الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة غرداية

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

قسم العلوم المالية والمحاسبة

محاضرات في الاقتصاد القياسي 01

مع تمارين محلولة

الدكتور عبدالرؤوف عباده

أستاذ محاضر صنف ب ، جامعة غرداية

abada.abderraouf@univ-ghardaia.dz

السنة الجامعية 2020/2019

المقدمة:

تعتبر الظواهر الاقتصادية ظواهر معقدة يصعب فهمها في كثير من الحالات، فسلوك بعض المتغيرات، والعلاقة بين هذه المتغيرات لا يمكن فهمها من الناحية النظرية، لكن بظهور علم الاقتصاد القياسي تمكن الكثير من الباحثين الاقتصاديين من تطوير علم الاقتصاد في وقت زمني قصير، كما أن غالبية العلاقات التي تقدمها لنا النظرية الاقتصادية، يمكن صياغتها في صور نماذج قياسية تقدر من واقع البيانات الفعلية، وهذا ما يمكننا من استخدام طرق الاقتصاد القياسي في شرح العلاقات بين مختلف المتغيرات كما تحدده النظرية الاقتصادية.

وبناءا على ما تقدم، فقد روعي في تقديم هذه المطبوعة كيفية تطبيق طرق الاقتصاد القياسي، من خلال تبسيط كيفية وضع المعادلات وصياغاتها بشكل قياسي مع إعطاء بعض التمرين وتوضيح طريقة حلها من خلال استخدام برنامج Eviews.9 للاقتصادي القياسي، وهذه المطبوعة مفيدة لطلبة الليسانس وطلبة الماستر تخصص مالية المؤسسة، اقتصاد كمي، اقتصاد قياسي، اقتصاد تطبيقي.

كما تم تقسيم هذه المطبوعة إلى خمسة أجزاء :- مدخل للاقتصاد القياسي؛

-الارتباط الخطي؛

-تحليل الانحدار الخطي البسيط؛

-تحليل الانحدار الخطي المتعدد؛

-مشاكل القياس الاقتصادي؛

مدخل للاقتصاد القياسي

أولا: الاقتصاد القياسي أهدافه وعلاقته بالفروع الأخرى

لقد أستخدم لفظ اقتصاد قياسي لأول مرة سنة 1926 م، ويرجع الفضل في ذلك للاقتصادي Ranger Frisch و يعرفه البعض بأنه القياس في الاقتصاد (أو القياس الاقتصادي)، و بصورة أكثر تفصيل هو العلم الذي يهتم بقياس العلاقات الاقتصادية من خلال بيانات واقعية، بغرض اختبار مدى صحة هذه العلاقات كما تقدمها النظرية، أو تفسير بعض الظواهر، أو رسم بعض السياسات، أو التنبؤ بسلوك بعض المتغيرات الاقتصادية.

و يلاحظ أن هذا التعريف يركز على نقطتين أساسيتين[1]:

1- **العلاقة بين الاقتصاد القياسي و الفروع الأخرى :** يعتبر الاقتصاد القياسي محصلة لثلاثة فروع من العلوم هي الإحصاء، النظرية الاقتصادية و الاقتصاد الرياضي، أما عن الإحصاء فهو يمدنا بأساليب و طرق القياس مثل الارتباط و الانحدار، بالإضافة إلى البيانات الواقعية المبوبة، أما بالنسبة للنظرية الاقتصادية فهي تحدد لنا العلاقات الاقتصادية المراد قياسها من خلال الفروض المفسرة التي تقدمها، بينما يصيغ لنا الاقتصاد الرياضي هذه العلاقات النظرية في صور معادلات رياضية قابلة للقياس. و لكن هذا لا يعني أن الاقتصاد القياسي ليس له صفة مستقلة عن هذه الفروع، و إنما هو فرع متميز عن كل واحد منهما.

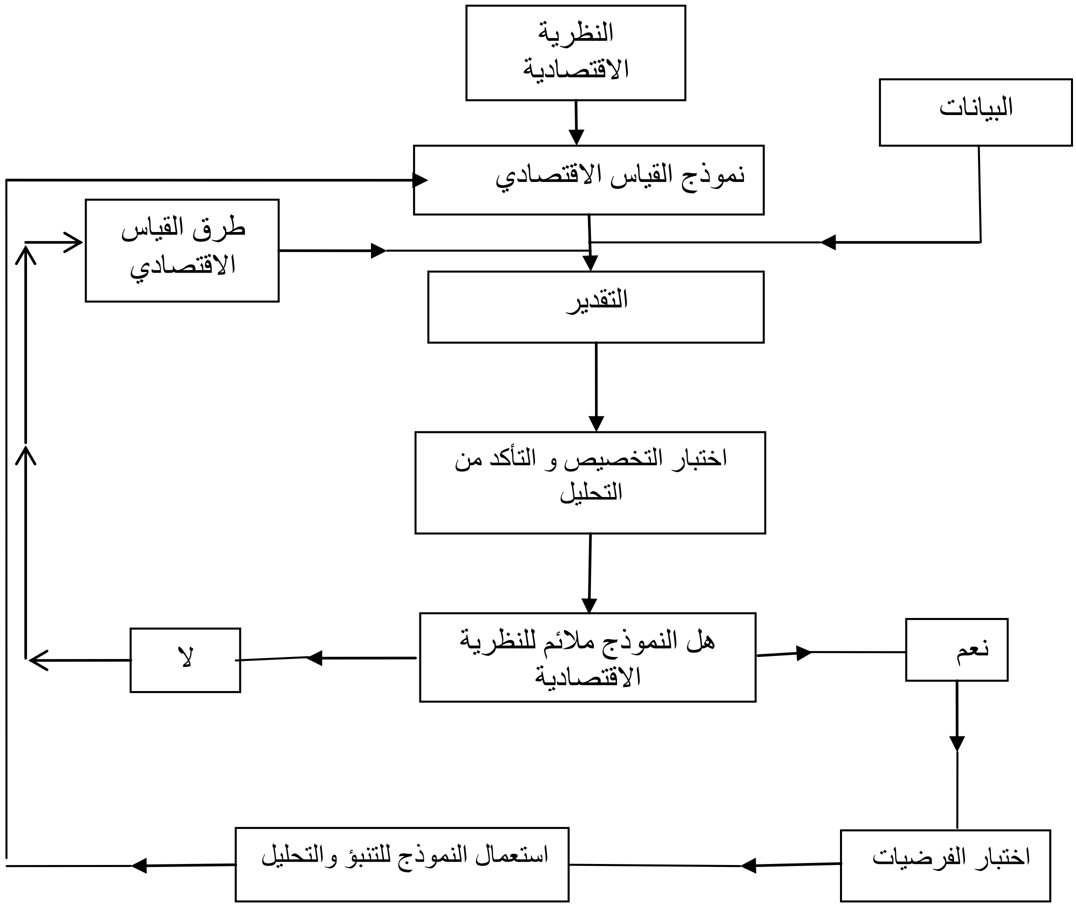
2- **أهداف الاقتصاد القياسي :** هناك ثلاثة أهداف رئيسية:

- بناء النماذج القياسية الاقتصادية في شكل قابل للإختبار الميداني، و تمثل هذه المرحلة مشكلة تصور الصياغة الرياضية في منهجية القياس الاقتصادي.
- تقدير و إختبار هذه النماذج مستعملين البيانات المتوفرة، و تمثل هذه العملية المرحلة الإحصائية في القياس الاقتصادي.
- استعمال النماذج المقدرة لغرض التنبؤ، التحليل الاقتصادي أو اتخاذ القرارات المناسبة.

لإيضاح أكثر نستطيع إدراج المخطط التالي، الذي يبين وضعية الاقتصاد القياسي من النظرية الاقتصادية.

[1]- عبد القادر محمد عبد القادر عطية، **الحديث في الاقتصاد القياسي بين النظرية و التطبيق**، الإسكندرية، الدار الجامعية،2004، ص4.

وضعية الاقتصاد القياسي من النظرية الاقتصادية



المصدر: تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الطبعة الثانية، الجزء الأول ، ديوان المطبوعات الجامعية، الجزائر، 2011، ص7.

ثانيا: منهجية البحث في الاقتصاد القياسي Methodology of Econometric Research:

لا يختلف الاقتصاد القياسي عن غيره من العلوم الاقتصادية والرياضية في منهجية بحثه. إلا أنه يركز في بحثه على إيجاد ثلاثة قيم رئيسية وهي[2]:

1- قيم المعلمات الاقتصادية القياسية (Econometrics Parameters)، وهي القيم التي تمثل حلقة الوصل والقياس بين المتغيرات المستقلة والمتغير التابع. ولهذا يتركز جل إهتمامها البحث عن أفضل الطرق التي يمكن أن نجد بواسطتها هذه المعلمات بأدق مستوى ممكن لتعبر أصدق تعبير عن طبيعة وقوة العلاقات الحقيقية الموجودة بين المتغيرات، إن كان ذلك المجتمع او العينة بحيث تصل إلى تطابق قيمها مع تلك

2- وليد إسماعيل السيفو، فيصل مفتاح شلوف، صائب جواد إبراهيم جواد، **أساسيات الاقتصاد القياسي التحليلي**، الطبعة الأولى، دار الاهلية، الأردن، 2006، ص 31.

القيم الخاصة بالمجتمع ، كما أنها تكون ممثلة لحقيقة الوضع الاقتصادي للظاهرة قيد الدراسة وسلوكها في الحاضر والماضي.

2- قيم المتغير العشوائي (ε_i)أو (μ_i) والذي يمثل حلقة الوصل بين القيم الحقيقية والقيم المقدرة، والذي يقارب فيما بينها وصولا إلى الصفر بقصد الوصول إلى أقرب دقة للقيم المقدرة (Exactness).

3- قيم المتغير التابع المقدرة او المستقبلية وكذلك قيم المتغيرات المستقلة من خلال الإسقاط والاستطالة (Interpolation & Extrapolation). ولفرض الوصول إلى هذه القيم يتطلب الأمر حسب المنهج القياسي إتباع الخطوات او المراحل الأتية:

المرحلة الأولى: مرحلة المعلوماتية (Information Stage):

وهي مرحلة بالغة الأهمية في القياس والوصف والتحليل فرغم انها تتعلق بالإحصاء وطبيعته، إلا أنها القاعدة الوحيدة والمنطق الحاسم للقياس، وتضم هذه المرحلة الإجراءات الأتية:

1- التحقق من البيانات المتوفرة والمجموعة عن تطور الظاهرة وسلوكها في الماضي وتدقيقها عبر دراسة المجتمع الإحصائي وأسلوب اختيار العينة وملاءمتها للمجتمع، ودقة المعالجة المعلوماتية وتقنيته ودقة جدولة المعلومات ووضعه في رسوم بيانية مناسبة، وعزل المعلومات المشوشة وتحديد التشويش ومصدره...الخ.

2- المعالجة الإحصائية للبيانات كإيجاد المتوسطات المتحركة والثابتة وإزالة المعلومات الشاذة وتصنيفها لأغراض القياس. ويقوم بهذه المهمة الإحصاء الاقتصادي وهو يعد بمثابة العجلات التي يسير بها الاقتصاد القياسي.

المرحلة الثانية: مرحلة الفروض الاقتصادية القياسية وتشخيص النموذج

Maintained Econometric Hypothesis & Specification of Model :

تعتبر الفروض الاقتصادية القياسية المعتمدة على النظرية الاقتصادية، فروض مؤكدة، لأن ذلك يعني افتراض وجود علاقة بين المتغيرات التابعة والمستقلة، ويتم تحديد ذلك كما يلي:

أولا: هو إسقاط أزواج المتغيرات على إحداثيات معينة ورسم الشكل الانتشاري الذي يمكن ان يؤكد لنا وجود علاقة او عدم وجود علاقة بين المتغيرات، والقوة التقديرية لهذه العلاقة وبناء النموذج الاقتصادي Economic Model على ضوئها.

ثانيا: هو صياغة هذه العلاقة باستخدام الرموز الرياضية كتصوير العلاقة بين السعر والطلب على شكل معادلة رياضية معينة كالآتي:

$$D = a - bP$$

حيث: D تمثل الكمية المطلوبة، P تمثل السعر، b,a تمثل قيم (المعلمات)

ثالثا: وهي صياغة النموذج الرياضي (Formulation of Mathematical Model) وهو نموذج افتراضي لخطية او لا خطية العلاقة بين المتغيرات وذلك اعتمادا على الشكل الانتشاري وكذلك تحديد المعلمات التي تحل العلاقة بين المتغيرات.

رابعا: بعد استكمال تحديد النموذج الرياضي يتم إدخال المتغير العشوائي (ε_i) لتقدير الأخطاء المعيارية للمعادلة ولصياغة النموذج القياسي (Econometric Model)

المرحلة الثالثة: مرحلة التقدير للمعلمات (Estimation of Parameters):

وفي هذه المرحلة يتم معالجة المعلومات المتوفرة عن المجتمع والعينة رياضيا وإحصائيا لاستخراج قيم المعلمات والمتغير العشوائي والتي تتفق منطقيا مع الفروض الاقتصادية ومنها نحصل على الصياغة الرقمية للنموذج باستخدام طريقة المربعات الصغرى أو أية تقنية أخرى مناسبة للنموذج.

المرحلة الرابعة: تقييم المقدرات (Evaluation of Estimates)

هي مطابقة مجموع القيم المقدرة مع القيم الحقيقية وفي حالة عدم التطابق يعاد الحل لحين الوصول إلى هذه النتيجة التي تعتبر عمليا أول اختبار لصحة التقدير وهي اختبارات الدرجة الأولى $\left(\sum Y_i = \sum \hat{Y}_i\right)$ وكذلك معلمات الارتباط والانحدار والتحديد والخطأ المعياري للتقدير، بحيث يكون الأخير في الحد الأدنى قياسيا لأي نموذج أخر.

المرحلة الخامسة: اختبار جودة الاستدلال (Test of Goodness of Inference)

يتم اختبار جودة الاستدلال للمعلمات والمعادلة الكلية للانحدار عبر اختبارات المعنوية الكلية للمعادلة والمعلمات المقدرة باستخدام الاختبارات المناسبة.

المرحلة السادسة: تقييم القوة التنبوئية للنموذج

(Evaluation of the Forecasting Power of the Estimated Model)

وذلك باختبار قدرة النموذج على التنبؤ وإجراء التنبؤ الفعلي بعد قبول النظرية إذا تطابقت البيانات.

ثالثا: أنواع المتغيرات الإحصائية

المتغيرات: هي الخاصية المدروسة أو الظاهرة الإحصائية المدروسة في المجتمع الإحصائي.

تنقسم المتغيرات الإحصائية إلى قسمين[3] :

1- **متغيرات كيفية:** هي تلك المتغيرات التي لا يمكن قياسها أو غير قابلة للقياس مثل الجنسية، الحالة العائلية، الحالة المدنية....إلخ.

2- **متغيرات كمية:** هي تلك الخصائص التي يمكن قياسها وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام، مثال لذلك: الإنتاج، الاستهلاك، الاستثمار، الوزن...الخ. وتنقسم المتغيرات الكمية بدورها إلى قسمين: متغيرات منفصلة (متقطعة)، ومتغيرات مستمرة.

أ- **متغيرات منفصلة (متقطعة):** وهي تلك المتغيرات التي تأخذ قيما صحيحة، لا يمكن تجزئتها مثلا عدد الأطفال في العائلة، عدد قطع الغيار المنتجة....إلخ.

ب-**متغيرات مستمرة:** وهي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المتناهي لهذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات.

رابعا: بنية البيانات الاقتصادية

تأخذ البيانات الاقتصادية عدة أشكال هي:

1- البيانات المقطعية Cross-Sectional Data

تتكون مجموعة البيانات المقطعية من عينة الأفراد، أو القطاع العائلي، أو الشركات، أو الدول، أو المناطق، أو المدن، أو أي نوع من الوحدات في نقطة محددة من الزمن. وفي بعض الحالات، لا تتماثل الفترة الزمنية للبيانات بالضبط، مثل مسح بيانات العائلات المختلفة خلال أيام مختلفة من الشهر، وفي هذه الحالة يتم إهمال فروق التوقيت في جمع البيانات وتسمى البيانات التي يتم جمعها بيانات مقطعية[4].

إن استخدام البيانات المقطعية واسع النطاق في الاقتصاد والعلوم الاجتماعية الأخرى. ويساهم تحليل البيانات المقطعية بالاقتصاد الجزئي التطبيقي: اقتصاديات العمل، والمالية العامة للدولة، واقتصاديات الأعمال، والاقتصاد الطبي، وبعض الحقول في الاقتصاد الجزئي: بيانات الأفراد، والعائلات، والشركات، والمدن، والمناطق في نقطة من الزمن وتستخدم تلك الحالات لاختبار فرضيات الاقتصاد الجزئي، وتقييم السياسات الاقتصادية.

---

[3]- جلاطو جيلالي، **الإحصاء مع تمارين ومسائل محلولة**، الطبعة الخامسة، ديوان المطبوعات الجامعية، الجزائر، 2005، ص 6.

[4]- خالد محمد السواعي، **مدخل إلى القياس الاقتصادي**، الطبعة الأولى، الدار العربية للعلوم ناشرون، لبنان، 2015، ص 30.

2- بيانات السلاسل الزمنية Time Series Data

تتكون بيانات السلاسل الزمنية من مشاهدات متغير واحد أو أكثر خلال الزمن، منظمة بترتيب تسلسلي زمني مثل: سنوي، نصف سنوي، فصلي، شهري، أسبوعي، يومي، كل ساعة. ومن أمثلة السلاسل الزمنية أسعار الأسهم، والناتج المحلي الإجمالي، وعرض النقد.

وطبيعة البيانات الزمنية تجعل من تحليلها أكثر صعوبة من تحليل البيانات المقطعية، وتعتمد البيانات الاقتصادية على الزمن، وهذا يعني أن أغلب بيانات السلاسل الزمنية الاقتصادية ترتبط بتاريخها الماضي، وتطبق أغلب الإجراءات القياسية على البيانات المقطعية وعلى البيانات السلاسل الزمنية، إلا أننا نحتاج في حالة بيانات السلاسل الزمنية إلى إجراءات لتحديد النموذج القياسي المناسب، كما أن البيانات الاقتصادية الزمنية تتضمن اتجاها زمنيا يقودنا إلى أساليب قياسية جديدة.

ومن مظاهر بيانات السلاسل الزمنية إتباعها نفس النسق وعرضها نمطا موسميا قويا، وهذا يتعلق بالبيانات الأسبوعية والشهرية والفصلية. أخيرا ومن المهم القول أن بيانات السلاسل الزمنية تستخدم بشكل رئيسي في تطبيقات الاقتصاد الكلي.

3- بيانات السلاسل الزمنية المقطعية Panel Data

تتكون بيانات Panel من سلاسل زمنية لكل طرف مقطعي في مجموعة البيانات مثل المبيعات وعدد العاملين لخمسين شركة خلال فترة خمس سنوات. ويمكن جمع بيانات Panel على الأساس الجغرافي، مثل الناتج المحلي الإجمالي وعرض النقد لعشرين دولة لفترة 20 سنة.

الارتباط الخطي Linear Correlation

يمثل الارتباط (أو على وجه التحديد معامل الارتباط) الأسلوب (أو المقياس) الذي يمكن من دراسة نوع وقوة العلاقة بين المتغيرات من حيث الاتجاه والقوة، هذه العلاقة بين المتغيرات التي قد تكون خطية أو غير خطية.

أولا: مفهوم الارتباط

الارتباط يعني وجود علاقة بين عاملين أو أكثر، كل منهما يؤثر على الأخر أي أن التأثير متبادل بينهما، وقد يكون التأثير موجب إذا كانت الزيادة في أحدهما يتبعه زيادة في الأخر وتسمى هذه الحالة بالارتباط الموجب. كما قد يكون التأثير سالبا إذا كانت الزيادة في أحدهما يتبعها نقص في الأخرى وتسمى هذه الحالة بالارتباط السالب[5].

إن الارتباط قد يكون بسيطا إذا كانت العلاقة بين متغيرين أو عاملين فقط وقد يكون متعدد إذا كانت العلاقة بين أكثر من متغيرين أو عاملين، وعادة ما يرمز لمعامل الارتباط بالرمز (r) إذا كانت محسوبا من العينة ويرمز له بالرمز (R) إذا كان محسوبا من المجتمع.

تتراوح قيمة الارتباط ما بين -1 و $+1$ فإذا كانت القيمة هي $(+1)$ دل ذلك على وجود ارتباط موجب، وكلما اقتربت قيمة معامل الارتباط من الصفر دل ذلك على وجود ارتباط تتناقص قيمته تدريجيا إلى ان يتلاشى وجوده، أما إذا وصلت قيمة الارتباط عند (-1) أعتبر ذلك على أنه ارتباط سالب، لذلك فإنه من الأهمية بمكان ملاحظة إشارة معامل الارتباط مع قيمته.

