



الجمهورية الجزائرية الديمقراطية الشعبية



وزارة التعليم العالي والبحث العلمي



المركز الجامعي بلحاج بوشعيب عين تموشنت



معهد العلوم الإقتصادية التسيير والعلوم التجارية

محاضرات في مقياس :

الاقتصاد القياسي 1

من إعداد الدكتور:

بن عامر عبد الكريم

موجهة لطلبة السنة الأولى ماستر: تخصص إدارة واقتصاد المؤسسة

السنة الجامعية 2019-2020

الصفحة	العنوان
I	الفهرس
1	مقدمة
2	I- الإطار النظري للاقتصاد القياسي
2	I-1- مفهوم الاقتصاد القياسي
2	I-2- التطور التاريخي لعلم الاقتصاد القياسي
5	I-3- المكونات العلمية للاقتصاد القياسي
6	I-4- أهداف الاقتصاد القياسي
7	I-5- منهجية الاقتصاد القياسي
9	I-6- الاقتصاد القياسي وجائزة نوبل
11	II- نظرية الارتباط البسيط
11	II-1- لوحة الانتشار
11	II-2- معامل الارتباط الخطي البسيط
15	II-3- معامل الارتباط المتعدد
17	II-4- معامل الارتباط للرتب
19	II-5- الارتباط النوعي
23	III- الانحدار الخطي البسيط
23	III-1- مفهوم الانحدار الخطي البسيط
23	III-2- استخدامات الانحدار الخطي البسيط
23	III-3- معادلة الانحدار الخطي البسيط
25	III-4- نموذج الانحدار الخطي البسيط وفرضياته
26	III-5- تقدير معاملات نموذج الانحدار الخطي البسيط
30	III-6- مؤشرات حساب جودة النموذج
32	III-7- الاختبارات الخاصة بملائمة النموذج وكذا معاملات الانحدار
40	IV- الانحدار الخطي المتعدد
40	IV-1- مفهوم الانحدار الخطي المتعدد
40	IV-2- تقديم نموذج الانحدار الخطي المتعدد

الفهرس:

40	3-IV- الشكل المصفوفي للنموذج
42	4-IV- فرضيات النموذج
42	5-IV- تقدير معلمات النموذج
46	6-IV- القدرة التفسيرية للنموذج
47	7-IV- إختبار صلاحية النموذج
56	V- مشاكل القياس
56	1-V- الارتباط الذاتي للأخطاء
61	2-V- عدم تجانس التباين
65	3-V- التعدد الخطي (Multicolinearty)
71	VI- السلاسل الزمنية
71	1-VI- تعريف السلسلة الزمنية
71	2-VI- أنواع السلاسل الزمنية
72	3-VI- أهداف تحليل السلاسل الزمنية
73	4-VI- مكونات السلسلة الزمنية
77	5-VI- خصائص السلسلة الزمنية
79	6-VI- نماذج تحليل السلسلة الزمنية
80	7-VI- كيفية الكشف عن مكونات السلسلة الزمنية
83	8-VI- إستقرارية السلسلة الزمنية
86	9-VI- اختبارات الاستقرارية (اختبارات جذر الوحدة)
94	10-VI- النماذج الخطية للسلسلة الزمنية
107	الخاتمة
108	المراجع

مقدمة عامة:

يعتبر الاقتصاد القياسي حالياً أحد أكثر فروع البحث الاقتصادي نشاطاً. بعد أن أصبح أداة لا غنى عنها والتي بدونها لا يمكن للاقتصاد أن يدعي مكانة علمية يمكن مقارنتها بمكانة العلوم التجريبية الرسمية التي يطمح إليها مثل الفيزياء وغيرها، وذلك كونه يتوقع مستقبل البحث الاقتصادي من خلال الكشف عن مشاكل جديدة. التي تتطلب تطوير المفاهيم المناسبة. ومن المفارقات أن ظهور الاقتصاد القياسي كنظام مستقل حديث جداً. ومع ذلك فإنه اليوم يستفيد استفادة كاملة من المعرفة الإحصائية مع مزايا وعيوب ذلك. على وجه الخصوص، نجد في الاقتصاد القياسي انعكاساً للوضع المعرفي المثير المبني على التوقعات.

وعلى الرغم من الجهود العديدة المبذولة من قبل الاقتصاديين لمنح علم الاقتصاد مكانة ضمن العلوم "الصعبة"، خاصة الفيزياء. إلا أنه يختلف بشكل واضح عن هذا الأخير لعدد من الأسباب. منها عدم وجود توحيداً تدرجياً للنظريات المختلفة ضمن أطر أكثر عمومية، بل نوعاً من التعارض الذي يتجلى في التمايز المفرط لمجالات التطبيق (اقتصاديات العمل، النقل، الصحة، الأسواق المالية، ...) ، مما يكشف عن خصوصية غير قابلة للاختزال على ما يبدو في كل مجال من هذه المجالات، كما لو كان مجالاً مستقلاً، على حساب نظرية أكثر عمومية. في الوقت نفسه، نرى ظهور نظريات متنافسة تهدف إلى شرح نفس الظواهر (نظريات البطالة، الدولة، العدالة الاجتماعية، ...) بطريقة متناقضة في كثير من الأحيان. تعكس هذه الخصائص الاقتصادية للاقتصاد عدم وجود توافق في الآراء بشأن الأسس المعرفية للمعرفة الاقتصادية.

ونظراً للأهمية المعترفة التي أولي بها مقياس الاقتصاد القياسي كمادة علمية في العلوم المختلفة جاءت

هذه المطبوعة لتسلط الضوء على هذا المقياس، فهي موجهة أساساً إلى طلبة الجامعات في ميدان العلوم الاقتصادية وعلوم التسيير والعلوم التجارية والمالية في دراسات الطور الجامعي، كما تعتبر أداة مهمة وفعالة للباحثين في مجال صناعة القرارات على المستوى الكلي أو الجزئي لتزويد معارفهم وتنمية كفاءاتهم في كيفية حل المشكلات الإدارية التي تواجههم في الحياة العملية.

I- الإطار النظري للاقتصاد القياسيI-1- مفهوم الاقتصاد القياسي:

إن مصطلح "اقتصاد قياسي" (économetrie) يحتوي على كلمة "اقتصاد" (économie) وهي جذر هذا المصطلح وذلك لأن ميدان استعماله الأساسي هو معالجة الظواهر الاقتصادية، الجزء الآخر من هذا المصطلح هو "القياسي" وتعني الحساب، القياس، أي التقدير الكمي للأشياء¹.

الاقتصاد القياسي هو نوع خاص من التحليل الاقتصادي الذي تمزج فيه النظرية الاقتصادية بعد صياغتها صياغة رياضية مع القياس العملي للظواهر الاقتصادية عن طريق الأساليب الإحصائية². كما عرفه سمويلسون (Samuelson) بأنه فرع من علوم الاقتصاد يبحث في التحليل الكمي للظواهر الاقتصادية الحقيقية مستعينا بتطور بالنظرية الاقتصادية والطرق الإحصائية³.

I-2- التطور التاريخي لعلم الاقتصاد القياسي:I-2-1- مرحلة الدراسات التجريبية:

تعود الدراسات التجريبية الأولى الخاصة بعلم الاقتصاد القياسي إلى القرنين السابع عشر والثامن عشر، ما يعرف بسلطة القانون الطبيعي في ذلك الوقت، والتي ظهرت تدريجياً من سلطة الدين وسلطة الأمير. حيث تم الاعتماد في ألمانيا على الإحصاء الوصفي المنبثق من الوصف الكلي للدول. وهنا يرى هرمان كنينق (Hermann CONRING 1681-1606) أن الإحصاءات هي وسيلة فعالة لتصنيف المعارف. أما في إنجلترا فتم اعتماد الحساب السياسي (فحص سجلات الزواج والوفيات). لبناء

¹ مكيد على، الاقتصاد القياسي دروس ومسائل محلولة، ديوان المطبوعات الجامعية، الطبعة الثانية، 2011، ص.09.

² محمد عبد السميع عناني، التحليل القياسي والإحصائي للعلاقات الاقتصادية، مدخل حديث باستخدام spss، الدار الجامعية، الاسكندرية، 2009، ص.6.

³ P.A.Samualson .T.C.Copmans, « Report Of Evaluative Committee For Econometrica , Econometric »,Vol 02, N 02,1954, Pp 141-146.

جداول الوفيات ومن ثم حساب متوسط العمر المتوقع. كما ساهم العديد من الباحثين القدامى في التطور التاريخي لعلم الاقتصاد القياسي نذكر منهم⁴:

- William PETTY: ساهم في تطوير الإحصاء الديموغرافي.

- Gregory KING: أول من رسم قانون الطلب

في القرن التاسع عشر تم تطوير الإحصاء الرياضي والتطور الاحتمالي لمجالات متعددة مثل: علم الفلك وعلم الاجتماع وعلم الأحياء والفيزياء.

I-2-2- مرحلة الاقتصاد القياسي⁵:

ظهر الاقتصاد القياسي في منتصف القرن التاسع عشر، مع تطور الاقتصاد الرياضي لأوغست كورنو (Auguste COURNOT) وليون والراس (Léon WALRAS). كما أن أول من ربط علاقة الانحدار بين متغيرين هو غالتون (GALTON). وفي الوقت نفسه، حدد فرانسيس إيدجورث (Francis EDGEWORTH) صيغة دالة الكثافة الطبيعية. كما قام كارل بيرسون (Karl PEARSON) بتطوير المعامل المتعدد والذي يظهر الارتباط مع معاملات الارتباط البسيطة، كما أن إيفلين هوكر (Evelyn HOCKER) هو أول من درس المتغيرات المتأخرة في نماذج الانحدار، مارسيل لينوير (Marcel LENOIR) هو أول من حاول شرح قوانين العرض والطلب. من خلال مقارنة تبين تداخل الاقتصاد الرياضي والإحصاء الوصفي والإحصاء الرياضي.

بدأ الاقتصاد القياسي "الحديث" بالفعل بتحليل سوق العمل الأمريكية: حين اهتم هنري مور (Henry MOORE) بمشكلة تحديد الأجور ووظائف الطلب ودورات التواتر، تحاول دراسته إثبات أن الرياضيات والإحصاء يمكن أن تكون بمثابة كاشف تجريبي وتسمح بتفسير ملموس للظواهر الاقتصادية.

في عام 1920، ظهر أول معهد اقتصادي. في الوقت نفسه تم إنشاء المكتب الوطني للبحوث الاقتصادية (NBER). والتي يعتمد دورها على البحوث الاقتصادية التجريبية. ومع بداية القرن العشرين

⁴ Françoise seyte, introduction a l'économétrie, cours magistrale, université Montpellier, 2011-2012, P06.

⁵ Françoise seyte, op cité, P07.

اتخذ التحليل الاقتصادي بعداً جديداً ساعد الاقتصاديون ورجال الأعمال والمهندسون على ربط الاقتصاد والرياضيات والإحصاء. وأصبح الاقتصاد معترفاً به كفرع في حد ذاته. في 29 ديسمبر 1930 أخذ الاقتصاد منحى آخر مع جان تينبرجن (Jan TINBERGEN) وراجنر فريش (Ragnar FRISCH). من خلال جمعية الاقتصاد القياسي. عقد المؤتمر الأول عام 1931 في لوزان. في نفس الوقت، أجرى ألفريد كولس (Alfred COWLES)، المستشار المالي والمتخصص في التنبؤ بسوق الأوراق المالية، اتصالات مع جمعية الاقتصاد القياسي واقترح مشروعين:

- تمويل نشر مجلة الاقتصاد القياسي
- تمويل منظمة بحثية تحت رعايته

في ثلاثينيات القرن العشرين، تم إنشاء مجلة *Econometrica*، وصدر العدد الأول في يناير 1933، وكان محررها في ذلك الوقت هو راجنر فريش (Ragnar FRISCH). وصدر أول نموذج للاقتصاد الكلي بدوره ويتضمن مبادئ احتمالية لورانس كلاين (Lawrence KLEIN) في السنوات 1950-1960. وشهد هذا العصر تطوير نماذج تأخير متدرجة من كويك (KOYCK) في عام 1954، وفي الوقت نفسه بدأت طرق التنبؤ قصيرة المدى في التطور، بما في ذلك نماذج BOX وJENKINS التي تقوم بتقدير عمليات أحادية المتغيرات للتنبؤ.

I-2-3- أزمة نمذجة الاقتصاد الكلي في التطور الحديث للاقتصاد القياسي:

ساهمت أزمة نمذجة متغيرات الاقتصاد الكلي في تطور علم الاقتصاد القياسي كما يلي⁶:

أولاً: في سنوات السبعينات:

شهدت سبعينيات القرن الماضي نهاية العصر الذهبي للاقتصاد الكلي، وتعارضت صدمة النفط مع جميع نماذج الاقتصاد الكلي السابقة. ومع ذلك، فإن هذا يفتح الطريق لنماذج جديدة، ولا سيما نماذج

⁶ Françoise seyte, PP 07-08.

شعاع الانحدار الذاتي (VAR)، مع تطوير في نفس الوقت تحليل السببية (Causality). analyses de la

ثانياً: الاقتصاد القياسي للسلسلة الزمنية منذ سنة 1980:

ظهرت اختبارات جذر وحدة لـ DICKEY-FULLER ونماذج ARMA ونمذجة ARCH ونماذج LOGIT-PROBIT ونماذج BOOSTRAP بالإضافة إلى نماذج BAYESIAN في الثمانينيات. حيث أصبح الاقتصاد حاليًا يتكون من عدة فروع ذات محتوى عالي جدًا.

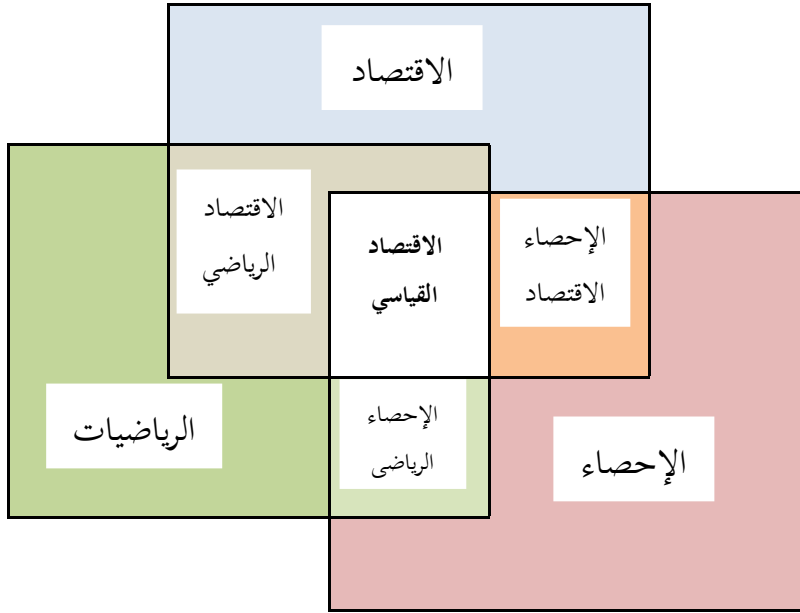
- الاقتصاد القياسي للسلسلة الزمنية (Séries temporelles)
- الاقتصاد القياسي للمتغيرات النوعية (variables qualitative)
- الاقتصاد القياسي لبيانات بانل (données de panel)

I-3- المكونات العلمية للاقتصاد القياسي : يتكون الاقتصاد القياسي من تفاعل ثلاثة عوم أساسية

وهي:

- ✓ علم الاقتصاد: الذي يدرس العلاقة التي تربط بين المتغيرات الاقتصادية
- ✓ علم الرياضيات: الذي يهتم بالتعبير عن هذه العلاقات في شكل رياضي، أي تحويل العلاقات التي تربط بين هذه المتغيرات إلى معادلات رياضية
- ✓ علم الإحصاء: والذي يتمثل دوره في تجهيز الاقتصاد القياسي بأدوات تحليلية تستخدم في دراسة العلاقة بين المتغيرات الاقتصادية وبطرق خاصة في معالجة الأخطاء في قياس تلك المتغيرات.

ويمكن أن نوضح العلاقة بين الموجود بين العلوم الثلاثة في الشكل التالي:



المصدر: محمد عبد السميع عناني، نفس المرجع، ص.14.

ويرى العالم Maurice Allais الحائز على جائزة نوبل في علم الاقتصاد سنة 1988 ، بأن غاية القياس الاقتصادي هي دراسة الوسائل الاقتصادية على المستويين النظري والتطبيقي بنفس منطق البناء المطبق في العلوم الفيزيائية وباستعمال نفس الطرق الكمية من رياضيات وإحصاء سواء كان ذلك على المستوى النظري أو على المستوى التطبيقي.

I-4- أهداف الاقتصاد القياسي:

يمكن أن نجمع أهداف الاقتصاد القياسي في ما يلي:

- تحليل واختبار النظرية الاقتصادية التي تضبط العلاقة الموجودة بين مختلف المتغيرات الاقتصادية، ووضع قوانين تحدد هذه العلاقة، وذلك من خلال اختبار فرضياتها وقدرتها التفسيرية.
- المساعدة على وضع السياسات الاقتصادية من خلال تحويل العلاقة التي تربط بين مختلف المتغيرات الاقتصادية إلى نماذج مقدرة المعالم الشيء الذي يساعد صناع القرار على رسم سياساتهم واتخاذ أفضل التدابير.
- التنبؤ بقيم مختلف المتغيرات الاقتصادية، وذلك بعد استنتاج السلوك الذي سوف تأخذه الظواهر الاقتصادية في المستقبل.

I-5-5- منهجية الاقتصاد القياسي:

لدراسة العلاقة بين ظاهرتين أو مجموعة من الظواهر الاقتصادية نتبع جملة الخطوات التالية⁷:

I-5-1- تكوين النموذج القياسي (نموذج الانحدار)

نعني بتشكيل النموذج القياسي تكوين الصيغة الرياضية (الكمية) للمشكلة المدروسة التي تربط بين الظاهرة المدروسة والعوامل التي تتحكم فيها وتحديد طبيعة هذه الصيغة. ينطلق تكوين النموذج من الفرضيات التي توفرها النظرية الاقتصادية عن العناصر التي تتحكم في الظاهرة وعلاقات التأثير فيما بينها (المستقل والتابع). نعبر عن هذه العلاقات النظرية في شكل علاقة دالية عامة أو معادلة أو مجموعة معادلات $y = f(x_1, x_2, \dots, x_n)$ بحيث نرمز لكل عنصر أو عامل بالمتغير.

