إذن تقبل H_0 و بالتالي ثبات تباين الأخطاء كما يمكننا استخدام إحصائيةX^2 كما يلي:

-نقوم بحساب $R^2_{(2)}$ للانحدار (2) أي معامل التحديد للانحدار الإضافي ما إذا كان n هو حجم العينة

P: عدد المتغيرات المستقلة $\quad x^2_{(p)} \sim n\, R^2_{(2)}$

-إذا كان $x^2_{(c)} = nR^2 < x^2_{(p)}$ إذن نقبل H_0 وبالتالي تباين الأخطاء في النموذج الأصلي ثابت ويمكن الحصول على هذا الاختبار في برنامج القياس الاقتصادي.

3-1-3 اختبار White للكشف عن مشكلة عدم ثبات تباين الأخطاء

اقترح White 1980 اختبار أكثر عمومية من اختبار BP حيث يعتمد على العلاقة بين المربعات البواقي وجميع المتغيرات المستقلة وكذا مربعاتها، وهو يصلح للعينات الكبيرة.

ولتطبيق هذا الاختبار نتبع الخطوات التالية:

-تقدير النموذج العام (1) .... $Y= X\beta + \varepsilon$ بطريقة المربعات الصغرى ثم حساب مربعات البواقي $\hat{\varepsilon}_i^{\,2}$

-تقدير المعادلة الوسيطة التالية :

$$\hat{\varepsilon}_i^{\,2}= \beta_0 +\beta_1 X_{i\,1} +\alpha_1 X^2_{i\,1}+\beta_2 X_{i\,2} +\alpha_2 X^2_{i\,2} +\ldots+ \beta_K X_{i\,K} +\alpha_K X^2_{i\,k}+\mu_i \ldots. \quad (2)$$

-نقوم بحساب معامل التحديد الخاص بهذا النموذج R^2

-طرح فرضية العدم لهذا الاختبار :

$$H_0 : \beta_0 = \beta_1 = \alpha_1 = \beta_2 = \alpha_2 =\ldots= \beta_k =\alpha_k =0$$

نحسب إحصائية مضاعف لاغرانج

$$LM= nR^2 \sim x^2_{(2k)}$$

درجة حرية= 2K

يصبح لدينا النتائج التالية للاختبار:

إذا كان $LM= nR^2 > x^2_{(2k)}$ عند معنوية α إذن نفرض H_0 و منه تباين الأخطاء غير ثابت .

4-1-3 ختبار ثبات التباين الشرطي للأخطاء (ARCH-LM)

تسمح نماذج ARCH بنمذجة المتغيرات التي تحوي على تباين شرطي غير ثابت للأخطاء العشوائية حيث أن التطاير الشرطي الذي يعبر في الغالب عن المخاطرة غير ثابت و هذا ما يسمح بنمذجة المتغيرات المالية التي تحتوي على ما ذكرناه سابقا.

يعتمد هذا الاختبار على مضاعف لاغرانج LM و خطوات هذا الاختبار كمايلي:

- تقدير النموذج العام $Y= X\beta + \varepsilon \ldots(1)$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_t^{\,2}$
- نقوم بتقدير النعادلة التالية:

$$\hat{\varepsilon}_t^{\,2}= \alpha_0 + \alpha_1\hat{\varepsilon}_{t-1}^{\,2} + \ldots\ldots + \alpha_q\hat{\varepsilon}_{t-q}^{\,2}+\mu_t$$

- نحسب معامل التحديد الخاص بهذه المعادلة R^2 حيث نفقد في هذه الحالة q مشاهدة.
- نطرح فرضية العدم التي مفادها ثبات التباين الشرطي للأخطاء أي:

$$H_0 = \alpha_0 = \alpha_1 = \ldots\ldots \alpha_q = 0$$

- نحسب إحصائية مضاعف لاغرانج $LM = (n - q).R^2$

حيث : $df = q$ و معنوية α

$$LM \sim x^2(q)$$

إذا كان $LM = (n-q)\, R^2 > x^2(q)$

نرفض H_0 أي هناك معامل واحد في معادلة ARCH على الأقل تختلف معنويا عن الصغر أي التباين الشرطي للأخطاء غير ثابت (غير متجانس).

3-2 طرق معالجة مشكلة تباين حد الخطأ:

من أهم الطرق المستخدمة لتصحيح مشكلة عدم تجانس تباين حد الخطأ هي طريقة المربعات الصغرى المرجحة، من خلال اعطاء وزن أو ترجيح أكبر للقيم ذات الانحراف الأقل من قيم ذات الانحراف الأكبر في تقدير النموذج، ويكون هذا النموذج المحول على حسب عدم ثبات التباين المكتشف في النموذج الأصلي.

كما يمكننا حل مشكلة عدم ثبات التباين إذا عرفنا التباينات σ_i^2، ففي هذه الحالة يمكننا تحويل النموذج الأصلي بواسطة قسمته على σ_i (أي الانحراف المعياري) و تقدير النموذج المحول بواسطة المربعات الصغرى العادية و هي أيضا طريقة تقدير المربعات الصغرى المرجحة (WLS).

أما إذا كان التباين σ_i^2 غير معلوم ففي هذه الحالة هناك عدة افتراضات:

<u>الافتراض الاول:</u>

إذا كان $E(\varepsilon_i^2) = \sigma^2 X_i^2$ يتم تحويل النموذج الأصلي إلى الشكل التالي:

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

$$= \beta_0 \frac{1}{X_i} + \beta_1 + \mu_i$$

و نجري انحدار ${Y_i}/{X_i}$ على ${1}/{X_i}$ باستخدام طريقة المربعات الصغرى العادية ونحصل على النموذج المقدر التالي:

$$\left[\frac{\widehat{Y}_i}{X_i}\right] = \hat{\beta}_0 \frac{1}{X_i} + \hat{\beta}_1$$

ثم نضرب المعادلة المحولة ب X_i ويتم بذلك الحصول على النموذج الأصلي $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1$ بعد معالجة عدم ثبات التباين $\sigma_{\varepsilon_1}^2$.

الافتراض الثاني:

إذا كان $E\left(\varepsilon_i^2\right)= \sigma^2 X_i$ ، إذن يتم تحويل النموذج الأصلي بقسمته على $\sqrt{X_i}$كما يلي:

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\beta_0}{\sqrt{X_i}} + \beta_2\sqrt{X_i} + \frac{\varepsilon_i}{\sqrt{X_i}}$$

$$= \beta_0 \frac{1}{\sqrt{X_i}} + \beta_1\sqrt{X_i} + \mu_i$$

و أيضا كما في الحالة الأولى نجري انحدار $Y_i/\sqrt{X_i}$ على $\sqrt{X_i}$ بواسطة MCO.

الافتراض الثالث :

إذا كان $E\left(\varepsilon_i^2\right)= \sigma^2 Y_i^2$ إذن تكون المعادلة المحولة من الشكل:

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

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

الافتراض الرابع

يتم تحويل النموذج الأصلي وذلك بإدخال اللوغاريتم أي تحويله إلى صيغته اللوغاريتمية كما يلي:

$$lnY_i = \beta_0 + \beta_1 lnX_i + \varepsilon_i$$

وهذا غالبا ما يؤدي إلى تقليل درجة عدم ثبات التباين لحد الخطأ.