ثانيا: أهمية دراسة الارتباط، أنواعه وطرق قياسه

1- أهمية دراسة الارتباط: تهدف دراسة الارتباط إلى وصف درجة التغير الاقتراني بين المتغيرات وتفيد في:

1- تحديد قوة الارتباط بين المتغيرين، أي بيان ما إذا كان الارتباط قوي أو ضعيف أو منعدم.

2- تحديد اتجاه العلاقة بين المتغيرين، أي بيان ما إذا كانت العلاقة طردية أو عكسية.

3- إن دراسة الارتباط تعد الأساس لدراسة وتحليل علاقات السببية.

4- تعطي مؤشرات لإمكان تقدير المتغيرات بدلالة أخرى.

---

[5] - إسماعيل بن قانة، **الإحصاء الوصفي الحيوي "دروس وتطبيقات"**، الطبعة الأولى، دار أسامة، الأردن، 2011، ص140.

2- **أنواع الارتباط:** يمكن تقسيم الارتباط حسب العديد من الخصائص التي تميزه، حيث يمكن إجمالها فيما يأتي:

أ/ من حيث قوته: ويمكن أن نميز:

1- ارتباط كامل (قيمته 1 أو -1) وهذا يعني أحد المتغيرات يتوقف كليا على تغير الأخر.

2- ارتباط جزئي وهذا يعني أنه يوجد ارتباط ولكن ليس بقوة الارتباط السابق.

ب/ من حيث عدد المتغيرات التي تؤخذ بعين الاعتبار: ويمكن أن نميز:

1- ارتباط بسيط أين ندرس العلاقة بين متغيرين فقط.

2- ارتباط متعدد أين ندرس العلاقة بين أكثر من متغيرين.

ج/ من حيث العلاقة الرياضية التي تربط بين المتغيرات: ونميز منها:

1- ارتباط خطي.

2- ارتباط غير خطي.

3- قياس الارتباط: يمكن قياس الارتباط بين المتغيرات على مرحلتين:

أ- **في المرحلة الأولى:** يتم التعرف عن طريق الوصف على مدى الترابط بين المتغيرين وذلك من خلال سحابة النقاط، إذ يمكن القول من خلال ملاحظة سحابة النقاط أن الارتباط قوي إيجابي أو سلبي ضعيف.

ب- في المرحلة الثانية: يتم تقدير قوة أو ضعف الارتباط من حيث قرب أو بعد النقاط عن الخط المستقيم فكلما اقتربت من الخط المستقيم كانت قوية وكلما ابتعدت عنه كان الارتباط ضعيفا وبشكل عام فإن قيمة الارتباط تتراوح بين 1، و -1 ولحساب هذه القيمة نلجأ إلى استخدام أحد المقاييس التالية:

1- معامل ارتباط بيرسون أو معامل ارتباط حاصل ضرب العزم:

يوضح معامل ارتباط بيرسون درجة العلاقة بين المتغيرات، التي تهدف إلى تحديد مدى جودة وصف معادلة خطية، فإذا كانت جميع قيم المتغيرات تحقق معادلة ما بالضبط فإنها ستحقق ارتباطا كاملا بين المتغيرات. وإذا كان شكل الانتشار أو سحابة النقاط تبدو كخط فإن الارتباط يسمى خطيا. ويعطى معامل ارتباط بيرسون بين متغيرين X و Y وفق الصيغة التالية:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

حيث وبعد إجراء بعض المعالجات الجبرية يمكن أن نستنتج:

$$r=\frac{n(\sum X_i Y_i)-(\sum X_i)(\sum Y_i)}{\sqrt{n(\sum X_i^2)-(X_i)^2}\sqrt{n(\sum Y_i^2)-(Y_i)^2}}\ or$$

$$r=\frac{\sum_{i=1}^{n} X_i Y_i-n\bar{X}\bar{Y}}{\sqrt{\sum X_i^2-n\bar{X}^2}\sqrt{\sum Y_i^2-n\bar{Y}^2}}$$

علما ان :

$\bar{X}$: تمثل الوسط الحسابي للبيانات X.

$\bar{Y}$: تمثل الوسط الحسابي للبيانات Y.

σ_x : تمثل الانحراف المعياري للبيانات X.

σ_y : تمثل الانحراف المعياري للبيانات Y.

ملاحظة:عند حساب معامل الارتباط لبيرسون يشترط أن يكون التوزيع لكلا المتغيرين طبيعي (أي يتبع التوزيع الاحتمالي الطبيعي أو لابلاس–قوس) وأن تكون العينة عشوائية وقيم الفرد لا تعتمد على قيم فرد أخر (استقلالية أفراد العينة). وفي حالة عدم خضوع المتغيرين للتوزيع الطبيعي نستخدم معامل ارتباط اخر هو معامل ارتباط سبيرمان أو كندال "تاو".

2– معامل الارتباط الرتبي Rank corrélation coefficient:

يستخدم معامل الارتباط الرتبي لدراسة الارتباط في حالة متغيرات كيفية، حيث نستعمل ترتيبا تصاعدية أو تنازلية عوضا عن القيم العديدية للمتغيرات المدروسة. فإذا كان الترتيب التصاعدي للمتغير المستقل يرافقه ترتيب تصاعدي للمتغير التابع نقول أن الارتباط موجب (علاقة طردية)، أما إذا كان الترتيب التصاعدي للمتغير المستقل يرافقه ترتيب تنازلي للمتغير التابع نقول أن الارتباط بين المتغيرين سالب (علاقة عكسية)

قياس الارتباط الرتبي:

لقياس الارتباط الرتبي بين مفردات المتغير التابع والمستقل نرتب كلاهما حسب أهمية خصائصهما ثم نحسب مجموع مربعات الفروق بين كل الرتب المتقابلة، لتصبح قياسات المتغيرين عبارة عن متوالية حسابية. ولاستخراج العلاقة الإحصائية لمعامل الارتباط الرتبي نعوض في علاقة الارتباط الخطي البسيط بقيم هاتين المتواليتين حسب التغيرات التالية[6]:

1– مجموع n عدد مرتبا ترتيبا تصاعديا هو: $\sum X_i=\sum Y_i=\frac{n(n+1)}{2}$

2– مجموع مربعات هذه الأعداد هو: $\sum X_i^2=\sum Y_i^2=\frac{n(n+1)(2n+1)}{2}$

[6]– جلاطو جيلالي، مرجع سبق ذكره، ص 152.

3- متوسط n عددا مرتبا ترتيبا تصاعديا هو: $\bar{X} = \bar{Y} = \frac{n+1}{2}$

4- أما تباين هذه القيم فهو:

$$V(X) = V(Y) = \frac{n(n+1)(2n+1)}{6n} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

5- مجموع مربعات الفروق بين رتب المتغيرين هو: $\sum d_i^2 = \sum (X_i - Y_i)^2$

حسب هذه المعطيات تصبح علاقة معامل الارتباط الرتبي بالشكل التالي:

$$r = \frac{\frac{(n+1)(2n+1)}{6} - \frac{\sum d_i^2}{2n} - \left(\frac{n+1}{2}\right)^2}{\frac{n^2-1}{12}}$$

$$r = \frac{\frac{\sum d_i^2}{2n}}{\frac{n^2-1}{12}} = 1 - \frac{\sum d_i^2 .12}{2n(n^2-1)}$$

إذن $\boldsymbol{r} = 1 - \frac{6.\sum d_i^2}{n(n^2-1)}$ ، تسمى هذه العلاقة بعلاقة: Spearman

3- معامل الارتباط المتعدد Multiple corrélation coefficient

يستخدم لقياس العلاقة بين أكثر من متغيرين، إلا أن إشارة معامل الارتباط لا تدل على اتجاه العلاقة هنا لأن هذا الاتجاه لا يكون موحدا لجميع المتغيرات. وصيغة حسابه في حالة 3 متغيرات لإيجاد العلاقة بين x_1 وكل من x_3 و x_2 هي[7]:

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{13}^2}}$$

حيث أن: r_{23}، r_{13}، r_{12} هي معاملات ارتباط يتم إيجادها بموجب صيغة الارتباط البسيط.

4- معامل الارتباط الجزئي Partial corrélation coefficient

يستخدم لقياس العلاقة بين زوج من المتغيرات عندما تكون باقي المتغيرات ثابتة. وبذلك فإن الفرق بين الارتباط البسيط والارتباط الجزئي هو أن الأول يقيس العلاقة بين متغيرين ضمن تأثير المتغيرات الأخرى، في حين يقيس الثاني العلاقة بين متغيرين مع استبعاد تأثير المتغيرات الأخرى.

وصيغة حساب معامل الارتباط الجزئي بين y و x_2 مع ثبات x_1 مثلا هي:

[7]- عبدالحميد عبدالمجيد البلداوي، **الأساليب التطبيقية لتحليل وإعداد البحوث العلمية مع حالات دراسية باستخدام برنامج SPSS** ، دار الشروق للنشر والتوزيع، عمان الأردن، 2008، ص177.

$$r_{y\,2.1} = \frac{r_{y2} - (r_{y1})(r_{12})}{\sqrt{(1 - r_{y1}^2)(1 - r_{12}^2)}}$$

حيث أن: r_{y2} و r_{y1} و r_{12} هي معاملات يتم إيجادها بموجب صيغة الارتباط البسيط.

ثالثا: اختبار المعنوية الإحصائية لمعامل الارتباط

بوضع ρ_{xy} القيمة معامل الإرتباط المحسوبة والذي هو تقدير للقيمة الحقيقة لمعامل الإرتباط r_{xy}

سوف نختبر الفرضية:
$$\begin{cases} H_0: r_{xy} = 0 \\ H_1: r_{xy} \neq 0 \end{cases}$$

نستطيع برهنة أن $\frac{\rho_{xy}}{\sqrt{\frac{(1-\rho_{xy}^2)}{n-2}}}$ تتبع قانون ستيودنت بدرجة حرية n−2

نحسب إحصائية t ستيودنت المحسوبة وفق العلاقة التالية[8]:

$$t^* = \frac{|\rho_{xy}|}{\sqrt{\frac{(1-\rho_{xy}^2)}{n-2}}}$$

إذا كانت[9] $t^* > t_{n-2}^{\alpha/2}$ نرفض فرضة العدم H_0 أي أن معامل الارتباط له معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha = 5\%$.

إذا كانت $t^* \leq t_{n-2}^{\alpha/2}$ نقبل فرضة العدم H_0 أي أن معامل الارتباط ليس له معنوية إحصائية عند مستوى معنوية $\alpha = 5\%$.

[8] - Régis Bourbonnais, **Econométrie**, 9e édition, Dunod, France, 2015, P 9.

[9] - $t_{n-2}^{\alpha/2}$: تسمى قيمة t ستيودنت المجدولة تستخرج من جدول t ستيودنت.

تمرين حول معامل الارتباط

يمثل الجدول التالي ساعات الدراسة لأحد الطلاب لخمسة امتحانات و العلامة التي حصل عليها.

عدد ساعات الدراسة (x)	6	5	7	8	3
العلامة (Y)	30	28	35	42	25

المطلوب : أحسب معامل الارتباط بين عدد ساعات الدراسة و العلامة المحصل عليها ؟ ماذا تستنتج ؟

- اختبر معنوية معامل الارتباط عند مستوى معنوية5 %؟

حل التمرين:

أولا: حساب معامل الارتباط بين عدد ساعات الدراسة (x) و العلامة المحصل عليها (y)

i	x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$
1	6	30	0.2	−2	−0.4	0.04	4
2	5	28	−0.8	−4	3.2	0.64	16
3	7	35	1.2	3	3.6	1.44	9
4	8	42	2.2	10	22	4.84	100
5	3	25	−2.8	−7	19.6	7.84	49
المجموع	29	160			48	14.8	178

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{N}=\frac{\sum_{i=1}^{5}x_i}{5}=\frac{29}{5}=5.8$$

$$\bar{x}=\frac{\sum_{i=1}^{n}y_i}{N}=\frac{\sum_{i=1}^{5}y_i}{5}=\frac{160}{5}=32$$

$$r_{x,y}=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}}=\frac{48}{\sqrt{14.8}\sqrt{178}}=0.9351$$

من $r_{x,y}=0.9351$ نستنتج أن هناك علاقة قوية طردية بين عدد ساعات الدراسة والعلامة المحصل عليها.

ثانيا: اختبار المعنوية الإحصائية لمعامل الارتباط

$$\begin{cases}H_0: r_{xy}=0\\ H_1: r_{xy}\neq 0\end{cases}$$

$$t^*=\frac{|\rho_{xy}|}{\sqrt{\frac{(1-\rho_{xy}^2)}{n-2}}}=\frac{|0.9351|}{\sqrt{\frac{(1-(0.9351)^2)}{5-2}}}=4.57 \qquad t_{n-2}^{\alpha/2}=t_3^{0.025}=3.182$$

نلاحظ أن $t^*=4.57$ أكبر من $t_3^{0.025}=3.182$ وبالتالي نرفض الفرضية H_0 إي أن معامل الإرتباط بين x و y له معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha=5\%$.

تحليل الانحدار الخطي البسيط

أولا: مفهوم الانحدار الخطي البسيط: يستخدم النموذج الخطي ذي المتغيرين، أو تحليل الانحدار الخطي البسيط، لاختبار الفروض حول العلاقة بين متغير تابع y، ومتغير مستقل أو مفسر x، والتنبؤ. ويبدأ الانحدار البسيط برسم مجموعة قيم XY في شكل انتشار تم التحديد بالنظر ما إذا كانت هناك علاقة خطية تقريبية [10].

$$Y_i = b_0 + b_1X_i \dots\dots\dots\dots\dots\dots\dots. (1)$$

وحيث أنه من غير المتوقع أن تقع النقاط تماما على الخط، فإن العلاقة الخطية التامة في المعادلة (1) يجب أن تعدل لكي تضم **حد تشويش عشوائي** أو **خطأ** أي "**عنصر عشوائي**"، ε_i.

$$Y_i = b_0 + b_1X_i + \varepsilon_i \dots\dots\dots\dots\dots\dots\dots. (2)$$

ثانيا: الفرضيات الخاصة بالمتغير العشوائي[11]:

1- أن المتغير العشوائي (ε_i) هو متغير تعتمد قيمته في أية فترة زمنية على عامل الصدفة، فقد تكون أكبر أو أصغر أو مساوية إلى الصفر، إلا أنها في المتوسط تساوي الصفر، أي $E(\varepsilon_i)=0$.

2- أن المتغير العشوائي يتوزع توزيعا طبيعيا، حول القيمة المتوقعة أو حول الوسط الحسابي المساوي للصفر عند كل قيمة من قيم المتغير المستقل x أي بشكل جرس.

3- أن تباين (Vriance)المتغير العشوائي (حد الخطأ)، حول الوسط الحسابي مقدار ثابت عند كل قيمة من قيم X أي: $var(\varepsilon_i) = E(\varepsilon_i)^2 = \sigma^2$

الفرضيات الثلاثة السابقة يمكن جمعها بشكل مختصر وتمثيلها كالآتي:

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

بمعنى أن الخطأ العشوائي،ε_i، يتوزع، ~، توزيعا طبيعيا،N ، بوسط حسابي مساوي للصفر، 0، وتباين ثابت قيمته σ^2 .

4- أن قيم ε_i غير مرتبطة بأي من المتغيرات المستقلة ، أي انعدام التباين المشترك Covariance بين ε_i و X_i أي : $Cov(\varepsilon_i X_i) = 0$.

5- القيم المختلفة للمتغير العشوائي (ε_i) تكون مستقلة عن بعضها البعض بعبارة أخرى قيم التباين المشترك ل ε_i مع ε_j مساوية للصفر، وعليه فإن قيمة المتغير العشوائي في أي فترة لا تعتمد على قيمته في فترة أخرى أي: $Cov(\varepsilon_i\varepsilon_j) = E(\varepsilon_i\varepsilon_j) = 0 \quad (i, j = 1,2,3, \dots, n,\ i \neq j)$

---

[10]- دومينيك سالفاتور، ترجمة سعدية حافظ منتصر، **الإحصاء والاقتصاد القياسي**، سلسلة ملخصات شوم، الطبعة السادسة ، الدار الدولية للإستثمارات الثقافية، 2012، ص138.

[11]– حسين علي بخيت، سحر فتح الله، **الاقتصاد القياسي**، الطبعة الأولى، دار اليازوري، الأردن، 2009، ص38– ص41.

6- انعدام العلاقة بين المتغيرات المستقلة.

ثالثا: تقدير نموذج الانحدار الخطي البسط بطريقة المربعات الصغرى العادية

طريقة المربعات الصغرى العادية (OLS) هي أسلوب لتوفيق " أفضل" خط مستقيم لعينة مشاهدات XY، وهو يتضمن تصغير مجموع المربعات لانحرافات النقاط (الرأسية) عن الخط إلى أدنى حد ممكن:

$$Min \sum (Y_i - \hat{Y}_i)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

حيث تشير Y_i إلى المشاهدات الفعلية، وتشير $\hat{Y}_i$ إلى القيم "الموافقة" المناظرة، بحيث تكون $Y_i - \hat{Y}_i = \varepsilon_i$ هي البواقي.

ويعطي هذا الأسلوب المعادلتين الطبيعتين التاليتين[12]:

$$\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_i$$

$$\sum X_i Y_i = \hat{b}_0 \sum X_i + \hat{b}_i \sum X_i^2$$

حيث n عدد المشاهدات، $\hat{b}_0$، $\hat{b}_1$ هي مقدرتان للمعلمتين الحقيقيتين b_0،b_1 .

وبحل المعادلتين أنيا، نحصل على:

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

ونحصل على قيمة $\hat{b}_0$ كما يلي :

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

ومن المفيد عادة استخدام صيغة مكافئة لتقدير $\hat{b}_1$: $\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2}$

حيث $x_i = X_i - \bar{X}$، $y_i = Y_i - \bar{Y}$ وتكون معادلة الإنحدار المقدرة بطريقة المربعات الصغرى (OLS) :

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

رابعا: خواص مقدرات طريقة المربعات الصغرى العادية

مقدرات المربعات الصغرى العادية (OLS) هي أفضل مقدرات خطية غير متحيزة (BLUE).

وعدم التحيز يعني $E(\hat{b}) = b$

بحيث أن: $Bias = E(\hat{b}) - b$

[12] – دومينيك سالفاتور، ترجمة سعدية حافظ منتصر، مرجع سبق ذكره، ص139.

أما وصف مقدر بأنه " **أفضل مقدر غير متحيز** " أو انه **مقدر كفؤ** فيعني أنه ذو أصغر تباين. وبالتالي فإن مقدرات OLS هي الأفضل من بين كل المقدرات الخطية غير المتحيزة. وتعرف هذه الخاصية بنظرية " جاوس ماركوف "، وهي تمثل أهم مبرر لاستخدام OLS.

أحيانا، قد يرغب الباحث ان يقبل بعض التحيز في مقابل تباين أصغر ، بتصغير متوسط مربع الخطأ، MSE

$$MSE(\hat{b}) = E(\hat{b}-b)^2 = var(\hat{b}) + (bias\hat{b})^2$$

ويكون المقدر **متسقا** إذا اقتربت قيمته من المعلمة الحقيقية مع اقتراب حجم العينة من ما لانهاية (بمعنى أنه غير متحيز في اللانهاية) وينتهي توزيعه إلى المعلمة الحقيقية.

خامسا: توزيع المعاينة للمقدرات والتقدير المجالي للمعالم

1- حساب تباينات المقدرات[13]:

لبناء مجال الثقة للمعالم، يتعين معرفة تباين كل من $\hat{b}_0$ ، $\hat{b}_1$ والبواقي.

تباين : $\hat{b}_0$

$$var(\hat{b}_0) = \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right]\sigma_\varepsilon^2 = \left[\frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i-\bar{X})^2}\right]\sigma_\varepsilon^2$$

تباين : $\hat{b}_1$

$$var(\hat{b}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

العلاقة بين $var(\hat{b}_1)$ و $var(\hat{b}_0)$ تعطى بالعلاقة التالية :

$$var(\hat{b}_0) = \frac{\sigma_\varepsilon^2}{n}\bar{X}^2 var(\hat{b}_1)$$

وبناءا على هذا التعريف تكون الانحرافات المعيارية (Standard déviation) هي الجذور التربيعية لتباينات المقدرات، أما الأخطاء المعيارية (Standard errors) فهي الجذور التربيعية لمقدرات الانحرافات المعيارية أي:

$$\hat{\sigma}_{\hat{b}_0} = \sqrt{var(\hat{b}_0)} = \sigma_\varepsilon\sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i-\bar{X})^2}}$$

$$\hat{\sigma}_{\hat{b}_1} = \sqrt{var(\hat{b}_1)} = \frac{\sigma_\varepsilon}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}}$$

[13]- شيخي محمد، **طرق الاقتصاد القياسي**، الطبعة الأولى، دار الحامد، الأردن، 2011،ص28-ص32.