من أجل تحديد نوع العلاقة الرياضية التي تجمع بين المتغير التابع والمتغير (المتغيرات) المستقل، أي من أجل تحديد أنسب الصيغ الرياضية التي تعبر عن هذه العلاقة بين المتغيرات تعبيراً دقيقاً (خطية أو غير خطية، بسيطة أم متعددة) يجب إتباع الخطوات التالية:

1-1- إجراء تحليل تمهيدي: نتناول في هذا التحليل قبل كل شيء تعريف الظاهرة الناتجة أو المفسرة (Y) ثم تحديد العوامل المؤثرة فيها (الظواهر المسببة أو المفسرة). يجب كذلك تحديد وحدات القياس التي نقيس بها تلك الظاهرة وكل من العوامل المؤثرة فيها على حدة.

1-2- التأكد من أن هناك علاقة جدلية واضحة بين الظاهرة المدروسة والعوامل المؤثرة فيها.

1-3- جمع المعلومات الأولية: يبدأ جمع المعطيات عن كل العوامل المرتبطة بالظاهرة المدروسة بواسطة إجراء القياسات اللازمة لكل منهم، مع مراعاة أن تكون القيم متقابلة مع بعضها البعض من حيث المكان والزمان ثم نقوم بترتيب هذه المعلومات في جدول خاص.

1-4- التعرف على الشكل البياني الحقيقي للعلاقة محل الدراسة ويتم ذلك بواسطة الشكل البياني للمعطيات المتعلقة بالمتغير التابع وكل متغير مستقل على حدة هذا الرسم يسمح إذن بتحديد الشكل البياني العام للعلاقة المدروسة بين مؤشرين (X) و (Y) وذلك من خلال تحديد الاتجاه العام لشكل انتشار

⁷ مكيد على، مرجع سبق ذكره، ص. 11-12.

النقط الخاصة بمهذين المؤشرين. للقيام بعملية الرسم البياني للعلاقة بين هذين المؤشرين نرسم محورين إحداثيين متعامدين ومتقاطعين في النقطة (0 المبدأ) نرسم للمحور الأفقي بـ X وتمثل عليه قيم الظاهرة المفسرة (X). ونرسم للمحور العمودي بـ Y ويمثل عليه قيم الظاهرة المدروسة (Y). بعد ذلك نضع القيم المتقابلة على شكل أزواج مرتبة كما يلي: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ ، نقوم بعدها برسم النقط الهندسية على المستوي (X, Y) والتي إحداثياتها تساوي الأزواج المرتبة السابقة فنحصل على جملة من هذه النقاط موزعة على مستوي (X, Y)، وتأخذ شكلا معيناً يسمى شكل الانتشار. نلاحظ أن شكل الانتشار لا يعطينا بوضوح ودقة شكل خط أو منحنى معروف لدينا إلا أننا نلاحظ أن هناك اتجاهها عاما للعلاقة (X, Y) يمكن أن نمثله بخط مستقيم أو منحنى أي آخر ملائم، وذلك بغض النظر عن بعض النقط (القيم) الشاذة التي تكون ناتجة عن أسباب أخرى (تأثير عوامل أخرى) غير المتغير المستقل (X).

1-5-5- اختيار أنسب الصيغ الرياضية التي تتلاءم مع الشكل البياني الحقيقي للعلاقة محل الدراسة وتمثيله بصفة عامة يمكن أن نميز الحالات التالية:

1-5-1- الاتجاه العام لشكل الانتشار يكون في شكل خط مستقيم أو قريبا منه، وبالتالي فالعلاقة بين (X, Y) هي علاقة خطية (تكون ممثلة بخط مستقيم). يمكننا أن ننشأ خط مستقيم يمر بين هذه النقاط ويتوسطها في آن واحد وذلك حسب شكل الاتجاه العام لشكل الانتشار.

1-5-2- الاتجاه العام لشكل الانتشار هو في شكل غير خطي (شكل منحنى) أو قريب منه، وبالتالي فالعلاقة بين المؤشرات محل الدراسة هي علاقة غير خطية. إن العلاقة غير الخطية بين عناصر الظاهرة المدروسة يمكن أن يكون في شكل معادلة من الدرجة الثانية (منحنى ذو نهاية واحدة عظمى أو صغرى). إذا كان المنحنى ذو نهايتين واحدة صغرى والأخرى عظمى فالعلاقة تكون في شكل معادلة من الدرجة الثالثة. إذا كان المنحنى يتقارب في اللانهاية من خط مستقيم ما فإن المنحنى يجب أن يكون في شكل قطع زائد، إلخ.

إن طريقة الرسم البياني تساعد كثيرا في الكشف عن نوعية العلاقة التي تربط بين المؤشرين (X) و(Y)، والصيغة الرياضية العامة التي تربطهما. لكن تصبح قاصرة عندما نريد دراسة العلاقة بين أكثر من ظاهرتين كما سنرى فيما بعد في نموذج الانحدار المتعدد.

I-5-2- تقدير النموذج القياسي (Estimation de modèle économétrique):

إن تقدير النموذج القياسي المشكل في المرحلة السابقة، يعني محاولة الوصول إلى تقديرات كمية لمعاملاته (أي معاملات المعادلة أو الدالة المقترحة). اختيار نوع المعادلة بدون تحديد الثوابت التي تتضمنها لا يفيدنا قطعاً في دراسة هذه العلاقة. لذلك كان لابد من القيام بإجراء حسابات معينة لتحديد ثوابت المعادلة المختارة. إن أي ثابت في المعادلة (الدالة) المختارة يمكن أن يأخذ قيمة متعددة وغير منتهية وبالتالي سنكون أمام لا نهاية من المعادلات التي لها نفس النوع المختار سابقاً. لذل فإن من الطبيعي جداً عند تحديد قيم الثوابت أن نبحث عن قيمة لكل منها بحيث نحصل على معادلة معينة (ذات ثوابت محددة) وتمثل العلاقة المدروسة أفضل تمثيل. إن عملية التقدير هذه تعني البحث عن تلك القيم بواسطة طريقة المربعات الصغرى.

I-5-3- تقييم النموذج القياسي (Evaluation du modèle économétrique):

قبل استخدام النموذج القياسي المقدر يجب التأكد من جودة تقدير هذا النموذج. يتم هذا التقدير من خلال بعض الاختبارات.

I-6- الاقتصاد القياسي وجائزة نوبل:

تم تكريم العديد من الباحثين في مجال الاقتصاد القياسي بجائزة نوبل وذلك من خلال المقالات المكتوبة في هذا الموضوع ومن بين الفائزين بهذه الجائزة نذكر⁸:

1- 1980: لورنس كلاين (Lawrence Klein) (أول نماذج الاقتصاد القياسي الكلي، تحليل التقلبات والنمو)

2- 1989: تريجف هافيلمو (Trygve Haavelmo) (مقدمة للنهج الاحتمالي، النماذج ذات المعادلات المتزامنة).

3- 2000: جيمس هيكممان (James Heckman) (نظريات وطرق تحليل العينة)

⁸ FRANÇOISE SEYTE, cours Introduction à l'économétrie, faculté d'économie, université Montpellier 1, 2011-2012, P. 05.

4- 2000: دانيال ماكفدن (Daniel Mcfadden) (تحليل الخيارات المنفصلة في الاقتصاد القياسي)

5- 2003: كليف أنجل (Clive Engle) وروبرت غرانجر (Robert Granger) (التكامل المشترك وتقلب السلاسل الزمنية ونماذج ARCH)

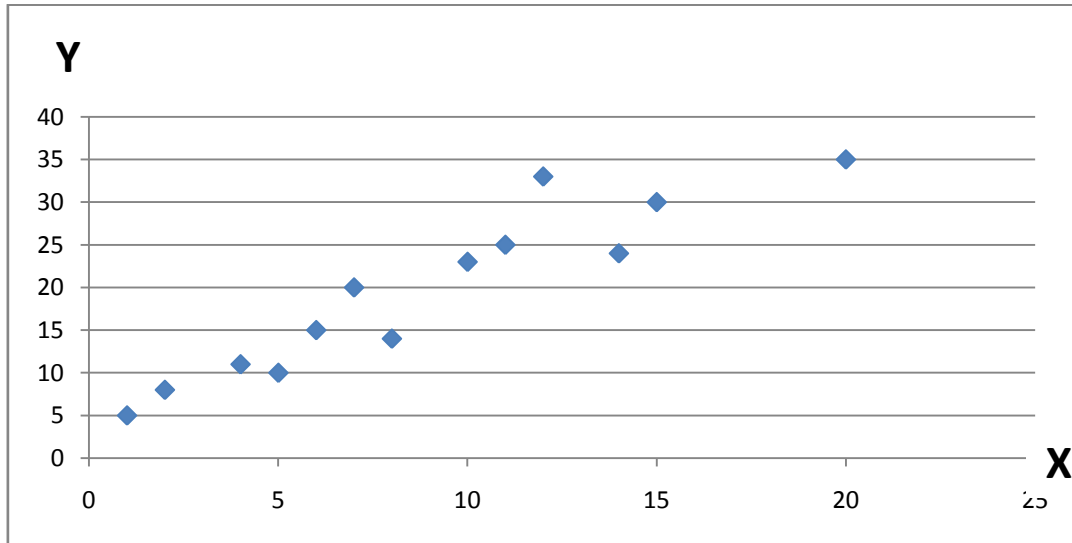
II-1- نظرية الارتباط البسيط:

تقوم نظرية الارتباط البسيط على دراسة العلاقة بين متغيرين أحدهما مستقل والآخر تابع، فمثلا قد نقول هل هناك علاقة بين الوزن وضغط الدم لدى الرجال؟ أو هل هناك علاقة بين دخل الفرد الشهري وقيمة الادخار؟ وإن وجدت علاقة فهل هي خطية أم غير خطية؟ فوجود علاقة يعني أن أحد المتغيرين مرتبط بالآخر. فموضوع الارتباط هو في كيفية قياس قوة العلاقة بين المتغيرات.

II-1-1- لوحة الانتشار:

إن أول طريقة تساعدنا على معرفة وجود علاقة بين متغيرين X و Y هي لوحة الانتشار التي تمثل شكل هندسي على معلم متعامد ومتجانس محوره الأفقي يمثل المتغير المستقل (المفسر) ومحوره العمودي المتغير التابع (المفسر)، وتبين لوحة الانتشار بشكل جيد ما إذا كان هناك علاقة بين المتغيرين X و Y أو عدم وجودها

الشكل: لوحة الانتشار

II-2- معامل الارتباط الخطي البسيط:II-2-1- معامل الارتباط البسيط بيرسون:

يستخدم لقياس العلاقة بين متغيرين ذات قيم رقمية (quantitative)، ويعتبر معامل ارتباط بيرسون من أهم الطرق المستخدمة في حالة الارتباط البسيط، ويحسب بالعلاقة التالية.

$$r_{x,y} = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

حيث:

$\text{Cov}(x, y)$: التباين المشترك بين x و y .

$\sigma_x \sigma_y$: الانحراف المعياري ل x والانحراف المعياري ل y .

n : عدد المشاهدات.

بتبسيط العلاقة أعلاه نحصل على العلاقة التالية:

$$r_{x,y} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n (x_i^2) (\sum_{i=1}^n x_i)^2} \sqrt{n \sum_{i=1}^n (y_i^2) (\sum_{i=1}^n y_i)^2}}$$

نستعمل x بدل x_i و y بدل y_i و $\sum x$ بدل $\sum_{i=1}^n x$ نحصل على العلاقة التالية:

$$r_{x,y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2} \sqrt{\sum y^2 - n\bar{y}^2}}$$

II-2-2- تفسير معامل الارتباط البسيط:

يمكن إثبات رياضياً أن معامل الارتباط يتراوح بين -1 و 1

✓ عندما يقترب $r_{x,y}$ من 1 المتغيرات ترتبط ارتباطاً إيجابياً.

✓ عندما يقترب $r_{x,y}$ من -1 المتغيرات ترتبط ارتباطاً سلبياً أو عكسياً.

✓ عندما يقترب $r_{x,y}$ من 0 ليس هنا ارتباط.

II-2-3- دلالة معامل الارتباط الخطي البسيط (person):

يمكن اختبار ما إذا كان معامل الارتباط له دلالة على وجود علاقة خطية بين المتغيرين باستعمال

الإحصائية t (student) للفرضية $H_0: P = 0$ مقابل الفرضية $H_0: P \neq 0$ بدرجة حرية n-2. قيمة الإحصائية المقاسية هي:

$$t_{cal} = \frac{r_{x,y}}{\sqrt{(1 - r^2)/(n - 2)}}$$

ملاحظة: وجود ارتباط خطي كامل بين متغيران لا يعني بالضرورة السببية لأنه يمكن أن نرجع هذا الارتباط إلى متغير ثالث يؤثر عليهما فيرتفعان أو ينخفضان معا.

مثال: الجدول التالي يمثل القياسات الخاصة بمتغيرين

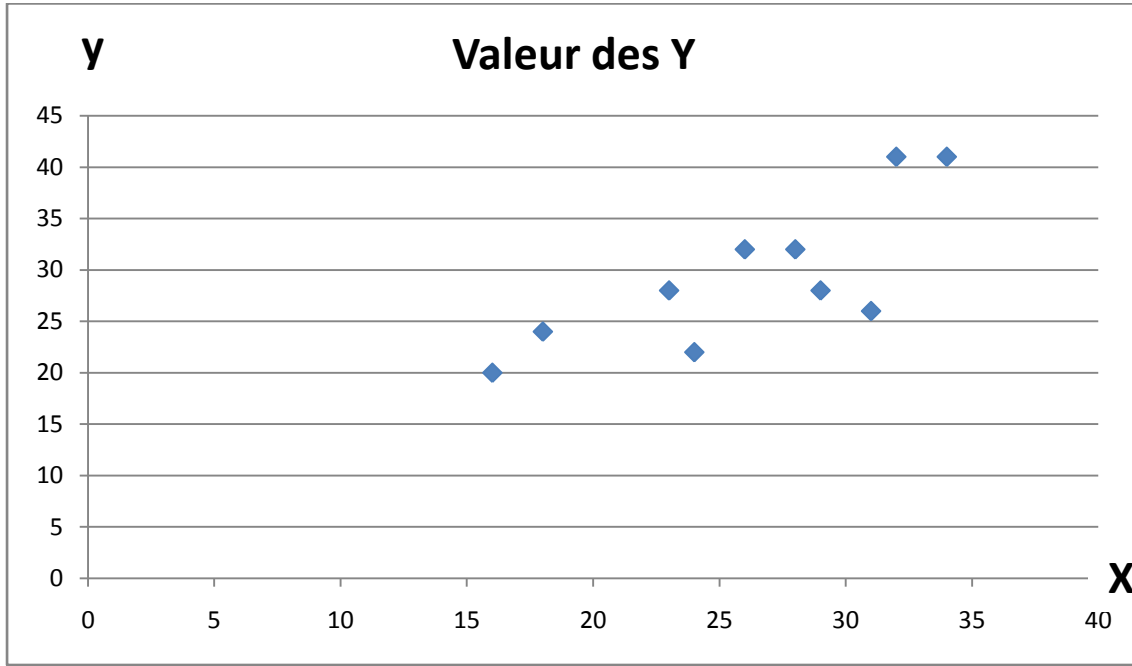
الأداء	16	18	23	24	28	29	26	31	32	34
كمية الطاقة	20	24	28	22	32	28	32	26	41	41

المطلوب:

- 1- رسم لوحة الانتشار؟
- 2- تحديد ما نوع العلاقة بين المتغيرين الأداء و كمية الطاقة ؟
- 3- حساب معامل الارتباط البسيط (person) ؟
- 4- اختبار دلالة معامل الارتباط عند مستوى الدلالة 5%؟

الحل:

1 - شكل الانتشار:



الشكل أدناه يبين العلاقة بين كمية الطاقة والأداء ويتضح من الشكل أنها علاقة خطية موجبة.

2 - حساب معامل الارتباط

	x	y			xy
	16	20	256	400	320
	18	24	324	576	432
	23	28	529	784	644
	24	22	576	484	528
	28	32	784	1024	896
	29	28	841	784	812
	26	32	676	1024	832
	31	36	961	1296	1116
	32	41	1024	1681	1312
	34	41	1156	1681	1394
المجموع	261	304	7127	9734	8286

$$r_{x,y} = \frac{8286 - (10 \times 26.1 \times 30.4)}{\sqrt{7127 - (10 \times 26.1^2)}\sqrt{9734 - (10 \times 30.4^2)}} = 0.89$$

$$r_{x,y} = 0.89$$

اختبار الدلالة لمعامل الارتباط البسيط:

القيمة الجدولية:

$$t_{tab} (n - 2, \alpha/2) = 1.85$$

القيمة الحسابية:

$$t_{cal} \frac{0.89}{\sqrt{(1 - 0.89^2)/(10 - 2)}} = 5.52$$

القيمة الجدولية أكبر من القيمة الحسابية إذن نرفض الفرضية العدمية ونقبل الفرضية البديلة وبالتالي معامل الارتباط يدل على وجود علاقة خطية بين المتغيرين.

II-3- معامِل الارتباط المتعدد:

II-3-1- تعريف معامِل الخطي المتعدد

ويستخدم لقياس العلاقة بين أكثر من متغيرين، إلا أن إشارة معامِل الارتباط هنا لا تدل على اتجاه العلاقة لان هذا الاتجاه لا يكون موحدًا لجميع المتغيرات، وأن عملية التحليل تقوم على فرض أن المتغيرات عشوائية متصلة ويدعى توزيعها بمتعدد المتغيرات، وصيغة حسابه هي امتداد لمعامل الارتباط البسيط، ففي حال 3 متغيرات مثلاً لإيجاد العلاقة بين x_1 و x_2 و x_3 و فإن صيغة الحساب هي⁹:

$$r_{2.13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - (2)r_{21}r_{23}r_{13}}{1 - r_{13}^2}}$$

ولإيجاد العلاقة بين x_1 و x_2 و x_3 و فإن صيغة الحساب هي:

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

⁹ عبد الحميد عبد المجيد البلداوي، أساليب الإحصاء للعلوم الاقتصادية وإدارة الأعمال مع استخدام برنامج SPSS، دار وائل للنشر والتوزيع، الطبعة الأولى، الأردن، 2009، ص. 201.