لاختبار فرضية عدم ثبات التباين وحل مشكلة عدم ثبات التباين إن وجد نقوم ببعض الأمثلة:

مثال 1:

لتكن لدينا البيانات التالية التي تعطينا الإنفاق الاستهلاكي C بدلالة الدخل التصرفي Y_dفي اقتصاد دولة ما لمدة 12 سنة كما يلي:

السنة	الإنفاقC	Y_d
1	26.1	38.3
2	29.3	43.5
3	35.6	53.5
4	39.4	60.8
5	42.7	66.4
6	46.3	71.2
7	50.1	77.2
8	54.5	86.1
9	60.1	94.6
10	64.9	102.4
11	69.2	109.9
12	73.1	115.6

المطلوب: قم باختبار عدم التجانس (ثبات) تباين الأخطاء باستخدام اختبار .Goldfeld-Quandt

الحل: نلاحظ أن قيم Y_d مرتبة من أصغر إلى أكبر قيمة له ولتطبيق هذا الاختبار نقوم بحذف المشاهدات التي تقع

في وسط العينة حيث تأخذ خمس مشاهدات $\frac{1}{5}\times 12 = 2.5$

إذن نحذف مشاهدتين في الوسط ثم نقسم العينة إلى مجموعتين يصبح لدينا:

المجموعة الأولى: تبدأ من 1 إلى 5

المجموعة الثانية: تبدأ من 8 إلى 12

أي :

المجموعة الأولى: $\bar{Y}_{d_1} = 52.5 \quad \bar{C}_1 = 34.62$

	C_{1i}	Y_{d_1}	$c_i - \bar{c}$	$Y_{d_i}-\bar{Y}_d$	$(C_1-\bar{C})(Y_{d_i}-\bar{Y}_d)$	$(Y_{d_i}-\bar{Y}_d)^2$	$\hat{C}_{i1}$	$C_1-\hat{C}_i$	$\hat{\varepsilon}_{1i}^2$
1	26.1	38.3	-8.52	-14.2	120.984	201.64	26.24	-0.14	0.0213
2	29.3	43.5	-5.32	-9.0	47.88	81.0	29.31	-0.012	0.0001
3	35.6	53.5	0.98	1.0	0.98	1.0	35.20	0.390	0.1523
4	39.4	60.8	4.78	8.3	39.674	68.89	39.51	-0.114	0.0131
5	42.7	66.4	8.08	13.9	112.312	193.21	42.81	-0.117	0.0137
Σ	17.31	262.5	0	0	321.83	454.75	173.1	0	0.2006

يصبح لدينا: $\hat{\beta}_1 = \frac{\sum(C_1-\bar{C})(Y_{d_i}-\bar{Y}_d)}{(Y_{d_i}-\bar{Y}_d)^2} = \frac{321.83}{454.74}$

$\hat{\beta}_1 = 0.589$

$\hat{\beta}_0 = \bar{c} - \hat{\beta}_1\bar{Y}d = 34.62 - (0.589)(52.5)$

$\hat{\beta}_0 = 3.660$

إذن النموذج المقدر الأول :

$$\hat{C}_{1i} = 3.66 + 0.589 Y_{d_{1i}}$$

$$\hat{\sigma}_{\hat{\varepsilon}_1}^2 = \frac{\sum \hat{\varepsilon}_{1i}^2}{n-2} = \frac{0.2006}{3} = 0.668$$

المجموعة الثانية:

$\bar{C_2} = 64.36$ ، $\bar{Y}d_2 = 101.72$

	C_{2i}	Y_{d_2}	$c_{2i}-\bar{c}$	$Y_{d_{2i}}-\overline{Y}_d$	$(C_{2i}-\bar{C})(Y_{d_i}-\overline{Y}_d)$	$(Y_{d_{2i}}-\overline{Y}_d)^2$	$\hat{C}_{i2}$	$C_{2i}-\hat{C}_i$	${\hat{\varepsilon}_{2i}}^2$
8	54.5	86.1	-9.86	-15.62	154.013	243.98			
9	60.1	94.6	-4.26	-7.12	30.331	50.69			
10	64.9	102.4	0.54	0.68	0.3672	0.64			
11	69.2	109.9	4.84	8.18	39.591	66.91			
12	73.1	115.6	8.74	13.88	121.311	192.65			
	321.8	508.6	0	0	354.614	554.70	215.33		0.13523

$$\hat{\beta}_1 = 0.623 \qquad \hat{\beta}_0 = 0.982$$

إذن

$$\hat{C}_{2i} = 0.982 + 0.623Y_{d_{2i}}$$

طريقة أخرى:

نظريا لدينا $\sum {\hat{\varepsilon}_{2i}}^2 = \sum {Y_i}^2 - \sum \hat{Y}_i^2$

$$\sum \hat{Y}_i^2 = \hat{\beta}_1 \sum (X_i - \overline{X})(Y_i - \overline{Y})$$

إذن يصبح لدينا $\sum {\hat{\varepsilon}_{2i}}^2 = \sum (C_i)^2 - \sum \hat{C}_i^2$

$$\sum \hat{C}_i^2 = \hat{\beta}_1 \sum (Y_d - \bar{Y})(C_i - \bar{C}) = 0.623(354.61) = 215.336$$

$$\Rightarrow \sum {\hat{\varepsilon}_{2i}}^2 = \sum {C_i}^2 - \sum \hat{C}^2 = 215.472 - 215.336$$

$$\sum {\hat{\varepsilon}_{2i}}^2 = 0.135$$

$$\hat{\sigma}_{\hat{\varepsilon}_2}^{\ 2} = \frac{\sum {\hat{\varepsilon}_{2i}}^2}{n-2} = \frac{0.135}{5-2} = 0.045$$

إذن نقوم بحساب إحصائية فيشر

$$F_c = \frac{\hat{\sigma}_{\hat{\varepsilon}_2}^{\ 2}}{\hat{\sigma}_{\hat{\varepsilon}_1}^{\ 2}} = \frac{0.0450}{0.0668} = 0.673$$

نستخرج قيمة فيشر الجدولية عند معنوية 5% و درجة حرية (3.3) للبسط و المقام $F_T = 9.2$

إذن نختبر الفرضية H_0 و التي مقامها ثبات التباين حد الخطأ.

لدينا $F_C < F_T$ إذن نقبل H_0 و بالتالي ثبات (تجانس) تباين حد الخطأ

مثال 2:

الفصل الرابع: مشاكل القياس الاقتصادي-Economic Measurement Problems

في تجربة على 30سيارة ضمن 05 ورشات لحساب الوقت المستغرق في مراجعة سلامة السيارات وكانت النتائج كما يلي:

الزمن المستغرق	عدد العيوب في كل ورشة				
	الورشة1	2	3	4	5
4	8	8	6	5	4
3.5	17	15	13	11	6
2	21	15	14	13	9
1.5	26	23	16	13	6
1	34	22	17	15	11
0.5	38	28	23	21	7

المطلوب: باستخدام اختبار white، قم بالكشف عن عدم ثبات التباين.