2- تباين البواقي :

$\hat{\sigma}_{\varepsilon}^2 = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-2}$ وهو تقدير غير متحيز

القيمة التي في البسط تعبر عن مجموع مربعات البواقي حيث $\sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$ أما $n-2$ فهي درجة الحرية، تعبر عن حجم العينة ناقص 2 وذلك لوجود معلمتين للتقدير في النموذج.

3- بناء مجال الثقة للمعالم:

بمعرفة توزيع $\hat{b}_0$ و $\hat{b}_1$ يمكن تكوين مجالات ثقة وإجراء إختبار الفرضيات الموضوعة حول معالم الإنحدار b_0 و b_1 على التوالي، نعطي مجالا للقيم التي يمكن أن تحتوي عليها معالم الإنحدار الحقيقية، مع كل مجال ثقة نضع مستوى إحصائيا للمعنوية، حيث أن إحتمال إحتواء المجال المذكور على معلمة الإنحدار الحقيقية يكون واحد مطروحا منه مستوى المعنوية، أي $(1-\alpha)$، ولتكوين مجال الثقة من التوزيع t بالنسبة للمعلمتين b_0 و b_1 نكتب القانون الخاص لكل معلمة:

في حالة $n \leq 30$ و σ^2 غير معروف[14]:

$$\frac{\hat{b}_0 - b_0}{\hat{\sigma}_{\hat{b}_0}} \sim t_{(n-2)}$$

$$\frac{\hat{b}_1 - b_1}{\hat{\sigma}_{\hat{b}_1}} \sim t_{(n-2)}$$

عند مستوى معنوية $(\%)$ يكون مجال الثقة لكلا المعلمتين[15]:

$$Pr\left[-t_{n-2,\frac{\alpha}{2}} \leq \frac{\hat{b}_0 - b_0}{\hat{\sigma}_{\hat{b}_0}} \leq +t_{n-2,\frac{\alpha}{2}}\right] = 1-\alpha$$

$$Pr\left[-t_{n-2,\frac{\alpha}{2}} \leq \frac{\hat{b}_1 - b_1}{\hat{\sigma}_{\hat{b}_1}} \leq +t_{n-2,\frac{\alpha}{2}}\right] = 1-\alpha$$

إذا ضربنا (داخل الاحتمال) كل الأطراف بواسطة $\hat{\sigma}_{\hat{b}_0}$ $(\hat{\sigma}_{\hat{b}_1})$ وأضفنا b_0 (b_1) لأطراف المتراجحة نجد:

$$b_0 \in \left[\hat{b}_0 - \hat{\sigma}_{\hat{b}_0} t_{n-2,\frac{\alpha}{2}}, \hat{b}_0 + \hat{\sigma}_{\hat{b}_0} t_{n-2,\frac{\alpha}{2}}\right]$$

$$b_1 \in \left[\hat{b}_1 - \hat{\sigma}_{\hat{b}_1} t_{n-2,\frac{\alpha}{2}}, \hat{b}_1 + \hat{\sigma}_{\hat{b}_1} t_{n-2,\frac{\alpha}{2}}\right]$$

في حالة $n > 30$ و σ^2 معروف:

14- – المرجع السابق.

15- $t_{n-2,\frac{\alpha}{2}}$: القيمة الحرجة لتوزيع Student بدرجة حرية n−2 ونسبة معنوية $(\%\alpha)$، تستخرج من جدول التوزيع ستيودنت.

$$\hat{b}_1 \sim N\left[b_1, \frac{\hat{\sigma}_\varepsilon^2}{\sum x_i^2}\right]$$

$$\hat{b}_0 \sim N\left[b_0, \hat{\sigma}_\varepsilon^2 \frac{\sum X_i^2}{n\sum x_i^2}\right]$$

$$\frac{\hat{b}_0 - b_0}{\hat{\sigma}_{\hat{b}_0}} \sim N(0,1)$$

$$\frac{\hat{b}_1 - b_1}{\hat{\sigma}_{\hat{b}_1}} \sim N(0,1)$$

عند مستوى معنوية $(\%)$ يكون مجال الثقة لكلا المعلمتين[16]:

$$Pr\left[-z_{\frac{\alpha}{2}} \leq \frac{\hat{b}_0 - b_0}{\hat{\sigma}_{\hat{b}_0}} \leq +z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

$$Pr\left[-z_{\frac{\alpha}{2}} \leq \frac{\hat{b}_1 - b_1}{\hat{\sigma}_{\hat{b}_1}} \leq +z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

إذا ضربنا (داخل الاحتمال) كل الأطراف بواسطة $\hat{\sigma}_{\hat{b}_0}$ $(\hat{\sigma}_{\hat{b}_1})$ وأضفنا b_0 (b_1) لأطراف المتراجحة نجد:

$$b_0 \in \left[\hat{b}_0 - z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{b}_0}, \hat{b}_0 + z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{b}_0}\right]$$

$$b_1 \in \left[\hat{b}_1 - z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{b}_1}, \hat{b}_1 + z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{b}_1}\right]$$

نبني أيضا مجال الثقة ل σ^2 لدينا[17]:

$$\chi_\alpha^2(n-2) \sim \frac{(n-2)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2}$$

يكون مجال الثقة :

$$Pr\left[\chi^2_{\alpha/2} \leq \frac{(n-2)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \leq \chi^2_{1-\alpha/2}\right] = 1-\alpha$$

$$\rightarrow Pr\left[\frac{(n-2)\hat{\sigma}_\varepsilon^2}{\chi^2_{1-\alpha/2}} \leq \sigma_\varepsilon^2 \leq \frac{(n-2)\hat{\sigma}_\varepsilon^2}{\chi^2_{\alpha/2}}\right] = 1-\alpha$$

يكون مجال الثقة لتباين الأخطاء:

16– $Z_{\frac{\alpha}{2}}$: القيمة الحرجة للتوزيع الطبيعي بنسبة $(\%\alpha)$، تستخرج من جدول التوزيع الطبيعي.

17– $\chi^2_\alpha(n-2)$: القيمة الحرجة لتوزيع χ^2 بدرجة حرية n–2 تستخرج من جدول التوزيع كي دو.

$$\sigma_\varepsilon^2 \in \left[\frac{(n-2)\hat{\sigma}_\varepsilon^2}{\chi^2_{1-\alpha/2}}, \frac{(n-2)\hat{\sigma}_\varepsilon^2}{\chi^2_{\alpha/2}}\right]$$

سادسا: اختبار جودة التوفيق والارتباط

كلما كانت المشاهدات أقرب إلى خط الانحدار (أي، كلما صغرت البواقي)، كلما زاد التغير في Y الذي "تفسره" معادلة الانحدار المقدرة. والتغير الاجمالي في Y يساوي التغير المفسر زائد تغير البواقي[18]:

$$\sum(Y_i-\bar{Y})^2 = \sum\left(\hat{Y}_i-\bar{Y}\right)^2 + \sum\left(Y_i-\hat{Y}\right)^2$$

$$TSS = ESS + RSS$$

حيث أن:

$\sum(Y_i-\bar{Y})^2$: التغير الإجمالي في Y. (مجموع مربعات الانحرافات الكلية (TSS))

: $\sum\left(\hat{Y}_i-\bar{Y}\right)^2$: التغير المفسر في Y. (مجموع مربعات الانحرافات المشرحة (ESS))

$\sum\left(Y_i-\hat{Y}\right)^2$: تغير البواقي في Y.(مجموع مربعات البواقي (RSS))

وبقسمة الطرفين على TSS نحصل على :

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

ومن هنا يمكن تعريف معامل التحديد R^2 بأنه النسبة في التغير الإجمالي في Y الذي " يفسره" انحدار Y_i على X_i ، وأن معامل التحديد يقيس ويشرح نسبة الانحرافات الكلية أو التغيرات التي تحدث في المتغير التابع Y_i، والمشروحة بواسطة تغيرات المتغير المستقل X_i فهي نسبة تأثير المتغير المستقل على المتغير التابع، فهو مقياس للقدرة التفسيرية للنموذج أي يختبر جودة التوفيق والارتباط[19].

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

ويمكن حساب R^2 كالأتي:

$$R^2 = \frac{\sum\left(\hat{Y}_i-\bar{Y}\right)^2}{\sum(Y_i-\bar{Y})^2} = 1 - \frac{\sum \varepsilon_i^2}{\sum(Y_i-\bar{Y})^2}$$

حيث[20] : $R^2 = r^2$

18- دومينيك سالفاتور، ترجمة سعدية حافظ منتصر، مرجع سبق ذكره، ص142.

19- شيخي محمد، مرجع سبق ذكره، ص39.

20- في الانحدار الخطي البسيط: معامل التحديد هو مربع معامل الارتباط الخطي البسيط.

وتتراوح قيمة R^2 بين 0 (عندما لا تفسر معادلة الانحدار أيا من التغير في y) و 1 (عندما تقع كل النقاط على خط الانحدار)[21]:

- فإذا كان $R^2 = 1$: فهذا يعني $\sigma_{\hat{y}}^2 = \sigma_y^2$ ، أي تباين القيم الفعلية عن الوسط الحسابي هو نفسه تباين القيم التقديرية عن نفس الوسط الحسابي. هذا يعني أن القيم التقديرية هي نفسها القيم الفعلية. في هذه الحالة النموذج المختار يعطي تطابقا تاما بين قيم (Y_i) الفعلية وقيمها التقديرية $(\hat{Y}_i)$، وهذا يعني أن معادلة التمثيل المختارة بشكل صحيح تماما، وأنه لا يؤثر على المتغير التابع (y) أي متغير أخر مستقل غير المتغير (x).

- أما إذا كان $R^2 = 0$: فهذا يعني $\sigma_{\hat{y}}^2 = 0$ في هذه الحالة تكون جميع المقادير $\sum(\hat{Y}_i - \bar{Y})^2 = 0$ وهذا يعني أن $\bar{Y} = \hat{Y}$. أي أن جميع القيم التقديرية ل Y متساوية وتساوي عددا ثابتا $(\bar{Y})$ وهو الوسط الحسابي. إن هذا يؤكد أن النموذج المقترح لتمثيل العلاقة (x, y) هو معادلة مستقيم موازي لمحور (X) وله الشكل التالي: $\hat{y}_i = \bar{y} = b_0$. وبمأن هذه المعادلة لا تتضمن قيم (x)، فنستنتج أنه عندما يتغير (x) فإن قيم (y) التقديرية لا تتأثر بذلك التغير مما يؤكد أن (y) غير مرتبط ب (x) نهائيا. بعبارة أخرى كل التغيرات التي تحدث في(y) هي بفعل عوامل أخرى غير (x).

- بصفة عامة إذا كان معامل التحديد ضعيفا فإن ذلك يكون ناتجا على أن النموذج الرياضي المختار لا يمثل الظاهرة المدروسة تمثيلا جيدا، قد يكون ذلك بسبب ان العوامل المستقلة (x_i) المختارة في النموذج لها تأثير ضعيف على (y_i) وأن عوامل أخرى لها تأثير كبير على (y_i) لم تؤخذ بعين الاعتبار في النموذج. وقد يكون معامل التحديد ضعيف بسبب أخطاء في القياس أو جمع المعطيات وإعدادها.

سابعا: اختبار الفرضيات

يمكن إجراء اختبار الفرضيات الموضوعة حول معالم النموذج b_0 و b_1، وذلك بمعرفة توزيع $\hat{b}_0$ و $\hat{b}_1$

1-اختبار المعنوية الإحصائية لمعالم النموذج :

إذا كان: $n \leq 30$: فنبغي إستعمال توزيع ستيودنت

1-1- اختبار المعنوية الإحصائية لb_0 :

$$\begin{cases} H_0: b_0 = 0 \\ H_1: b_0 \neq 0 \end{cases}$$

21- مكيد علي، **الاقتصاد القياسي دروس وتمارين محلولة**، الطبعة الثانية، ديوان المطبوعات الجامعية، الجزائر، 2011، ص49-50.

$\frac{\hat{b}_0 - b_0}{\hat{\sigma}_{\hat{b}_0}}$ تتبع توزيع ستيودنت بدرجة حرية n−2

تحت الفرضية H_0 $(b_0 = 0)$ تصبح نسبة ستيودنت $\frac{\hat{b}_0 - 0}{\hat{\sigma}_{\hat{b}_0}}$ تتبع توزيع ستيودنت بدرجة حرية n−2.

إذا اختبار فرضية العدم يرتكز على مقارنة القيمة المحسوبة لستيودنت $t^* = \frac{|\hat{b}_0|}{\hat{\sigma}_{\hat{b}_0}}$ ، مع القيمة المجدولة لتوزيع ستيودنت بدرجة حرية $n-2$ عند مستوى معنوية α، $t_{n-2}^{\alpha/2}$.

- إذا كان $t^* > t_{n-2}^{\alpha/2}$ نرفض فرضة العدم H_0 أي أن المعلمة b_0 لها معنوية إحصائية فهي تختلف معنويا عن الصفر عند مستوى معنوية $\alpha = 5\%$.

- إذا كان $t^* \leq t_{n-2}^{\alpha/2}$ نقبل فرضة العدم H_0 أي أي أن المعلمة b_0 ليس لها معنوية إحصائية أي يساوي معنويا الصفر عند مستوى معنوية $\alpha = 5\%$.

1-2- اختبار المعنوية الإحصائية لـb_1 :

$$\begin{cases} H_0: b_1 = 0 \\ H_1: b_1 \neq 0 \end{cases}$$

$\frac{\hat{b}_1 - b_1}{\hat{\sigma}_{\hat{b}_1}}$ تتبع توزيع ستيودنت بدرجة حرية n−2

تحت الفرضية H_0 $(b_1 = 0)$ تصبح نسبة ستيودنت $\frac{\hat{b}_1 - 0}{\hat{\sigma}_{\hat{b}_1}}$ تتبع توزيع ستيودنت بدرجة حرية n−2.

إذا اختبار فرضية العدم يرتكز على مقارنة القيمة المحسوبة لستيودنت $t^* = \frac{|\hat{b}_1|}{\hat{\sigma}_{\hat{b}_1}}$ ، مع القيمة المجدولة لتوزيع ستيودنت بدرجة حرية $n-2$ عند مستوى معنوية α، $t_{n-2}^{\alpha/2}$.

- إذا كان $t^* > t_{n-2}^{\alpha/2}$ نرفض فرضة العدم H_0 أي أن المعلمة b_1 لها معنوية إحصائية فهي تختلف معنويا عن الصفر عند مستوى معنوية $\alpha = 5\%$.

- إذا كان $t^* \leq t_{n-2}^{\alpha/2}$ نقبل فرضة العدم H_0 أي أي أن المعلمة b_1 ليس لها معنوية إحصائية أي يساوي معنويا الصفر عند مستوى معنوية $\alpha = 5\%$.

إذا كان $n - 2 > 30$ ، فنأخذ $t_{\infty}^{0.05/2} = 1.96$ عند مستوى معنوية $\alpha = 5\%$ [22].

- إذا كان $t^* > t_{\infty}^{0.05/2} = 1.96$ نرفض فرضة العدم H_0 أي أن المعلمة b_0 (b_1) لها معنوية إحصائية فهي تختلف معنويا عن الصفر عند مستوى معنوية $\alpha = 5\%$.

[22] - Régis Bourbonnais, Op. Cit., P 28.

ثامنا: جدول تحليل التباين واختبار المعنوية الكلية للنموذج

يهدف جدول تحليل التباين، ANOVA Table، إلى توضيح تأثير المتغير المستقل في المتغير التابع، وتزداد أهمية هذا الجدول عند دراسة الانحدار المتعدد حيث يستفاد منه في معرفة تأثير كل متغير من المتغيرات المستقلة في المتغير التابع وبالتالي اعتماد المتغيرات المؤثرة في النموذج.

ويمكن بناء جدول تحليل التباين كالتالي:

جدول تحليل التباين للانحدار الخطي البسيط

مصدر التغير	مجموع المربعات	درجة الحرية	المربعات المتوسطة
المتغير المستقل	ESS	1	ESS/1
البواقي	RSS	n-2	RSS/n-2
المجموع	TSS	n-1	

يمكننا الحصول على إحصائية فيشر F من حاصل قسمة ESS/1 على RSS/n-2 أي أن:

$$F^* = \frac{ESS/_1}{RSS/_{n-2}} = \frac{\sum(\hat{Y}_i - \bar{Y})^2/_1}{\sum \varepsilon_i^2/_{n-2}}$$

أيضا إذا كان التباين المفسر له معنوية أكبر من تباين البواقي يمكن اعتبار أن المتغير x_t كمتغير مفسر حقيقي

حيث نستخدم إحصائية فيشر F لاختبار الفرضية:

$$\begin{cases} H_0:\ b_1 = 0 \\ H_0:\ b_1 \neq 0 \end{cases}$$

حيث أن إحصائية فيشر تتبع توزيع فيشر بدرجتي حرية 1 ، و n−2 عند مستوى معنوية α أي أن[23]:

$$F^* \sim F(1, n-2, \alpha)$$

إذا كان $F^* > F(1, n-2, \alpha)$ نرفض فرضية العدم H_0 أي النموذج له معنوية إحصائية عند مستوى معنوية α .

إذا كان $F^* \leq F(1, n-2, \alpha)$ نقبل فرضية العدم H_0 أي النموذج ليس له معنوية إحصائية عند مستوى معنوية α .

[23] – $F(1, n-2, \alpha)$: تستخرج من جدول توزيع فيشر عند مستوى معنوية α.

العلاقة بين توزيع فيشر F ومعامل التحديد R^2 :

من العلاقة: $R^2 = r^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$ نجد أن:

$$ESS = R^2.TSS = R^2.\sum(Y_i - \bar{Y})^2 \dots\dots\dots\dots\dots (1)$$

$$RSS = (1 - R^2).TSS = (1 - R^2)\sum(Y_i - \bar{Y})^2 \dots\dots (2)$$

وبتعويض العلاقتين (1) و(2) في العلاقة $F^* = \frac{ESS/_1}{RSS/_{n-2}}$ نجد أن :

$$F^* = \frac{R^2/_1}{(1 - R^2)/_{n-2}} = \frac{R^2}{(1 - R^2)}(n - 2)$$

تاسعا: التنبؤ Prediction

إن أحد الأدوار الرئيسية للقياس الاقتصادي هو التنبؤ بتأثر أحد المتغيرات من طرف المتغيرات الأخرى. فمثلا، نفرض أننا نريد اختبار أثر تخفيض الضريبة على مستوى الإنفاق الاستهلاكي للعائلات، أو أثر زيادة الإنفاق الحكومي على ذلك، فإذا عرفنا بأي مقدار يمكن للضريبة البديلة أن تزيد من الدخل المتاح، نستطيع استعمال دالة الاستهلاك المقدرة للتنبؤ بأثار الضريبة المخفضة على الاستهلاك.

لنأخذ نموذجنا البسيط، ولنفرض أننا نعرف قيمة x في دورة التنبؤ Forecasting period ، ونرمز لها بالرمز x_f . فإذا فرضنا أن البناء الهيكلي للمعادلة لا يتغير في المستقبل، تكون قيمة المتغير التابع y في هذه الفترة f كما يلي[24]:

$$y_f = b_0 + b_1 x_f + \varepsilon_f$$

عندما نستعمل علاقة ما للتنبؤ بالقيمة y_f ، هناك مصدران لعدم الوضوح والدقة في تنبؤاتنا وهما:

1- لا نعرف المعلمتين b_0، b_1، وبالتالي يجب الإعتماد على مقدرتي العينة $\hat{b}_0$، $\hat{b}_1$ لكي تقدر القيمة y_f، إن هذه القمة هي وسط y الموافق ل y_f ، أي:

$$E(y_i) = b_0 + b_1 x_i$$
$$E(y_f/x_f) = b_0 + b_1 x_f$$

[24]- تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الجزء الأول، الطبعة الثانية ، ديوان المطبوعات الجامعية، الجزائر، 2011، ص 65-69.

2- بالإضافة إلى أن الخطأ ε_f هو متغير رعشوائي غير مشاهد، ولهذا ، حتى وإن عرفنا قيمتي b_0، b_1، وبالتالي استطعنا حساب $E(y_f/x_f)$ ، نبقى غير قادرين على التنبؤ بقيمة y_f تماما بسبب الخطأ ε_f.