حيث أن r_{23} , r_{13} , r_{12} هي معاملات ارتباط يتم حسابها صيغة الارتباط البسيط لبيرسون

المبينة أعلاه

مثال: لدينا المعطيات التالية:

$$\begin{aligned} \sum y &= 15, & \sum x_1 &= 33, & \sum x_2 &= 28, \\ \sum y^2 &= 47, & \sum x_1^2 &= 235, & \sum x_2^2 &= 170, \\ \sum yx_1 &= 103, & \sum yx_2 &= 88, & \sum x_1x_2 &= 188, \\ r_{12} &= 0.763 & r_{y1} &= 0.936 & r_{y2} &= 0.931 \end{aligned}$$

المطلوب: حساب معامل الارتباط المتعدد؟

الحل:

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

$$r_{y.12} = \sqrt{\frac{(0.876) + (0.866) - (2)(0.936)(0.931)(0.763)}{1 - (0.613)}}$$

$$r_{y.12} = \sqrt{\frac{(0.876) + (0.866) - (2)(0.936)(0.931)(0.763)}{1 - (0.582)}}$$

$$r_{y.12} = \sqrt{\frac{0.412}{0.418}}$$

$$r_{y.12} = 0.99$$

II-3-2- دلالة معامل الارتباط المتعدد:

يمكن اختبار ما إذا كان معامل الارتباط المتعدد له دلالة على وجود علاقة خطية بين المتغيرات

باستعمال الإحصائية F (fischer) للفرضية $H_0: P_{y,12} = 0$ مقابل الفرضية $H_0: P_{y,12} \neq 0$

بدرجة حرية $n-k-1$ للبسط و k للمقام. قيمة الإحصائية المقاسية هي:

$$f_{cal} = \frac{r_{y,12}^2}{1 - r_{y,12}^2} \cdot \frac{n - k - 1}{k}$$

مثال: في المثال السابق اختبر دلالة معامل الارتباط المتعدد

القيمة الجدولية:

$$f_{tab}(n - k - 1, k, \alpha/2) = 16.04$$

القيمة الحسابية:

$$f_{cal} = \frac{0.98}{1 - 0.98} \cdot \frac{6 - 3 - 1}{3}$$

$$f_{cal} = 32.667$$

قيمة f_{cal} المحسوبة أكبر من قيمة f_{tab} الجدولية ومنه نرفض الفرضية العدمية ونقبل الفرضية

البديلة وبالتالي معامل الارتباط يدل على وجود علاقة خطية بين المتغيرات.

II-4- معامل الارتباط للرتب:

يهتم الباحث أحيانا بإيجاد معامل الارتباط بين رتب البيانات والمشاهدات وليس بين المشاهدات

نفسها. وتبرز أهمية هذا المعامل عندما مع متغيرات في مستوى القياس الرتبي. إن أشهر العلاقات التي

وضعت لهذا الغرض هي العلاقة التي وضعها الإحصائي البريطاني تشارلز سبيرمان عام 1904م بإسم

معامل ارتباط سبيرمان للرتب والتي تعطى بالعلاقة التالية¹⁰:

¹⁰ سالم عيسى بدر، عماد غصاب عبابنة، مبادئ الإحصاء الوصفي والاستدلالي، دار المسيرة للنشر والتوزيع والطباعة، الطبعة الثانية،

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

حيث:

n: هي عدد المشاهدات

d: هي الفرق بين رتبة x و y

ملاحظة: يكون هذا المعامل كفوً إذا كان عدد المشاهدات أقل من 30، وأكثر كفاءة عندما تكون عدد المشاهدات بين 25 و30.

مثال:

الجدول التالي يبين العلاقة بين معدلات الطلبة في مقياسي المحاسبة (X) والرياضيات (Y) لعشرة طلاب:

x	12	9	15	13	5	6	18	14	3	11
y	14	12	16	11	8	10	17	13	9	5

المطلوب:

- أحسب معامل الارتباط للرتب سيبيرومان ؟

الحل:

نقوم بترتيب المعدلات من الأكبر إلى الأصغر

x	y	رتبة x	رتبة y	الفرق بين الرتب d	مربع الفرق d^2
12	14	5	3	2	4
9	12	7	5	2	4
15	16	2	2	0	0
13	11	4	6	-2	4
5	8	9	9	0	0
6	10	8	7	1	1
18	17	1	1	0	0
14	13	3	4	-1	1
3	9	10	8	2	4
11	5	6	10	4	16
المجموع					34

$$r_s = 1 - \frac{6(34)}{10(100 - 1)}$$

$$r_s = 0.80$$

عندما تكون لدينا بيانات متساوية تأخذ نفس الرتبة فإننا نأخذ الوسط الحسابي للرتب

II-5- الارتباط النوعي:

II-5-1- معامل الاقتران:

هو وصف درجة العلاقة بين ظواهر غير كمية لا يمكن التعبير عنها بالأرقام بل بالوصف فقط مثل الارتباط بين الجنسين (الذكر والأنثى)، أو بين الحالة التعليمية لمجموعة من الأفراد، أو دراسة التطعيم بمصل واقعي والإصابة بمرض معين وغير ذلك من الأمثلة ولدراسة هذا النوع من العلاقات نستعمل معامل الاقتران¹¹. ويعرض وفق الجدول التالي لمتغيرين X و Y:

¹¹ وليد إسماعيل السيفو، عيد أحمد أبو بكر، غالب عوض الرفاعي، أساسيات الأساليب الإحصائية للأعمال وتطبيقاتها في العلوم المالية والإدارية والاقتصادية، زمزم ناشرون وموزعون، الطبعة الأولى، 2010، ص. 2016.

	X	
Y	A	B
	C	D

$$T_n = \frac{AD - BC}{AD + BC}$$

- تتراوح قيمة هذا المعامل بين -1 و 1 أي $-1 \leq T_n \leq 1$

مثال:

في دراسة أجريه لمعرفة العلاقة بين الدروس الإضافية ونجاح الطلبة أجريت دراسة على عينة من 400 طالب من الطلبة المقبلين على امتحان البكالوريا فكانت النتائج كما في الجدول التالي:

المجموع	لم يقوموا بدروس إضافية	قاموا بدروس إضافية	
300	80	220	نجحوا في امتحان البكالوريا
100	40	60	لم ينجحوا في امتحان البكالوريا
400	120	280	المجموع

المطلوب:

- حساب معامل الاقتران؟

الحل:

$$T_n = \frac{AD - BC}{AD + BC}$$

$$T_n = \frac{(220)(40) - (80)(60)}{(220)(40) + (80)(60)}$$

$$T_n = 0.29$$

معنى هذا أن هناك علاقة طردية ضعيفة بين الدروس الإضافية ومستوى النجاح في البكالوريا.

II-5-2- معامل ارتباط فاي:

له نفس خصائص معامل الاقتران حيث يتعامل مع متغيرات منفصلة ثنائية التصنيف (dichotomous) تم تطويره من قبل يدني يول (Udny yule) في عام 1912، وهو معطى بالعلاقة التالية:

$$r_f = \frac{AD - BC}{\sqrt{(A + B)(A + C)(A + D)(C + D)}}$$

مثال: في المثال السابق أحسب معامل الارتباط لفاي

الحل:

$$r_f = \frac{(220)(40) - (80)(60)}{\sqrt{(220 + 80)(220 + 60)(220 + 40)(60 + 40)}}$$

$$r_f = 0.08$$

معنى هذا أن هناك علاقة ضعيفة جدا بين الدروس الإضافية ومستوى النجاح في البكالوريا.

II-5-3- معامل التوافق لبيرسون:

يدرس هذا المعامل العلاقة بين الظواهر الوصفية التي تنقسم إلى أكثر من نوعين. حيث لا يساعد معامل الاقتران السالف الذكر بل يستخدم معامل التوافق الذي يصلح أيضا لقياس العلاقة بين ظواهر كمية قابلة للقياس وأخرى وصفية لا يمكن قياسها¹². ويحسب بموجب العلاقة التالية:

$$T_p = \sqrt{\frac{G - 1}{G}}$$

¹² وليد إسماعيل السيفو، عيد أحمد أبو بكر، غالب عوض الرفاعي، نفس المرجع السابق، ص 218.

G: مجموع خارج قسمة مربع كل تكرار في الجدول التكراري على حاصل ضرب مجموع بيانات الصف في مجموع بيانات العمود التي يقع فيها التكرار. حيث نقوم بتربيع كل تكرار وارد في الجدول التكراري. ثم نقوم بقسمته على حاصل ضرب التكرار الكلي الرأسي في التكرار الكلي الأفقي.

مثال:

أحسب درجة إرتباط بين تقدير الطالب في المادة (X) والمادة (Y) من البيانات التالية:

Y \ X	مقبول	جيد	جيد جدا	ممتاز	المجموع
مقبول	10	18	23	14	65
جيد	45	103	52	50	280
جيد جدا	27	50	30	21	128
ممتاز	82	171	135	85	473

الحل:

$$G = \frac{(10)^2}{82 \times 65} + \frac{(45)^2}{82 \times 280} + \frac{(27)^2}{28 \times 128} + \frac{(18)^2}{171 \times 65}$$

$$+ \frac{(103)^2}{171 \times 280} + \frac{(50)^2}{171 \times 128} + \frac{(23)^2}{135 \times 65} + \frac{(52)^2}{135 \times 280}$$

$$+ \frac{(30)^2}{135 \times 128} + \frac{(14)^2}{85 \times 65} + \frac{(50)^2}{85 \times 280} + \frac{(21)^2}{85 \times 128}$$

$$G = 1.0066$$

$$T_p = \sqrt{\frac{1.0066 - 1}{1.0066}}$$

$$T_p = 0.08$$

III-1- الانحدار الخطي البسيطIII-1- مفهوم الانحدار الخطي البسيط:

يعتبر الانحدار الخطي البسيط من أهم التطبيقات الإحصائية في الاقتصاد، فهو يبنى على وجود علاقة خطية بين متغيرين سواء كانت طردية أو عكسية، أحدهما مستقل (Independent) والآخر تابع (Dependent).

III-2- استخدامات الانحدار الخطي البسيط:

أهم استخدامات معادلة الانحدار الخطي البسيط هو التنبأ (Prévision)، بقيم المتغير المعتمد (y_i) والتي تسمى بالقيم التقديرية (Estimate Value) والتي يرمز لها بالرمز وذلك من خلال تحديد قيم المتغير المستقل (x_i)¹³.

إذا كانت هناك علاقة خطية بين متغيرين فكيف نعبّر عن هذه العلاقة بواسطة معادلة تربط بين المتغيرين؟

III-3- معادلة الانحدار الخطي البسيط:

إذا كان لدينا عينة من الأزواج المرتبة $[(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)]$ حيث x_1, x_2, \dots, x_n قيم للمتغير X و y_1, y_2, \dots, y_n القيم المقابلة لها للمتغير Y ، وقمنا برسم لوحة الانتشار للنقاط XY يمكن الحكم فيما إذا كانت هناك علاقة خطية بين المتغيرين X و Y أم لا.

- ✓ إذا كان الشكل الناتج يشبه خط مستقيم فالعلاقة بين المتغيرين خطية
- ✓ إذا كان الشكل الناتج لا يشبه تماما خط مستقيم فالعلاقة غير خطية

¹³ وليد اسماعيل السيفو، عيد أحمد أبوبكر، غالب عوض الرفاعي، مرجع سبق ذكره، ص. 219.

مثال تقديمي:

تقوم النظرية الكينزية على وجود علاقة خطية بين الاستهلاك والدخل تكتب العلاقة على الشكل التالي¹⁴:

$$C = a_0 + a_1y$$

حيث:

C : هو الاستهلاك

Y : هو الدخل

a_0 : حد الكفاف

a_1 : الميل الحدي

نسمي متغير الاستهلاك بالمتغير التابع، ونسمي متغير الدخل بالمتغير المستقل، وبالتالي الدخل هو الذي يفسر الاستهلاك.

a_0 و a_1 تسمى بمعلمات النموذج أو معلمات الانحدار.

- النماذج المعروضة في شكل سلاسل زمنية. المتغيرات تمثل الظواهر على شكل مشاهدات في مجال زمني محدد ومتوازن. وكمثال على ذلك الاستهلاك و الدخل السنوي لـ 20 سنة بالنسبة لدولة ما. وبالتالي يكتب النموذج كالتالي:

$$C_t = a_0 + a_1y_t \quad t = 1,2,3, \dots \dots \dots 20$$

حيث:

C_t : الاستهلاك في الفترة t

y_t : الدخل في الفترة t

- النماذج المعروضة على أشكال أزواج متقابلة، المتغيرات تمثل الظواهر المشاهدة في نفس الوقت، لكن للعديد من الأزواج. وكمثال على ذلك الاستهلاك والدخل المشاهد لعينة من 20 دولة في نفس السنة، وبالتالي يكتب النموذج كالتالي:

¹⁴ Régie Bourbonnais, économétrie cours et exercices corrigés, DUNOD, 9eme édition, 2015, P13-14.

$$C_i = a_0 + a_1 y_i \quad t = 1, 2, 3, \dots \dots \dots 20$$

حيث:

C_i : الاستهلاك للدولة i للسنة المعنية.

y_i : الدخل للدولة i للسنة المعنية.

المتغير العشوائي:

يعتبر الدخل المتغير الأساسي الذي يؤثر على الاستهلاك لكن لا يمكن التسليم بأنه المتغير الوحيد الذي يؤثر على الاستهلاك، بل هناك متغيرات أخرى تؤثر على الاستهلاك. وعليه نضيف للمعادلتين السابقتين المتغير ε_t (للسلاسل الزمنية) أو ε_i (للعينة) فيصبح النموذجين أعلاه:

$$C_t = a_0 + a_1 y_t + \varepsilon_t$$

$$C_i = a_0 + a_1 y_i + \varepsilon_i$$

حيث:

ε_t ، ε_i يمثلان المتغيرات الأخرى التي تؤثر على الاستهلاك وغير متعلقة بالدخل، أي ε_t يمثل الفرق بين المشاهدات الحقيقية لـ C_t والقيم التي يمكن تحديدها من العلاقة أعلاه وعليه هو عبارة عن مجموع ثلاثة أنواع من الأخطاء:

1- الخطأ في التحديد: المتغير المستقل ليس الوحيد الذي يؤثر في المتغير التابع بل هناك متغيرات أخرى لم تؤخذ بعين الاعتبار.

2- خطأ في القياس: البيانات لا تمثل بدقة الظاهرة المدروسة.

3- خطأ عدم الاستقرار: في المعاينة التقدير يتغير من عينة إلى أخرى.

III-4- نموذج الانحدار الخطي البسيط وفرضياته:

إليك نموذج الانحدار التالي:

$$y_t = a_0 + a_1 x_t + \varepsilon_t$$

y_t : المتغير التابع.

x_t : المتغير المستقل

a_0 و a_1 معلمات النموذج

ε_t : خطأ التحديد (الفرق بين النموذج الحقيقي والنموذج المحدد) هذا الخطأ غير معلوم ويبقى غير

معلوم.

n: عدد المشاهدات

فرضيات النموذج:

1- النموذج خطي

2- المتغير المستقل x_t محدد وثابت وعدم وجود أخطاء في قياس المتغير المستقل.

3- الأخطاء العشوائية مستقلة ولها توزيع طبيعي متوسطه الحسابي صفر $E(\varepsilon_t) = 0$ ، أي أن قيم

تتمركز حول الصفر وتباينها يساوي قيمة ثابتة وموجبة $V(\varepsilon_t) = \sigma^2$.

4- المتغير المستقل x_t مستقل عن الخطأ العشوائي ε_t أي $COV(x_t, \varepsilon_t) = 0$

5- قيم الخطأ العشوائي غير مرتبطة فيما بينها $E(\varepsilon_t, \varepsilon_f) = 0$.

III-5- تقدير معلمات نموذج الانحدار الخطي البسيط:

III-5-1- طريقة المربعات الصغرى.

هي طريقة في التقدير تعتمد على التقليل قدر الإمكان من الفوارق بين القيم النظرية والقيم

المشاهدة، وتستعمل بكثرة في الاقتصاد القياسي لتقدير نموذج الانحدار الخطي البسيط.

سنقوم بتقليل قدر الإمكان من الفوارق بين القيم المشاهدة والقيم النظرية التي يمنحها لنا مستقيم

الانحدار وتكون هذه الفوارق موجبة أحيانا وسالبة أحيانا أخرى.

$$\text{MIN} \sum \varepsilon_i^2 = \sum (Y_i - a_0 - a_1 x_i)^2 = \text{MIN } S$$

في الرياضيات التوصل إلى نهاية صغرى يعني أن المشتقة الأولى تنعدم

بالنسبة لـ a_0 :

$$\begin{aligned}\frac{dS}{da_0} &= -2 \sum (Y_i - \hat{a}_0 - \hat{a}_1 x_i) = 0 \\ \Rightarrow \sum Y_i - n\hat{a}_0 - \hat{a}_1 \sum x_i &= 0 \\ \Rightarrow n\hat{a}_0 &= \sum Y_i - \hat{a}_1 \sum x_i\end{aligned}$$

نقسم المعادلة على n فتصبح:

$$\begin{aligned}\hat{a}_0 &= \frac{\sum Y_i}{n} - \hat{a}_1 \frac{\sum x_i}{n} \\ \hat{a}_0 &= \bar{Y} - \hat{a}_1 \bar{x}\end{aligned}$$

بالنسبة لـ a_1 :

$$\begin{aligned}\frac{dS}{da_1} &= -2 \sum x_i (Y_i - \hat{a}_0 - \hat{a}_1 x_i) = 0 \\ -2 \left(\sum x_i y_i - \hat{a}_0 \sum x_i - \hat{a}_1 \sum x_i^2 \right) &= 0\end{aligned}$$

نعوض القيمة \hat{a}_0 بقيمتها

$$\sum x_i y_i - \left(\frac{\sum y_i}{n} - \hat{a}_1 \frac{\sum x_i}{n} \right) \sum x_i - \hat{a}_1 \sum x_i^2 = 0$$

نضرب طرفي المعادلة في n

$$\begin{aligned}n \sum x_i y_i - \sum x_i \cdot \sum y_i + \hat{a}_1 \left(\sum x_i \right)^2 - \hat{a}_1 n \sum x_i^2 &= 0 \\ n \sum x_i y_i - \sum x_i \cdot \sum y_i &= \hat{a}_1 \left(\sum x_i \right)^2 - \hat{a}_1 n \sum x_i^2\end{aligned}$$

$$n \sum x_i y_i - \sum x_i \cdot \sum y_i = \hat{a}_1 (n \sum x_i^2 - (\sum x_i)^2)$$

نقسم البسط والمقام على n فنحصل على:

$$\hat{a}_1 = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2}$$

تمكنا في الأخير من تقدير معاملات نموذج الانحدار البسيط \hat{a}_0 و \hat{a}_1 ، وبالتالي تقدير نموذج الانحدار.