الحل

1-نقوم أولا بتقدير النموذج العام لعدد العيوب Y_i على الزمن المستغرق X_i لعينة من 30 سيارة كما يلي:

$$\hat{Y}_i = 24.094 - 4.125X$$

$$R^2 = 0.386$$

$$F_C = 17.63$$

$$DW = 1.86$$

$$\sum_{1=1}^{30} \hat{\varepsilon}^2 = 1311.40$$

ثم نقوم بتقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}^2{}_i = \beta_0 + \beta_1 X_i + X_i{}^2\beta_2 + \mu_i$$

تحصلنا على المعادلة التالية

$$\hat{\varepsilon}^2{}_i = -78.58X_i + 11.98X_i{}^2 + 136.02 + \mu_i$$

مع $R^2 = 0.226$

$$F_C = 3.956$$

نختبر الفرضية H_0 و التي مقامها ثبات التباين حد الخطأ و ذلك بحساب إحصائية معامل لانجراج

$$LM = n\ R^2 \sim X^2_{(0.05\ ,2)}$$

$$LM = 30*0.226 = \quad 6.78$$

$$X^2_{T(0.05\times 2)} = 5.99$$

إذن لدينا$LM>X_T^2$ إذن نفرض H_0 و بالتالي وجود مشكلة عدم تجانس تباين الأخطاء.

(هذه الاختبارات يمكن تطبيقها على البرامج الإحصائية و هذا ما سنراه على برنامج Eviews).

المثال الثالث:

نريد تقدير معادلة توضح أثر كل من عدد الغرف و المساحة على ثمن الشقق في مدينة ما كما يلي:

$$P= \beta_0 + \beta_1 R + \beta_2 S + \varepsilon_i$$

حيث:

P: ثمن الشقة

R: عدد الغرف

S: مساحة الشقة

وكان عدد الشقق 88 شقة ، و تم تقدير النموذج وتحصلنا على النتائج التالية:

$$\hat{P}= -19315 + 15198.19 R + 128.43 S$$

$$R^2= 0.63 \quad ,\sum \hat{\varepsilon}_i^{\,2} = 3.38\times 10^{11}, \quad F_C= 72.96$$

نقوم بتقدير النموذج الوسيطي (المساعد)

$$\hat{\varepsilon}_i^2= \alpha_0 + \alpha_1 R + \alpha_2 S + \mu_i$$

تحصلنا على المعادلة التالية

$$\hat{\varepsilon}_i^2 = -8.22*10^9+1.19*10^9R+3881720S$$

$$R^2 = 0.120 \quad , \quad F_C = 5.80$$

إذن نختبر الفرضية : H_0 $\alpha_0 = \alpha_1 = \alpha_2 = 0$(ثبات التباين حد الخطأ)

إذن:

-نقارن F_C ب F_T عند $\alpha = 0.05$ ودرجة حرية $F_T(2.85)=3.07$

لدينا: $F_C>F_T$ومنه نفرض H_0 إذن وجود مشكلة تجانس تباين الأخطاء.

-نحسب إحصائية لاغرانج

$$LM= nR^2 \sim X^2_{(0.05\ ,2)}$$

$$LM= 88\times 0.120= 10.56$$

نستخرج إحصائية X^2 الجدولية$X^2_{(5\%,2)}=5.99$

إذن$LM>X^2_{(5\%\ ,2)}$ ومنه نفرض H_0إذن وجود مشكلة عدم ثبات التباين.

إذن قمنا باختبار عدم ثبات التباين باستخدام اختبارBreusch-Pagan

ولتصحيح النموذج وبما أن σ^2 معلوم حسب نتائج تقدير النموذج باستخدام برنامج eviews σ=62315.95 (الانحراف المعياري)

نقوم بتقدير النموذج بواسطة MCO

$$\frac{P_i}{\sigma_{\varepsilon_i}} = \beta_0 \frac{1}{\sigma_{\varepsilon_i}} + \beta_1 \frac{R_i}{\sigma_{\varepsilon_i}} + \beta_2 \frac{S}{\sigma_{\varepsilon}} + \frac{\varepsilon_i}{\sigma_{\varepsilon_i}}$$

بقسمة المعطيات أو البيانات للمتغيرات على σ ثم نعيد اختبار الكشف عن إشكالية عدم التباين.

4–مشكلة ارتباط المتغير المستقل مع الاخطاء:

4–1–سبب هذه المشكلة: هو نقض الفرضية السادسة في نموذج الانحدار الخطي و التي تتمثل في استقلالية المتغيرات المستقلة و الاخطاء، حيث $cov(X_i, \varepsilon_i) = 0$ و سبب هذا الاشكال هو خطا في قياس المشاهدات للمتغيرات المستقلة او المتغير التابع مما يجعلها لا تمثل بدقة المجتمع المدروس.

اذا كان لدينا البيانات الحقيقية غير المعلومة التالية $y^*, x_1^*, x_2^*, x_3^*, \dots x_k^*$ للمشاهدات المتحصل عليها من القياس و هي مشاهدات معلومة $y, x_1, x_2, \dots x_k$ حيث نقوم بدراسة خصائص مقدرات المربعات الصغرى كما يلي:

لدينا النموذج العام التالي: $Y^* = X^*\beta + \varepsilon$، حيث $X = X^* + \mu$ و $Y = Y^* + \vartheta$ مع افتراض مايلي:

$$E(\mu) = 0, \quad E(\vartheta) = 0, \quad E(X^*\mu) = 0, \quad E(Y^*\vartheta) = 0, E(X^*\vartheta) = 0, E(Y^*\mu) = 0$$

ويمكننا إثبات مايلي:

$$E(\acute{\varepsilon}\mu) = E\left[(Y^* -\acute{} X^*\beta)\mu\right] = E\left(\acute{Y}^*\mu\right) - \beta E\left(\acute{X}^*\mu\right) = 0$$

$$E(\acute{\varepsilon}\vartheta) = E\left[(Y^* -\acute{} X^*\beta)\vartheta\right] = E\left(\acute{Y}^*\vartheta\right) - \beta E\left(\acute{X}^*\vartheta\right) = 0$$

و هذا ما يعني استقلالية خطأ القياس للنموذج عن اخطاء المتغيرات μ، ν و منه العلاقة بين المتغيرات المشاهدات X، Y تكون كالتالي:

$$Y^* = Y - \vartheta = (X - \mu)\beta + \varepsilon \Rightarrow Y = X\beta + \vartheta - \mu\beta + \varepsilon = X\beta + \gamma$$

حيث: $\gamma = \vartheta - \mu\beta + \varepsilon$ و الذي له الخصائص التالية:

$$E(\gamma) = E(\vartheta - \mu\beta + \varepsilon) = E(\vartheta) - E(\mu)\beta + E(\varepsilon) = 0$$

$$E\left(\acute{X}^*\gamma\right) = E\left(\acute{X^*}\theta\right) - E\left(\acute{X^*}\mu\right)\beta + E\left(\acute{X^*}\varepsilon\right) = 0$$

$$E\left(\acute{X}\gamma\right) = E[(X^* + \mu)\gamma] = E(\acute{\mu}\gamma) = E(\acute{\mu}\vartheta) - E(\acute{\mu}\mu)\beta + E(\acute{\mu}\varepsilon) = -E(\acute{\mu}\mu)\beta \neq 0$$

و منه نتأكد أن المتغيرات X مرتبطة بالأخطاء γ مما ينقض الفرضية السادسة و يجعل من مقدرات طريقة المربعات الصغرى غير دقيقة.