إذن نأخذ، أولا، مقدرا للمقدر $E(y_f/x_f)$، ونستعين به في التنبؤ بقيمة y_f نفسها، ثم نضع مجالا للتنبؤ ب y_f. ومادام:

$$E(y_f/x_f) = b_0 + b_1 x_f$$

فيكون المقدر الطبيعي لها على الشكل:

$$\hat{y}_f = \hat{b}_0 + \hat{b}_1 x_f$$

ويمكن أن نبين بأن هذا المقدر هو مقدر غير متحيز ل$E(y_f/x_f)$. وأنه عبر المقدرات الخطية غير المتحيزة الأخرى يعتبر هذا الأخير أحسنها (أي أصغر تباين). ويعرف بإسم أفضل تنبؤ خطي غير متحيز أي Best Linear Unbiased Predictor (BLUP). وتأتي هذه الخاصية من كون $\hat{b}_0$، $\hat{b}_1$ لهما خاصية أفضل مقدر خطي غير متحيز BLUE. وإذا فرضنا أن مستقلة. يكون تباين $\hat{y}$ على الشكل:

$$var(\hat{y}_f) = var(\hat{b}_0 + \hat{b}_1 x_f) = var(\hat{b}_0) + x_f^2 var(\hat{b}_1) + 2x_f cov(\hat{b}_0, \hat{b}_1)$$

وبإستعمال المعادلات التالية[25]:

$$var(\hat{b}_0) = \sigma_\varepsilon^2 \sum_{i=1}^{n} \left[\frac{1}{n} - \bar{x} w_i\right]^2 = \sigma_\varepsilon^2 . \frac{\sum x_i^2 \sum w_i^2}{n}$$

$$var(\hat{b}_1) = E\left[\sum w_i \varepsilon_i\right]^2 = \sigma_\varepsilon^2 \sum w_i^2 = \frac{\sigma_\varepsilon^2}{\sum x_i^2}$$

$$cov(\hat{b}_0, \hat{b}_1) = -\bar{x}\sigma_\varepsilon^2 \sum w_i^2 = \frac{-\bar{x}\sigma_\varepsilon^2}{\sum x_i^2}$$

حيث : $w_i = \frac{x_i}{\sum x_i^2}$

نجد أن:

$$var(\hat{y}_f) = \sigma_\varepsilon^2 \sum w_i^2 \left[\frac{\sum x_i^2}{n} + (x_f - \bar{x})^2\right]$$

[25] – المرجع السابق، ص 41.

كما يمكن كتابها على الشكل:

$$var(\hat{y}_f) = \sigma_\varepsilon^2 \left[\frac{1}{n} + (x_f - \bar{x})^2 . \sum w_i^2\right]$$

ونلاحظ أن تباين مقدر التنبؤ ينخفض كلما:

1- انخفضت القيمة $(x_f - \bar{x})^2$ أي كلما إقتربت x_f من وسط العينة $\bar{x}$.

2- ازدادت n أي حجم العينة يزداد.

3- ازدادت القيمة $\sum x_i^2$، أي كلما انخفضت القيمة $\sum w_i^2$.

لنأخذ الأن هذا المقدار ل $E(y_f/x_f)$ كقيمة متنبأ بها Predicted Value بواسطة تقدير وسطها. إن مقدار الخطأ الداخل في هذا التنبؤ معطى بالعبارة:

$$\hat{\varepsilon}_f = y_f - \hat{y}_f$$

ونسميه مقدر خطأ التنبؤ prediction error أو Forecast error ثم نلاحظ أن:

$$E(\hat{\varepsilon}_f) = E(y_f - \hat{y}_f) = 0$$

ويصبح تباين خطأ التنبؤ هو:

$$var(\hat{\varepsilon}_f) = var(y_f - \hat{y}_f) = var(y_f) + var(\hat{y}_f)$$

إن قيمة y_f تعتمد مباشرة على ε_f، بينما تعتمد $\hat{y}_f$ على أخطاء العينة $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ بواسطة المقدرتين $\hat{b}_0$، $\hat{b}_1$ وبالتالي يكون $cov(y_f, \hat{y}_f) = 0$[26]

ونلاحظ أنه إذا كانت x مستقلة فإن $var(y_f) = var(\hat{\varepsilon}_f)$ كما وجدنا من قبل. ولدينا

$$var(\hat{y}_f) = \sigma_\varepsilon^2 \left[\frac{1}{n} + (x_f - \bar{x})^2 . \sum w_i^2\right]$$

لنجد:

$$var(\hat{\varepsilon}_f) = \sigma_\varepsilon^2 \left[1 + \frac{1}{n} + (x_f - \bar{x})^2 . \sum w_i^2\right]$$

ولنعرف $var(\hat{\varepsilon}\ \ ) = \sigma_{\varepsilon f}^2$ لنجد:

$$\sigma_{\varepsilon f}^2 = \sigma_\varepsilon^2 \left[1 + \frac{1}{n} + (x_f - \bar{x})^2 . \sum w_i^2\right]$$

ومنه فإن المقدر غير المتحيز لتباين خطأ التنبؤ $v\hat{a}r(\hat{\varepsilon}_f)$ هو

[26] -$cov(y_f, \hat{y}_f) = E\left[\left(y_f - E(y_f)\right)\left(\hat{y}_f - E(\hat{y}_f)\right)\right]$
$cov(y_f, \hat{y}_f) = E\left[\varepsilon_f \left((\hat{b}_0 - b_0) - (\hat{b}_1 - b_1)\right) x_f\right]$
$cov(y_f, \hat{y}_f) = E\left[\varepsilon_f \left(\sum \left(\frac{1}{n} - \bar{x}x_i\right)\varepsilon_i\right)\right] - E\left[x_f\varepsilon_f - \sum w_i\varepsilon_i\right] = 0$

$$\hat{\sigma}_{\varepsilon f}^2 = \hat{\sigma}_\varepsilon^2\left[1+\frac{1}{n}+\left(x_f-\bar{x}\right)^2.\sum w_i^2\right]$$

ونلاحظ أنه كلما كبر حجم العينة فإن $\hat{\sigma}_\varepsilon^2$ تقترب من $\hat{\sigma}_\varepsilon^2$ أي:

$$\lim_{n\to\infty}\left(\hat{\sigma}_{\varepsilon f}^2\right) = \hat{\sigma}_\varepsilon^2$$

ولهذا فعندما يكون حجم العينة كبيرا يمكن استعمال $\hat{\sigma}_\varepsilon$ كتقريب ل $\hat{\sigma}_{\varepsilon f}$.

نأمل الأن في إيجاد مقياس لتحديد دقة هذا التنبؤ ل y_f. وللقيام بذلك نفرض توزيعا إحتماليا معينا للإضطرابات العشوائية، وهو التوزيع الطبيعي، ثم مادام ε_f موزع طبيعيا وكذلك y_f، كما أن أخطاء العينة $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ موزعة طبيعيا وكذلك $\hat{b}_0$، $\hat{b}_1$ فإن $\hat{y}_f$ تكون موزعة طبيعيا أيضا. ولهذا فإن خطأ التنبؤ، $\hat{\varepsilon}_f = y_f - \hat{y}_f$ يكون متغيرا عشوائيا موزعا توزيع طبيعيا بوسط مساوي للصفر وتباين هو $\sigma_{\varepsilon f}^2$ أما المقدر $\hat{\sigma}_{\varepsilon f}^2$.

ومنه فإن:

$$Z = \frac{\hat{\varepsilon}_f}{\sigma_{\varepsilon f}} \sim N(0,1)$$

كما أن $\sigma_{\varepsilon f}^2$ تعتمد على القيمة غير المعروفة σ_ε^2 كما في المعادلة

$$\sigma_{\varepsilon f}^2 = \sigma_\varepsilon^2\left[1+\frac{1}{n}+\left(x_f-\bar{x}\right)^2.\sum w_i^2\right]$$

فعمليا نعوض بمقدرها $\hat{\sigma}_{\varepsilon f}^2$ لتعطي المتغير العشوائي للتوزيع t كما يلي:

$$\frac{\hat{\varepsilon}_f}{\hat{\sigma}_{\varepsilon f}} = \frac{y_f - \hat{y}_f}{\hat{\sigma}_{\varepsilon f}} \sim t_{n-2}$$

إذا كانت $t_{n-2,\frac{\alpha}{2}}$ هي القيمة الحرجة (من جدول) لتوزيع t بحيث تحقق:

$$pr\left[t_{n-2,\frac{\alpha}{2}} \le \frac{y_f-\hat{y}_f}{\hat{\sigma}_{\varepsilon f}} \le t_{n-2,\frac{\alpha}{2}}\right] = 1-\alpha$$

فإن مجال الثقة للتنبؤ يكون:

$$\hat{y}_f - \hat{\sigma}_{\varepsilon f}t_{n-2,\frac{\alpha}{2}} \le y_f \le \hat{y}_f + \hat{\sigma}_{\varepsilon f}t_{n-2,\frac{\alpha}{2}}$$

$$y_f \in \left[\hat{y}_f - \hat{\sigma}_{\varepsilon f}t_{n-2,\frac{\alpha}{2}}, \hat{y}_f + \hat{\sigma}_{\varepsilon f}t_{n-2,\frac{\alpha}{2}}\right]$$

$$\text{C.I}\left(y_f\right) = \hat{y}_f - \hat{\sigma}_{\varepsilon f}t_{n-2,\frac{\alpha}{2}}$$

تمرين حول الانحدار الخطي البسيط

الجدول يوضح معدل الدخل السنوي x ومقدار التوفير السنوي (بالدينار) y لعينة تتكون من 10 أسر:

18	25	11	10	36	16	22	9	15	12	Xi
1.1	1.7	0.6	0.4	3.6	1.2	2.4	0.2	1.1	0.6	Yi

$$\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y}) = 75.54 \text{ ، } \sum_{i=1}^{10} y_i = 12.9 \text{ ، } \sum_{i=1}^{10} x_i = 174 \text{ ، } \sum_{i=1}^{10} x_i y_i = 300$$

$$\sum_{i=1}^{10}(y_i-\bar{y})^2 = 9.749 \text{ ، } \sum_{i=1}^{10} x_i^2 = 3656 \text{ ، } \sum_{i=1}^{10} y_i^2 = 26.39 \text{ ، } \sum_{i=1}^{10}(x_i-\bar{x})^2 = 628.4$$

المطلوب:

1- كتابة النموذج الخطي ؟

2- تقدير معادلة الانحدار الخطية باستخدام طريقة المربعات الصغرى مستخدما معدل الدخل السنوي كمتغير مستقل ؟

3- حساب تباين البواقي ؟

4- حساب تباين المقدرات؟

5- إيجاد مجال الثقة لمعالم النموذج ؟

6- حساب معامل التحديد ؟

7- إيجاد جدول تحليل التباين ؟

8- اختبار المعنوية الإحصائية لمعالم النموذج عند مستوى معنوية α=5 % ؟

9- اختبار المعنوية الكلية للنموذج عند مستوى معنوية α=5 % ؟

الحل:

1- كتابة النموذج الخطي

$$Y_i = b_0 + b_1 X_i + \varepsilon_i \qquad i = 1, \cdots, 10$$

2- تقدير معادلة الانحدار الخطية باستخدام طريقة المربعات الصغرى مستخدما معدل الدخل السنوي كمتغير مستقل

$$\hat{Y}_i = \hat{b}_0 - \hat{b}_1 X_i$$

$$\hat{b}_1 = \frac{\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{10}(x_i-\bar{x})^2} = \frac{75.54}{628.4} = 0.12$$

$$\bar{X} = \frac{\sum_1^{10} x_i}{N} = \frac{174}{10} = 17.4 \qquad \bar{Y} = \frac{\sum_1^{10} y_i}{N} = \frac{12.9}{10} = 1.29$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X} = 1.29 - (0.12) * 17.4 = -0.798$$

إذن المعادلة المقدرة هي:

$$\hat{Y}_i = -0.798 + 0.12\, X_i \qquad i = 1, \cdots, 10$$

3- حساب تباين البواقي

$\hat{\varepsilon}_i^2 = (\hat{Y}_i - Y_i)^2$	$\hat{\varepsilon}_i = \hat{Y}_i - Y_i$	$\hat{Y}_i$	Y_i	X_i
0.001764	-0.042	0.642	0.6	12
0.009604	0.098	1.002	1.1	15
0.006724	-0.082	0.282	0.2	9
0.311364	0.558	1.842	2.4	22
0.006084	0.078	1.122	1.2	16
0.006084	0.078	3.522	3.6	36
4E-06	-0.002	0.402	0.4	10
0.006084	0.078	0.522	0.6	11
0.252004	-0.502	2.202	1.7	25
0.068644	-0.262	1.362	1.1	18
0.66836				المجموع

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-2} = \frac{\sum_{i=1}^{10} \hat{\varepsilon}_i^2}{10-2} = \frac{0.668}{8} = 0.08$$

4- حساب تباين والإنحراف المعياري للمقدرات:

تباين : $\hat{b}_0$

$$var(\hat{b}_0) = \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]\sigma_\varepsilon^2 = \left[\frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]\sigma_\varepsilon^2$$

$$var(\hat{b}_0) = \left[\frac{3656}{10 * 628.4}\right]0.08 = 0.046$$

الانحراف المعياري ل $\hat{b}_0$

$$\boldsymbol{\sigma}_{\hat{b}_0} = \sqrt{var(\hat{b}_0)} = \sqrt{0.046} = \mathbf{0.214}$$

تباين : $\hat{b}_1$

$$var(\hat{b}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{0.08}{628.4} = 0.000127$$

الانحراف المعياري ل $\hat{b}_1$

$$\sigma_{\hat{b}_1} = \sqrt{var(\hat{b}_1)} = \sqrt{0.000127} = 0.011$$

5- إيجاد مجال الثقة لمعالم النموذج

مجال الثقة ل b_0

$$b_0 \in \left[\hat{b}_0 - \hat{\sigma}_{\hat{b}_0} t_{n-2,\frac{\alpha}{2}}\,, \hat{b}_0 + \hat{\sigma}_{\hat{b}_0} t_{n-2,\frac{\alpha}{2}}\right]$$

$$t_{n-2}^{\alpha/2} = t_8^{0.025} = 2.306$$

إذن:

$$b_0 \in [-0.798 - (0.214 * 2.306)\,, -0.798 + (0.214 * 2.306)]$$

$$b_0 \in [-1.291\,, -0.304]$$

مجال الثقة ل b_1

$$b_1 \in \left[\hat{b}_1 - \hat{\sigma}_{\hat{b}_1} t_{n-2,\frac{\alpha}{2}}\,, \hat{b}_1 + \hat{\sigma}_{\hat{b}_1} t_{n-2,\frac{\alpha}{2}}\right]$$

$$b_1 \in [0.12 - (0.011 * 2.306)\,, 0.12 + (0.011 * 2.306)]$$

$$b_1 \in [0.094\,, 0.145]$$

6- تقدير معامل التحديد :

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{10}\hat{\varepsilon}_i^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

$$R^2 = 1 - \frac{0.668}{9.749} = 0.9314$$

الاستنتاج: نلاحظ أن:$R^2 = 0.9314$ نستنتج أن القدرة التفسيرية للنموذج عالية وأن جودة عالية للتوفيق والإرتباط، وأن المتغيرات المفسرة x_{1i}، x_{2i} تفسر المتغير التابع بنسبة 93.14%.

7- إيجاد جدول تحليل التباين

مصدر التغير	مجموع المربعات	درجة الحرية	متوسط المربعات
x_i	$ESS = 9.081$	1	$ESS/_1 = 9.081$
البواقي	$RSS = 0.668$	$n-2=8$	$RSS/_{n-2} = 0.08$
المجموع	$TSS = 9.749$	$n-1=9$	

$$TSS = \text{ESS} + \text{RSS}$$
$$ESS = \text{TSS} - \text{RSS}$$

$$ESS = 9.749 - 0.668 = 40$$

7-اختبار المعنوية الإحصائية لمعالم النموذج

اختبار المعنوية الاحصائية للمعلمة b_0 :

$$\begin{cases} H_0:\ b_0 = 0 \\ H_1:\ b_0 \neq 0 \end{cases}$$

$$t^* = \frac{|\hat{b}_0|}{\hat{\sigma}_{\hat{b}_0}} = \frac{0.798}{0.214} = 3.728 \qquad t_{n-2}^{\alpha/2} = t_8^{0.025} = 2.306$$

نلاحظ أن $t^* = 3.728$ أكبر من $t_8^{0.025} = 2.306$ وبالتالي نرفض الفرضية H_0 وأي أن المعلمة b_0 لها معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha = 5\%$.

اختبار المعنوية الاحصائية للمعلمة a_1 :

$$\begin{cases} H_0:\ b_1 = 0 \\ H_1:\ b_1 \neq 0 \end{cases}$$

$$t^* = \frac{|\hat{b}_1|}{\hat{\sigma}_{\hat{b}_1}} = \frac{0.12}{0.011} = 10.909$$

نلاحظ أن $t^* = 10.909$ أكبر من $t_8^{0.025} = 2.306$ وبالتالي نرفض الفرضية H_0 وأي أن المعلمة b_1 لها معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha = 5\%$.

8- اختبار المعنوية الكلية للنموذج:

$$\begin{cases} H_0:\ a_1 = 0 \\ H_1:\ a_1 \neq 0 \end{cases}$$

$$F^* = \frac{ESS/_1}{RSS/_{n-2}} = \frac{9.081/_1}{0.668/_8} = 113.51$$

$$F(1,n-2,\alpha) = F(1,8,0.05) = 5.32$$

نلاحظ أن $F^* = 113.51$ أكبر من $F(1,8,0.05) = 5.32$ وبالتالي نرفض الفرضية H_0 أي أن النموذج له معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha = 5\%$.

تحليل الانحدار الخطي المتعدد

أولا: مفهوم الانحدار الخطي المتعدد: يوضح الانحدار الخطي المتعدد العلاقة بين متغير تابع Y_i و مجموعة من المتغيرات التفسيرية Xij ، هذا ما يعني أن أي تغير في المتغيرات التفسيرية يتبعها تغير في المتغير التابع.

وتشير خطية العلاقة بين المتغيرات التفسيرية و المتغير التابع إلى أن أثر المتغير التفسيري على المتغير التابع لا يختلف عن أثر متغير أخر، فيفترض أن جميع الأفراد يتصرفون بنفس الطريقة، أو أن تفضيلات الأفراد متماثلة، و نظرا لأن هذا الافتراض لا يمثل الحقيقة فإن استخدام الانحدار الخطي المتعدد ينطوي على وجود نوع من الخطأ في التقدير، و لذا فإننا ندخل في علاقة الانحدار حدا يعرف بالحد العشوائي ε_t [27].

ثانيا: تعيين نموذج الانحدار الخطي المتعدد

تأخذ علاقة الانحدار الخطي المتعدد الشكل التالي:

$$Y_t = B_1 + B_2X_{2t} + B_3X_{3t} + \cdots + B_kX_{kt} + \varepsilon_t$$

و يمكن كتابته على الشكل المصفوفات التالي:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{21} & x_{31} & \cdots & x_{k1} \\ 1 & x_{22} & x_{32} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \\ 1 & x_{2n} & x_{3n} & \cdots & x_{kn} \end{pmatrix} . \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ \vdots \\ B_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_t \end{pmatrix}$$

$$Y_t \quad = \qquad X_{ij} \qquad B_{ij} \; + \; \varepsilon_t$$

حيث: Y_{ij} : شعاع مشاهدات المتغير التابع (n×1) .

X_{ij} : مصفوفة مشاهدات المتغيرات التفسيرية (n×(k+1)) .

B_{ij} : شعاع المعاملات (k×1) .

ε_t : شعاع الحد العشوائي (n×1) .

[27]– فروخي جمال: **نظرية الاقتصاد القياسي**، ديوان المطبوعات الجامعية، الجزائر 1993، ص 51

ثالثا: فرضيات النموذج

إن الطريقة المستعملة في تقدير معاملات نموذج الانحدار الخطي المتعدد هي طريقة المربعات الصغرى العادية" MCO "[28]، و لهذا فإن هذه الفرضيات تتعلق بهذه الطريقة و كلها تدور حول طبيعة و شكل المتغير العشوائي و هي:[29]

1- الفرضيات الاحتمالية:

الفرضية 1: قيم المتغيرات X_{it} محددة بدون أخطاء.

الفرضية 2: التوقع الرياضي للخطاء معدوم، $E(\varepsilon_t) = 0$.

الفرضية 3: $\forall t$ تباين الخطأ ثابت، $E(\varepsilon_t^2) = \sigma_t^2$

الفرضية 4: استقلالية الأخطاء إذا كان $t \neq t'$ فإن $E(\varepsilon_t, \varepsilon_{t'}) = 0$.

الفرضية 5: حد الخطأ مستقل عن المتغيرات المفسرة، $E(x_{it}, \varepsilon_t) = 0$.

2- فرضيات هيكلية:

الفرضية 6: عدم وجود مشكل التعدد الخطي بين المتغيرات التفسيرية، هذا يدل على أن المصفوفة $(X'X)$ نظامية و أن المصفوفة $(X'X)^{-1}$ موجودة.

الفرضية 7: $(X'X)/n$

الفرضية 8: $n > k + 1$ عدد الملاحظات أكبر من عدد سلاسل المتغيرات التفسيرية.