ملاحظة: كلما كانت مجموع مربعات الأخطاء $\sum (Y - \hat{Y})^2$ أصغر كانت القوة التفسيرية للنموذج أكبر.

مثال:

إليك الجدول التالي والذي يوضح مجموعة من القيم لمتغيرين X و Y حيث Y هو المتغير التابع و

X هو المتغير المستقل؟

X	4	10	9	12	8	5
Y	2	6	8	11	5	4

1- أوجد معادلة الانحدار للمتغيرين؟

2- بماذا تقدر القيمة y المقابلة للقيمة $x=9$ ؟

3- ماهو الخطأ في تقدير y عندما $x=9$ ؟

الحل:

1- معادلة الانحدار البسيط هي:

$$\hat{Y} = a_0 + a_1 x$$

حيث:

$$\hat{a}_1 = \frac{\sum xy - n \bar{x} \bar{y}}{\sum x^2 - n \bar{x}^2}$$

و

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

	x	y	xy	
	4	2	8	16
	10	6	60	100
	9	8	72	81
	12	11	132	144
	8	5	40	64
	5	4	20	25
المجموع	48	36	332	430

بالتعويض:

$$\bar{x} = \frac{48}{6} = 8, \bar{y} = \frac{36}{6} = 6$$

$$\hat{a}_1 = \frac{332 - 6 \times 8 \times 6}{430 - 6(8)^2} = 0.96$$

$$\hat{a}_0 = 6 - 0.96 \times 8 = -1.68$$

إذن معادلة الانحدار هي:

$$\hat{Y} = -1.68 + 0.96x$$

2- القيمة التقديرية للمتغير y عند x=9 هي:

$$\hat{Y} = -1.68 + 0.96 \times 9 = 6.96$$

3- تحديد حد الخطأ عند x=9

$$e = (Y - \hat{Y})$$

$$e = 8 - 6.96 = 1.04$$

ملاحظة:

تعرض النظرية الاقتصادية بعض الحالات التي تكون فيها $a_0 = 0$ وكمثال على ذلك العلاقة بين الإنتاج والعمل، فنعلم أنه عندما يكون العمل 0 يكون الإنتاج 0 فيصبح تقدير \hat{a}_1 في هذه الحالة هو:

$$\hat{a}_1 = \frac{\sum x_t y_t}{\sum x_t^2}$$

III-6- مؤشرات حساب جودة النموذج:

يتم قياس جودة النموذج بمؤشرين هما:

III-6-1- معامل التحديد:

ويقاس المدى الذي يفسر به المتغير المستقل x_t المتغير التابع y_t وتتراوح قيمته بين 0 و 1. فكلما اقترب من الواحد نقول x_t بأن يفسر نسبة كبيرة من التغيرات التي تحدث في المتغير التابع y_t . والعكس صحيح.

يمكن إرجاع الاختلاف الكلي في المتغير التابع بالنسبة لوسطه الحسابي إلى الاختلاف بسبب الانحدار $(\hat{y}_t - \bar{y})^2$ والاختلاف بسبب الخطأ e_t^2 .

1- مجموع المربعات الكلي (SST):

مجموع المربعات الإجمالي للتغيرات التي تحدث في المتغير التابع وهي تتكون من جزئين:

$$SST = SSR + SSE$$

$$\sum (Y_t - \bar{Y})^2 = \sum (\hat{Y}_t - \bar{Y})^2 + \sum e_t^2$$

2- مجموع مربعات الانحدار (SSR):

هو جزء من تباين قيمة المتغير الذي تم تفسيره بواسطة الانحدار: أي جزء من التغيرات التي تحدث في المتغير التابع والذي تم تفسيره بواسطة النموذج المقدر.

$$SSR = \sum (\hat{Y}_t - \bar{Y})^2$$

3- مجموع مربعات الأخطاء (SSE):

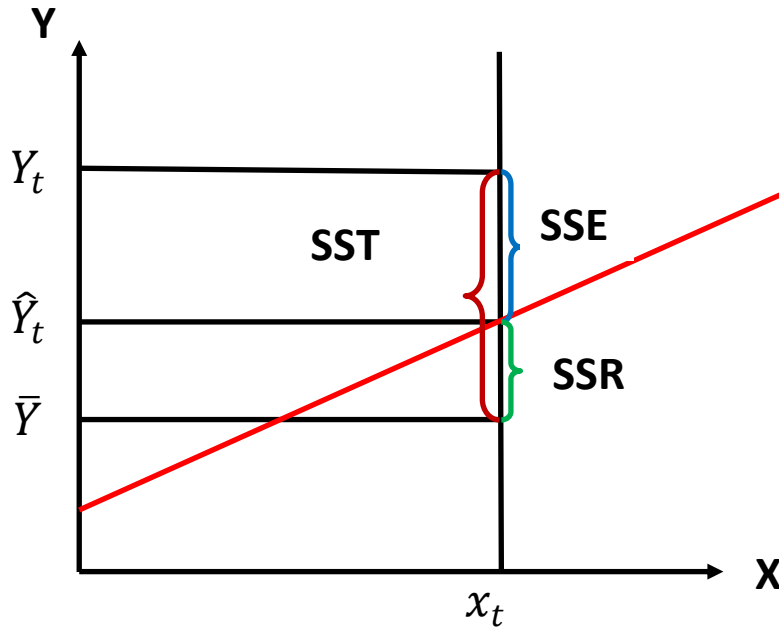
أي مجموع مربعات البواقي وهو الجزء الذي فشل النموذج في تفسيره.

$$SSE = \sum e_t^2$$

4- معامل التحديد (معامل التفسير):

يمثل النسبة بين مجموع مربعات الانحدار (SSR) ومجموع المربعات الكلي (SST)

$$R^2 = \frac{SSR}{SST}$$

III-6-2- تباين الخطأ العشوائي

تقدير المربعات الصغرى لتباين الخطأ العشوائي هو $\hat{\sigma}_e^2$ ويحسب بالمعادلة التالية:

$$\hat{\sigma}_e^2 = \frac{\sum e_t^2}{n-2} = \frac{(Y_t - \hat{Y}_t)}{n-2}$$

يقيس هذا المؤشر مدى انحراف القيم الفعلية عن القيم المقدرة. فإذا كان هذا التقدير كبيراً فإن

انحرافات القيم الفعلية للمتغير التابع عن القيم المقدرة لها كبير أي أن النموذج غير كفاً.

III-7- الاختبارات الخاصة بملائمة النموذج وكذا معلمات الانحدار:III-7-1- اختبار مدى ملائمة النموذج:

يوضح هذا الاختبار مدى ملائمة النموذج في تمثيل العلاقة بين المتغيرين التابع والمستقل. أو إن كان هذا التمثيل جيد أم لا. ويعتمد على إحصائية فيشر F . ولاختبار مدى ملائمة النموذج نختبر الفرضيتين التاليتين:

$$H_0: a_1 = 0$$

$$H_1: a_1 \neq 0$$

أو

$$H_0: SSR = 0$$

$$H_1: SSR \neq 0$$

أو صياغتها على النحو التالي:

H_0 : النموذج الحالي **ملائم** لتمثيل العلاقة بين المتغير المستقل X والمتغير التابع Y .

H_1 : النموذج الحالي **غير ملائم** لتمثيل العلاقة بين المتغير المستقل X والمتغير التابع Y .

إحصائية فيشر المحسوبة تستخرج بموجب العلاقة التالية:

$$F^* = \frac{MSR}{MSE}$$

حيث:

$$MSR = \frac{SSR}{1}$$

$$MSE = \frac{SSR}{n - 2}$$

MSR: متوسط مربعات الانحدار

MSE: متوسط مربعات الأخطاء

جدول تحليل التباين الأحادي:

قيمة F	متوسط المربعات	درجات الحرية	مجموع المربعات	مصدر التغير
$F^* = \frac{MSR}{MSE}$	$MSR = \frac{SSR}{1}$ $MSE = \frac{SSE}{n-2}$	1 n-2	$SSR = \sum (\hat{Y}_t - \bar{Y})^2$ $SSE = \sum e_t^2$	- المتغير المستقل X - البواقي
		n-1	$SST = \sum (Y_t - \bar{Y})^2$	التغير الكلي

يخضع الاختبار إلى توزيع فيشر بدرجات حرية كما يلي:

- درجة حرية البسط هو عدد المتغيرات المستقلة. وهو 1 في الانحدار الخطي البسيط.
- درجة حرية المقام هو عدد المشاهدات منقوص منها عدد المتغيرات المستقلة.

III-7-2- اختبار معلمات الانحدار:

لاختبار معلمات الانحدار لدينا ثلاث أنواع من الاختبارات وهي:

$$H_0: a_0 = 0$$

$$H_0: a_1 = 0$$

1- اختبار ذو جناحين:

$$H_1: a_0 \neq 0$$

$$H_1: a_1 \neq 0$$

2- اختبار ذو جناح أيمن:

$$H_0: a_0 = 0$$

$$H_0: a_1 = 0$$

$$H_1: a_0 > 0$$

$$H_1: a_1 > 0$$

3- اختبار ذو جناح أيسر:

$$H_0: a_0 = 0$$

$$H_0: a_1 = 0$$

$$H_1: a_0 < 0$$

$$H_1: a_1 < 0$$

يخضع الاختبار إلى الإحصائية t بدرجة حرية (n-2)

$$t_{cal(\hat{a}_1)} = \frac{\hat{a}_1 - a_1}{\sigma(\hat{a}_1)}$$

$$t_{cal(\hat{a}_0)} = \frac{\hat{a}_0 - a_0}{\sigma(\hat{a}_0)}$$

مع:

$$\sigma(\hat{a}_1) = \frac{\hat{\sigma}_e^2}{\sum(x_t - \bar{x})^2}$$

$$\sigma(\hat{a}_0) = \hat{\sigma}_e^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_t - \bar{x})^2} \right)$$

تمرين:

التمرين الثاني:

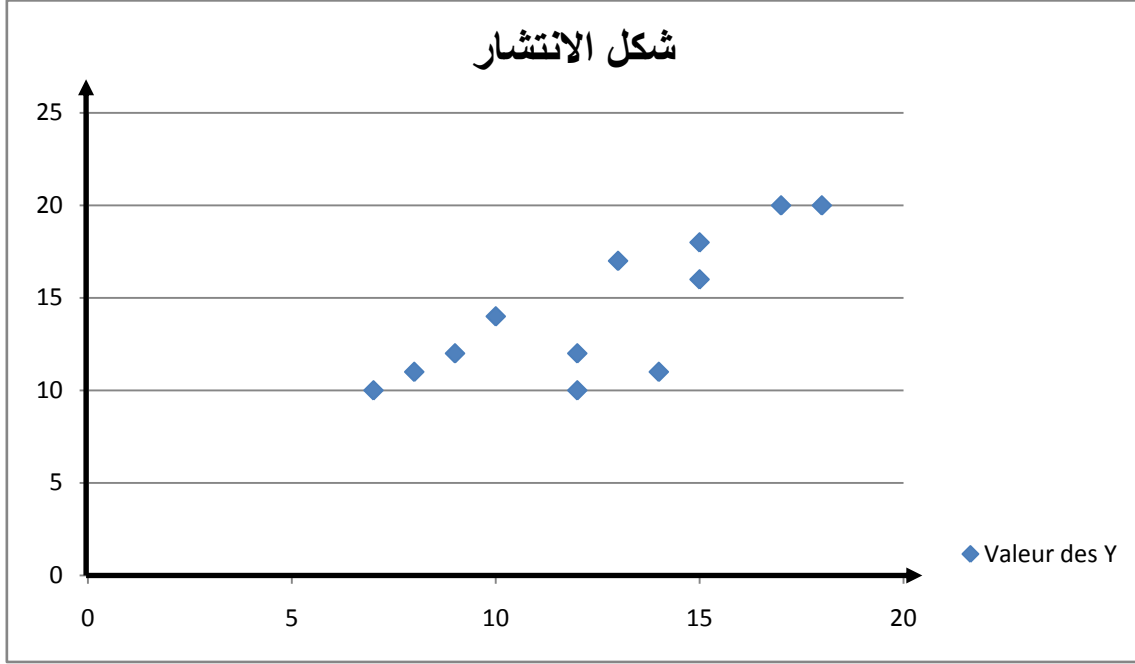
يعطي الجدول التالي علامات 12 طالبا في الامتحان الأول x والامتحان الثاني y.

x	18	14	10	15	7	12	13	8	9	17	15	12
y	20	11	14	16	10	10	17	11	12	20	18	12

- 1- أرسم لوحة الانتشار؟
- 2- أوجد معادلة خط انحدار y على x؟
- 3- أخذ طالب في الاختبار الأول العلامة 16 وغاب عن الاختبار الثاني، ما هي العلامة التقديرية التي يحصل عليها الطالب في الاختبار الثاني؟
- 4- ما هو الخطأ في تقديرك للعلامة في الاختبار الثاني إذا حصل على 10 في الاختبار الأول موضحا ذلك برسم؟
- 5- قدر تباين الخطأ العشوائي؟
- 6- أحسب معامل التحديد؟
- 7- إختبر مدى ملائمة النموذج المقترح في تمثيل العلاقة بين X و Y عند مستوى معنوية 5%؟
- 8- إختبر مدى ملائمة معلمات الانحدار للنموذج المقترح؟

حل التمرين:

1 - رسم شكل الانتشار:



2 - إيجاد معادلة خط انحدار y على x

$$\hat{a}_1 = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$\hat{a}_1 = 0.87$$

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

$$\hat{a}_0 = 3.375$$

$$\hat{Y} = 3.375 + 0.87x$$

3 - العلامة التقديرية التي يحصل عليها الطالب في الاختبار الثاني:

لدينا $X=16$

$$\hat{Y} = 3.375 + 0.87 \quad (16)$$

$$\hat{Y} = 17.295$$

4 - حساب الخطأ التقديري

$X=10$

$$\hat{Y} = 3.375 + 0.87 \quad (10)$$

$$\hat{Y} = 12.075$$

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

$$e_i = 14 - 12.075$$

$$e_i = 1.925$$

5 - تقدير تباين الخطأ العشوي:

$$\hat{\sigma}_e^2 = \frac{\sum e_i^2}{n-2} = \frac{(Y_i - \hat{Y}_i)}{n-2}$$

$$\hat{\sigma}_e^2 = \frac{56.03}{10}$$

$$\hat{\sigma}_e^2 = 5.603$$

6 - حساب معامل التحديد:

$$R^2 = \frac{SSR}{SST}$$

$$R^2 = \frac{102.19}{158.22}$$

$$R^2 = 0.64$$

يمكن تفسيره بأن 64% من التغيرات التي تحصل في Y سببها X

7 - إختبار مدى ملائمة النموذج المقترح في تمثيل العلاقة بين Y و X :

- وضع الفرضيات:

$H_0: a_1 = 0$ النموذج الحالي غير ملائم لتمثيل العلاقة بين المتغير المستقل X نقاط الطلبة للامتحان الأول والمتغير التابع Y نقاط الطلبة للامتحان الثاني.

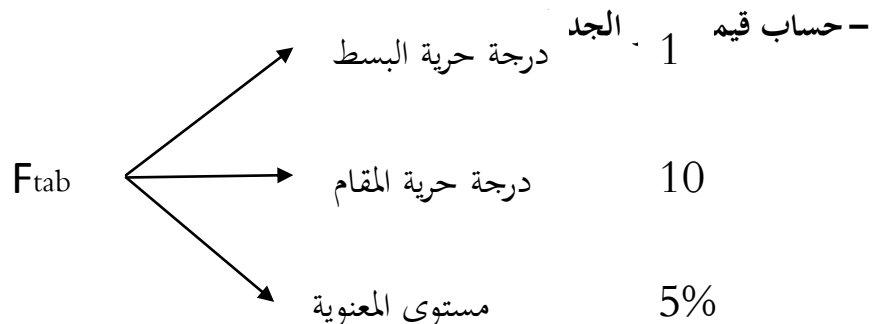
$H_1: a_1 \neq 0$ النموذج الحالي ملائم لتمثيل العلاقة بين المتغير المستقل X طول الطلبة والمتغير التابع Y وزن الطلبة.

- حساب قيمة فيشر الحسابية F^* :

يمكن إيجاد جدول تحليل التباين (ANOVA). وذلك من خلال حساب مختلف المجموع

والذي يظهر على الشكل التالي:

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة F
- المتغير المستقل X	$SSR = \sum (\hat{Y}_i - \bar{Y})^2$ $SSR = 102.19$	1	$MSR = \frac{SSR}{1}$ $MSR = 102.19$	$F^* = \frac{MSR}{MSE}$ $F^* = \frac{102.19}{5.603}$ $F^* = 18.23$
- البواقي	$SSE = \sum e_i^2$ $SSE = 56.03$	$n-2$	$MSE = \frac{SSR}{n-2}$ $MSE = 5.603$	
التغير الكلي	$SST = \sum (Y_i - \bar{Y})^2$ $SST = 158.22$	$n-1$		



$$F_{tab} = 4.96$$

-قاعدة القرار:

بما أن قيمة F_{cal} المحسوبة أكبر من قيمة F_{tab} الجدولية إذن نقول أن النموذج المحصل عليه ملائم لتمثيل العلاقة بين نقط الطلبة للامتحان الأول من جهة ونقاط الطلبة للامتحان الثاني من جهة أخرى.