2-4 تصحيح مشكلة ارتباط المتغير المستقل بالأخطاء:

1-2-4 طريقة التقدير بالمتغيرات الوسيطة *Instrumental Variables*:

تأكدنا من قبل أن وجود مشكلة ارتباط المتغيرات المستقلة بالأخطاء يجعل من مقدرات المربعات الصغرى $\hat{\beta}$ لن تتجه نحو القيم الحقيقية للمعلمات β، الأمر الذي يتطلب طرق جديدة للتقدير.

ففي هذه الطريقة يجب تحديد k متغير بحيث يكون لدينا:

$$E(\acute{Z}\gamma) = 0;\ Z = (Z_1, Z_2, \ldots\ldots, Z_k); Cov(\acute{Z}X) \neq 0$$

كما لدينا:

$$E(\acute{Z}Y) = E\left(\acute{Z}(X\beta + \gamma)\right) = E(\acute{Z}X)\beta + E(\acute{Z}\gamma) = E(\acute{Z}X)\beta$$

و تكون المقدرات على النحو التالي

$$\hat{\beta} = (\acute{Z}X)^{-1}\acute{Z}Y$$

يمكننا ان نلاحظ ان تباين المقدرات يكون ضعيف كلما زادت قوة الارتباط بين Z' و X و ذلك من خلال مصفوفة التباينات المشتركة للمقدرات التي تكون على النحو التالي:

$$\hat{\Omega}_{\hat{\beta}} = \hat{\sigma}_{\varepsilon}^2(Z'X)^{-1}(Z'Z)(X'Z)^{-1}$$

تشمل صعوبة هذه المنهجية في تحديد المتغير الوسيطي Z بحيث يكون غير مرتبط بالأخطاء γ و يكون قوي الارتباط بالمتغيرات X، لكن في بعض الحالات يمكن اعتبار المتغير الوسيطي هو المتغير المستقل بتأخر فترة زمنية.

2-2-4 اختبار التجانس (1978) **Hausman**:

يستخدم هذا الاختبار في فحص إمكانية وجود علاقة بين المتغير المستقل X_{it} و الأخطاء ε_t وفق الفرضيتين التاليتين:

$$\begin{Bmatrix} H_0: Cov\ (x_t, \varepsilon_t) = 0 \\ H_1: Cov\ (x_t, \varepsilon_t) \neq 0 \end{Bmatrix}$$

لدينا في هذه الحالة طريقتين لإجراء الاختبار و ذلك كالتالي:

الطريقة الأولى: اختبار الفروقات:

تقوم هذه الطريقة من خلال حساب الإحصائية التالية:

$$H = (\hat{\beta}_{VI} - \hat{\beta}_{MCO})'\left[Var\,(\hat{\beta}_{VI}) - Var(\hat{\beta}_{MCO})\right]^{-1}(\hat{\beta}_{VI} - \hat{\beta}_{MCO})$$

حيث $\hat{\beta}_{VI}$ مقدرات طريقة المتغيرات الوسيطة في حين $\hat{\beta}_{MCO}$ مقدرات طريقة المربعات الصغرى.

إحصائية H_0 تتبع توزيع كاي تربيع χ^2 ب K درجة حرية، حيث يكون القرار كالتالي:

- ✓ إذا كان $H<\chi^2(K)$ نقبل الفرضية العدمية و نقول أن مقدرات طريقة المربعات الصغرى غير متحيزة مما يعني أنه لا يوجد ارتباط بين المتغير المستقل X_{it} و الأخطاء ε_t عند درجة معنوية $\alpha\%$.
- ✓ إذا كان $H \geq \chi^2(K)$ نرفض الفرضية العدمية و نقول أن مقدرات طريقة المربعات الصغرى متحيزة مما يعني أنه يوجد ارتباط بين المتغير المستقل X_{it} و الأخطاء ε_t عند درجة معنوية $\alpha\ \%$.

الطريقة الثانية: طريقة الانحدار المتصاعد أو الموسع:

تتمثل هذه الطريقة في أربعة مراحل:

المرحلة الأولى: تقدير نموذج المربعات الصغرى حيث يكون المتغير التابع هو المتغير المستقل الذي نريد دراسة ارتباطه مع الأخطاء إضافة و المتغير الوسيط حيث عموما يكون المتغير المستقل متأخر بفترة زمنية واحدة.

$$XX_t = \beta_{01} + \beta_{11}Z_t + \varepsilon_{1t}$$

المرحلة الثانية: استخراج القيم ل $\widehat{XX}_t$ من خلال معادلة التقدير المتحصل إليها.

المرحلة الثالثة: تتمثل في إدراج القيم المقدرة للمرحلة الأولى ضمن النموذج التالي:

$$Y_t = \beta_{02} + \beta_{12}x_t + \beta_{22}\widehat{XX}_t + \varepsilon_{2t}$$

حيث تمثل $\widehat{XX}_t$ القيم المقدرة من خلال المرحلة الأولى.

المرحلة الرابعة: نقوم باختبار معنوية المعلمة β_{22} حيث إذا كانت معنوية نرفض الفرضية العدمية و نقول بوجود ارتباط بين المتغير المستقل قيد الدراسة و الأخطاء مما يعني $Cov\ (X_t, \varepsilon_t) \neq 0$.

3-2-4. طريقة العزوم المعممة للتقدير La méthode des moments généralisées:

طريقة العزوم المعممة للتقدير GMM (La méthode des moments généralisées) تستعمل في حالة الإخلال بالفرضيتين السادسة ($E(\varepsilon_t, \varepsilon'_t) = \sigma_\varepsilon^2 I \quad et \quad Cov(X_t, \varepsilon_t) = 0$)، حيث تتميز هذه الطريقة بدمج كلا من طريقتي المربعات الصغرى المعممة و طريقة التقدير بالمتغيرات الوسيطية في إطار مقدرة واحدة تحسب كالتالي:

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

حيث Y : المتغير التابع، X: المتغيرات المستقلة، Z: المتغيرات الوسيطية، Ω: مصفوفة التباين و التباين المشترك المقدرة في المرحلة الأولى بواسطة طريقة المتغيرات الوسيطية.

مع الإشارة إلى أنه في حالة تحقق الفرضية $E(\varepsilon_t, \varepsilon'_t) = \sigma_\varepsilon^2 I$ فمقدرة GMM سوف تصبح هي نفسها المقدرة VI للتقدير بواسطة المتغيرات الوسيطية $\hat{\beta} = (Z'X)^{-1}Z'Y$.

المثال الأول:

لدينا معادلة التقدير التالية التي تربط بين متغير تابع و ثلاثة متغيرات مستقلة على النحو التالي:

$$y_t = 112.25 + 23.1x_{1t} - 10.23x_{2t} + 32.56x_{3t} + e_t$$

(3.22) (5.3) (4.2) (4.6)

كما لدينا:

$$R^2 = 0.74; n = 85, (.) = t\ de\ student\ ; r^2_{y,x_1} = 0.63$$

المطلوب:

1. حساب معامل الارتباط الجزئي $r^2_{x_2,x_1}$.