رابعا: تقدير نموذج الانحدار الخطي المتعدد

يتم تقدير نموذج الانحدار الخطي المتعدد بطريقة المربعات الصغرى العادية، التي تهدف إلى إيجاد الشعاع β

الذي يصغر مجموع مربعات الأخطاء $\hat{\varepsilon}$ (الفرق بين القيمة المقدرة $\hat{Y}$ والقيمة الحقيقية Y)[30]:

$$\Rightarrow \hat{\varepsilon} = Y - \hat{Y} = \begin{pmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \vdots \\ \hat{\varepsilon}_n \end{pmatrix}$$

[28] - **M**oindres des **C**arrées **O**rdinnaires.
[29] - Régis Bourbonnais**,** Op. Cit., P 51.

[30]- شيخي محمد، مرجع سبق ذكره، ص60.

$$Min\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = Min\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$$

$$\hat{\varepsilon}_i = y_i - \hat{y}_i \qquad i = 1 \dots\dots n$$

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

نسمي :

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

حيث: $\hat{Y} = X\hat{\beta}$. ومنه الهدف هو $\min\Gamma(Y,X,\hat{\beta})$

إذا كان $\hat{B}$ موجود فيجب أن يحقق الشرط الضروري:

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

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

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

نضرب طرفي المعادلة ب $(X'X)^{-1}$ لنحصل على :

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

وهو تقدير β.

خامسا: تقدير تباين الأخطاء σ^2 ومصفوفة التباين- التباين المشترك للمقدارات $\Omega_{\hat{\beta}}$

1- تقدير تباين الأخطاء[31]:

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

2- مصفوفة التباين والتباين المشترك:

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

[31] - Régis Bourbonnais, Op. Cit., P 53

ملاحظة: القطر الرئيسي لمصفوفة التباين والتباين المشترك يعطينا تباين المقدرات $\hat{\sigma}^2_{\hat{b}_0}, \hat{\sigma}^2_{\hat{b}_1}, \cdots, \hat{\sigma}^2_{\hat{b}_k}$ وعند إدخال الجدر التربيعي نجد الانحراف المعياري للمقدرات $\hat{\sigma}_{\hat{b}_0}, \hat{\sigma}_{\hat{b}_1}, \cdots, \hat{\sigma}_{\hat{b}_k}$

سادسا: قياس جودة التقدير : Goodness of fit (R^2)

كما في الانحدار الخطي البسيط، يقيس معامل التحديد R^2 نسبة فروقات المتغير التابع Y المفسرة بفروقات المتغير التفسيري، ونقدم في الانحدار المتعدد نفس المقياس ونفس الصيغة، لكننا سنتكلم عن نسبة فروقات المتغير التابع المفسرة بجميع المتغيرات التفسيرية الداخلة بالنموذج الخطي، وبذلك يكون معامل التحديد Coefficient of determination كما يلي:

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

أو

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{n}\varepsilon_i^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

حيث أن ESS هي فروقات Y " المفسرة " في النموذج (مجموع مربعات الانحدار)، و TSS هي مجموع فروقات Y حول وسطها (مجموع المربعات الكلي)، و RSS هي مجموع المربعات الصغرى (الأخطاء) وهي نسبة فروقات Y غير المفسرة في النموذج.

تشير $\hat{Y}$ إلى القيمة المتنبأ بها predicted value للمتغير التابع Y لكل المتغيرات التفسيرية، حيث أن:

$$\hat{Y}_i = b_1 + b_2X_{2i} + b_3X_{3i} + \cdots + b_kX_{ki}$$

الوسط الحسابي للعينة $\bar{Y}$ هو وسط كل من Y_i و $\hat{Y}_i$ يزودنا بنموذج يتضمن المقطع b_1 في هذه الحالة.

يعرض معامل التحديد كذلك مقياسا لقدرة النموذج على التنبؤ خلال فترة العينة، أو قياس جودة الانحدار المقدر، وتساوي قيمة R^2 مربع معامل إرتباط العينة بين Y_i و $\hat{Y}_i$، حيث أن الارتباط يقيس العلاقة الخطية المناسبة بين المتغيرين. فإذا كان R^2 مرتفعا، فهذا يعني أن الترابط بين القيمة Y_i والقيمة المتنبأ بها $\hat{Y}_i$ يكون قريبا، وفي هذه الحالة يكون نموذج تقدير البيانات جيدا، فإذا كان R^2 منخفضا لا يكون الترابط قريبا بين قيم Y_i وقيم $\hat{Y}_i$ المتنبأ بها في النموذج، ولا يكون النموذج مناسبا[32].

سابعا: معامل التحديد المصحح $\bar{R}^2$

إحدى صعوبات استخدام R^2 كمقياس لجودة التقدير هي زيادة حجمه عند إضافة متغيرات جديدة حتى لو كانت المتغيرات المضافة ليس لها أي قيمة إقتصادية. وجبريا فإن إضافة المتغيرات تخفض من مجموع

[32]– خالد محمد السواعي، مرجع سبق ذكره، ص177.

مربعات الأخطاء RSS وبالتالي يرتفع R^2، كما يمكن أن تصبح درجة الحرية ضعيفة[33] ، وكمقياس بديل لجودة التقدير هو R^2 المصحح ويرمز له ب$\overline{R}^2$ حيث تم تصحيح R^2 بإدراج درجة الحرية في حالة استعمال عدد مشاهدات ضعيفة مقارنة بعدد المتغيرات المستقلة ويحسب كما يلي:

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

حيث $R^2 > \bar{R}^2$ وإذاكان n كبير يكون $R^2 \simeq \bar{R}^2$

ثامنا: اختبار الفرضيات

سوف نوضح كيفية تنفيذ اختبار المعنوية الإحصائية في إطار علاقة الانحدار المتعدد، وذلك باستعمال اختبار ستيودنت واختبار فيشر حيث[34]:

1- اختبار فيشر F يستعمل لتحديد ما إذا كانت هناك علاقة معنوية بين المتغير التابع ومجموع المتغيرات المستقلة والتي تسمى اختبار المعنوية الكلية.

2- اختبار ستيودنت t يستعمل لتحديد ما إذا كان كل واحد من المتغيرات المستقلة له معنوية إحصائية، اختبار ستيودنت يطبق على كل متغير مستقل في النموذج، ويسمى اختبار المعنوية الفردية.

1-اختبار المعنوية الإحصائية لمعالم النموذج :

إذا كان: $n \leq 30$: فنبغي إستعمال توزيع ستيودنت

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

نكتب $t_c = \frac{\hat{b} \ -b_j}{\hat{\sigma}_{\hat{b}_j}}$ وهي القيمة المحسوبة

ما دمنا نختبر فرضية العدم، نكتب $t_c = \frac{\hat{b}_j}{\hat{\sigma}_{\hat{b}_j}}$، حيث نقبل H_0 بمستوى معنوية α إذا كانت

$\left|\frac{\hat{b}_j}{\hat{\sigma}_{\hat{b}_j}}\right| \leq t_{n-2}^{\alpha/2}$ ففي هذه الحالة، المعلمة b_j ليس لها معنوية إحصائية أي يساوي معنويا الصفر، حيث

[33]- في حالة نموذج يحتوي على عدد المشاهدات يساوي عدد المتغيرات المستقلة (درجة الحرية تساوي الصفر) أي $R^2 = 1$ إذا ليس هناك قدرة تفسيرية للنموذج. (Régis Bourbonnais, Op.Cit., P55 .)

[34] - Anderson,Sweeney, Williams,Camm, Cochran, traduction de la 7eédition américaine par Claire Borsenberger, **Statistiques pour l'economie et la gestion**, 5e édition, De Boeck supérieur, Paris,2015, p776

$t_{n-2}^{\alpha/2}$ مأخوذة من جدول التوزيع t، ونرفض H_0 بمستوى معنوية α إذا كانت $\left|\frac{\hat{b}_j}{\hat{\sigma}_{\hat{b}_j}}\right| > t_{n-2}^{\alpha/2}$ ففي هذه الحالة، المعلمة b_j لها معنوية إحصائية أي تختلف معنويا عن الصفر[35].

عندما يكون حجم العينة كبيرا $(n > 30)$ فينبغي إستعمال التوزيع الطبيعي ويمكن أخذ القيمة الحرجة $z_{\alpha/2}$ وذلك بحساب المضللة للتوزيع.

2- اختبار المعنوية الكلية

فرضيات اختبار فيشر تركز على معالم نموذج الانحدار المتعدد

$$\begin{cases} H_0: b_1 = b_2 = \cdots = b_K = 0 \\ H_1: \text{يوجد على الأقل واحد من المعالم لا يساوي الصفر} \end{cases}$$

$$F^* = \frac{ESS/k}{RSS/n-k-1} \sim F(k, n-k-1, \alpha)$$

إذا كان $F(k, n-k-1, \alpha) \leq F^*$ نرفض الفرضية H_0 أي أنه يوجد واحد من معالم النموذج لا يساوي الصفر، وأن العلاقة بين المتغير التابع y والمتغيرات المستقلة $x_1, x_2, \cdots, x_k$ معنوية.

إذا كان $F(k, n-k-1, \alpha) > F^*$ نقبل الفرضية H_0 أي أن معالم النموذج كلها تساوي الصفر، وأن العلاقة بين المتغير التابع y والمتغيرات المستقلة $x_1, x_2, \cdots, x_k$ غير معنوية.

3- جدول تحليل التباين:

جدول ANOVA لنموذج الانحدار المتعدد ل k متغير مستقل

مصدر التغير	مجموع المربعات	درجة الحرية	متوسط المربعات	F
الانحدار	ESS	k	ESS/k	$F = \frac{ESS/k}{RSS/n-k-1}$
البواقي	RSS	$n-k-1$	$RSS/n-k-1$	
المجموع	TSS	$n-1$		

الجدول أعلاه المتعلق بجدول تحليل التباين (ANOVA) الذي يقدم نتائج اختبار فيشر لنموذج الانحدار الخطي المتعدد، القيمة الإحصائية المحسوبة لاختبار فيشر تظهر في العمود الأخير نستطيع مقارنتها

35- شيخي محمد، مرجع سبق ذكره، ص 72.

بقيمة فيشر Fالمجدولة مع k درجة حرية و n−k−1، لكي نستطيع اتخاذ القرار فيما يخص اختبار الفرضيات[36].

ثامنا: التنبؤ العلمي باستعمال الانحدار الخطي المتعدد

نتطرق في هذه الفقرة إلى قضية التنبؤ بالملاحظات المستقبلية، (أو خارج العينة) لشعاع الملاحظات الخاصة بالمتغير التابع $Y_t : t = 1,2,\cdots,T$ وذلك بمعرفتنا لمصفوفة الملاحظات المستقبلية للمتغيرات المستقلة، فليكن النموذج الخطي العام خلال العينة T والمقدر على الشكل التالي: $\hat{Y} = X\hat{\beta}$ ومقدر المربعات الصغرى العادية $\hat{\beta} = (X'X)^{-1}X'Y$ ، ويكون المقدر بملاحظة واحدة في المستقبل هو:

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

التنبؤ بعد فترتين في المستقبل:

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

التنبؤ بعد الفترة h في المستقبل:

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

حيث $h = 1,2,\cdots,H$ يسمى بأفق التنبؤ.

وعليه نصل إلى التنبؤ بالفترة H في المستقبل:

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

إذن إذا أردنا التنبؤ بمجموعة من الملاحظات المستقبلية بفترة تساوي H ملاحظة مرة واحدة يكون شعاع القيم التقديرية[37]:

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

أما مصفوفة ملاحظة المتغيرات المستقلة المستقبلية فهي:

[36] - Anderson,Sweeney, Williams,Camm, Cochran, traduction de la 7eédition américaine par Claire Borsenberger, Op. Cit., p779

[37] – شيخي محمد، مرجع سبق ذكره، ص 82-85.

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

ومنه يمكن كتابة النموذج الخطي العام المتنبأ به على الشكل:

$$Y_{T+H} = X_{T+H}\beta + \varepsilon_{T+H}$$

كما يكون متوسط مقدر التنبؤ هو:

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

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

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

ليكون التباين:

$$var\left(\hat{Y}_T(H)\right) = \left[\left(\hat{Y}_T(H) - X_{T+H}\beta\right)\left(\hat{Y}_T(H) - X_{T+H}\beta\right)'\right]$$

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

لنعرف شعاع أخطاء التنبؤ:

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

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

أما تباين شعاع أخطاء التنبؤ فهو:

$$var(\hat{\varepsilon}_{T+H}) = var\left(Y_{T+H} + \hat{Y}_T(H)\right)$$

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

لنجد في الأخير:

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

ويكون هذا التنبؤ هو أحسن تنبؤ خطي غير متحيز (BLUE) يمكن الحصول عليه، حيث إذا عرفنا $\tilde{Y}_T(H)$ تنبؤ اخر خطي لعينة ملاحظات المتغير التابع مع متوسط خطأ التنبؤ مساوي للصفر $E(\tilde{\varepsilon}_{T+H}) = E\left(Y_{T+H} + \tilde{Y}_T(H)\right) = 0$ تكون المترجحة :

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

ومنه يمكن القول أن $\hat{Y}\ \ (H) = X_{T+H}\hat{\beta}$ 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،

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

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

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

وإذا كان H=1 يصيح التوزيع أعلاه على الشكل:

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

تمرين حول الانحدار الخطي المتعدد

ليكن لدينا النموذج التالي:

علما أن:

$$(X'X)^{-1} = \begin{pmatrix} 6 & -2 & -5 \\ -2 & 4 & 1 \\ -5 & 1 & 5 \end{pmatrix} \text{،} \; X'Y = \begin{pmatrix} 10 \\ 3 \\ 4 \end{pmatrix} \text{،} \; \sum_{i=1}^{50}(y_i - \bar{y})^2 = 260 \text{،} \; \sum_{i=1}^{50}(\hat{y}_i - \bar{y})^2 = 220$$

المطلوب:

1- تقدير معالم النموذج القياسي؟

2- حساب قيمة معامل التحديد؟ ماذا تستنتج ؟

3- أنشئ جدول تحليل التباين؟

1- اختبار المعنوية الإحصائية لمعالم النموذج عند مستوى معنوية 5%؟

2- اختبار المعنوية الكلية للنموذج عند مستوى معنوية 5%؟

الحل:

1- تقدير معالم النموذج القياسي:

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_{1i} + \hat{b}_2 x_{2i} \qquad i = 1,\dots\dots\dots,50$$

$$\hat{B} = (\acute{X}X)^{-1}\acute{X}Y$$

$$\hat{B} = \begin{pmatrix} 6 & -2 & -5 \\ -2 & 4 & 1 \\ -5 & 1 & 5 \end{pmatrix}\begin{pmatrix} 10 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 34 \\ -4 \\ -27 \end{pmatrix}$$

$$\hat{B} = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \end{pmatrix} = \begin{pmatrix} 34 \\ -4 \\ -27 \end{pmatrix}$$

$$\hat{y}_i = 34 - 4x_{1i} - 27x_{2i} \qquad i = 1, \dots\dots\dots, 50$$

2-تقدير معامل التحديد :

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

$$R^2 = \frac{220}{260} = 0.8461$$

الإستنتاج:

نلاحظ أن:$R^2 = 0.8461$ نستنتج أن القدرة التفسيرية للنموذج عالية وأن جودة عالية للتوفيق والإرتباط، وأن المتغيرات المفسرة x_{1i} ،x_{2i} تفسر المتغير التابع بنسبة 84.61%.

3-إيجاد جدول تحليل التباين

مصدر التغير	مجموع المربعات	درجة الحرية	متوسط المربعات
$x_{1i}x_{2i}$ ،	$ESS = 220$	$K = 2$	$ESS/_k = 110$
البواقي	$RSS = 40$	$n - K - 1 = 47$	$RSS/_{n-K-1} = 0.85$
المجموع	$TSS = 260$	$n - 1 = 49$	

$$TSS = ESS + RSS$$
$$RSS = TSS - ESS$$

$$RSS = 260 - 220 = 40$$

4-اختبار المعنوية الإحصائية لمعالم النموذج

أولا: حساب تباين مقدرات النموذج

إيجاد تباين البواقي

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

$$\hat{\sigma}_{\varepsilon}^{2}=\frac{40}{47}=0.85$$

إيجاد مصفوفة التباين والتباين المشترك:

$$\widehat{\Omega}_{\hat{a}}=\hat{\sigma}_{\varepsilon}^{2}\left(\acute{X}X\right)^{-1}$$

$$\widehat{\Omega}_{\hat{a}}=0.85\begin{pmatrix}6&-2&-5\\-2&4&1\\-5&1&5\end{pmatrix}$$

$$\widehat{\Omega}_{\hat{a}}=\begin{pmatrix}5.1&-1.7&-4.25\\-1.7&3.4&0.85\\-4.25&0.85&4.25\end{pmatrix}$$

إذن:

$$\hat{\sigma}_{\hat{b}_0}^{2}=5.1 \qquad ،\hat{\sigma}_{\hat{b}_1}^{2}=3.4 \qquad ،\hat{\sigma}_{\hat{b}_2}^{2}=4.25$$

اختبار المعنوية الاحصائية للمعلمة b_0 :

$$\begin{cases}H_0:\ b_0=0\\H_1:\ b_0\neq 0\end{cases}$$

لدينا

$$\hat{\sigma}_{\hat{b}_0}^{2}=5.1 \quad\Rightarrow\quad \hat{\sigma}_{\hat{b}_0}=\sqrt{\hat{\sigma}_{\hat{b}_0}^{2}}=\sqrt{5.1}=2.25$$

$$t^{*}=\frac{|\hat{b}_0|}{\hat{\sigma}_{\hat{b}_0}}=\frac{34}{2.25}=15.11$$

نلاحظ أن $t^{*}=15.11$ أكبر من 1.96 وبالتالي نرفض الفرضية H_0 أي أن المعلمة b_0 لها معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha=5\%$.

اختبار المعنوية الاحصائية للمعلمة b_1 :

$$\begin{cases}H_0:\ b_1=0\\H_1:\ b_1\neq 0\end{cases}$$

لدينا

$$\hat{\sigma}_{\hat{b}_1}^{2}=3.4 \quad\Rightarrow\quad \hat{\sigma}_{\hat{b}_1}=\sqrt{\hat{\sigma}_{\hat{b}_1}^{2}}=\sqrt{3.4}=1.84$$

$$t^{*}=\frac{|\hat{a}_1|}{\hat{\sigma}_{\hat{b}_1}}=\frac{4}{1.84}=2.17$$

نلاحظ أن $t^{*}=2.17$ أكبر من 1.96 وبالتالي نرفض الفرضية H_0 أي أن المعلمة b_1 لها معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha=5\%$.

اختبار المعنوية الاحصائية للمعلمة b_2 :

$$\begin{cases} H_0:\ b_2 = 0 \\ H_1:\ b_2 \neq 0 \end{cases}$$

لدينا

$$\hat{\sigma}^2_{\hat{b}_2} = 4.25 \quad \Rightarrow \hat{\sigma}_{\hat{b}_2} = \sqrt{\hat{\sigma}^2_{\hat{b}_2}} = \sqrt{4.25} = 2.06$$

$$t^* = \frac{|\hat{b}_2|}{\hat{\sigma}_{\hat{b}_2}} = \frac{27}{2.06} = 13.10$$

نلاحظ أن $t^* = 13.10$ أكبر من 1.96 وبالتالي نرفض الفرضية H_0 أي أن المعلمة b_2 لها معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha = 5\%$.

اختبار المعنوية الكلية للنموذج:

$$\begin{cases} H_0:\ b_1 = b_2 = 0 \\ H_1: \exists\ b_i \neq 0 \quad i = 1,2 \end{cases}$$

$$F^* = \frac{ESS/_k}{RSS/_{n-K-1}} = \frac{110}{0.85} = 129.14$$

$$F(k, n-k-1, \alpha) = F(2,47,0.05) = 3.23$$

نلاحظ أن $F^* = 129.14$ أكبر من 3.23 وبالتالي نرفض الفرضية H_0 أي أن النموذج له معنوية إحصائية خلافا للصفر عند مستوى معنوية $\alpha = 5\%$.