8 - إختبر مدى ملائمة معلمات الانحدار للنموذج المقترح:

لاختبار مدى ملائمة معلمات النموذج للنموذج المقترح نستعمل الإحصائية t (student) بدرجة حرية $n-2$

-وضع الفرضيات:

$$a_1 = 0 : H_0$$

$$a_1 \neq 0 : H_1$$

2- حساب قيمة t المحسوبة:

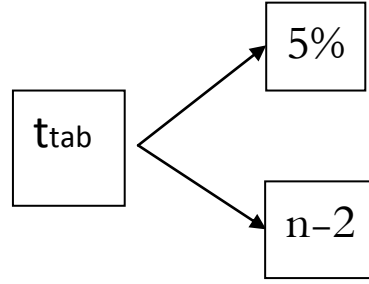
$$t_{cal} = \frac{\hat{a}_1 - a_1}{\sigma_{\hat{a}_1}} = \frac{\hat{a}_1}{\sigma_{\hat{a}_1}}$$

$$\sigma_{\hat{a}_1}^2 = \frac{\sigma_e^2}{\sum(x_i - \bar{x})^2} = \frac{5.603}{135} = 0.041$$

$$\sigma_{\hat{a}_1}^2 = 0.041 \rightarrow \sigma_{\hat{a}_1} = 0.20$$

$$t_{cal} = \frac{0.87}{0.20} = 4.35$$

3- حساب قيمة t الجدولية:



$$t_{tab} = 2.228$$

– قاعدة القرار:

قيمة t المحسوبة أكبر من قيمة t الجدولة وبالتالي نرفض الفرضية العدمية ونقبل الفرضية البديلة أي المتغير X له أثر معنوي على المتغير Y.

IV- الانحدار الخطي المتعددIV-1- مفهوم الانحدار الخطي المتعدد

يهدف استخدام تحليل الانحدار المتعدد في البحث عن العلاقة ما بين أكثر من متغير مستقل. عبارة أخرى العوامل المؤثرة على الظاهرة المدروسة، فهناك حالات علمية عديدة يكون المتغير التابع معتمد في تفسيره على مجموعة من المتغيرات المستقلة.

IV-2- تقديم نموذج الانحدار الخطي المتعدد

يظهر الشكل العام للمعادلة الخطية المتعددة التي هي الأساس لكافة الأشكال الأخرى للانحدار كما يلي:

$$Y_t = a_0 + a_1x_{1t} + a_2x_{2t} \dots + a_kx_{kt} + \varepsilon_t$$

Y_t : المتغير التابع في الفترة

x_{1t} : المتغير المستقل الأول في الفترة t

x_{2t} : المتغير المستقل الثاني في الفترة t

x_{kt} : المتغير المستقل الثالث في الفترة t

$a_0, a_1, a_2, \dots, a_k$: معاملات النموذج

ε_t : خطأ التحديد (الفرق بين النموذج الحقيقي والنموذج المقدر) هذا الخطأ غير معلوم ويبقى غير معلوم.

n: عدد المشاهدات

IV-3- الشكل المصفوفي للنموذج:

يعتبر الشكل المصفوفي (الجبر الخطي) الأسهل في تقدير نموذج الانحدار الخطي المتعدد بالإضافة إلى إجراء الاختبارات اللازمة. وكون النموذج يحتوي على عدد من المتغيرات المستقلة. بالإضافة إلى عدد المشاهدات n ويمكن كتابته على الشكل التالي¹⁵:

¹⁵ Régie Bourbonnais, économétrie cours et exercices corrigés, op cité P 48

$$Y_1 = a_0 + a_1x_{11} + a_2x_{21} \dots + a_kx_{k1} + \varepsilon_1$$

$$Y_2 = a_0 + a_1x_{12} + a_2x_{22} \dots + a_kx_{k2} + \varepsilon_2$$

$$Y_t = a_0 + a_1x_{1t} + a_2x_{2t} \dots + a_kx_{kt} + \varepsilon_t$$

$$Y_n = a_0 + a_1x_{1n} + a_2x_{2n} \dots + a_kx_{kn} + \varepsilon_n$$

الشكل المصفوفي لهذا النموذج هو:

$\mathbf{Y} =$	\mathbf{a}	\mathbf{x}	$+$	$\boldsymbol{\varepsilon}$
$(n, 1)$	$n, k + 1$	$(k + 1, 1)$		$(n, 1)$

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_t \\ \vdots \\ Y_n \end{pmatrix}, X = \begin{pmatrix} 1 & X_{11} & X_{21} & \dots & X_{K1} \\ 1 & X_{12} & X_{22} & \dots & X_{K2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1t} & X_{2t} & \dots & X_{Kt} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{Kn} \end{pmatrix}$$

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_t \\ \vdots \\ a_k \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_t \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

إن توفر البرامج الإحصائية الجاهزة قد تمكن الباحث من صعوبة القيام بالعمليات الحسابية المعقدة

في تحليل الانحدار المتعدد، وخاصة بالنسبة للنماذج التي تحتوي على عدد كبير من المتغيرات المستقلة.

ومن أهم البرامج الإحصائية المستعملة في تحليل الانحدار المتعدد نجد (eviews, spss, stata)

و غيرها).

IV-4- فرضيات النموذج:

يمكن إجمال فروض هذا النموذج في الخمس فرضيات التالية:

$$1- \text{التوقع الرياضي لشعاع الخطأ العشوائي يساوي الصفر } E(\varepsilon_i) = 0$$

$$E(\varepsilon_i) = E \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

2- المتغيرات المفردة مستقلة إحصائياً إلى وجود استقلال خطي بين المتغيرات وعدد المشاهدات يجب أن يكون أكبر من عدد المتغيرات.

3- تباين الخطأ العشوائي ثابت ويساوي σ^2 وهو ثابت من مشاهدة لأخرى أي:

$$V(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$$

والتباين المشترك بينهما يساوي صفراً أي:

$$COV(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0 \quad V \quad i \neq j$$

4- لا يوجد ارتباط بين قيم حد الخطأ العشوائي وقيم المتغيرات المستقلة، أي أن أعمدة المصفوفة X مستقلة خطياً عن شعاع الأخطاء العشوائية E.

5- الخطأ العشوائي ε_i يتبع توزيعاً طبيعياً أي $\varepsilon \rightarrow N(0, \sigma)$

IV-5- تقدير معاملات النموذج:

الشكل المصفوفي للنموذج متكون من k متغيرات مستقلة و n مشاهدة.

$$Y = Xa + \varepsilon$$

$$MIN \sum \varepsilon_t^2 = MIN \varepsilon' \varepsilon \quad MIN (Y - Xa)' (Y - Xa)$$

مع ε' هي منقول المصفوفة ε

$$\text{و } \sum \varepsilon_t^2 = \varepsilon' \varepsilon \text{ ومنه:}$$

$$\varepsilon_t^2 = \varepsilon_1^2 + \varepsilon_2^2 \dots \dots \dots \varepsilon_n^2$$

$$\varepsilon' \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} (\varepsilon_1 \quad \varepsilon_2 \quad \dots \quad \varepsilon_n)$$

$$S = (Y - Xa)'(Y - Xa) = Y'Y - Y'Xa - a'X'Y + a'X'Xa$$

ملاحظة:

$a'X'Y$: عدد ثابت وبالتالي هو نفسه $Y'Xa$ فتصبح المعادلة كالتالي:

$$S = Y'Y - 2a'X'Y + a'X'Xa$$

$$\frac{\partial S}{\partial a} = -2X'Y + 2X'Xa \rightarrow \hat{a} = \frac{X'Y}{X'X}$$

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

$(X'X)^{-1}$: هي مقلوب المصفوفة $X'X$

مثال:

البيانات التالية تمثل العلاقة بين المتغير التابع (Y) والمتغيرين المستقلين (X_1 و X_2) لعينة مكونة

من خمسة مشاهدات.

4	6	5	8	7	Y
1	3	1	2	1	X1
2	1	3	1	2	X2

المطلوب: بافتراض أن العلاقة بين المتغيرات هي علاقة خطية قدر معادلة الانحدار Y على X_1 و

X_2 باستعمال الشكل المصفوفي؟

الحل:

$$\hat{Y} = a_0 + a_1x_1 + a_2x_2$$

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

-1 حساب (x'x)

$$x = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}, y = \begin{pmatrix} 7 \\ 8 \\ 5 \\ 6 \\ 4 \end{pmatrix}$$

$$x'x = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 & 1 \\ 2 & 1 & 3 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

$$x'x = \begin{pmatrix} 5 & 8 & 9 \\ 8 & 16 & 12 \\ 9 & 12 & 19 \end{pmatrix}$$

-2 حساب (x'x)⁻¹

• حساب المحدد:

$$\Delta = \begin{vmatrix} 5 & 8 & 9 \\ 8 & 16 & 12 \\ 9 & 12 & 19 \end{vmatrix}$$

$$= 5 \begin{vmatrix} 16 & 12 \\ 12 & 19 \end{vmatrix} - 8 \begin{vmatrix} 8 & 12 \\ 9 & 19 \end{vmatrix} + 9 \begin{vmatrix} 8 & 16 \\ 9 & 12 \end{vmatrix}$$

$$\Delta = 16$$

• حساب المصفوفة المرافقة (adjoint matrix):

$$adj(x'x) = \begin{vmatrix} + & \begin{vmatrix} 16 & 12 \\ 12 & 19 \end{vmatrix} & - & \begin{vmatrix} 8 & 12 \\ 9 & 19 \end{vmatrix} & + & \begin{vmatrix} 8 & 16 \\ 9 & 12 \end{vmatrix} \\ - & \begin{vmatrix} 8 & 9 \\ 12 & 19 \end{vmatrix} & + & \begin{vmatrix} 5 & 9 \\ 9 & 19 \end{vmatrix} & - & \begin{vmatrix} 5 & 8 \\ 9 & 12 \end{vmatrix} \\ + & \begin{vmatrix} 8 & 9 \\ 16 & 12 \end{vmatrix} & - & \begin{vmatrix} 5 & 9 \\ 8 & 12 \end{vmatrix} & + & \begin{vmatrix} 5 & 8 \\ 8 & 16 \end{vmatrix} \end{vmatrix}$$

$$adj(x'x) = \begin{vmatrix} 160 & -44 & -48 \\ -44 & 14 & 12 \\ -48 & 12 & 16 \end{vmatrix}$$

$$(x'x)^{-1} = \frac{adj(x'x)}{\Delta} = \begin{vmatrix} 10 & -2.75 & -3 \\ -2.75 & 0.875 & 0.75 \\ -3 & 0.75 & 1 \end{vmatrix}$$

$$x'y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 & 1 \\ 2 & 1 & 3 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 8 \\ 5 \\ 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 30 \\ 50 \\ 51 \end{pmatrix}$$

• حساب المصفوفة \hat{a}

$$\begin{aligned} \hat{a} &= (x'x)^{-1} \cdot x'y = \begin{pmatrix} 10 & -2.75 & -3 \\ -2.75 & 0.875 & 0.75 \\ -3 & 0.75 & 1 \end{pmatrix} \cdot \begin{pmatrix} 30 \\ 50 \\ 51 \end{pmatrix} = \\ &= \begin{pmatrix} 9.5 \\ -0.5 \\ -1.5 \end{pmatrix} \end{aligned}$$

$$\hat{a} = (x'x)^{-1}x'y = (9.5 \quad -0.5 \quad -1.5)$$

ومنه

$$\hat{Y} = 9.5 - 0.5X_1 + 1.5X_2$$

IV-6- القدرة التفسيرية للنموذج:

يمكن قياس القدرة التفسيرية للنموذج من خلال:

IV-6-1- معامل التحديد (المتعدد)

هو مقياس يوضح نسبة مساهمة المتغيرات المستقلة في تفسير التغير الحاصل في المتغير التابع، وتتراوح قيمته بين 0 و 1 وكلما اقترب من الواحد كلما كانت القدرة التفسيرية للنموذج كبيرة في حين أن انخفاضه أو اقترابه من الصفر تصبح القدرة التفسيرية للنموذج ضعيفة ويمكن حسابه من خلال العلاقة التالية:

$$R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y_t - \bar{Y})^2} = 1 - \frac{\sum e_t}{\sum(Y_t - \bar{Y})^2}$$

IV-6-2- معامل التحديد المصحح:

وهو أكثر كفاءً من معامل التحديد العادي في تفسير النموذج: وهو معطي بالعلاقة التالية:

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

ملاحظة:

يجب استعمال معامل التحديد المصحح للمقارنة بين نماذج الانحدار المختلفة ذات المتغير التابع الواحد. ويفضل استخدام معامل \bar{R}^2 وذلك لأن معامل التحديد R^2 يتزايد بتزايد عدد المتغيرات المفسرة

أي أنه في نموذج الانحدار الخطي البسيط $\bar{R}^2 = R^2$ أما في نموذج الانحدار الخطي المتعدد $\bar{R}^2 \leq R^2$.

IV-7- اختبار صلاحية النموذج:

IV-7-1- اختبار المعنوية الإجمالية للنموذج (اختبار F):

الفرضية العدمية: النموذج الحال غير ملائم لتمثيل العلاقة بين المتغير التابع من جهة والمتغير المستقل من جهة أخرى.

الفرضية البديلة: النموذج الحال ملائم لتمثيل العلاقة بين المتغير التابع من جهة والمتغير المستقل من جهة أخرى.

ولإجراء هذا الاختبار نقوم بحساب الإحصائي نقوم بحساب الإحصائية F فيشر كما يلي:

$$F_{CAL} = \frac{\frac{SSR}{K-1}}{\frac{SSE}{n-K}}$$

$$F_{CAL} = \frac{\frac{R^2}{K-1}}{\frac{1-R^2}{n-K}}$$

ومقارنتها بالقيمة الجدولية لفisher F_{TAB} عند مستوى معنوية (1% ، 5% ، 10%)، بدرجة حرية البسط (K-1) ودرجة حرية المقام (n-K).

قاعدة القرار:

إذا كانت F_{CAL} أكبر من F_{TAB} فإننا نرفض الفرضية العدمية ونقبل الفرضية البديلة، وبالتالي نقول أن النموذج ملائم لتمثيل العلاقة بين المتغير التابع والمتغيرات المستقلة. والعكس صحيح.

مثال 2: في المثال السابق:

- 1 - أحسب معامل التحديد ومعامل التحديد المصحح؟
 2 - اختبر مدى ملائمة النموذج السابق في تفسير العلاقة بين Y و $X1$ و $X2$

الحل:

1 - حساب معامل التفسير

$$R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y_t - \bar{Y})^2}$$

$$R^2 = \frac{3.5}{10}$$

$$R^2 = 0.35$$

ومعنى هذا أن 35% من التغيرات التي تحدث في المتغير التابع (Y) سببها المتغيرات المستقلة ($X1$) و ($X2$)، وأما 65% من التغيرات التي تحدث في المتغير التابع (Y) سببها متغيرات أخرى لم تأخذ بعين الاعتبار في النموذج.

2 - حساب معامل التفسير المصحح

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

$$\bar{R}^2 = 1 - \left[\frac{5-1}{5-2} (1 - 0.35) \right]$$

$$\bar{R}^2 = 0.13$$

نفسر معامل التحديد المصحح بنفس تفسير معامل التحديد العادي وبالتالي 13% من التغيرات التي تحدث في المتغير التابع (Y) سببها المتغيرات المستقلة ($X1$) و ($X2$)، وأما 87% من التغيرات التي تحدث في المتغير التابع (Y) سببها متغيرات أخرى لم تأخذ بعين الاعتبار في النموذج.

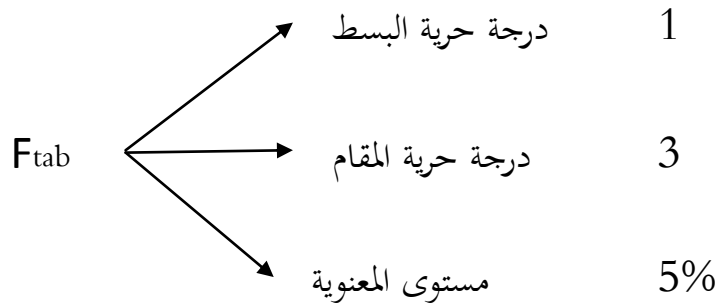
IV-7-2- اختبار مدى ملائمة النموذج السابق في تفسير العلاقة بين Y و X1 و X2

- وضع الفرضيات:

H0: النموذج الحالي غير ملائم لتمثيل العلاقة بين المتغير التابع (Y) من جهة، والمتغيرات المستقلة (X1) و (X2) من جهة أخرى.

H1: النموذج الحالي ملائم لتمثيل العلاقة بين المتغير التابع (Y) من جهة، والمتغيرات المستقلة (X1) و (X2) من جهة أخرى.

- حساب F قيمة الحرجة (الجدولية):



$$F_{tab} = 10.1$$

- حساب F قيمة القياسية

$$F_{CAL} = \frac{\frac{R^2}{K-1}}{\frac{1-R^2}{n-K}}$$

$$F_{CAL} = \frac{\frac{0.35}{2-1}}{\frac{1-0.35}{5-2}}$$

$$F_{CAL} = 1.65$$

- قاعدة القرار:

بما أن قيمة F القياسية (F_{cal}) أصغر من قيمة F الجدولية (F_{tab}) فإننا نقبل الفرضية العدمية ونرفض الفرضية البديلة والتالي النموذج المحصل غير ملائم لتمثيل العلاقة بين المتغير التابع (Y) والمتغيرات المستقلة (X_1) و (X_2).

IV-7-3- طريقة اختيار عدد متغيرات النموذج:

هناك عدة طرق لاختيار عدد المتغيرات المستقلة (أفضل توليفة من المتغيرات المستقلة) التي يمكن أن يضمها النموذج، وبالتالي نبدأ بالمتغير الذي يضيف أكبر زيادة ممكنة للقدرة التفسيري للنموذج، وإذا كان علينا حذف متغير فيجب أن تأثير حذفه أقل ما يمكن على القدرة التفسيرية للنموذج، وفيما يلي أهم الطرق التي يمكن استعمالها لاختيار أفضل تركيبة من المتغيرات:

1- طريقة شمول كافة المتغيرات (All Possible Regression):

وتستخدم إذا كان عدد المتغيرات ليس كبيراً، وأبرز عيوبها حاجتها لعمليات حسابية ووقت كبيرين.