الحل:

لغرض حساب هذا المعامل لدينا العلاقة التالية:

$$1 - R^2{}_{y,x_1,x_2,x_3} = \left(1 - r^2_{y,x_1}\right)\left(1 - r^2_{y,x_{2,}x_1}\right)\left(1 - r^2_{y,x_3,x_1,x_2}\right)$$

$$(1 - r^2_{y,x_{2,}x_1}) = \frac{1 - R^2{}_{y,x_1,x_2,x_3}}{(1 - r^2_{y,x_1})(1 - r^2_{y,x_3,x_1,x_2})}$$

لكن بداية نحسب مايلي:

$$\left(1 - r^2_{y,x_3,x_1,x_2}\right) = \frac{t^2_{\hat{a}_3}}{t^2_{\hat{a}_3} + (n - k - 1)} = \frac{4.6^2}{4.6^2 + (85 - 3 - 1)} = 0.2071$$

$$(1 - r^2_{y,x_{2,}x_1}) = \frac{1 - R^2{}_{y,x_1,x_2,x_3}}{(1 - r^2_{y,x_1})(1 - r^2_{y,x_3,x_1,x_2})} = \frac{1 - 0.74}{(1 - 0.63)(1 - 0.2071)} = 0.886$$

$$r^2_{y,x_{2,}x_1} = 0.1137$$

المثال الثاني:

لدينا الجدول التالي للمتغير التابع Y_i و المتغيرات المستقلة X_1, X_2, X_3 كالتالي:

i	Y_i	X_1	X	X_3
1	4	7	1	8
2	5	2	5	5
3	2	3	6	1
4	3	9	7	3
5	4	4	8	6
6	5	8	2	4
7	8	6	3	8
8	6	5	0	2

9	1	1	4	0
10	5	2	8	1
11	6	3	9	4
12	4	5	3	6
13	5	7	5	8
14	8	6	1	5
15	2	5	6	2
16	2	6	8	7
17	6	0	2	9
18	5	2	6	5
19	4	4	4	3
20	2	8	2	1

حيث أن معادلة التقدير في هذه الحالة هي:

$$y_i = 4.324 - 0.113x_{1i} - 0.198x_{3i} + e_t$$
$$(3.22) \quad (5.3) \quad (4.2) \quad (4.6)$$
$$R^2 = 0.3115\,; (.) = t\ de\ student$$

$$r^2_{x_1x_2} = 0.0519;\ r^2_{x_1x_3} = 0.00951;\ r^2_{x_2x_3} = 0.0181$$

المطلوب:

1. اختبر فرضية التعدد الخطي للنموذج المقدر باستعمال اختبارين.
2. اختبر فرضية الارتباط الذاتي للأخطاء للنموذج باستعمال اختبار داربن واتسن.
3. اختبر مشكلة الارتباط الذاتي للأخطاء عند درجة أكبر من واحد حيث لدينا:

$$\hat{\varepsilon}_{\mathrm{i}} = -0.284 + 0.117\mathrm{X}_{1\mathrm{i}} - 0.060\ \mathrm{X}_{2\mathrm{i}} + 0.038\ \mathrm{X}_{3\mathrm{i}} - 0.055\hat{\varepsilon}_{\mathrm{t}-1} - 0.445\hat{\varepsilon}_{\mathrm{t}-2} + \mu_{1\mathrm{t}}$$
$$R^2 = 0.211$$

$$\hat{\varepsilon}_{\mathrm{i}} = -0.549 + 0.252\mathrm{X}_{1\mathrm{i}} - 0.020\ \mathrm{X}_{2\mathrm{i}} + 0.074\ \mathrm{X}_{3\mathrm{i}} - 0.103\ \hat{\varepsilon}_{\mathrm{t}-1} - 0.566\ \hat{\varepsilon}_{\mathrm{t}-2} + 0.526\ \hat{\varepsilon}_{\mathrm{t}-3} + \mu_{1\mathrm{t}}$$
$$R^2 = 0.469$$

الحل:

1. اختبار التعدد الخطي:

بداية مع إجراء اختبار Klein الذي يتضح جليا من خلال المعطيات أن معامل التحديد أكبر من مربعات معاملات الارتباط البسيطة كلها حيث $R^2 > r^2_{x_ix_j}$ مما يدل على عدم وجود مشكل التعدد الخطي في النموذج المقدر من خلال هذا الاختبار.

ننتقل الآن إلى الطريقة الثانية المتمثلة في اختبار Farrar-Glauber حيث بداية لابد من تحديد مصفوفة معاملات الارتباط البسيطة على النحو التالي:

$$D = \begin{pmatrix} 1 & \sqrt{0.0519} & \sqrt{0.00951} \\ \sqrt{0.0519} & 1 & \sqrt{0.0181} \\ \sqrt{0.00951} & \sqrt{0.0181} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.227 & 0.0975 \\ 0.0227 & 1 & 0.134 \\ 0.0975 & 0.134 & 1 \end{pmatrix}$$

حيث قيمة المحدد لهذه المصفوفة هو $|D| = 0.9263$

المرحلة الموالية تتمثل في حساب إحصائية كاي تربيع المحسوبة وفق العلاقة التالية:

$$x^{2*} = -\left[\mathrm{n} - 1 - \frac{1}{6}(2\mathrm{K} + 5)\right].\ln|D| = -\left[20 - 1 - \frac{1}{6}(2 \times 4 + 5)\right].\ln(0.09263) = 1.288$$

و أخيرا نستخرج قيمة x^2 الجدولية عند $6 = (1 - 4)4) \times \frac{1}{2} = \frac{1}{2K(K+1)}$ منه $x^2(6) = 12.59$.

بالتالي لدينا $x^2 > x^{2*}$ الأمر الذي يعني قبول الفرضية العدمية و القول بعدم وجود تعدد خطي في النموذج المقدر أعلاه عند درجة معنوية 5% الأمر الذي يعزز نتيجة اختبار Klein.

2. اختبار الارتباط الذاتي للأخطاء:

لغرض تطبيق اختبار داربن واتسن لابد من حساب الإحصائية المحسوبة للاختبار وفق العلاقة التالية:

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

التي بداية تتطلب حساب كلا من $\sum_{t=1}^{n} (\hat{\varepsilon} - \hat{\varepsilon}_{t-1})^2$ و $\sum_{t=1}^{n} \hat{\varepsilon}_t^2$ من خلال الجدول أدناه:

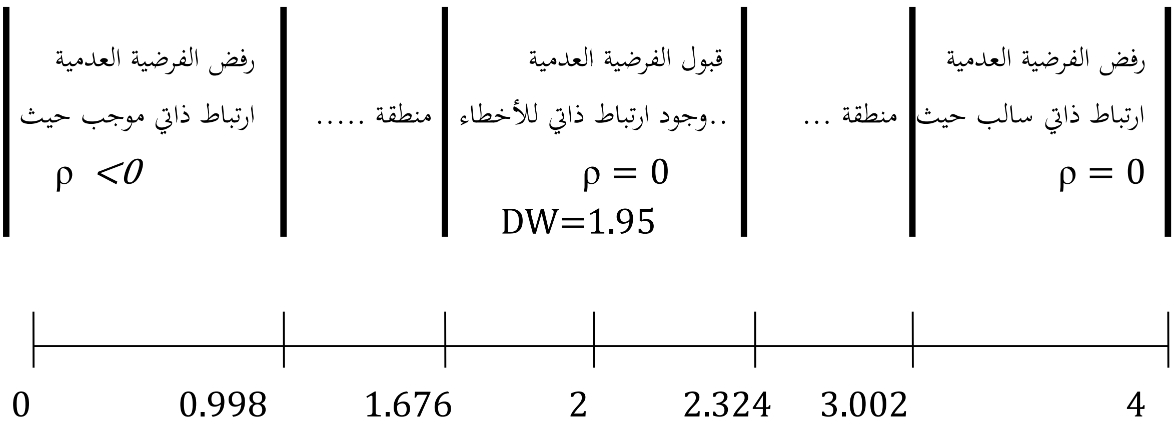
i	y_i	$\hat{y}_i$	$\hat{\varepsilon}_i$	$\hat{\varepsilon}_i^2$	$\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1}$	$(\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2$
1	4	5.959	-1.959	3.837681		
2	5	4.748	0.252	0.063504	2.211	4.888521
3	2	3.125	-1.125	1.265625	-1.377	1.896129
4	3	2.905	0.095	0.009025	1.22	1.4884
5	4	4.256	-0.256	0.065536	-0.351	0.123201
6	5	4.336	0.664	0.440896	0.92	0.8464
7	8	5.676	2.324	5.400976	1.66	2.7556
8	6	4.415	1.585	2.512225	-0.739	0.546121
9	1	3.419	-2.416	5.851561	-4.004	16.032016
10	5	2.842	2.158	4.656964	4.577	20.948929
11	6	3.515	2.485	6.175225	0.327	0.106929
12	4	5.133	-1.133	1.283689	-3.618	13.089924
13	5	5.167	-0.167	0.027889	0.966	0.933156
14	8	5.088	2.912	8.479744	3.079	9.480241
15	2	3.227	-1.227	1.505529	-4.139	17.131321
16	2	4.358	-2.358	5.560164	-1.131	1.279161
17	6	6.88	-0.88	0.7744	1.478	2.184484

18	5	4.55	0.45	0.2025	1.33	1.7689
19	4	4.064	-0.064	0.004096	-0.514	0.264196
20	2	3.352	-1.352	1.827904	-1.288	1.658944
Σ				49.945133		97.422573

و منه نطبق العلاقة أعلاه حيث:

$$DW = \frac{\sum_{t=1}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = \frac{97.422573}{49.945133} = 1.950$$

أما فيما يخص إحصائيتي داربن واتسن الجدوليتين فلدينا عند درجة معنوية 5% ، 20 مشاهدة و 3 متغيرات القيم التالية من الجدول في الملاحق $d_1 = 0.998$ و $d_2 = 1.676$ و لتبسيط القرار نقوم برسم الشكل التالي:



من خلال الشكل نلاحظ أن قيمة داربن واتسن المحسوبة تقع ضمن المجال d_2 و d_2 -4 مما يدل على عدم وجود مشكل الارتباط الذاتي للأخطاء في النموذج المقدر من الدرجة الأولى عند درجة معنوية 5%.

3. اختبار Breusch-Godfery:

لهذا الغرض نقوم بحساب إحصائية مضاعف لاغرانج للمعادلة الأولى وفق العلاقة التالية:

$$LM = (n-p)\, R^2 = (20-2)0.211 = 3.798$$

أما إحصائية كاي تربيع عند 2 درجات حرية فهي $x^2(2) = 5.99$، بالتالي بما أن $LM < x^2(2)$ نقبل الفرضية العدمية و نقول بعدم وجود ارتباط ذاتي للأخطاء من الدرجة الثانية للنموذج المقدر عند درجة معنوية 5%.

أما فيما يخص الارتباط الذاتي من الدرجة الثالثة فبنفس الطريقة لدينا:

$$LM = (n-p)\, R^2 = (20-3)0.496 = 7.97$$

أما احصائية كاي تربيع عند 3 درجات حرية فهي $x^2(3) = 7.81$، بالتالي بما أن $LM > x^2(2)$ نرفض الفرضية العدمية و نقول بوجود ارتباط ذاتي للأخطاء من الدرجة الثالثة للنموذج المقدر عند درجة معنوية 5%.

المثال الثالث:

لدينا المعطيات التالية ل 31 سنة من سنة 1989 إلى غاية 2019 لمتغيرين على النحو التالي:

t	y	x	t	y	x	t	y	x
1989	11.08667	9.521641	2000	10.86597	9.422625	2011	11.05917	9.572341
1990	11.13071	9.536907	2001	10.5072	9.199583	2012	10.37748	9.24619
1991	10.39385	9.161465	2002	10.35003	9.105202	2013	10.35704	9.334768
1992	9.525005	8.579229	2003	10.65408	9.24619	2014	10.39385	9.391411
1993	10.88303	8.940105	2004	10.88211	9.391161	2015	10.31311	9.30556
1994	10.3951	8.644707	2005	10.07103	9.276596	2016	11.29385	9.739909
1995	9.524202	8.372167	2006	11.4432	9.700453	2017	9.639652	9.680281
1996	11.49489	9.754523	2007	11.24155	9.572828	2018	9.968994	9.392662
1997	10.88979	9.569412	2008	10.88208	9.547027	2019	11.05491	9.595126
1998	10.39381	9.213535	2009	11.49915	9.818039			
1999	10.51797	9.213834	2010	9.586171	9.522081			

المطلوب:

-اختر تباين الخطأ إن كان ثابتا ومتجانسا عبر الزمن بطرق مختلفة.

الحل:

1) طريقة Goldfeld-Quandt:

تتم هذه الطريقة وفق عدة مراحل كما سبق و أشرنا و أولى هذه المراحل هي تقدير معادلة الانحدار للمتغيرين و التي باستعمال برنامج Eviews 10 تحصلنا على النتائج التالية:

Variable	Coefficient	Std. Error	t.Statistic	Prob.
X	1.037598	0.24563	4.227106	0.0002
C	0.910358	2.294331	0.396786	0.6944
R-squared	0.381247		Mean dependent var	10.60244
Adjudted R-squared	0.359910		S.D . dependent var	0.575197
S.E. of regression	0.460190		Akaike info criterion	1.347985
Sum squares resid	6.141460		Schwarz criterion	1.440500
Log likelihood	-18.89376		Hannan-quinn criter	1.378142
F-statistic	17.86842		Durbin-Watson stat	2.015557
Prob(F-statistic)	0.000215			

نلاحظ من خلال النتائج أن المتغير المستقل يؤثر معنويا على المتغير التابع نظرا لأن قيمة ستيودنت المحسوبة له (4.227) أكبر من القيمة الجدولية ب 29 درجة معنوية (2.045) عند مستوى احتمال 5%، كما أن النموذج المقدر ككل معنوي أيضا بسبب كبر إحصائية فيشر المحسوبة (17.868) على الجدولية (4.18) عند مستوى احتمال 5%، أيضا النموذج خالي من الارتباط الذاتي للأخطاء نظرا لأن قيمة داربن واتسن قريبة من 2 (2.015) حيث

$d_1 = 1.363$ و$d_2 = 1.496$، هذا ما يعني لحد الساعة النموذج المقدر مقبول من الناحية الإحصائية لكن يتبقى مشكل عدم ثبات تباين الأخطاء.