مشاكل القياس الاقتصادي

تصنف مشاكل الانحدار تحت خمسة عناوين رئيسية هي[38]:

- **عدم التجانس (hetroskedasticity or unequal variance):** أي ان تكون قيمة تباين الخطأ العشوائي غير ثابتة.
- **الإرتباط الذاتي (autocorrelation):** أي أن التباين المشترك بين الأخطاء العشوائية غير مساوي للصفر، ويكون إرتباطها مع بعضها البعض مختلفا عن الصفر.
- **التداخل الخطي (multicolinearity):** أي عدم إستقلال المتغيرات المستقلة عن بعضها بعضا، وتنشأ هذه الحالة عندما تعتمد بعض أو كل المتغيرات المستقلة على بعضها بعضا بشكل خطي (linear dependence)، وتسمى هذه المشكلة في بعض الأحيان الترابط الخطي المتعدد.
- **خطأ التوصيف (specification error):** يحدث هذا الخطأ عندما لا يتم تحديد النموذج بطريقة سليمة، نتيجة لحذف متغير أو متغيرات ضرورية، أو أن يشمل النموذج متغير أو متغيرات غير ضرورية، أو أن يأخذ شكلا غير مناسب، رياضيا أو اقتصاديا.
- **الأنية (simultaneity):** أي أن يتأثر النموذج الموصوف بنموذج أخر، فلا يعرف مسار السببية بين المتغير (المتغيرات) في الطرف الأيمن والمتغير (الظاهرة) في الطرف الأيسر من الاقتران، أو أن توجد أثار أنية من أي نوع، أو أن تتأثر قيمة المتغير المستقل بالخطأ العشوائي، مما يترتب على ذلك انتهاكا لفرضية ثبات ذات المتغير.

سوف نقوم فيما يلي بتحليل، أسباب وعرض الاختبارات الكشف والحلول المقترحة لمشكلة عدم تجانس التباين، الارتباط الخطي بين الأخطاء، والتعدد الخطي.

[38]– عبدالرزاق بني هاني، **الاقتصاد القياسي نظرية الانحدار البسيط والمتعدد**، الطبعة الأولى، دار وائل، الأردن،2014 ،ص136.

مشكل الارتباط الذاتي بين الأخطاء

أولا: مفهوم الارتباط الخطي بين الأخطاء

تحدث مشكلة الارتباط الذاتي بين الأخطاء عندما ترتبط البواقي العشوائية مع بعضها البعض أي أن قيمة المتغير العشوائي ε_t في السنة t يرتبط مع قيمته السابقة ε_{t-1} أو مع قيمته اللاحقة ε_{t+1} وأن هذا الارتباط يحدث بسبب العلاقات الاقتصادية بين الظاهرة المدروسة[39]، مما يخرق أحد الفروض المحيطة بالمتغير العشوائي:

استقلالية الأخطاء إذا كان $t \neq t'$ فإن $E(\varepsilon_t, \varepsilon_{t'}) = 0$ أي أن التباين المشترك ما بين المتغيرات العشوائية يجب أن تساوي صفرا.

ثانيا: أسباب ظهور الارتباط الذاتي بين الأخطاء

يظهر الارتباط الذاتي بين الأخطاء للأسباب التالية[40]:

1- الأثار الممتدة لبيانات السلاسل الزمنية : إن بعض العوامل العشوائية الطارئة وغير المتكررة قد ينتج عنها ترابط في قيم العنصر العشوائي ε_t لأكثر من فترة زمنية واحدة، فالحروب والفيضانات والزلازل تمتد بأثارها على فعالية الاقتصاد لعدة سنوات متتالية، مما يتسبب في حصول إرتباط ذاتي بين قيم ε_t المتلاحقة، حيث أن القيم الحالية تتأثر بالقيم الأخرى للفترات السابقة.

2- حذف بعض المتغيرات المستقلة من العلاقة المدروسة لسبب أو لأخر: مثل عدم توفر البيانات المناسبة عنها أو لغرض تبسيط هيكل النموذج، وقد يكون من بين هذه المتغيرات المحذوفة متغير أو أكثر مرتبط ذاتيا، الأمر الذي يؤدي إلى جعل العنصر العشوائي يتضمن تلك المتغيرات المرتبطه، ومن ثم فإن ε_t لا يعكس الخطأ العشوائي في النموذج فحسب، إنما يعكس أيضا المتغيرات المحذوفة.

3- معالجة البيانات: تجرى على البيانات المنشورة أحيانا عمليات تشذيب وقد يتم تقدير قيم بعض المشاهدات اعتمادا على قيم مشاهدات أخرى، ذلك أن عمليات التشذيب والتقدير تعتمد في العادة على أخذ معدلات قيم المشاهدات المتتالية، مما يخلق علاقة ما بين أخطاء تلك المشاهدات وبالتالي التأثير على طبيعة توزيعها.

[39]- عدنان داود العذاري، صادق على الجبوري، **الاقتصاد القياسي نظرية وحلول تطبيق باستخدام Minitab, Relase 14**، الطبعة الأولى، دار جرير، الأردن، 2010، ص 99.

[40]- حسين علي بخيت، سحر فتح الله، مرجع سبق ذكره، ص189.

4- الصياغة غير الدقيقة للنموذج: بمعنى أن شكل العلاقة الدالية المستخدمة لا يتطابق مع الشكل الحقيقي، فإذا افترضنا علاقة خطية بين المتغيرين y و x، في حين أن العلاقة الحقيقية غير خطية فإنه بمكن أن ينتج عن ذلك ترابط ذاتي في العنصر العشوائي.

ثالثا: النتائج المترتبة على وجود الارتباط الذاتي

تبقى مقدرات طريقة المربعات الصغرى العادية (OLS) تتسم بالخطية وعدم التحيز، إلا أنها تفقد خاصية أفضل وأصغر تباين (التباين بوجود الارتباط الذاتي يفوق التباين في غياب الارتباط الذاتي)، كما يؤثر الارتباط الذاتي على نتائج تحليل الانحدار فتعطي الاختبارات t ، F نتائج أقل دقة من تلك في حالة عدم وجوده، حيث تستند تلك الاختبارات على التباين والخطاء المعياري وعدم دقة التنبؤات المتحصلة عليها بطريقة OLS، مما يتطلب استخدام طرق أخرى للتقدير كطريقة المربعات الصغرى المعممة (GLS) Generalized Least Squares Method لأنها تعطي أفضل تقدير خطي غير متحيز في حالة وجود الارتباط الذاتي شريطة أن تكون قيمة (ρ) معلومة[41].

رابعا: اختبارات الكشف عن الارتباط الذاتي بين الأخطاء

1- اختبار Durbin-Watson:

هذا الاختبار يخص الارتباط الذاتي من الدرجة الأولى، و تأخذ الفرضية الشكل التالي:

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

نقوم بحساب الإحصائية DW التي تعطى بالعلاقة التالية: $DW = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})}{\sum_{t=1}^{n}\varepsilon_t^2}$.

بعد حساب DW نقارنها مع القيمتين المجدولتين، d_1 التي تمثل الحد الأدنى لانعدام الارتباط الذاتي، و d_2 التي تمثل الحد الأقصى لانعدام الارتباط الذاتي، و ذلك حسب عدد المشاهدات n ، و عدد المتغيرات التفسيرية في النموذج عند درجة معنوية 5%.

و يتم قبول و رفض الفرضيتين حسب المخطط التالي:

0		d_1		d_2		2		$4-d_2$		$4-d_1$		4
	P>0		?		P=0		P=0		?		P<0	

قيمة DW الوسطية هي 2 و عندها ينعدم الارتباط الذاتي، أي: P=0.

[41] - حسين علي بخيت، سحر فتح الله، مرجع سبق ذكره، ص 198.

ويتم قبول و رفض H_0 حسب الحالات التالية:

$0 < DW < \mathrm{d}_1$ وجود ارتباط ذاتي موجب.

$\mathrm{d}_1 < DW < \mathrm{d}_2$ مجال غير محسوم، هناك شك في وجود أو عدم وجود ارتباط ذاتي.

$\mathrm{d}_2 < DW < 4 - \mathrm{d}_2$ عدم وجود ارتباط ذاتي.

$4 - \mathrm{d}_2 < DW < 4 - \mathrm{d}_1$ مجال غير محسوم، هناك شك في وجود أو عدم وجود ارتباط ذاتي.

$4 - d_1 < DW < 4$ وجود ارتباط ذاتي سالب.

سلبيات اختبار دربن واتسون:

لا يمكن استعمال هذا الاختبار إلا في ظل الشروط التالية:

- يجب أن يكون النموذج متضمنا للمعلم الثابت b_0 .
- النموذج المقدر لا يتضمن متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة.
- لا يختبر دربن واتسون إلا الارتباط الذاتي من الدرجة الأولى.
- لا يمكن اتخاذ القرار بوجود أو عدم وجود ارتباط ذاتي بين الأخطاء في وقوع إحصائية دربن واتسون $\widehat{DW}$ في منطقة الشك.

2- اختبار Breusch-Godfrey

يرتكز هذا الاختبار على مضاعف لاغرنج والذي يسمح بإختبار وجود ارتباط ذاتي من درجة أكبر من الواحد. نموذج الانحدار الذاتي للأخطاء من الدرجة p يكتب على الشكل التالي[42]:

$$\varepsilon_t = \rho_t \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots + \rho_p \varepsilon_{t-p} + \mu_t$$

ليكن النموذج العام حيث أن الأخطاء مرتبطة ذاتيا:

$$Y_t = B_0 + B_1 X_{t1} + \cdots + B_k X_{tk} + \rho_t \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots + \rho_p \varepsilon_{t-p} + \mu_t$$

هناك ثلاث خطوات لإجراء هذا الإختبار:

- تقدير النموذج العام بطريقة المربعات الصغرى ثم حساب البواقي $\hat{\varepsilon}_t$
- تقدير المعادلة الوسيطية التالية:
- $\hat{\varepsilon}_t = B_0 + B_1 X_{t1} + \cdots + B_k X_{tk} + \rho_t \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \cdots + \rho_p \hat{\varepsilon}_{t-p} + \mu_t$

[42]- شيخي محمد، مرجع سبق ذكره، ص 100.

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

- فرضية استقلالية الأخطاء H_0 التي ينبغي إختبارها هي:

$$H_0 = \rho_1 = \rho_2 = \cdots = \rho_P = 0$$

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

خامسا: الحلول المقترحة للارتباط الذاتي

إن حل كل حالة يعتمد على مصدر الارتباط الذاتي[43]:

فإذا كان المصدر هو حذف متغيرات أساسية من النموذج، يكون الحل بإعادة إدخال هذه المتغيرات في مجموعة المتغيرات المستقلة. وتكون أبسط طريقة لإكتشاف ما إذا كان سبب الارتباط الذاتي هو حذف متغيرات من النموذج بواسطة تقدير بواقي المربعات الصغرى $\hat{\varepsilon}_t$ ضد المتغيرات التي تكون مسبقا لها علاقة بالمتغيرات المستقلة للظاهرة الموجودة تحت الدراسة.

عندما نكتشف ظاهرة الارتباط الذاتي (بإجراء الاختبارات المناسبة) يمكن أن نصحح النموذج وذلك بالحصول على مقدر ρ بواسطة تطبيق المربعات الصغرى العادية لمجموعة البيانات المحولة.

إن تحويل البيانات الأصلية يعتمد على الشكل أو البناء الهيكلي الأصلي للارتباط الذاتي. حيث إذا كان الارتباط الذاتي من المرتبة الأولى AR(1)، يكون التحويل بواسطة طرح النموذج المؤخر بفترة واحدة ومضروبا بمقدر المعلمة ρ أي $\hat{\rho}$ من النموذج الأصلي أي:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots\cdots + \beta_k X_{kt} + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + \mu_t$$

$$Y_{t-1} = \beta_1 + \beta_2 X_{2t-1} + \cdots\cdots + \beta_k X_{kt-1} + \varepsilon_{t-1}$$

ثم نضرب المعادلة الأخيرة ب $\hat{\rho}$ ثم نطرحها من المعادلة الممثلة ل Y_t لنجد:

$$Y_t - \hat{\rho}Y_{t-1} = (\beta_1 - \hat{\rho}\beta_1) + \beta_2(X_{2t} - \hat{\rho}X_{2t-1}) + \cdots\cdots + \beta_k(X_{kt} - \hat{\rho}X_{kt-1}) + \mu_t$$

$$Y_t^* = \beta_1^* + \beta_2 X_{2t}^* + \cdots\cdots + \beta_k X_{kt}^* + \mu_t$$

[43] – تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الجزء الثاني، الطبعة الثانية، ديوان المطبوعات الجامعية، الجزائر، 2010، ص69–ص71.

حيث أن:

$$\beta_1^* = \beta_1(1-\hat{\rho})$$

$$X_{jt}^* = X_{jt} - \hat{\rho}X_{jt-1} \qquad j = 2,3,\cdots\cdots,k$$

ومنه إذا عرفنا قيمة $\hat{\rho}$ نستطيع تطبيق قانون المربعات الصغرى العادية على النموذج المحول. كما نلاحظ أن النموذج المحول يحتوى على (n−1) ملاحظات بسبب سقوط الملاحظة الأولى.

ولتحاشي هذا الضياع في الملاحظات فإن بعض الإحصائيين أمثال Prais-Winsten، يقترحون كتابة الشكل المحول للملاحظات الأولى كما يلي:

$$Y_1^* = \sqrt{1-\hat{\rho}^2}.Y_1$$

$$X_1^* = \sqrt{1-\hat{\rho}^2}.X_1$$

ويلاحظ بعض الكتاب أنه إذا كانت إحصائية DW أقل من معامل التحديد المضاعف أي $DW < R^2$، فإنه يمكن إستعمال الفروقات الأولى في ملاحظات متغيرات النموذج. حيث نقدر الفرق الأول $(Y_t - Y_{t-1})$ في مجموعة الفروقات للمتغيرات المستقلة $(X_t - X_{t-1})$، وذلك تحت الفرضية القائلة بأن الفروقات الأولى للأخطاء $(\varepsilon_t - \varepsilon_{t-1})$ تكون مستقلة.

تمرين حول مشكل الارتباط الذاتي للأخطاء

لتكن لدينا البيانات التالية الموضحة في الجدول التالي:

السنوات	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Y	87.4	97.9	96.7	98.2	99.8	100.5	103.2	107.8	96.6	88.9
X1	98.6	101.2	102.4	100.9	102.3	101.5	101.6	101.6	99.8	100.3
X2	99.1	99.1	98.9	110.8	108.2	105.6	109.8	108.7	100.6	81
X3	108.5	110.1	110.4	104.3	107.2	105.8	107.8	103.4	102.7	104.1
السنوات	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Y	75.1	76.9	84.6	90.6	103.1	105.1	96.4	104.4	110.7	127.1
X1	97.6	97.2	97.3	96	99.2	100.3	100.3	104.1	105.3	107.6
X2	68.6	70.9	81.4	102.3	105	110.5	92.5	89.3	93	106.6
X3	99.2	99.7	102	94.3	97.7	101.1	102.3	104.4	108.5	111.3

المطلوب:

1− تقدير معالم النموذج ؟

2- اختبار وجود الارتباط الذاتي بين الأخطاء :أ- اختبار دربن واتسون ؟

ب- اختبار Breusch –Godfrey ؟

3- صحح النموذج من مشكل الارتباط الذاتي للأخطاء إن وجد ؟

الحل:

1- تقدير معالم النموذج

Dependent Variable: Y
Method: Least Squares
Date: 01/25/20 Time: 10:39
Sample: 1 20
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-242.7146	26.93755	-9.010268	0.0000
X1	3.887240	0.402380	9.660612	0.0000
X2	0.404491	0.061666	6.559355	0.0000
X3	-0.869804	0.241450	-3.602423	0.0024

R-squared	0.938269	Mean dependent var	97.55000
Adjusted R-squared	0.926694	S.D. dependent var	11.83076
S.E. of regression	3.203188	Akaike info criterion	5.343027
Sum squared resid	164.1666	Schwarz criterion	5.542173
Log likelihood	-49.43027	Hannan-Quinn criter.	5.381902
F-statistic	81.06246	Durbin-Watson stat	1.060806
Prob(F-statistic)	0.000000		

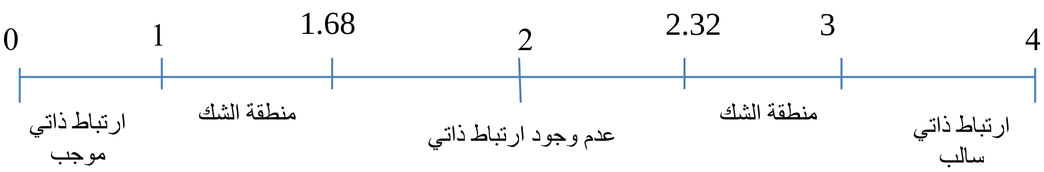
من الجدول نلاحظ أن معالم النموذج (معاملات المتغيرات المفسرة) لها معنوية إحصائية خلافا للصفر($t_{16}^{0.05} = 2.12$)،و قيمة فيشر المحسوبة أكبر من (F(3, 16, 0.05)=3.24).

2- اختبار وجود الارتباط الذاتي للأخطاء:

أ- اختبار دربن واتسون

من الجدول نلاحظ أن إحصائية دربن واتسون المحسوبة DW= 1.0608

لدينا أن K=3 و N=20 من جدول دربن واتسون نجد أن d1= 1.00 و d2=1.68



وبالتالي فإن DW=1.0608 محصورة بين d1= 1.00 و d2=1.68أي نحن في منطقة الشك لا يمكن اتخاذ قرار بوجود أو عدم مشكل الارتباط الذاتي بين الأخطاء.

ب- اختبار Breusch –Godfrey

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.782095	Prob. F(2,14)	0.0961
Obs*R-squared	5.688137	Prob. Chi-Square(2)	0.0582

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 01/25/20 Time: 10:43
Sample: 1 20
Included observations: 20
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.632449	24.80630	-0.347994	0.7330
X1	0.101819	0.370069	0.275135	0.7872
X2	-0.025021	0.056883	-0.439867	0.6667
X3	0.008226	0.231046	0.035601	0.9721
RESID(-1)	0.653275	0.282721	2.310670	0.0366
RESID(-2)	-0.383934	0.290205	-1.322976	0.2070

R-squared	0.284407	Mean dependent var	-2.35E-15
Adjusted R-squared	0.028838	S.D. dependent var	2.939447
S.E. of regression	2.896753	Akaike info criterion	5.208383
Sum squared resid	117.4765	Schwarz criterion	5.507103
Log likelihood	-46.08383	Hannan-Quinn criter.	5.266697
F-statistic	1.112838	Durbin-Watson stat	2.011463
Prob(F-statistic)	0.397368		

من خلال الجدول نلاحظ أن $R^2=0.289762$ و لدينا n=20:

نحسب إحصائية مضاعف لاغرنج

$$LM=n.R^2=20X0.284407=5.688137$$

نستخرج إحصائية كي دو بدرجة حرية 2 من جدول كي دو

$$\aleph^2_{0.05}(2) = 5.99$$

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.782095	Prob. F(2,14)	0.0961
Obs*R-squared	5.688137	Prob. Chi-Square(2)	0.0582

المقارنة :

نلاحظ أن

$\aleph^2_{0.05}(2) = 5.99$ أكبرمن LM=5.795238 وبالتالي نقبل الفرضية Ho أي عدم وجود مشكل الإرتباط ذاتي للإخطاء

3- في هذه الحالة لا يوجد مشكل الارتباط الذاتي بين الأخطاء لكن سوف نوضح كيفية تصحيح النموذج مشكل الارتباط الذاتي للأخطاء في حالة وجوده وذلك بتحويل المتغيرات:

نحسب raul :

$$\hat{\rho}_1 = 1 - \frac{DW}{2} = 1 - \frac{1.060}{2} = 0.470$$

نحول المتغيرات:

$$dy = (y_t - \hat{\rho}_1 * y_{t-1}) = (y_t - 0.47 * y_{t-1})$$
$$dx_1 = \left(x_{1,t} - \hat{\rho}_1 * x_{1,t-1}\right) = \left(x_{1,t} - 0.47 * x_{1,t-1}\right)$$
$$dx_2 = \left(x_{2,t} - \hat{\rho}_1 * x_{2,t-1}\right) = (x_{2,t} - 0.47 * x_{2,t-1})$$
$$dx_3 = (x_{3,t} - \hat{\rho}_1 * x_{3,t-1}) = (x_{3,t} - 0.47 * x_{3,t-1})$$

نحسب قيمة المشاهدة الأولى للمتغيرات

$$Y^*_{1981} = \sqrt{1-\hat{\rho}^2}.Y_{1981} = \sqrt{1-(0.47)^2}.(87.4) = 68.093$$
$$x^*_{1,1981} = \sqrt{1-\hat{\rho}^2}.x_{1,1981} = \sqrt{1-(0.47)^2}.(98.6) = 76.819$$
$$x^*_{2,1981} = \sqrt{1-\hat{\rho}^2}.x_{2,1981} = \sqrt{1-(0.47)^2}.(99.1) = 77.208$$
$$x^*_{3,1981} = \sqrt{1-\hat{\rho}^2}.x_{3,1981} = \sqrt{1-(0.47)^2}.(108.5) = 84.532$$

ثم نقدر النموذج الجديد

Dependent Variable: DY
Method: Least Squares
Date: 01/25/20 Time: 11:38
Sample: 1 20
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.478704	17.84695	-0.082855	0.9350
DX1	0.795992	1.058123	0.752269	0.4628
DX2	0.484086	0.159842	3.028522	0.0080
DX3	-0.247851	0.753566	-0.328905	0.7465

R-squared	0.621326	Mean dependent var	53.72300
Adjusted R-squared	0.550324	S.D. dependent var	9.386202
S.E. of regression	6.294187	Akaike info criterion	6.693987
Sum squared resid	633.8686	Schwarz criterion	6.893133
Log likelihood	-62.93987	Hannan-Quinn criter.	6.732862
F-statistic	8.750889	Durbin-Watson stat	1.096600
Prob(F-statistic)	0.001149		

ثم حساب الثابت:

$$\hat{b}_0{}^* = \hat{b}_0 / (1-\hat{\rho}_1) = -1.478704 / (1-0.47) = -2{,}790007$$

مشكل التعدد الخطي (التداخل الخطي)

أولا: مفهوم التداخل الخطي المتعدد

التداخل الخطي مصطلح معرب من الأصل الإنجليزي (Multicolinearty) وهو مصطلح مركب من ثلاثة مقاطع

الأول: Multi ومعناه تعدد أو متعدد.