2- طريقة الإضافات المتتالية (Forward Selection Regression):

وفيها إذا كانت قيمة F الجدولية أقل من قيمة F المحسوبة عندها يتوقف البحث عن متغير، وإذا كان العكس يتم إدخال متغير جديد إلى المعادلة وإعادة الحساب أي:

$$H_0: a_i = 0$$

$$H_1: a_i \neq 0$$

3- طريقة الحذف التنازلي (Backward Elimination Selection):

إذا كانت قيمة F المحسوبة لكافة المتغيرات أقل من قيمة F الجدولية، يتم حذف متغير من المعادلة، وإعادة حساب قيمة F المحسوبة من جديد إلى غاية أن تصبح قيم F المحسوبة أكبر من قيمة F الجدولية.

4- طريقة الخطوات المتتالية (Backward Selection Regression):

تجمع هذه الطريقة بين طريقتين، طريق الإضافة المتتالية (FS) وطريقة الحذف التنازلي (BE) وفي كل خطوة يتم اختبار أحد المتغيرات انطلاقاً من الأكثر أهمية ولغاية عدم هبوط قيمة F المحسوبة عن قيمة الجدولية، بعبارة أخرى إجراء اختبارات معاملات المتغيرات لمعرفة معنويتها من عدمها. وتعتبر هذه الطريقة أكثر الطرق استخداماً من الناحية العلمية لقلة الوقت الذي تحتاجه في عملية الحساب بالإضافة إلى أنها تعرض النتائج في كل خطوة بصورة واضحة ومرضية ومبكرة من دون الحاجة إلى إجراء الخطوات غير المعنوية

تمرين:

قامت إحدى الدراسات بدراسة العلاقة بين الكمية المطلوبة من سلعة معينة كمتغير تابع وسعر Y_i السلعة X_1 والدخل X_2 كمتغيرين مستقلين. وكانت البيانات الفعلية التي تم جمعها من عينة مكونة من 5 مشاهدات عن هذه المتغيرات كما في الجدول التالي:

100	70	40	30	10	الكمية (Y_i)
1	2	3	4	5	السعر (X_1)
100	80	70	60	40	الدخل (X_2)

المطلوب:

- 1- باستعمال الشكل المصفوفي قم بتقدير نموذج الانحدار لـ Y على X_1 و X_2 ؟
- 2- أحسب معامل التفسير ومعامل التفسير المصحح وفسرهما إقتصادياً ؟
- 3- إختبر مدى كفاءة النموذج عند مستوى معنوية 5% ؟

حل التمرين

- 1 - تقدير نموذج الانحدار Y على X_1 و X_2 باستعمال الشكل المصفوفي:

$$x = \begin{pmatrix} 1 & 5 & 40 \\ 1 & 4 & 60 \\ 1 & 3 & 70 \\ 1 & 2 & 80 \\ 1 & 1 & 100 \end{pmatrix}, y = \begin{pmatrix} 10 \\ 30 \\ 40 \\ 70 \\ 100 \end{pmatrix}$$

$$x'x = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 4 & 3 & 2 & 1 \\ 40 & 30 & 40 & 70 & 100 \end{pmatrix} \cdot \begin{pmatrix} 1 & 5 & 40 \\ 1 & 4 & 60 \\ 1 & 3 & 70 \\ 1 & 2 & 80 \\ 1 & 1 & 100 \end{pmatrix}$$

$$x'x = \begin{pmatrix} 5 & 15 & 350 \\ 15 & 55 & 910 \\ 350 & 910 & 26500 \end{pmatrix}$$

• حساب المحدد:

$$\Delta = \begin{vmatrix} 5 & 15 & 350 \\ 15 & 55 & 910 \\ 350 & 910 & 26500 \end{vmatrix} = 5 \begin{vmatrix} 55 & 910 \\ 910 & 26500 \end{vmatrix}$$

$$-15 \begin{vmatrix} 15 & 910 \\ 350 & 26500 \end{vmatrix} + 350 \begin{vmatrix} 15 & 55 \\ 350 & 910 \end{vmatrix}$$

$$\Delta = 2000$$

• حساب المصفوفة المرافقة (adjoint matrix):

$$adj(x'x) = \begin{vmatrix} \begin{vmatrix} 55 & 910 \\ 910 & 26500 \end{vmatrix} & \begin{vmatrix} 15 & 910 \\ 350 & 26500 \end{vmatrix} & \begin{vmatrix} 15 & 55 \\ 350 & 910 \end{vmatrix} \\ \begin{vmatrix} 15 & 350 \\ 910 & 26500 \end{vmatrix} & \begin{vmatrix} 5 & 350 \\ 350 & 26500 \end{vmatrix} & \begin{vmatrix} 5 & 15 \\ 350 & 910 \end{vmatrix} \\ \begin{vmatrix} 15 & 350 \\ 55 & 910 \end{vmatrix} & \begin{vmatrix} 5 & 350 \\ 15 & 910 \end{vmatrix} & \begin{vmatrix} 5 & 15 \\ 15 & 55 \end{vmatrix} \end{vmatrix}$$

$$adj(x'x) = \begin{vmatrix} 629400 & -79000 & -5600 \\ -79000 & 10000 & 700 \\ -5600 & 700 & 50 \end{vmatrix}$$

$$(x'x)^{-1} = \frac{adj(x'x)}{\Delta} = \begin{vmatrix} 314.7 & -39.5 & -2.8 \\ -39.5 & 5 & +0.35 \\ -2.8 & 0.35 & 0.025 \end{vmatrix}$$

• حساب المصفوفة \hat{a}

$$\hat{a} = (x'x)^{-1}x'y = \begin{pmatrix} 314.7 & -39.5 & -2.8 \\ -39.5 & 5 & +0.35 \\ -2.8 & 0.35 & 0.025 \end{pmatrix} \cdot \begin{pmatrix} 250 \\ 530 \\ 20600 \end{pmatrix}$$

$$\hat{a} = (x'x)^{-1}x'y = (60 \quad -15 \quad 0.5)$$

ومنه

$$\hat{Y} = 60 - 15X_1 + 0.5X_2$$

3 - حساب معامل التفسير

$$R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y_t - \bar{Y})^2}$$

$$R^2 = \frac{4850}{5000}$$

$$R^2 = 0.97$$

ومعنى هذا أن 97% من التغيرات التي تحدث في المتغير التابع (Y) سببها المتغيرات المستقلة (X1) و (X2)، وأما 0.03% من التغيرات التي تحدث في المتغير التابع (Y) سببها متغيرات أخرى لم تأخذ بعين الاعتبار في النموذج.

4 - حساب معامل التفسير المصحح

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

$$\bar{R}^2 = 1 - \left[\frac{5-1}{5-2} (1 - 0.97) \right]$$

$$\bar{R}^2 = 0.96$$

نفس معامل التحديد المصحح بنفس تفسير معامل التحديد العادي وبالتالي 96% من التغيرات التي تحدث في المتغير التابع (Y) سببها المتغيرات المستقلة (X1) و (X2)، وأما 0.04% من التغيرات التي تحدث في المتغير التابع (Y) سببها متغيرات أخرى لم تأخذ بعين الاعتبار في النموذج.

5 - اختبار مدى ملائمة النموذج السابق في تفسير العلاقة بين Y و X1 و X2

- وضع الفرضيات:

H0: النموذج الحالي غير ملائم لتمثيل العلاقة بين المتغير التابع الكمية (Y) من جهة، والمتغيرات المستقلة السعر (X1) والدخل (X2) من جهة أخرى.

H1: النموذج الحالي ملائم لتمثيل العلاقة بين المتغير التابع الكمية (Y) من جهة، والمتغيرات المستقلة السعر (X1) والدخل (X2) من جهة أخرى.

- حساب F قيمة الحرجة (الجدولية):

F _{tab}	↗	درجة حرية البسط	1
	→	درجة حرية المقام	3
	↘	مستوى المعنوية	5%

$$F_{\text{tab}} = 10.1$$

- حساب F قيمة القياسية

$$F_{\text{CAL}} = \frac{\frac{R^2}{K-1}}{\frac{1-R^2}{n-K}}$$

$$F_{\text{CAL}} = \frac{\frac{0.97}{2-1}}{\frac{1-0.97}{5-2}}$$

$$F_{\text{CAL}} = 97$$

- قاعدة القرار:

بما أن قيمة F القياسية (F_{cal}) أكبر من قيمة F الجدولية (F_{tab}) فإننا نرفض الفرضية البديلة ونقبل الفرضية البديلة والتالي النموذج المحصل ملائم لتمثيل العلاقة بين المتغير التابع الكمية (Y) والمتغيرات المستقلة السعر (X1) والدخل (X2).

V- مشاكل القياس

V-1- الارتباط الذاتي للأخطاء:

كما رأينا في نموذج الانحدار الخطي البسيط فإن تطبيق النموذج يتطلب تحقق الفرضيات التالية:

H_1 - النموذج خطي (وجود علاقة خطية بين المتغير التابع والمتغير المستقل).

H_2 - عدم وجود أخطاء في قياس مشاهدات المتغير المستقل x_t .

$$E(\varepsilon_t) = 0 \text{ و } E(\varepsilon_t^2) = \sigma_\varepsilon^2 \quad -H_3$$

$$E(\varepsilon_t, \varepsilon_{\hat{t}}) = 0 \text{ si } t \neq \hat{t} \quad -H_4$$

$$\text{Cov}(x_t, \varepsilon_t) = 0 \quad -H_5$$

مواصفات التباين-التباين المشترك للخطأ العشوائي هي:

$$\Omega_\varepsilon = E(\varepsilon\varepsilon'_{t'}) = \begin{bmatrix} E(\varepsilon_1, \varepsilon_1) & E(\varepsilon_1, \varepsilon_2) & \dots & E(\varepsilon_1, \varepsilon_n) \\ E(\varepsilon_2, \varepsilon_1) & E(\varepsilon_2, \varepsilon_2) & \dots & E(\varepsilon_2, \varepsilon_n) \\ \dots & \dots & \dots & \dots \\ E(\varepsilon_n, \varepsilon_1) & E(\varepsilon_n, \varepsilon_2) & \dots & E(\varepsilon_n, \varepsilon_n) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \dots & 0 \\ 0 & \sigma_\varepsilon^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_\varepsilon^2 \end{bmatrix}$$

عندما لا تتحقق الفرضية H_4 أعلاه يصبح $E(\varepsilon\varepsilon'_{t'}) = \Omega_\varepsilon \neq \sigma_\varepsilon^2 I$ وبالتالي

$\text{Cov}(\varepsilon_t, \varepsilon_{\hat{t}}) \neq 0$ ، والمقدرات المستخرجة بالاعتماد على طريقة المربعات الصغرى تكون غير متحيزة

ولكنها لم تعد للحد الأدنى من التباين أي:

$$\Omega_{\hat{a}} = E\{(\hat{a} - a)(\hat{a} - a)'\} = (x'x)^{-1}x'E(\varepsilon\varepsilon')x(x'x)^{-1}$$

$$= (x'x)^{-1}(x'\Omega_\varepsilon x)(x'x)^{-1}$$

معنى هذا أن \hat{a} هو مقدر للقطر الأول لمصفوفة التباين والتباين المشترك وأكبر من $\sigma_\varepsilon^2(x'x)^{-1}$

ومن هنا يمكن طرح ثلاث أسئلة

- كيف يمكننا تحديد مقدر جديد ل a ؟

- كيف يمكننا الكشف عن الارتباط الذاتي للأخطاء؟

- ما هي الطريقة المستعملة في ذلك ؟

V-1-1- اختبارات الكشف عن الارتباط الذاتي للأخطاء:

هناك العديد من الاختبارات التي يمكن الاعتماد عليها لكشف عن وجود ارتباط ذاتي بين حدود

الخطأ العشوائي أهمها:

أولاً: إختبار 1950 Durbin watson

هو أحد أشهر الاختبارات المستعملة في الكشف عن الارتباط الذاتي بين الأخطاء من الدرجة

الأولى بالاعتماد على الشكل التالي¹⁶:

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t \quad \text{avec} \quad v_t \rightarrow N(0, \sigma_v^2)$$

فرضيات الاختبار تعرض على الشكل التالي:

$$H_0: \rho = 0$$

(مع $\rho > 0$ و $\rho < 0$ العتبة ليست هي نفسها)

لاختبار الفرضية العدمية H_0 ، نقوم بحساب إحصائية Durbin و watson

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

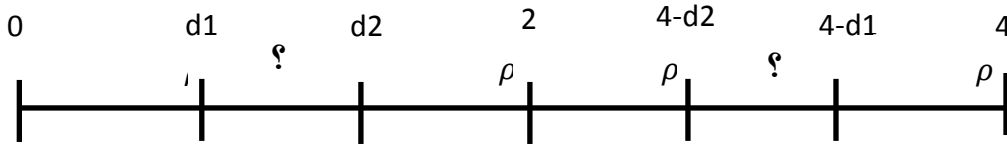
مع:

e_t : هي البواقي عند تقدير النموذج

إذا كانت n عند مستوى معنوية 5% وعدد المتغيرات المستقلة k نقوم بتحديد القيمتين d_1 و d_2

المتضمنة أربع مجالات حسب الشكل التالي:

¹⁶ Régie bourbonnais, op cite, P.129.



قاعدة القرار:

- $\rho = 0$ وبالتالي: $d_2 < DW < 4 - d_2$ نقبل الفرضية H_0
- $\rho > 0$ وبالتالي: $0 < DW < d_1$ نرفض الفرضية H_0
- $\rho < 0$ وبالتالي: $4 - d_1 < DW < 4$ نرفض الفرضية H_0
- $d_1 < DW < d_2$ أو $4 - d_2 < DW < 4 - d_1$ نحن في منطقة عدم التحديد أو منطقة مشكوك فيها وبالتالي اعتماد على هذا الاختبار في هذه الحالة قد يؤدي إلى نتائج مضللة.

ثانياً: إختبار **Breusch-Godfrey**:

يعتمد هذا الاختبار على أسس اختبار فيشر (fisher)، يستعمل لكشف الارتباط الذاتي بين الأخطاء من رتبة أكبر من الواحد، الفكرة العامة لهذا الاختبار هو البحث عن علاقة تفسيرية بين بواقى معادلة الانحدار ونفس البواقى متأخرة¹⁷.

تكتب علاقة الارتباط الذاتي بين الأخطاء من الرتبة P بالعلاقة التالية:

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

أما النموذج العام للأخطاء المرتبطة ذاتياً من الرتبة P هو:

$$y_t = a_1 x_{1t} + a_2 x_{2t} + \dots + a_k x_{kt} + a_0 + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \dots + \rho_p \varepsilon_{t-p} + v_t$$

يتم إجراء هذا الاختبار من خلال ثلاث مراحل:

¹⁷ Rumana rois, muhd muzibur rahman, tapati basak, ajit majumder, **Modified breusch-godfrey test for restricted higher order autocorrelation in dynamic linear model- a distance based approach**, Researchgate, international journal of business and management, august 2012. P.90

- تقدير النموذج بالاعتماد على طريقة المربعات الصغرى الاعتيادية وحساب البواقي e_t بسبب أن الأخطاء غير معروفة لان الاختبار متعلق بالبواقي.

- تقدير المعادلة الوسيطة باستعمال طريقة المربعات الصغرى

$$y_t = a_1x_{1t} + a_2x_{2t} + \dots + a_kx_{kt} + a_0 + \rho_1e_{t-1} + \rho_2e_{t-2} + \dots + \rho_p e_{p-2} + v_t$$

- اختبار الفرضيات للمعادلة الوسيطة:

H_0 : هي الفرضية العدمية وتنص على عدم وجود ارتباط ذاتي بين الأخطاء أي:

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

إذا قمنا برفض الفرضية العدمية، معناه إننا نواجه خطر الارتباط الذاتي بين الأخطاء عند الرتبة P .

لإجراء هذا الاختبار، لدينا إمكانيتان وهما إما استعمال اختبار فيشر العادي (fisher)، أو حساب

الإحصائية LM التي تخضع إلى توزيع ذي الحدين بدرجة حرية P

إذا كان $n \times R^2 > \chi^2(P)$ عند مستوى معنوية α نرفض فرضية استقلالية الأخطاء.

ثالثاً: إختبارات **Box-Pierce** و **Ljung-Box**: تعتبر أشهر إختبارات المستعملة في الكشف

عن الارتباط الذاتي للأخطاء وهي¹⁸:

1- إختبار Box-Pierce (1970): تم تطوير هذا الإختبار من قبل الباحثان Box وPierce

سنة 1970 ويعرف عموماً بإحصائية Q حيث إذا كان الوسط الحسابي مساوياً للصفر فإن التوزيع

المقارب للارتباطات الذاتية r_k هو التوزيع الطبيعي الذي وسطه الحسابي صفر وتباينه $\frac{n-k}{n(n+2)}$ وتحسب

إحصائية B-P بالصيغة التالية:

¹⁸ ظافر رمضان مطر، وسن رعد ذنون، مقارنة بين الإختبارات الـ **Portmaunteau** باستخدام المحاكاة، المجلة العراقية للعلوم الإحصائية (22) 2012، ص.29.

$$Q_{B-P} = n \sum_{k=1}^m r_k^2 \sim \chi_{m-p-q}^2$$

حيث:

n: حجم العينة.

r_k : معاملات الارتباط الذاتي المقدرة.

k: الارتدادات الزمنية وتأخذ القيم من 1 إلى m.

2- إختبار Ljung-Box (1978): جاء هذا الإختبار ليعالج النقص الذي عرفه إختبار Box

وPierce الذي لوحظ أن تطبيق إحصائية B-P تعطي قيم منخفضة جدا تدعو إلى الريبة فحاول Ljung و Box إيجاد إختبارات أكثر كفاءة وعليه ادخلا تعديلا بسيط على إختبار B-P وفق

الشكل التالي:

$$r_k \sim \left[0, \frac{(n-k)}{n(n+2)} \right]$$

$$\sqrt{\frac{(n-k)}{n(n+2)}} r_k \sim N(0,1)$$

بالتربيع:

$$\frac{(n-k)}{n(n+2)} r_k^2 \sim \chi_{(1)}^2$$

بإدخال مجموع m من الارتباطات الذاتية بعد تقدير المعلمات p و q للنموذج فإن

$$\left[n(n+2) \sum_{k=1}^m \frac{r_k^2}{(n-k)} \right] \sim \chi_{m-p-q}^2$$

وهذا ما يعرف باسم إحصائية L-B أو Q_{L-B}

$$Q_{L-B} = n(n+2) \sum_{k=1}^m \frac{r_k^2}{(n-k)} \sim \chi_{m-p-q}^2$$

بعد هذا التعديل وجد الباحثان أن مستوى معنوية الإحصائية المحصل عليها أكثر قربا من القيم النظرية وتوزيعها أكثر قربا من توزيع χ^2 .