كمرحلة ثانية نقوم بترتيب المشاهدات تصاعديا حسب قيم المتغير المستقل و ذلك على النحو التالي:

t	y	x	t	y	x	t	y	x
1	9.524202	8.372167	12	10.071034	9.276596	23	10.889789	9.569412
2	9.525005	8.579229	13	10.313111	9.305560	24	11.059173	9.572341
3	10.395100	8.644707	14	10.357044	9.334768	25	11.241550	9.572828
4	10.883035	8.940105	15	10.882114	9.391161	26	11.054914	9.595126
5	10.350031	9.105202	16	10.393845	9.391411	27	9.639652	9.680281
6	10.393845	9.161465	17	9.968994	9.392662	28	11.443200	9.700453
7	10.507202	9.199583	18	10.865975	9.422625	29	11.293849	9.739909
8	10.393815	9.213535	19	11.086671	9.521641	30	11.494894	9.754523
9	10.517970	9.213834	20	9.586171	9.522081	31	11.499151	9.818039
10	10.654078	9.246190	21	11.130713	9.536907			
11	10.377483	9.24190	22	10.882077	9.547027			

ثم نقوم باستبعاد بعض القيم من وسط السلسلة المرتبة، وبما أنه لدينا 31 مشاهدة سوف نقوم بحذف 9 مشاهدات وسطى كي تتبقى لنا 22 مشاهدة كليا بحيث يكون الجزء الأول مكونا من 11 مشاهدة من 1 إلى 11 والجزء الثاني من 21 إلى 31 بحيث m=9.

ثم نقوم كمرحلة رابعة بتقدير المعادلتين لكل جزء من المشاهدات أعلاه على النحو التالي:

المشاهدات من 1 إلى 11.

$$\hat{y}_1 = 1.411 + 0.990\mathrm{x}_1$$
$$R_1^2 = 0.543; SCR_1 = 0.817461; n = 11$$

المشاهدات من 21 إلى 31.

$$\hat{y}_2 = -1.783 + 1.331\mathrm{x}_2$$
$$R_2^2 = 0.0625; SCR_2 = 2.530609; n = 11$$

بعد ذلك نقوم بحساب إحصائية فيشر المحسوبة التي تساوي:

$$F^* = \frac{\mathrm{SCR}^1/_{\mathrm{n}-\mathrm{m}-2(\mathrm{K}+1)}}{\mathrm{SCR}^2/_{\mathrm{n}-\mathrm{m}-2(\mathrm{K}+1)}} = \frac{\sum_{t=1}^{n} e_{1t}^2}{\sum_{=1t}^{n} e_{2t}^2} = \frac{2.530609}{0.817461} = 3.095$$

و لغرض استخراج قيمة فيشر الجدولية عند درجة معنوية 5% لا بد من تحديد عدد درجات الحرية للإختبار وفق العلاقة التالية:

$$\mathrm{Ddl} = \mathrm{ddl}_1 = \mathrm{ddl}_2 = \frac{\mathrm{n}-\mathrm{m}-2(\mathrm{k}+1)}{2} = \frac{31-9-2(2+1)}{2} = 9$$

و من الجدول الملاحق إحصائية فيشر هي $F_{9,9}^{0.05} = 3.18$ و بما أن القيمة المحسوبة أصغر من القيمة الجدولية نقبل الفرضية العدمية و نقول بثبات تباين الأخطاء في النموذج المقدر عند درجة معنوية 5%.

2) اختبار White:

كما أشرنا اختبار White يعتمد على تقدير معادلة وسيطية مبنية على المعادلة العامة للتقدير و التي من الشكل:

$$\hat{\varepsilon}_t^2 = a_0 + a_1 x_{1t} + b_1 x_{1t}^2 + a_2 x_{2t} + b_2 x_{2t}^2 + \cdots + a_k x_{kt} + b_k x_{kt}^2 + \mu_t$$

حيث كانت نتائج التقدير:

$$\hat{\varepsilon}_t^2 = 28.076 - 6.340 x_{1t} + 0.358 x_{1t}^2$$
$$R_1^2 = 0.057156 \,; n = 31$$

ومن المعادلة يمكننا حساب إحصائية مضاعف لاغرانج كالتالي:

$$\text{LM} = n \times R^2 = 0.057156 \times 31 = 1.771$$

أما قيمة كاي تربيع الجدولية عند درجة حركية 2K هي $x^2(2) = 5.99$ و بما أن القيمة المحسوبة لمضاعف لاغرانج أصغر من القيمة الجدواية لكاي تربيع$x^2(2)$ >LM نقبل الفرضية العدمية و نقول أن بعدم وجود مشكلة عدم ثبات تباين الأخطاء مما يعني أن خطأ التقدير في المعادلة ذو تباين ثابت و متجانس عند درجة معنوية 5%.

3) اختبار ARCH-LM:

كما أشرنا اختبار ARCH-LM يعتمد على تقدير معادلة وسيطية مبنية على المعادلة العامة للتقدير و التي من الشكل:

$$\varepsilon_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + a_2 \varepsilon_{t-2}^2 + \cdots + a_p \varepsilon_{t-p}^2 + \mu_t$$

و كانت نتائج التقدير لثلاثة معادلات بثلاثة تأخرات متتالية على النحو التالي:

$$\hat{\varepsilon}_t^2 = 0.187 + 0.070 \varepsilon_{t-1}^2$$

$$R_1^2 = 0.005024; n = 30; \text{AIC} = 1.098$$

$$\hat{\varepsilon}_t^2 = 0.220 + 0.082 \varepsilon_{t-1}^2 - 0.167 \varepsilon_{t-2}^2$$

$$R_2^2 = 0.031929; n = 29; \text{AIC} = 1.76$$

$$\hat{\varepsilon}_t^2 = 0.238 + 0.072 \varepsilon_{t-1}^2 - 0.171 \varepsilon_{t-2}^2 - 0.054 \varepsilon_{t-3}^2$$

$$R_3^2 = 0.034488; n = 28; \text{AIC} = 1.278$$

حيث سنقوم بحساب مضاعف لاغرانج لكل معادلة وفق العلاقة التالية:

$$LM=(n-p)\times R^2$$

$$LM_1=(31-p_1)\times R_t^2=(31-1)\times 0.005024=0.15072$$

$$LM_2=(31-p_1)\times R_2^2=(31-2)\times 0.031929=0.92594$$

$$LM_3=(31-p_1)\times R_3^2=(31-3)\times 0.034488=0.96566$$

أما قيم إحصائية كاي تربيع الجدولية عند p درجة حرية على التوالي $\chi^2(p=1)=3.84$ ،χ^2(p=2)=5.99، $\chi^2(p=3)=7.81$ و نلاحظ أن كل قيم كاي تربيع الجدولية أكبر من القيم المحسوبة.