الثاني: Co وتعني التلازم أو التبادل أو الاشتراك أو التداخل.

الثالث: Linearity وتعني الخطية.

وقد اختلفت الترجمات العربية له، فمنهم من يسميه (الارتباط الخطي المتعدد). وهو مفهوم قد يخلق خلطا له مع الارتباط الخطي المتعدد (Multiple Correlation) ولهذا سنستبعده من الاستخدام. ومنهم من يترجمه ب(تعدد العلاقات الخطية) وهو يفقد الوضوح والتحديد. ومنهم من يسميه (عدم استقلالية المتغيرات مسبقا عن بعضها). وأخرون يترجمونه إلى (الامتداد الخطي المتعدد). ويترجمه مؤلف أخر ب (الاشتراك الخطي) أو (تعدد العلاقات الخطية) وغيرهما من الترجمات.

أما أقرب ترجمة لهما ضمن المفهوم العام فهي المصطلحات الأتية[44]:

التزامن– التلازم الخطي – الخطية المتلازمة للمتغيرات المفسرة – (الترابط الخطي) – التشابك الخطي المتعدد أو التداخل الخطي المتعدد.

ويعتبر النرويجي (راكنرفريش) أول من اكتشف وجود هذه الظاهرة الإحصائية وذلك في معرض تحليله لبيانات السلاسل الزمنية. فقد اتضح له وجود تداخل أو ترابط خطي متبادل بين المتغيرات الاقتصادية التفسيرية بسبب التطور المتناغم أو المتوافق لهذه المتغيرات عبر الزمن أي أنها تزداد وتنخفض بوتيرة واحدة أو متقاربة عبر الزمن مما يجعل أحدهما غير ذو مقدرة تفسيرية في النموذج ويجعل بالتالي معلمة او أكثر من المعلمات الإنحدارية أو المقطعية مساوي للصفر أو مقاربة له، أو تتساوى المعلمتان الانحداريتان مما يجعل معامل الارتباط الجزئي بين المتغيرين المستقلين مساويا للواحد عدد صحيح.

والمشكلة هنا هي بين المتغيرات التفسيرية ذاتها وليس بين المتغير التفسيري والمتغير التابع.

وهذا يعني أن المشكلة تظهر فقط عندما يكون هناك أكثر من متغير تفسيري واحد مثل X و Z و W.

ولهذا فإن هذه المشكلة غير موجودة في النموذج الخطي البسيط بل في النموذج الخطي العام والمتعدد.

44– وليد إسماعيل السيفو، فيصل مفتاح شلوف، صائب جواد إبراهيم جواد، مرجع سبق ذكره، ص 93.

ثانيا: أسباب ظهور التداخل الخطي المتعدد

هناك عدة أسباب لتطور مشكلة التلازم الخطي أو التعدد الخطي بين المتغيرات المستقلة وخاصة في السلاسل الزمنية وأهمها:

1- اشتراك بعض المتغيرات المستقلة في التحرك باتجاه زمني واحد ودون تخلف زمني.

2- التغير المتداخل لعدم جمع بيانات كافية ومن عينات أكثر.

3- التحرك بإتجاه واحد أو متعاكس بمعدل متزامن أو واحد ولنفس الفترة الزمنية.

4- استخدام المتغيرات المتخلفة زمنيا مثل استخدام دخل سنة (t) المعتمد على دخل سنة (t-1).

5- عدم إمكانية التحكم ببيانات المشاهدات لأنها لا تخضع للسيطرة والتجارب.

ثالثا: مؤشرات وجود التداخل الخطي

هناك بعض المؤشرات لوجود تداخل خطي بين المتغيرات المستقلة ومن أهمها الأتي:

1- ظهور قيمة واحدة أو أكثر من قيم المعلمات الانحدارية والثوابت مساوية للصفر.

2- تضخم قيمة الخطأ المعياري للتقدير وقد يصل إلى ما لا نهاية.

3- انعدام المعنوية الإحصائية لإحدى المعلمات المقدرة او أكثر من واحدة رغم أن (R^2) يكون عاليا، إضافة إلى أنها غير متحيزة.

4- ظهور ارتباط جزئي تام بين متغيرين تفسيريين أو أكثر، أو بينه وبين المتغير التابع.

5- تكون قيم بعض المعلمات الانحدارية والثوابت غير صفرية لكنها متساوية القيم مما يجعل أحد المتغيرية التفسيريين فائضا وليس بذي أهمية في النموذج لعدم إمكان عزل تأثير أحدهما على الأخر، لهذا فإن القدرة التفسيرية للنموذج والمتغير تنخفض.

6- عدم قبول الحاسب الإلكتروني للبيانات أو عدم احتساب قيم الثوابت.

رابعا: اختبارات الكشف عن التداخل الخطي المتعدد

إن مشكلة التداخل الخطي وتأثيره تظهر في درجة الترابط المتبادل (R_{xz}) بين المتغيرات المفسرة، وكذلك معامل الارتباط المتعدد الكلي للنموذج (Ryxz). ويعتقد الكثيرون أن المعامل الأخير يمكن أن يستخدم لاختبار وجود تداخل خطي بين المتغيرات المفسرة.

غير ان الممارسة العملية قد أثبتت أن استخدام أي من هذه المتغيرات لوحده كمعيار لوجود التداخل الخطي المتعدد بين المتغيرات المفسرة هو غير كاف للأسباب الأتية:

1- أن وجود قيمة عالية للخطأ المعياري للتقدير لا تظهر دائما مع التداخل الخطي، بدليل أن هناك نماذج اقتصادية مأخوذة من عينات كثيرة مثل دالة (كوب- دوغلاس) تظهر فيها أخطاء معيارية كبيرة.

2- أن R^2 يمكن أن يكون عاليا لكن النتائج لا تكون دقيقة ولا معنوية وقد تحصل على إشارات خاطئة أو خطأ معياري عال.

مع ذلك فإن استخدام مثل هذه المؤشرات يمكن أن ينفع في الإحساس بهذه المشكلة والكشف عن وجوده. ولهذا فإننا نستخدم هذه الطرق إضافة إلى طرق أخرى مبتكرة لكتاب قياسيين معتمدين.

وفيما يأتي عرض لأهم الاختبارات الخاصة بالكشف عن هذه العلاقة.

أولا: طريقة (فريش) Frich's Confluence Analysis

تقوم طريقة فريش على الأتي:

1- إجراء انحدار للمتغير التابع على كل متغير تفسيري على حدى. بهذه الطريقة نحصل على الانحدار الأساسي ونفحصها في ضوء معلومات مسبقة Apriori ومعايير إحصائية.

2- نختار النتائج الانحدارية التي تعطي نتائج أفضل في ضوء المعايير أعلاه.

3- ندخل بالتدريج متغير تفسيري بعد أخر على أفضل انحدار في ضوء ما جاء في (2) ونفحص أثره على المعاملات الارتباطية الفردية وعلى الأخطاء المعيارية وعلى (R^2).

4- وعلى أساس ذلك يمكننا تصنيف المتغيرات التفسيرية على أنها متغيرات تحديدية أو فائضة في النموذج، أي ليست لها أهمية قياسية واقتصادية وعلى الأسس الأتية:

أ- إذا ما حسن المتغير التفسيري المضاف (R^2) دون أن يؤثر في دقة المعلمات بحيث يحدث فيها أخطاء غير مقبولة في ضوء المعلومات النظرية الاقتصادية، عند ذلك سيكون المتغير هذا مفيد في التحليل القياسي.

ب- عندما لا يغير التغير التفسيري المضاف القدرة التفسرية للنموذج و(R^2) كما لا يؤثر في قيمة المعلمات إيجابيا سيعتبر هذا المتغير زائدا او فائضا ولا يدخل ضمن المتغيرات التفسيرية للنموذج.

ج- إذا ما أثر إدخال المتغير التفسيري الجديد على قيمة معلمات النموذج وحول إشاراتها لتكون مقبولة اقتصاديا عند ذلك سيكون المتغير تحديديا.

د- إذا ما أثر إدخال المتغير التفسيري على قيمة معلمات النموذج، وحول إشارتها بإتجاه يحولها إلى قيم غير مقبولة اقتصاديا عند ذلك سيكون لدينا مشكلة تداخل خطي متعدد ذات اهمية جدية.

ولهذا ولأجل ألا نهمل او نرمي خارجا متغير تفسيري تحديدي علينا أن نلجأ إلى حلول لهذه المشكلة بعد تحديدها، بخلافه فإن تأثير هذا المتغير سيمتص من قبل المتغيرات التفسيرية والخطأ العشوائي.

ثانيا: طريقة فارارا– كلاوبير للإختبار Glauber Test- The Farrar

وهذه الطريقة مؤلفة من ثلاثة اختبارات وهي:

1- اختبار مربع كاي Square-Chi للكشف عن وجود درجة شدة التداخل الخطي المتعدد من خلال اختبار الاستقلالية الإحصائية للمتغيرات المستخدمة في العينة.

2- اختبار (F) لتحديد المتغيرات المرتبطة خطيا.

3- اختبار إيجاد نموذج التداخل الخطي المتعدد وتحديد المتغيرات المسؤولة عن التداخل الخطي المتعدد.

4- اختبار (t) لتحديد المتغيرات المسببة للتداخل الخطي.

ويحدد فارار الأتي:

إذا كان التداخل الخطي المتعدد تاما فإن المعلمات ستكون غير محددة. ويمكن قياس الارتباط بين المتغيرات التفسيرية من خلال معاملات الارتباط الجزئية.

خامسا: حلول التداخل الخطي المتعدد

رغم ان التداخل الخطي المتعدد يعتبر عند البعض مشكلة خطيرة إلا أنها عند الأخرين لا تؤلف مشكلة ويمكن التعايش معها وخاصة عندما لا يؤثر بشكل خطير على المعلمات المقدرة.

ومع ذلك يفضل الكثير من الكتاب إيجاد حلول لهذه المشكلة والتي يمكن ان تتمثل بالأتي:

1- حذف المتغير التفسيري المسبب لهذه الظاهرة إن كان تأثيره في القدرة التفسيرية للنموذج ليست بذات أهمية وخاصة تلك التي تكون معلماتها مقاربة للصفر أو مساوية له حتى لا يقع في مطب التوصيف بعدم إدخال المتغيرات المطلوبة وأن تصاب المعلمات المقدرة الأخرى بالتحييز. والمعيار هو هل سيساوي تخفيض التباين بالمقارنة مع زيادة تحييز المعلمات؟ وقد يفضل بعض الكتاب معيار (OLS) ليس أصغر تباين بل (وسط مربع الخطأ) أي S بدلا من σ.

2- قد تصيب هذه الظاهرة بعض المعلمات ولا تصيب الأخرى بالتحيز. فيمكن في هذه الحالة استخدام غير المصابة لأغراض التحليل والتنبؤ ووضع السياسة. وأن استخدمت للتنبؤ يفضل أن يتواصل النموذج من الماضي إلى المستقبل.

3- قد يكون صغر العينة سببا في ظهور هذه المشكلة عند ذلك يصار إلى زيادة حجم العينة بالحصول على بيانات أكثر لا تتصف بالتداخل الخطي.

4- استخدام معلومات معطاة مسبقا عن المعلمات مثل الميل الحدي للإستهلاك أي استخدام معلومات خارجية تدمج في النموذج.

5- تحويل بعض المتغيرات بقسمتها على (k) مثلا أو (z) أو ضرب المتغيرات والمعلمات في (yw) وقسمتها على (z).

6- استخدام طرق مختلفة لتكوين معلومات مقطعية خارجية مثل، طريقة المربعات الصغرى المقيدة، وتعديل (دربن) للمربعات الصغرى العامة، وطريقة التقدير المختلط التي اقترحها (ثايل وكولد بركر).

7- إحلال المتغيرات المتخلفة زمنيا بدل الأنية.

8- إدخال معاملات إضافية للنموذج كقلبها إلى معادلة أنية.

تمرين حول مشكل التعدد الخطي

لتكن لدينا البيانات الموضحة في الجدول التالي:

1971	1970	1969	1968	1967	1966	1965	1964	annee
64	58.5	52.9	47.7	40.6	37.7	32	28.4	y
1063.4	982.4	935.5	868.5	796.3	753	688.1	635.7	X1
121.3	116.3	109.8	104.2	100	97.2	94.5	92.9	X2
1979	1978	1977	1976	1975	1974	1973	1972	annee
260.9	217.5	185.8	155.4	126.9	131.9	94.4	75.9	y
2368.5	2127.6	1899.5	1702.2	1528.8	1412.9	1306.6	1171.1	X1
217.4	195.4	181.5	170.5	161.2	147.7	133.1	125.3	X2

المطلوب :

1- اختبر وجود مشكل التعدد الخطي بتطبيق الاختبارات التالية:

أ- اختبار كلاين klien ؟

ب-اختبار فارار كلاوبير Fara klauber ؟

الحل:

أ- اختبار كلاين klien:

1-تقدير النموذج

Dependent Variable: Y
Method: Least Squares
Date: 05/12/14 Time: 19:36
Sample: 1964 1979
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-101.4885	33.08031	-3.067943	0.0090
X1	0.078534	0.055958	1.403450	0.1839
X2	0.758554	0.761246	0.996464	0.3372

R-squared	0.987366	Mean dependent var	100.6563
Adjusted R-squared	0.985422	S.D. dependent var	71.92798
S.E. of regression	8.684477	Akaike info criterion	7.328312
Sum squared resid	980.4618	Schwarz criterion	7.473172
Log likelihood	-55.62650	F-statistic	507.9814
Durbin-Watson stat	1.186381	Prob(F-statistic)	0.000000

2- - حساب معاملات الارتباط البسيط بين المتغيرات:

Correlation Matrix		
	X1	X2
X1	1.000000	0.997165
X2	0.997165	1.000000

من مصفوفة الارتباط نجد أن: $r^2_{x1\,x2} = (0.997165)^2 = 0.994338$

نلاحظ أن معامل الارتباط بين المتغيرات المفسرة أكبر من معامل التحديد $R^2= 0.987366$ وهذا يدل على وجود مشكل التعدد الخطي

ب-اختبار فارار كلاوبير Fara klauber:

1-حساب محدد مصفوفة الارتباط:

$$D = \begin{vmatrix} 1 & r_{x1\,x2} \\ r_{x1\,x2} & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0.997165 \\ 0.997165 & 1 \end{vmatrix} = 1 - 0.994328 = 0.005662$$

2-حساب إحصائية كي دو:

$$\aleph^2 = -\left[n - 1 - \frac{1}{6}(2K+7)\right] \times LnD$$

$$\aleph^2 = -\left[16 - 1 - \frac{1}{6}(2x2+7)\right] \times (-2.247) = 34.0795$$

3- إيجاد قيمة كي دو المجدولة بدرجة حرية $\frac{1}{2}K(K+1) = \frac{1}{2}2(2+1) = 3$

$$\aleph^2(3) = 7.815$$

4- المقارنة: نلاحظ أن قيمة كي دو المحسوبة 34.0795 أكبر من كي دو المحسوبة $\aleph^2(3) = 7.815$ وهذا يدل على وجود مشكلة التعدد الخطي.

مشكلة عدم تجانس التباين
The Hetroscedasticity Problm

أولا: مفهوم مشكل عدم تجانس التباين

أحد الفرضيات الأساسية التي يقوم عليها النموذج الخطي (العام والبسيط) هو ثبات التباين لحدود الخطأ (تجانس تباين الخطأ)، Homoscedasticity، لجميع المشاهدات (i) أي:

$$E\left(\varepsilon_i^2\right) = \sigma_i^2$$

وفي النموذج الخطي العام فإن الفرض المناظر:

$$E(\varepsilon_i \varepsilon_i') = \sigma_i^2 I_n$$

وفي واقع الأمر، فإننا كثيرا مانواجه حالات يتعسر فيها إستيفاء الشرط أعلاه، ومن ثم فإن التباين لا يكون ثابتا بل يختلف لكل مشاهدة من مشاهدات العينة، وتصبح لدينا قيم مختلفة وغير ثابتة لتباينات حدود الخطأ العشوائية، وعليه فإن القطر الرئيسي لمصفوفة التباين والتباين المشترك الخاصة بحدود الخطأ يحتوى على قيم مختلفة وغير ثابتة، أي أن[45]:

$$E(\varepsilon_i \varepsilon_i') \neq \sigma_i^2 I_n$$

حيث أن : $\sigma_1^2 \neq \sigma_2^2 \neq \cdots \neq \sigma_n^2$

ويحدث هذا أساسا في حالة الدراسات المعتمدة على البيانات الإحصائية المقطعية Cross-Section Data

ثانيا: أسباب مشكلة عدم تجانس حد الخطأ

هناك عدة أسباب لظهور هذه المشكلة ومن أهمها ما يأتي:

1- سلوكية وتصرف الأفراد التي تقل الأخطاء فيها بمرور الزمن، وعليه فإن تباين حد الخطأ () يتناقص هو الأخر خلال الفترة الزمنية.

2- مع تحسن أساليب جمع البيانات يقل حد الخطأ وذلك لأن جمع البيانات الدقيقة والواقعية تقلل من الأخطاء، مثال ذلك أن الأخطاء التي ترد في المستندات في المؤسسات الحكومية والتي تستخدم الحاسب الألي لتحليل البيانات تكون أقل من مثيلتها في المؤسسات التي لا تستخدم الحاسب الألى.

---

[45]– حسين علي بخيت، سحر فتح الله، مرجع سبق ذكره، ص 260.

3- عادة ما تحدث هذه المشكلة في البيانات المقطعية، مثلا عند اعتبار عينة ممثلة لإنفاق الأسر ذات الدخل المنخفض والمرتفع يلاحظ أن تباين الخطأ الخاص بإنفاق الأسر ذات الدخل المنخفض (على السلع الضرورية) تكون أقل من إنفاق الأسر ذات الدخل المرتفع[46].

ثالثا: اكتشاف ظاهرة عدم تجانس تباينات الأخطاء

لاكتشاف ظاهرة عدم تجانس تباينات الأخطاء يستعمل بحثو الاقتصاد القياسي عدة اختبارات مختلفة من بينها:

1- اختبار Goldfed-Quandt

يصلح هذا الاختبار للعينات الصغيرة والكبيرة، حيث يشترط أن يكون حجم العينة على الأقل ضعف عدد المعالم المطلوب تقديرها في النموذج. أستعمل هذا الاختبار لأول مرة من طرف الباحثين Goldfeld-Quandt. حيث إذا رأينا أن أحد المتغيرات المستقلة x_{ji} هو المتسبب في ظاهرة عدم تجانس تباين الأخطاء نعتبر في هذه الحالة، أن σ_i^2 مرتبطة ومتناسبة إيجابيا (طرديا) مع أحد المتغيرات x_{ji} . ثم تقوم بتقسيم الملاحظات إلى عينتين، العينة الأولى تناسب قيم x_{ji} الكبيرة، والعينة الثانية تخص قيم x_{ji} الصغيرة. ثم نوفق إنحدارين مختلفين (منفصلين) على عينتين مختلفين (لكن متساويتين في الحجم)، وذلك لهدف تطبيق اختبار مساواة تبايني أخطاء العينتين المنفصلتين. ويقترح هذين الباحثين حذف بعض الملاحظات المركزية من العينة الأصلية لكي نفرق مابين تباين أخطاء العينتين[47].

مثال:

لنعتبر النموذج الخطي البسيط: $y_i = a_0 + a_1 x_i + \varepsilon_i \ , i = 1,2, \dots \dots, n$

ولنفترض أن الأخطاء ε_i لها التوزيع الطبيعي وغير مرتبطة فيما بينها

أي أن $E(\varepsilon_i \varepsilon_j) = 0$ ، $i \neq$. فإذا كانت لدينا:

$$E(\varepsilon_i^2) = \sigma_i^2 = \sigma_\varepsilon^2 x_i^2$$

ويجرى اختبار (G-Q) وفقا للخطوات التالية:

1- نرتب الملاحظات y_i و x_i وفقا لقيم x_i (من القيمة الصغيرة إلى القيمة الكبيرة)

---

[46]- حمداوي الطاوس، **مدخل للاقتصاد القياسي**، الطبعة الأولى، دار هومه، الجزائر، 2016، ص101.