يتم مقارنة القيمة الحسابية لـ Q_{L-B} الجدولية لـ χ^2 بدرجة حرية $(m-p-q)$. فإذا كانت القيمة المحسوبة أصغر من القيمة الجدولية فهذا يعني أن البواقي أو الأخطاء تشويش أبيض والعكس صحيح فإذا كانت القيمة المحسوبة أكبر من القيمة الجدولية فهذا يعني أن البواقي أو الأخطاء ليست تشويش أبيض.

رابعا: مجموعة إختبارات أخرى:

رغم أن الإختبارات سالفة الذكر هي الأكثر شيوعا واستعمالا في الحالات العملية في الكشف عن الارتباط الذاتي للأخطاء، إلا أن هناك إختبارات أخرى أكثر حداثة ومنها إختبار Li-)L-M (Macleod, 1981)، بالإضافة إلى إختبار MT (Monti, 1994)، وإختبار H-M (Huitema-Mcleod, 2007). والتي جاءت كلها كامتداد وتطوير لإختبار L-M.

V-2 - عدم تجانس التباين:

كما أشرنا سابقا فإن مصفوفة التباين والتباين المشترك للأخطاء عند عدم تحقق فرضية تجانس التباين عرف كالتالي:

$$\Omega_{\varepsilon} = E(\varepsilon \varepsilon'_{t}) = \begin{bmatrix} \sigma_{\varepsilon}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\varepsilon}^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_{\varepsilon}^2 \end{bmatrix} \neq \sigma_{\varepsilon}^2 I$$

وهنا يمكن القول بأن تباين الأخطاء مرتبط بقيم المتغير المستقل، وعدم تجانس هذا التباين يترتب عنه العديد من الآثار والتي تتمثل في:

تبقى المعالم المقدرة بطريقة المربعات الصغرى متصفة بعدم التحيز والاتساق.

تصبح التباينات المقدرة وكذلك التباينات المشتركة الخاصة بالمعالم المقدرة متحيزة وغير متسقة، ولذا فإن اختبارات الفرضيات لا تصبح دقيقة أو ملائمة. وبالرغم من التنبؤات القائمة على أساس المعالم المقدرة، باستخدام المربعات الصغرى العادية تضل غير متحيزة، إلا أنها تفقد صفة الكفاءة، وهو ما يعني أنها تكون أقل مصداقية من التنبؤات الأخرى.

V-2-1- إختبارات الكشف عدم تجانس تباين الخطأ:

على عكس الارتباط الذاتي للأخطاء، لا توجد منهجية موحدة لتصحيح عدم ثبات تباين هذه الأخطاء، ويستعمل الباحث الإحصائي العديد من الاختبارات للكشف عن هذا المشكل، ومن أهم هذه الاختبارات نجد¹⁹:

أولاً: إختبار Goldfeld-Quands:

بافتراض النموذج التالي:

$$Y_t = \beta_0 + \beta_1 x_t + \varepsilon_i, \quad t = 1, 2, \dots, n$$

يمكن تبيان كيفية استخدام اختبار *Goldfeld-Quands* في اكتشاف عدم تباين الخطأ من خلال الخطوات التالية:

❖ ترتيب مشاهدات X ترتيباً تصاعدياً

❖ استبعاد المشاهدات الوسطى لكل من X و Y، ثم تكوين مجموعتين من المشاهدات بحيث يكون

لكل مجموعة على حدا معادلة خاصة بها كما يلي:

1 - المجموعة الأولى: وتمثل في المشاهدات الخاصة بكل من X و Y الواردة قبل المشاهدات التي

تم استبعادها والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{1t} = a + bx_{1t} + \varepsilon_{1t}$$

2 - المجموعة الثانية: وتمثل في المشاهدات الخاصة بكل من X و Y الواردة بعد المشاهدات التي

تم استبعادها والمعادلة الخاصة بهذه المجموعة هي:

¹⁹ شبيخي محمد، طرق الاقتصاد القياسي محاضرات وتطبيقات، دار الحامد للنشر والتوزيع، الطبعة الأولى، 2012، ص. 114-115.

$$Y_{2t} = a + bx_{2t} + \varepsilon_{2t}$$

❖ تقدير معاملات المعادلتين السابقتين باستعمال طريقة المربعات الصغرى:

$$\hat{Y}_{1t} = \hat{a} + \hat{b} x_{1t}$$

$$\hat{Y}_{2t} = \hat{c} + \hat{d} x_{2t}$$

❖ الحصول على القيم المقدرة لحد الخطأ

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

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

❖ إيجاد القيمة المحسوبة للإحصائية F كما يلي:

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

❖ إيجاد درجات الحرية:

$$DF = \frac{n - m - 2(k + 1)}{2}$$

حيث k عدد المتغيرات المستقلة و m عدد المشاهدات المستبعدة.

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

❖ مقارنة القيمة المحسوبة ل F والقيمة الجدولية لها.

✓ إذا كانت قيمة F المحسوبة أكبر من قيمة F الجدولية نرفض فرضية العدم ونقبل الفرضية البديلة التي تنص على عدم تجانس أو ثبات التباين.

✓ إذا كانت قيمة F المحسوبة أصغر من قيمة F الجدولية نقبل فرضية العدم التي تنص على تجانس أو ثبات التباين.

ملاحظة: اختبار لا يمكن تطبيقه إلا في حالة ما إذا كانت إحدى المتغيرات المستقلة هي المسببة لمشكلة

عدم تجانس تباين حد الخطأ.

ثانياً: اختبار **white**:

إقترح White(1980) اختباراً يعتمد على العلاقة بين مربعات البواقي وجميع المتغيرات المستقلة وكذا مربعاتها. ويمكن إبراز خطوات هذا الاختبار كما يلي:

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

❖ تقدير المعادلات الوسيطة التالية

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

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2

❖ فرضية ثبات تباين الخطأ التي ينبغي اختبارها هي:

$$H_0: \beta_0 = \alpha_1 = \beta_1 = \dots = \alpha_K = \beta_K = 0$$

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

$n \times R^2$ أكبر من $\chi^2(2K)$ (القيم الحرجة لتوزيع χ^2 بمستوى معنوية α) فإننا نرفض H_0 أي إذا

كان هناك على الأقل معامل واحد من معاملات المعادلة الوسيطة يختلف معنوياً عن الصفر فإن تباين الأخطاء غير متجانس.

ثالثاً: 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 الاختبار كالتالي:

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

❖ تقدير المعادلة التالية:

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

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

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

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

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

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

V-3- التعدد الخطي (Multicolinearty):

V-3-1- تعريف التعدد الخطي (Multicolinearty)

هي كلمة مركبة من ثلاث أجزاء (Multi متعدد، Co)، مشترك أو متداخل أو مرتبط، (Linearty) خطي، وهو عبارة عن وجود علاقة خطية بين المتغيرات المستقلة في النموذج المدروس (X_1, X_2, \dots, X_n) ، وهو ما يتعارض مع إحدى الفرضيات التي يقوم عليها الانحدار الخطي المتعدد، حيث يفترض في هذا الأخير أن

تكون أن المتغيرات المستقلة تكون مستقلة خطياً أي لا يوجد بينها علاقة.

3-2- أنواع العلاقات الخطية:

هناك نوعين من العلاقات الخطية هما²⁰:

أولاً: العلاقة الخطية التامة:

تتحقق هذه العلاقة عندما تكون هناك علاقة بين قيم اثنين أو أكثر من المتغيرات التوضيحية

فتكون نتيجة محدد المصفوفة تساوي صفر $|X'X| = 0$ وهذا يؤدي إلى انتهاك شرط الرتبة Rank

²⁰ دجلة ابراهيم مهدي، نذير عباس إبراهيم، الاقتصاد القياسي أسلوب كمي باستعمال SPSS، EViews، MINITAB، الطبعة الأولى، 2014.

الذي يعرف بأنه رتبة المصفوفة عندما المحدد لا يساوي صفر إذ تكون المصفوفة غير كاملة الرتبة أي أن $rank(x) < P$ الرتبة أقل من عدد المتغيرات وعليه لا يمكن إيجاد معكوس المصفوفة وبالتالي لا يمكن تقدير معاملات النموذج وهنا يظهر التعدد الخطي التام.

ثانياً: العلاقة الخطية الجزئية:

تظهر هذه الحالة عندما يكون محدد مصفوفة المعلومات لا تساوي صفر $|X'X| \approx 0$ وإنما قريب منه. وتميل المتغيرات للتحرك سوية بالزيادة أو النقصان كما في حالة استخدام المتغيرات المرتدة ففي هذه الحالة يكون تقدير معاملات النموذج غير دقيق وغير ممثل للمشكلة المدروسة في الواقع لأن تباين المعلمات سيكون كبير نتيجة لكبر حجم الأخطاء المعيارية وهذا يؤثر على ظهور بعض أو جميع قيم الإحصاءة T-test لمعاملات النموذج صغيرة نسبياً وهذا يؤدي إلى عدم معنوية هذه مقدرات وإن كانت قيمة معامل التحديد R^2 للنموذج كبير.

V-3-2 أسباب التعدد الخطي

ينشأ التعدد الخطي من عدة أسباب منها ما يلي:

- ❖ اتجاه المتغيرات الاقتصادية معا للتغير مع مرور الزمن. فبمرور الزمن سوف تتزايد المتغيرات الاقتصادية التالية معا: الدخل، الاستهلاك، الادخار، الاستثمار، المستوى العام للأسعار والعمالة، مع أم هناك ارتباط بين هذه المتغيرات فإن التعدد الخطي سوف يتحقق.
- ❖ استخدام متغيرات مستقلة ذات فترة إبطاء في المعادلة المراد تقديرها: فالدخل في الفترة الزمنية الحالية يتحدد جزئياً بواسطة قيمته في الفترة الزمنية السابقة، مع وجود ارتباط بين القيم المتتالية لمتغير ما فإن التغير الخطي سوف يتحقق.
- وفي وجود التعدد الخطي سوف يترتب عنه:
- ❖ زيادة التباين والتباين المشترك للمقدرات بدرجة كبيرة دون التأثير على التنبؤات المستمدة من الانحدار.
- ❖ القيم المقدرة للمعاملات الانحدار سوف تكون غير محددة وغير دقيقة.

$C_{jj} \sigma^2$ يمكننا النظر إلى C_{jj} على أنها عامل يؤدي إلى زيادة التباين نتيجة لقرنها من الاعتماد خطيا على بعض المتغيرات التفسيرية المضمنة في النموذج، والمعادلة أدناه تقيس ما يسمى بعامل تضخم التباين.

$$VIF = C_{jj} = 1/(1 - R_j^2)$$

حيث:

VIF : عبارة عن عامل تضخم التباين للمتغير التفسيري j .

R_j^2 : معامل تحديد نموذج الحدار المتغير التفسيري j على المتغيرات التفسيرية المتبقية (P-1)

مقياس التحمل (Tolérance):

كذلك يتم استخدام معكوس عامل تضخم التباين في بعض حزم البرامج الإحصائية الحاسوبية ويعرف بمقياس التحمل $Tolérance$ للكشف عن وجود التعدد الخطي بين المتغير X_j والمتغيرات المستقلة الأخرى. وبحسب بالعلاقة التالية:

$$Tolérance = \frac{1}{VIF_j} = 1 - R_j^2$$

وقيم التحمل الحدية الأدنى المستخدمة في هذه البرامج لدخول المتغير المستقل النموذج هي (0.01، 0.001، 0.0001).

ثانيا: القيم المميزة الكامنة Eigenvalues:

تعتبر القيم المميزة لمصفوفة معاملات الارتباط بين معلمات نموذج الانحدار مؤشرات جيدة لقياس التعدد الخطي، ولإيجاد القيم المميزة يتم تنفيذ الخوارزمية التالية:

■ إيجاد المصفوفة Z حيث:

$$Z = (X'X)$$

■ إيجاد المصفوفة القطرية S بحيث:

$$S = \text{diag}(X'X)^{-\frac{1}{2}}$$

■ إيجاد المصفوفة التي تكون عناصرها القطرية عبارة عن واحد صحيح:

$$SZS = S(X'X)S$$

يتم حساب القيم المميزة للمصفوفة SZS ، وإذا كانت القيمة المميزة قريبة من الصفر دل ذلك على وجود تعدد خطي عال، أما إذا كانت قيمة أحد القيم المميزة تساوي الصفر دل ذلك على تام خطي تعدد وجود $\text{Perfect Multicollinearity}$.

مؤشر الحالة (CI) Condition index

يستخدم مؤشر الحالة للكشف عن وجود التعدد الخطي، ويتم حسابه بواسطة القيم المميزة وفقا

للصيغة:

$$CI_j = \sqrt{\frac{\lambda_{\max}}{\lambda_j}}$$

حيث:

λ_{\max} : أكبر قيمة مميزة

λ_j : القيمة المميزة رقم j

رقم الحالة (CN) Condition Number

يستخدم رقم الحالة أيضا للكشف عن وجود التعدد الخطي ويحسب باستخدام القيم المميزة وفقا

للصيغة:

$$CN = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$$

إقترح Johnston أن قيمة CN إذا كانت ما بين 20 إلى 30 فإنها تكون مؤشراً لوجود تعدد خطي مرتفع، بينما اقترح Belsley وآخرون إذا كانت قيمة CN ما بين 30 إلى 100 يعتبر ذلك مؤشراً إلى تعدد خطي مرتفعا جدا.

VI- السلاسل الزمنية:

يعتبر تحليل السلاسل الزمنية من الموضوعات الهامة لرجل الأعمال والمسؤولين والمخططين والباحثين في المجالات المختلفة وبصفة خاصة في المجال الاقتصادي، وترجع تلك الأهمية إلى أن تحليل السلاسل الزمنية يتعرض لتفسير وتحليل حدث في ماضي لمعرفة العوامل التي تؤثر في المتغيرات المختلفة مما يمكن من التنبؤ بما قد يحدث في المستقبل. وترتبط الظواهر الاقتصادية بالزمن حيث ترتبط به ارتباطاً وثيقاً حيث يعتبر أحد أبعاد الظاهرة. فقولنا مثلاً أن ظاهرة ما تتغير عبر الزمن فهذا يعني أن الزمن يؤثر على هذه الظاهرة ومن هنا يمكن اعتباره متغير مستقل يفسر الظاهرة محل الدراسة.

VI-1- تعريف السلسلة الزمنية:

هي مجموعة من القيم الخاصة بمؤشر ما مأخوذة خلال فترات زمنية متتالية وهي تعكس تطور ذلك المؤشر عبر الزمن. كل قيمة حد (y_t) من حدود السلسلة الزمنية يتشكل نتيجة لتفاعل عدد كبير من العوامل المؤثرة في الظاهرة المدروسة والتي يمكن اصطلاحاً تقسيمها إلى أربع مجموعات²²:

- 1- العوامل التي يؤدي تفاعلها إلى تكوين اتجاه عام لمسار تطور السلسلة.
- 2- العوامل التي تنشأ عنها التقلبات الموسمية في السلسلة.
- 3- العوامل التي تؤدي إلى تكوين التقلبات الدورية في السلسلة.
- 4- العوامل ذات التأثير العشوائي على قيم السلسلة.

VI-2- أنواع السلاسل الزمنية:

يمكن تقسيم السلاسل الزمنية عدة أنواع منها:

سلاسل زمنية عقدية (عشرية): وتحدث أو تصمم أو تدرس لكل عشرة سنوات مثل التعدادات السكانية والزراعية والصناعية.

²² مكيد على، مرجع سبق ذكره، ص. 279.

سلاسل زمنية سنوية: حيث تكون المشاهدات سنوية كإنتاج الحبوب والفواكه والدخل القومي والمبيعات والإنتاج الصناعي.

سلاسل زمنية فصلية: تكون فيها المشاهدات فصلية مثل استهلاك المشروبات الغازية والملابس القطنية أو الصوفية أو استهلاك الكهرباء والعصائر.

سلاسل زمنية شهرية: تكون البيانات المشاهدة فيها شهرية مثل الإحصاءات السياحية وغيرها.

سلاسل زمنية أسبوعية: وتتضمن بيانات أسبوعية مثل أسعار الصرف والأسهم والسندات وغيرها.

VI-3- أهداف تحليل السلاسل الزمنية:

تتيح دراسة سلسلة زمنية تحليل الظاهرة ووصفها وشرحها بمرور الوقت واستخلاص العواقب على عملية صنع القرار. من بين الأهداف الرئيسية لدراسة سلسلة زمنية هو التنبؤ بالقيم المستقبلية لظاهرة ما إنطلاقاً من القيم المشاهدة. ومن بين الأهداف الأخرى هي تحديد المنحى أو الاتجاه الذي تسلكه هذه الظاهرة وتأثير ذلك على المتغير المدروس نفسه. ومنه يمكن أن نلخص هذه الأهداف في:

VI-3-1- النمذجة: وتتكون من

❖ تطوير نماذج لوصف سلسلة زمنية واحدة أو أكثر.

❖ وضع منهجية لـ

- تحديد

- تقدير

- التحقق

نموذج مناسب لبيانات معينة

VI-3-2- التنبؤ:

هو تحديد قيمة أو مجموعة من القيم غير المشاهدة من قيم المتغير المدروسة y_{T+h} بالاعتماد على بيانات مشاهدة (y_1, y_2, \dots, y_n) ، هذا التنبؤ يمكن أن يكون نقطي (قيمة معينة)، كما يمكن أن يكون عبارة عن فترة (مجال ثقة للتنبؤ).