تمارين للحل:

التمرين الاول

الجدول الآتي يتضمن المعطيات الخاصة بالواردات كمتغير تابع (M) والدخل الوطني كمتغير مستقل اول (RN)و أسعار الواردات كمتغير مستقل ثاني (PM) في احد الدول للفترة (1985-1993) و تحصلنا على التالي

PM	RN	M	السنة
100	100	100	1985
99	104	106	1986
110	106	107	1987
126	111	120	1988
113	111	110	1989
103	115	116	1990
102	120	124	1991
103	124	133	1992
98	126	137	1993

المطلوب:

الفصل الرابع: مشاكل القياس الاقتصادي-Economic Measurement Problems

-تقدير معالم النموذج مع دراسة:

– جودة ومعنوية النموذج عند مستوى معنوية 5%

– معنوية مقدرات النموذج عند مستوى معنوية 5%

– الكشف عن مشكلة التعدد الخطي باستخدام اختبار Farrar Glauber

-اختبر وجود مشكلة الارتباط الذاتي بالنموذج عند مستوى معنوي α =5%

التمرين الثاني:

نبحث في تقدير دالة استهلاك السلع الزراعية y_i انطلاقا من x_{1i} الدخل المتأتي من الزراعة و x_{2i} الدخل الغير متأتي من الزراعة، و ذلك انطلاقا من عينة 10 سنوات لبلد ما موضحة كما يلي:

n	1	2	3	4	5	6	7	8	9	10
yi	3	4	5	6	7	8	8	9	10	10
x1i	4	5	8	9	12	14	16	18	20	24
x2i	10	8	6	7	5	5	4	2	2	1

المطلوب:

1/ تقدير دالة الاستهلاك

2/ تقييم جودة النموذج

3/ اختبار معنوية مقدرات النموذج عند مستوى معنوية 5%

4/ اختبار المعنوية الكلية للنموذج عند مستوى معنوية 5%

5/الكشف عن مشكلة التعدد الخطي باستخدام اختبار Farrar Glauber

6– الكشف على مشكلة عدم تجانس تباين حد الخطأ بأحد الطرق التي تعرضنا لها في المحاضرة

التمرين الثالث: قمنا بتقدير دالة الواردات بدلالة الناتج الوطني الخام لبلد ما للفترة (1960–1979)، تحصلنا على النموذج التالي:

$$\widehat{M} = -56,13 + 0,13PNB$$

$$\sum_{2}^{T}(e_t - e_{t-1})^2 = 1279,825 \;\wedge\; \sum_{1}^{T} e_t^2 = 2468,96 \;\wedge\; \sum_{2}^{T} e_{t-1}^{T} = 2385,69$$

الفصل الرابع: مشاكل القياس الاقتصادي-Economic Measurement Problems

المطلوب: اختبر وجود مشكلة الارتباط الذاتي بالنموذج عند مستوى معنوي $\alpha = 5\%$

التمرين الرابع:

لتكن لدينا البيانات التالية والمطلوب هو الكشف عن مشاكل القياس ثم توضيح كيفية معالجتها (الارتباط الذاتي، التعدد الخطي وعدم ثبات التباين) بمختلف الطرق التي تعرضنا لها في المحاضرة

الفترة i	(1) y_i	(2) x_i	(3) e_i	(4) $e_i e_{i-1}$	(5) e_i^2
1	3.63	0.97	0.2812	-	0.0791
2	4.20	0.95	0.3654	0.1028	0.1335
3	3.33	0.99	0.4670	0.1706	0.2181
4	4.54	0.91	-0.2662	-0.1243	0.0709
5	2.89	0.98	-0.2159	0.0575	0.0466
6	4.87	0.90	-0.1791	0.0387	0.0321
7	4.90	0.89	-0.3920	0.0702	0.1537
8	5.29	0.86	-0.7307	0.2864	0.5339
9	6.18	0.85	-0.0836	0.0611	0.0070
10	7.20	0.82	0.2077	-0.0174	0.0431
11	7.25	0.79	-0.4710	-0.0978	0.2218
12	6.09	0.83	-0.6594	0.3106	0.4348
13	6.80	0.81	-0.4352	0.2870	0.1894
14	8.65	0.77	0.4432	-0.1929	0.1964
15	8.43	0.76	-0.0197	-0.0087	0.0004
16	8.29	0.80	0.8119	-0.0160	0.6592
17	7.18	0.83	0.4306	0.3496	0.1854
18	7.90	0.79	0.1790	0.0771	0.0320
19	8.45	0.76	0.0003	0.0001	0.0000
20	8.23	0.78	0.2661	0.0001	0.0708

المـــــــراجع

-حمداوي الطاوس(2016)، مدخل للاقتصاد القياسي: دروس و تمارين مرفقة بالحل، دار هومي للنشر، الجزائر

-تومي صالح(2011)، مدخل لنظرية القياس الاقتصادي: دراسة نظرية مدعمة بأمثلة و تمارين- الجزء الاول، ديوان المطبوعات الجامعية، الجزائر.

-حسين بخيت، سحر فتح الله(2019)، الاقتصاد القياسي، ذار اليازوري العلمية للنشر و التوزيع، عمان

-دامودار جوجارات، ترجمة هند عبد الغفار عودة و عفاف علي حسن الدش(2015)، الاقتصاد القياسي- الجزء الاول، دار المريخ للنشر، الرياض.

-عبد القادر محمد عبد القادر عطية(2004)، الحديث في الاقتصاد القياسي بين النظرية و التطبيق، الدار الجامعية للنشر، الاسكندرية.

-محمد شيخي(2011)،طرق الاقتصاد القياسي –محاضرات و تطبيقات، الطبعة الاولى، دار حامد للنشر، عمان.

-مكيد علي(2011)، الاقتصاد القياسي: دروس و مسائل محلولة- الطبعة الثانية، ديوان المطبوعات الجامعية، الجزائر.

- وليد إسماعيل السيفو، فيصل مفتاح و آخرون(2007)، مشاكل الاقتصاد القياسي التحليلي، الأهلية للنشر و التوزيع، عمان.

-Ben M'BAREK ALaya, Séries d'exercices Corrigés d'économétrie, Imprimere Officielle de la République Tunisienne, 2002,

-Ben Vogelvang(2005), Econometrics Theory and Application with Eviews, Pearson Education Limited, 1 st Edition, England.

-Brigitte Dormont(1999), Introduction à l économétrie, édition Montchrestien, Paris

--Bruce E. Hansen, (2019), Econometrics, University of Wisconsin.

- Bruno crépon(2005), économétrie linéaire, poly-cours, Crest. En ligne « http://www.crest.fr/ckfinder/userfiles/files/pageperso/crepon/poly20052006.pdf »

-Bruno crépon, Nicolas Jacquemet(2018), Econométrie : Méthodes et applications, DeBoeck, 2ème édition, Paris

-Farouk Kriaa(2008), Econométrie-Modélisation Econométrique-cours et exercices corrigés, Université El Manar2, Tunisie

-Rachid Bendib(2001), Econométrie : théorie et application, OPU, Alger.

-Régis Bourbonnais, (2015), Econométrie: Cours et exercices corrigés, DUNOD, 9eme édition.

-William H. Greene(2003), Econometric Analysis, Fifth Editon, Pearson Education, United states.