[47]- تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الجزء الثاني، مرجع سبق ذكره، ص32.

2- نحذف (نسقط)[48] m ملاحظات مركزية ليكون الباقي هو (n-m) ملاحظات متبقية ثم نقسم الباقي (n-m) إلى عينتين متساويتين n-m)/2).

تحتوي الأولى على n-m)/2) من قيم x_i الصغيرة، وتحتوي الثانية على n-m)/2) ملاحظات من قيم x_i الكبيرة.

3- نكون انحدارين منفصلين بواسطة المربعات الصغرى العادية شريطة أن يكون $\left(\frac{n-m}{2}\right) > k$ على الصيغة التالية:

a) $y_i = a_0 + a_1 x_i + \varepsilon_i \ , i = 1,2, \dots \dots , \frac{n-m}{2}$.

b) $y_i = a_0 + a_1 x_i + \varepsilon_i \ , i = \frac{n-m}{2} + 1, \frac{n-m}{2} + 2, \dots \dots , n$

نحصل من الانحدار الأول على مجموع مربعات البواقي RSS_a، من الانحدار الثاني على RSS_b.

4- نكون الاختبار الإحصائي التالي:

$$\frac{RSS_a}{\sigma_\varepsilon^2} \sim \chi^2_{(n-m-2k)/2}$$
$$\frac{RSS_b}{\sigma_\varepsilon^2} \sim \chi^2_{(n-m-2k)/2}$$

ومنه نجد التوزيع:

$$Q = \frac{RSS_a / v_1}{RSS_b / v_2}$$
$$Q = \frac{RSS_a}{RSS_b} \sim F_{v_1, v_2}$$

حيث أن : $v_1 = v_2 = (n - m - 2k)/2$

فإذا كانت Q=1 يعنى هذا تساوي التباينين ومنه نقبل H_0 القائلة بتجانس تباينات الأخطاء، وكلما كان المقدر Q كبيرا، كلما كان عدم التجانس أقوى. وتعتمد قوة الاختبار كذلك على عدد الملاحظات المركزية والمحذوفة من الحجم الكلي للعينة، حيث يقترح Johnston أن يكون هذا العدد في حدود ثلث حجم العينة الأصلية.

2- اختبار White

اقترح White (1980) اختبارا يعتمد على العلاقة بين مربعات البواقي وجميع المتغيرات المستقلة وكذا مربعاتها. يمكن إبراز خطوات هذا الاختبار كما يلي[49]:

[48] - حيث ان m عادة يساوي على الأقل ربع العينة (n/4). فمثلا العدد الأمثل ل m لما تكون n=30 هو m=8. ولما تكون n=60، فإن m=16، وهكذا.

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

ثانيا: تقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}_i^2 = \beta_0 + \beta_1 X_{i1} + \alpha_1 X_{i1}^2 + \cdots + \beta_k X_{ik} + \alpha_k X_{ik}^2 + \mu_i$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2.

ثالثا: فرضية ثبات تباين الأخطاء H_0 التي ينبغي إختبارها هي:

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

إحصائية مضاعف لاغرانج $LM = n \times R^2$ تتبع توزيع χ^2 بدرجة حرية $2k$.

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

3– اختبار ثبات التباين الشرطي للأخطاء ARCH-LM

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

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

ثانيا: تقدير المعادلة التالية:

$$\hat{\varepsilon}_i^2 = \theta_0 + \theta_1 \hat{\varepsilon}_{t-1}^2 + \cdots + \theta_q \hat{\varepsilon}_{t-q}^2 + \mu_t$$

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

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

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

إحصائية مضاعف لاغرانج $LM = (n-q) \times R^2$ تتبع توزيع χ^2 بدرجة حرية q.

[49]– شيخي محمد، مرجع سبق ذكره، ص115.

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

رابعا: معالجة مشكلة عدم التجانس

الحال الأولى: المربعات الصغرى الموزونة (weighted least squares)

تستخدم هذه الطريقة عندما يكون التباين σ_i^2 معلوما. لتوضيح ذلك، دعنا نفترض أننا نرغب بتقدير معادلة الانحدار البسيط التالية[50]:

$$Y_i = b_0 + b_1X_i + \varepsilon_i$$

علمنا بإستخدام أحد الاختبارات أن هناك مشكلة " عدم التجانس " في البيانات المستخدمة للتقدير، وبنفس الوقت، تعرفنا على قيمة التباين (σ_i^2). في هذه الحال، يتم تحويل الدالة أعلاه كما يلي: بالقسمة على الانحراف المعياري نحصل على :

$$\frac{Y_i}{\sigma_i} = \frac{b_0}{\sigma_i} + \frac{b_1X_i}{\sigma_i} + \frac{\varepsilon_i}{\sigma_i}$$

$$(Y_i/\sigma_i) = b_0(1/\sigma_i) + b_1X_i(1/\sigma_i) + \nu_i$$

حيث ترمز (ν_i) للخطأ العشوائي الجديد. بتقدير معلمتي هذه الدالة نحصل على ما يسمى بمعادلة انحدار المربعات الصغرى الموزونة (weighted least squares) (WLS) وتكون صيغة المعادلة المقدر كما يلي:

$$(Y_i/\sigma_i) = (\hat{b}_0/\sigma_i) + b_1(X_i/\sigma_i)$$

نلاحظ أن الطرف الأيمن من المعادلة يتكون من حاصل قسمة $(\hat{b}_0)$ على (σ_i) وهو متغير مستقل، ويتم الحصول عليه بقسمة الواحد الصحيح على قيمة الانحراف المعياري. ويتكون الطرف الأيمن من حاصل قسمة (X_i) على (σ_i)، وهما متغيران مستقلان. وبذلك فإن المعادلة البسيطة الأصلية (ذات متغير مستقل واحد) قد تحولت إلى معادلة انحدار متعددة المتغيرات (من متغيرين) ولا يوجد لها نقطة تقاطع ويتم تفسير معامل التحديد (R^2) بطريقة تختلف عن الطريقة العادية.

مثال: إذا كانت لدينا نتائج تقدير دالة الانحدار التالية:

$$Y_i = 292.1798 + 10.4042X_i$$

[50] – عبدالرزاق بني هاني، مرجع سبق ذكره ، ص150.

$$(33.6705) \qquad (2.1235)$$

$$R^2 = 0.75 \qquad\qquad n = 10$$

بإستخدام المربعات الصغرى الموزونة بقسمة كل من (Y_i) و (X_i) على الانحراف المعياري (σ_i) وتقدير معلمتي دالتي الانحدار من البيانات المحولة نحصل على النتائج التالية:

$$(Y_i/\sigma_i) = 307.4466\,(1/\sigma_i) + 9.2143(X_i/\sigma_i)$$

$$(25.0225) \qquad (1.800)$$

$$R^2 = 0.66 \qquad\qquad n = 10$$

يتضح من هذه النتائج مدى تحسن التقديرات وأخطائها المعيارية رغم أن قيمة (R^2) قد إنخفضت، لكن مقدرة المعادلة الثانية على التنبؤ تبقى أفضل من بديلتها الأولى.

الحالة الثانية: المربعات الصغرى الموزونة في حال عدم معرفة التباين

لنفترض أننا نرغب بتقدير معلمتي معادلة انحدار من الشكل التالي:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$

وكان:

$$Var(\varepsilon_i) = E\left(\varepsilon_i^2\right) = \sigma^2 X_i^n$$

أي أن تباين الخطأ العشوائي هو دالة في المتغير (X_i)، ويتناسب طرديا معه.

في هذه الحال يتم تحويل معادلة الانحدار أعلاه، بتقسيم طرفيها على المتغير $\left(X_i^{n/2}\right)$، أي

$$Y_i/(X_i)^{n/2} = \frac{b_0}{X_i^{n/2}} + \frac{b_1 X_i}{X_i^{n/2}} + \frac{\varepsilon_i}{X_i^{n/2}} = b_0/(X_i)^{n/2} + b_1 X_i/(X_i)^{n/2} + \nu_i$$

حيث ترمز (ν_i) للخطأ العشوائي الجديد. ومنه يكون

$$Var(\varepsilon_i) = Var(\varepsilon_i)\left(\frac{1}{X_i^{n/2}}\right)^2 = \sigma^2 X_i^n/(X_i)^n = \sigma^2$$

وبالتالي نكون قد حققنا تجانسا في تباين الخطأ العشوائي.

أما في حال أن:

$$Var(\varepsilon_i) = E\left(\varepsilon_i^2\right) = \sigma^2 (b_0 + b_1 X_i)^n$$

فإن تحويل الدالة يتم بقسمة أطرافها كما يلي:

$$\frac{Y_i}{(b_0 + b_1 X_i)^{n/2}} = \frac{b_0}{(b_0 + b_1 X_i)^{n/2}} + \frac{b_1 X_i}{(b_0 + b_1 X_i)^{n/2}} + \frac{\varepsilon_i}{(b_0 + b_1 X_i)^{n/2}}$$

حيث يتم الحصول على $(Y_i = b_0 + b_1 \quad _i)$ من تقدير معلمتي دالة الانحدار، ثم تستخدم الدالة المقدرة $\left(\hat{Y}_i = \hat{b}_0 + b_1 X_i\right)$ للتحويل كما يلي:

$$\frac{Y_i}{(\hat{Y}_i)^{n/2}} = \frac{b_0}{(\hat{Y}_i)^{n/2}} + \frac{b_1 X_i}{(\hat{Y}_i)^{n/2}} + \frac{\varepsilon_i}{(\hat{Y}_i)^{n/2}}$$

الحالة الثالثة: يمكننا التخلص من مشكلة عدم التجانس بإستخدام اللوغاريتم الطبيعي لتحويل معادلة الانحدار كما يلي:

لنفترض ان لدينا دالة الانحدار التالية :

$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$

وعلمنا، من خلال الاختبارات، بوجود مشكلة " عدم التجانس ". يمكننا استخدام اللوغاريتم الطبيعي في تحويل البيانات للحصول على الصيغة التالية، ثم تقدير معلماتها للتخلص من المشكلة:

$$Ln\, Y_i = \mathrm{A} + B\, Ln\, X_i + \omega_i$$

يتم تقدير معلمتى الدالة الجديدة بإستخدام البيانات المحولة باللوغاريتم الطبيعي، إذ أن اللوغاريتم يقوم بتخفيض الأوزان الكبيرة (المشاهدات الكبيرة) كي تكون قريبة من المشاهدات ذات الأوزان الصغيرة.

تمرين حول مشكل عدم تجانس التباين

لتكن لدينا البيانات التالية:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y	4	5	6	8	8	6	11	13	15	17	9	13	14	15	21
X	4	4	4	4	4	3.5	3.5	3.5	3.5	3.5	2	2	2	2	2
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
y	6	13	16	23	26	11	15	17	22	34	7	21	23	28	38
X	1.5	1.5	1.5	1.5	1.5	1	1	1	1	1	0.5	0.5	0.5	0.5	0.5

المطلوب:

1- قدر معالم النموذج ؟

2- من خلال الرسم البياني لسحابة النقاط ،هل هناك تجانس لتباين الأخطاء ؟

3- اختبر وجود تجانس تباين الخطأ بإجراء الاختبارات التالية:

أ- اختبار white ؟

ب- اختبار ثبات التباين الشرطي ARCH-LM ؟

4-قم بتصحيح النموذج من مشكل عدم تجانس تباين الأخطاء إن وجد ؟

الحل:

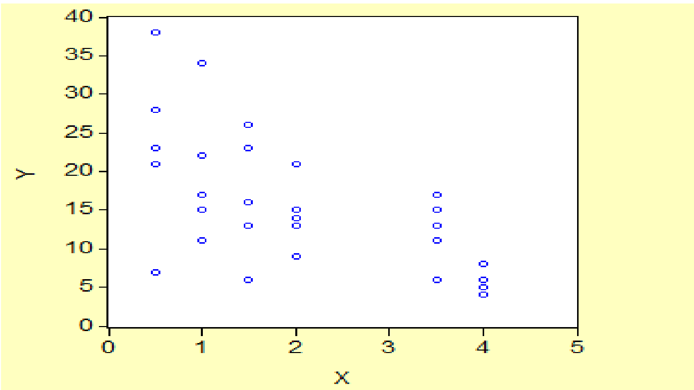
1- تقدير معالم النموذج

$$y = a_0 + a_0 x + \varepsilon_i$$

Dependent Variable: Y
Method: Least Squares
Date: 05/12/14 Time: 20:43
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	24.09442	2.397697	10.04898	0.0000
X	-4.125322	0.982272	-4.199774	0.0002

R-squared	0.386478	Mean dependent var	15.50000
Adjusted R-squared	0.364566	S.D. dependent var	8.585272
S.E. of regression	6.843674	Akaike info criterion	6.748867
Sum squared resid	1311.404	Schwarz criterion	6.842280
Log likelihood	-99.23300	F-statistic	17.63810
Durbin-Watson stat	1.808528	Prob(F-statistic)	0.000245



من الرسم البياني نلاحظ تشتت سحابة النقاط x,y وبالتالي يوجد عدم ثبات للتباين (عدم تجانس تباين الأخطأ)

2-اختبارات الكشف عن وجود تجانس تباين الأخطاء:

أ- اختبار White:

White Heteroskedasticity Test:

F-statistic	3.956388	Probability	0.031126
Obs*R-squared	6.799324	Probability	0.033385

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 05/12/14 Time: 20:46
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	136.0182	43.70370	3.112280	0.0044
X	-78.58151	48.29367	-1.627160	0.1153
X^2	11.98436	10.25804	1.168289	0.2529

R-squared	0.226644	Mean dependent var	43.71348
Adjusted R-squared	0.169359	S.D. dependent var	69.19018
S.E. of regression	63.05961	Akaike info criterion	11.22068
Sum squared resid	107365.9	Schwarz criterion	11.36080
Log likelihood	-165.3102	F-statistic	3.956388
Durbin-Watson stat	1.589880	Prob(F-statistic)	0.031126

نحسب إحصائية مضاعف لغرنج :

LM=n.R²=30x0.226644=6.79932

إيجاد قيمة كي دو المجدولة بدرجة حرية 2K

لدينا K=1 إذا درجة الحرية هي 2

$$\aleph_{\alpha}^{2}(2K) = \aleph_{\alpha}^{2}(2) = 5.991$$

المقارنة :

نلاحظ ان فيمة LM =6.79932 أكبر من قيمة كي دو بدرجة حرية 2 التي هي 5,991 وبالتالي نرفض الفرضية H0 ونقبل الفرضية H1 القائلة بعدم تجانس تباين الأخطاء

ب- اختبار ثبات التباين الشرطي ARCH-LM:

ARCH Test:

F-statistic	1.887374	Probability	0.180798
Obs*R-squared	1.894733	Probability	0.168669

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 05/12/14 Time: 20:49
Sample (adjusted): 2 30
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	33.41826	15.26088	2.189800	0.0374
RESID^2(-1)	0.311774	0.226940	1.373817	0.1808

R-squared	0.065336	Mean dependent var	44.77564
Adjusted R-squared	0.030718	S.D. dependent var	70.16552
S.E. of regression	69.07943	Akaike info criterion	11.37486
Sum squared resid	128843.1	Schwarz criterion	11.46916
Log likelihood	-162.9355	F-statistic	1.887374
Durbin-Watson stat	1.454612	Prob(F-statistic)	0.180798

نحسب إحصائية مضاعف لغرنج :

LM=(n-q).R²=(30-1)x0.065336=1.8947

إيجاد قيمة كي دو المجدولة بدرجة حرية q

لدينا q=1 إذا درجة الحرية هي 1

$$\aleph_{\alpha}^{2}(2K) = \aleph_{\alpha}^{2}(1) = 3.841$$

المقارنة :

نلاحظ ان فيمة LM =1.8947 أكبر من قيمة كي دو بدرجة حرية 1 التي هي 3.841 وبالتالي نرفض الفرضية H1 ونقبل الفرضية H0 التي تدل على وجود تجانس تباين الأخطاء

ARCH Test:

F-statistic	1.887374	Probability	0.180798
Obs*R-squared	1.894733	Probability	0.168669

2-تصحيح النموذج:

نقوم بحساب تباين الأخطاء

$$\sigma_\varepsilon^2 = \frac{\sum \hat{\varepsilon}_i^2}{n-k-1} = \frac{1311.404}{30-1-1} = 46.835$$

ثم نقوم بتحويل المتغيرات :

$$a_{01} = \frac{a_0}{\sqrt{\sigma_\varepsilon^2}} = \frac{a_0}{46.835}$$

$$Y1 = \frac{Y}{\sqrt{\sigma_\varepsilon^2}} = \frac{Y}{46.835}$$

$$X1 = \frac{X}{\sqrt{\sigma_\varepsilon^2}} = \frac{X}{46.835}$$

ثم نقوم بتقدير النموذج المصحح:

Dependent Variable: Y1
Method: Least Squares
Date: 05/15/14 Time: 06:59
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.520718	0.350356	10.04898	0.0000
X1	-4.125322	0.982272	-4.199774	0.0002

R-squared	0.386478	Mean dependent var	2.264886
Adjusted R-squared	0.364566	S.D. dependent var	1.254495
S.E. of regression	1.000009	Akaike info criterion	2.902236
Sum squared resid	28.00052	Schwarz criterion	2.995649
Log likelihood	-41.53354	F-statistic	17.63810
Durbin-Watson stat	1.808528	Prob(F-statistic)	0.000245

قائمة المراجع

أولا: المراجع باللغة العربية

1- عبد القادر محمد عبد القادر عطية، **الحديث في الاقتصاد القياسي بين النظرية و التطبيق،** الإسكندرية، الدار الجامعية،2004.

2-تومي صالح، **مدخل لنظرية القياس الاقتصادي،** الطبعة الثانية، الجزء الأول ، ديوان المطبوعات الجامعية، الجزائر، 2011.

3- وليد إسماعيل السيفو، فيصل مفتاح شلوف، صائب جواد إبراهيم جواد، **أساسيات الاقتصاد القياسي التحليلي،** الطبعة الأولى، دار الاهلية، الأردن، 2006.

4-إسماعيل بن قانة، **الإحصاء الوصفي الحيوي "دروس وتطبيقات"**، الطبعة الأولى، دار أسامة، الأردن،2011.

5- - دومينيك سالفاتور، ترجمة سعدية حافظ منتصر، **الإحصاء والاقتصاد القياسي،** سلسلة ملخصات شوم، الطبعة السادسة ، الدار الدولية للإستثمارات الثقافية، 2012.

6- حسين علي بخيت، سحر فتح الله، **الاقتصاد القياسي،** الطبعة الأولى، دار اليازوري، الأردن، 2009.

7- شيخي محمد، **طرق الاقتصاد القياسي،** الطبعة الأولى، دار الحامد، الأردن، 2011.

8-مكيد علي، **الاقتصاد القياسي دروس وتمارين محلولة،** الطبعة الثانية، ديوان المطبوعات الجامعية، الجزائر،2011.

9-فروخي جمال: **نظرية الاقتصاد القياسي،** ديوان المطبوعات الجامعية، الجزائر 1993.

10- خالد محمد السواعي، **مدخل إلى القياس الاقتصادي،** الطبعة الأولى، الدار العربية للعلوم، لبنان، 2015.

11- عبدالرزاق بني هاني، **الاقتصاد القياسي نظرية الانحدار البسيط والمتعدد،** الطبعة الأولى، دار وائل، الأردن،2014.

12- عدنان داود العذاري، صادق على الجبوري، **الاقتصاد القياسي نظرية وحلول تطبيق باستخدام Minitab, Relase 14**، الطبعة الأولى، دار جرير، الأردن، 2010.

13-تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الجزء الثاني، الطبعة الثانية، ديوان المطبوعات الجامعية، الجزائر،2010.

14-حمداوي الطاوس، **مدخل للاقتصاد القياسي**، الطبعة الأولى، دار هومه، 2016، الجزائر.

15-عبدالحميد عبدالمجيد البلداوي، الأساليب التطبيقية لتحليل وإعداد البحوث العلمية مع حالات دراسية باستخدام برنامج SPSS ، دار الشروق للنشر والتوزيع، عمان الأردن، 2008.

16-جلاطو جيلالي، **الإحصاء مع تمارين ومسائل محلولة**، الطبعة الخامسة، ديوان المطبوعات الجامعية، الجزائر، 2005.

المراجع باللغة الأجنبية :

17- Régis Bourbonnais**, Econométrie**, 9e édition, Dunod, France, 2015.

18- Anderson,Sweeney, Williams,Camm, Cochran, traduction de la 7eédition américaine par Claire Borsenberger, **Statistiques pour l'economie et la gestion**, 5e édition, De Boeck supérieur, Paris,2015,