VI-4- مكونات السلسلة الزمنية:

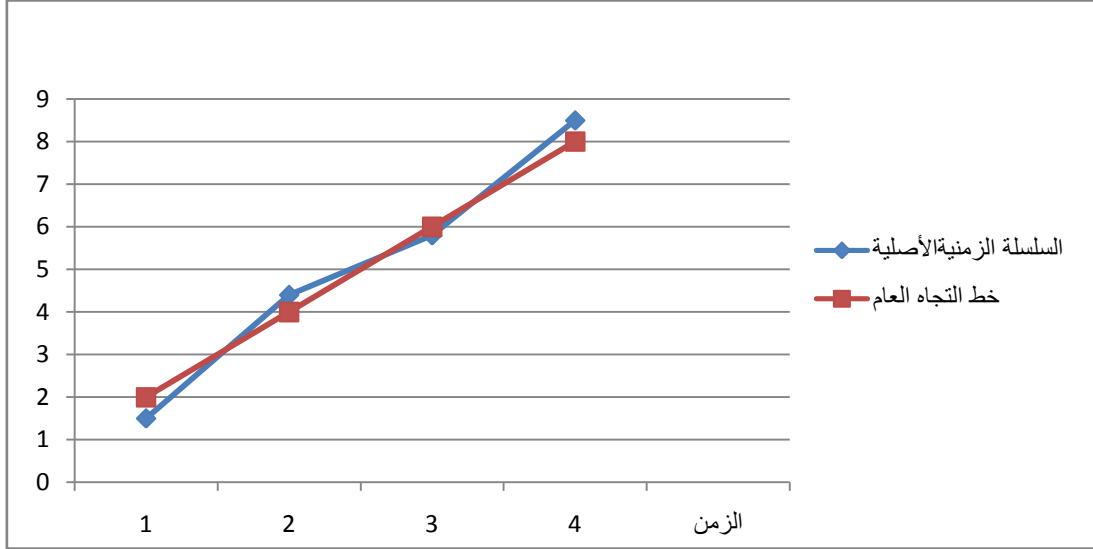
إن الهدف الأساسي لدراسة وتحليل السلاسل الزمنية هو توضيح وتحديد المكونات الهيكلية للسلسلة الزمنية (مركبات السلسلة الزمنية) وتقدير وقياس نموذج الانحدار الذي تتطور وفقه هذه السلسلة عبر الزمن وكذلك استخدام المعلومات المحصل عليها من أجل إجراء الاستطلاع والحصول على قيم تقديرية للسلسلة في المستقبل. وتتكون السلسلة من أربع مركبات أساسية هي²³:

VI-4-1- تغيرات الاتجاه العام (مركبة الاتجاه) T :

يشير الاتجاه العام إلى ميزات خصائص الحركة الممهدة للسلسلة الزمنية صعودا أو هبوطا على مدى فترة طويلة من الزمن تحتاج مثل هذه التغيرات في حدها الأدنى إلى فترة زمنية تتراوح بين 15-20 سنة لوصفها. والتي تعزي إلى عوامل مختلفة مثل التغيرات السكانية، أو تطور تقني، أو تبديلات كبيرة في سلوك الأفراد وغيرها.

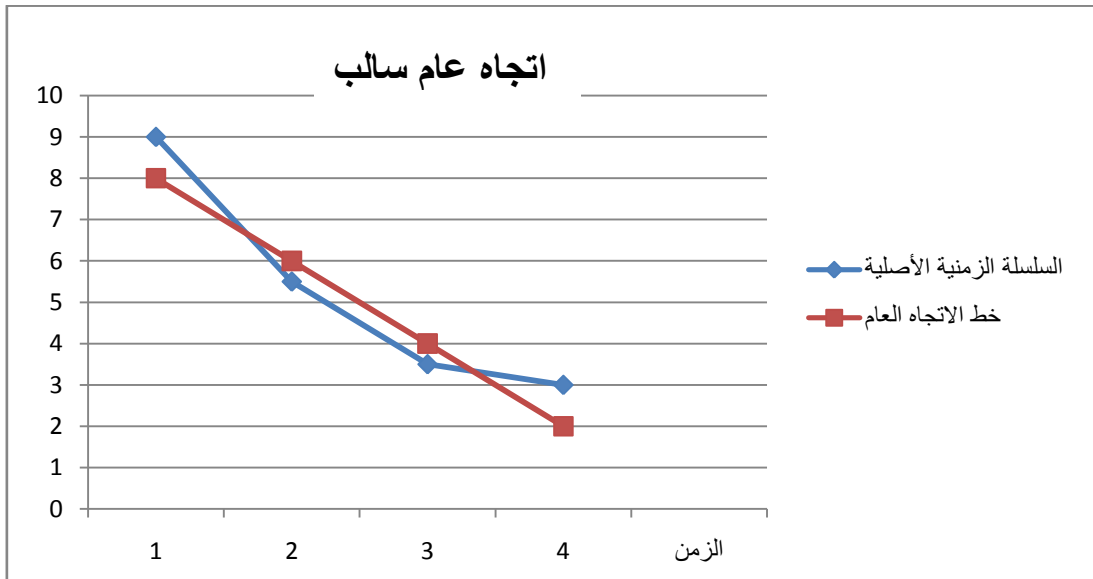
وفي العديد من الأحيان يشار إليها بالتحركات طويلة الأجل، وإن كان في العديد من الأحيان يوجد تذبذب في المنحنى التاريخي للظاهرة، ونقول أن الاتجاه العام للسلسلة موجب إذا كان الاتجاه نحو الارتفاع بمرور الزمن كما هو الحال عند دراسة النمو الديمغرافي للدول العالم الثالث. أي أن عدد السكان يتزايد بمرور الزمن، والشكل التالي يعطي اتجاه عام موجب للسلسلة زمنية.

²³ سالم عيسى بدر، عماد غصاب عباينة، مرجع سبق ذكره، ص. 189-190.



اتجاه عام موجب

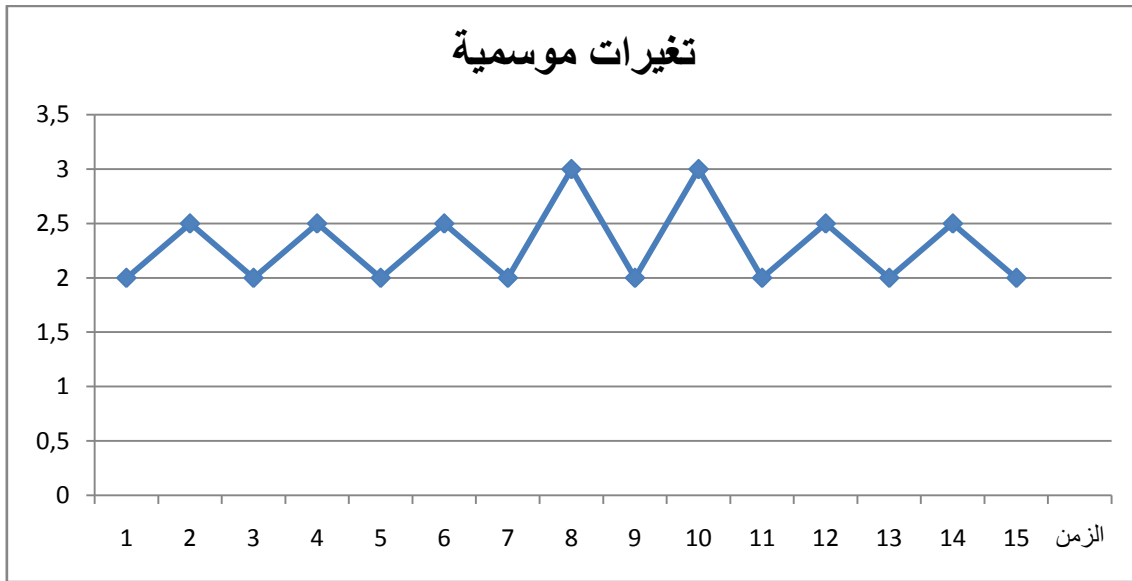
ونقول أن الاتجاه العام للسلسلة الزمنية سالب إذا كان الاتجاه نحو التناقص بمرور الزمن كما هو الحال في معدل الأمية في غالبية دول العالم، والشكل التالي يبين اتجاه عام سالب لسلسلة زمنية معينة:



VI-4-2- التغيرات الموسمية (المركبة الموسمية) S :

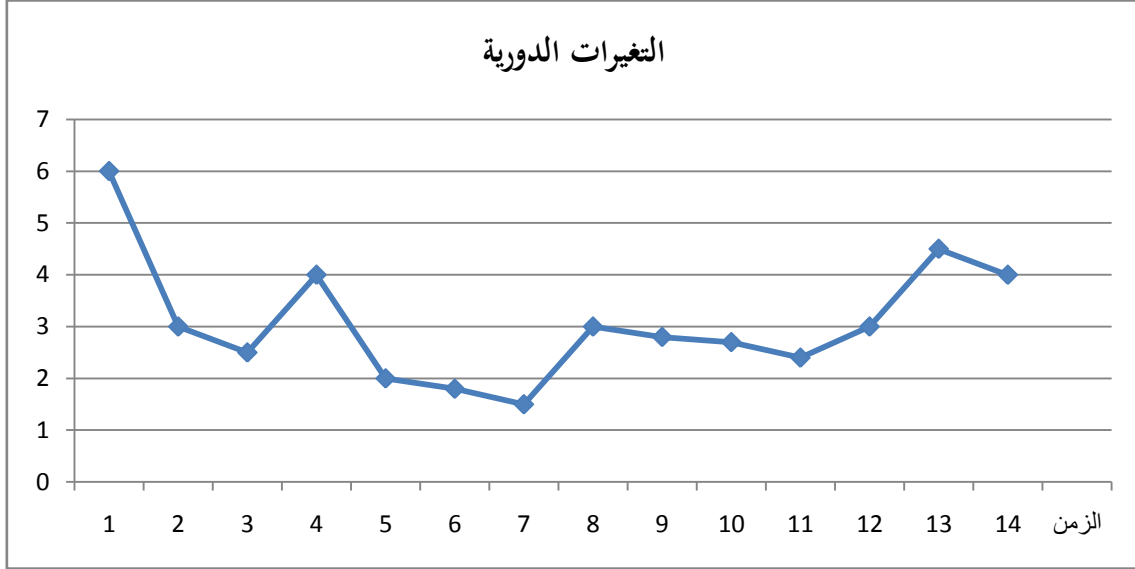
وتكون نوعاً ما من الدورة التي تكمل نفسها خلال فترة السنة التقويمية ثم تستمر في تكرار هذا النمط الأساسي. إن العوامل الأساسية في ظهور هذا النمط السنوي من التغيرات الموسمية ينشأ عادة من

العطل المختلفة وعادات المناسبات والأعياد. تستخدم سلاسل البيانات الشهرية وربيع السنوية تفسير هذه التغيرات الموسمية. فاستهلاك المرطبات يزداد في فصل الصيف أكثر من فصل الشتاء. كما أن استهلاك وقود التدفئة المنزلية يزداد في فصل الشتاء أكثر من الشهور الأخرى. توضح هاتان الحالتان تأثير عوامل الطقس أو المناخ في تحديد النماذج الموسمية. وبالتالي فإن مبيعات الملابس تزداد خلال مناسبات الأعياد مما يعكس العادات الشرائية المرتبطة بهذه المناسبات.



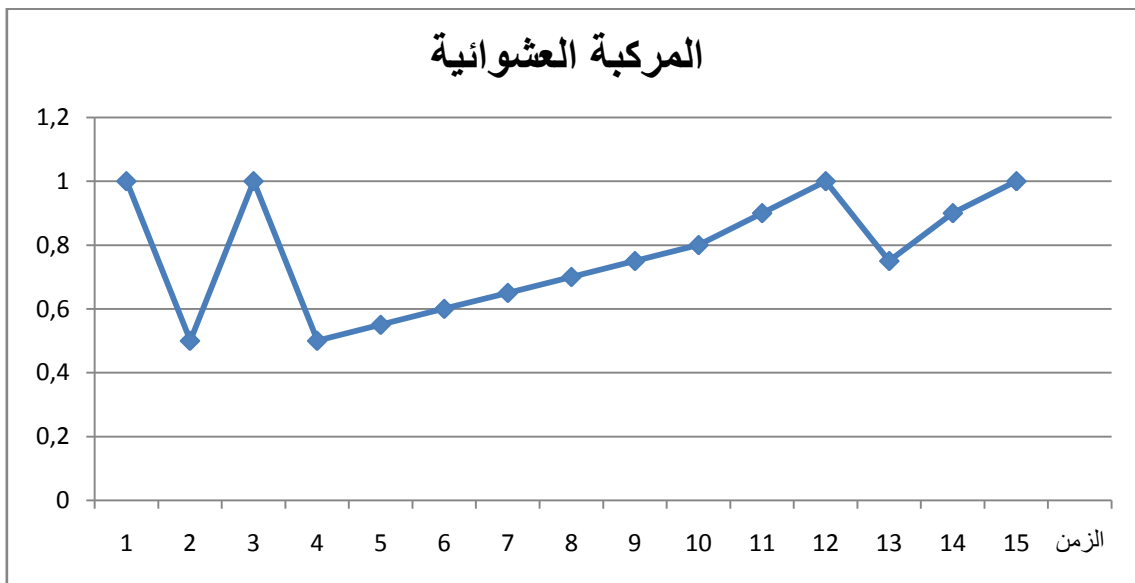
VI-4-3- التغيرات الدورية (المركبة الدورية) C :

وتشير إلى التذبذبات طويلة المدى حول خط الاتجاه العام، وقد تحدث وقد تحدث في فترات زمنية محددة قد تتراوح بين 2-12 سنة ولا يوجد تفسير واحد بسيط لنشاط مركبة الدورة وقد تكون هذه الدورات على فترات متتالية أو قد لا تكون، أي أنها قد تتبع نفس النمط من التغير بعد كل فترة زمنية متساوية وقد لا تتبع نفس النمط. وتعتبر التحركات الدورية في مجال الأعمال إذا تكررت بعد فترات زمنية تزيد عن السنة، ومن الأمثلة على ذلك دورات الأعمال التي تمثل فترات الرخاء ثم الركود ثم الكساد ثم الانتهاء من الأزمة ثم الرخاء وهكذا، وكذلك دورات الأمطار في منطقة جغرافية من حيث الهطول ثم الاعتدال ثم الجفاف ثم الاعتدال ثم كثرة الهطول مرة أخرى وهكذا.



VI-4-4- التغيرات العشوائية (المركبة العشوائية) I :

هي تحركات أو تغيرات السلسلة الزمنية تستمر لفترة قصيرة من الزمن، تكون عرضية في طبيعتها ولا تتبع أي نمط منتظم. ويشار إلى هذه التحركات أحيانا بأنها متبقية (Residual Variation) لأنها تمثل وفقا لتعريفها ما تبقى من للسلسلة الزمنية بعد أن يؤخذ في الاعتبار العناصر الاتجاهية والفصلية والدورية. وتنشأ مثل هذه التغيرات غير المنتظمة بسبب عوامل غير منتظمة كالزلازل والكوارث وأيضا الحروب وما شابه ذلك في مجال الاقتصاد وأما في قياس السمات النفسية والسلوك الاجتماعي فقد يرتبط بعوامل دخيلة أخرى.



VI-5-1- خصائص السلسلة الزمنية:VI-5-1- المتوسط:

الصيغة العامة لمتوسط سلسلة زمنية Y_t بصفة عامة عندما يكون لدينا عدد T من المشاهدات

هي:

$$\bar{Y} = \sum_{t=1}^T \frac{y_t}{T}$$

VI-5-2- التباين:

يقيس التشتت حول الوسط الحسابي، الصيغة العامة للتباين عندما يكون لدينا عدد T من

المشاهدات هي:

$$VAR(Y) = \frac{\sum_{t=1}^T (Y_t - \bar{Y})^2}{T}$$

الجذر التربيعي للتباين يسمى الانحراف المعياري $\sigma(Y)$

$$\sigma(Y) = \sqrt{VAR(Y)}$$

VI-5-3- دالة التباين المشترك:

يقيس التباين المشترك بين متغير ونفس المتغير في فترتين مختلفتين ولمرة واحدة.

$$\gamma(h) = COV(y_t, y_{t+h})^2 \rightarrow \forall h, t \in z$$

Cov: التباين المشترك

تتحقق دالة التباين المشترك في السلاسل الزمنية المستقرة من:

$$\begin{cases} |\gamma(0)| = VAR(Y_t) \\ |\gamma(h)| \leq \gamma(0) \rightarrow \forall h \\ \gamma(-h) = \gamma(h) \rightarrow \forall h \in z \end{cases}$$

يتم تقدير دالة التباين المشترك بالعلاقة التالية:

$$\hat{Y}_t(h) = \frac{1}{T} \sum_{t=1}^{T+|h|} (y_t - \bar{y})(y_{t+|h|} + \bar{y})$$

VI-5-4- دالة الارتباط الذاتي (ACF):

يقيس الارتباط بين السلسلة (Y_t) والسلسلة نفسها مزاحة بفترة زمنية h (Y_{t+h}) ، وهي معرفة

بالعلاقة التالية:

$$\rho(h) = \rho(y_t, y_{t+h}) = \frac{\gamma(h)}{\gamma(0)}$$

$\rho(h)$: تقوم بقياس الارتباط بين y_t و y_{t+h} ، وتحقق دالة الارتباط الذاتي في السلاسل الزمنية

المستقرة من:

$$\begin{cases} \rho(0) = 1 \\ |\rho(h)| \leq 1 \rightarrow \forall h \\ \rho(-h) = \rho(h) \forall h \in z \end{cases}$$

يتم تقدير دالة الارتباط الذاتي بالعلاقة التالية:

$$\hat{\rho}(h) = \frac{\bar{\gamma}(h)}{\bar{\gamma}(0)} = \frac{\sum_{t=1}^{T-|h|} (y_t - \bar{y})(y_{t+|h|} - \bar{y})}{\sum_{j=1}^T (y_t - \bar{y})^2}$$

VI-5-5- دالة الارتباط الذاتي الجزئي (PACF):

يمكننا تعريف دالة الارتباط الذاتي الجزئي للتأخير h حيث يدخل معامل الارتباط الجزئي بين y_t

و y_{t+h} تأثير المتغيرات الأخرى التي تأخرت بـ h فترة $(y_{t-1}, \dots, y_{t+h-1})$ ، بعد سحبها،

يمكن حسابها كنسبة بين محددين بالصيغة التالية:

$$r(h, h) = \frac{\begin{vmatrix} \rho(1) & 1 & \dots & \rho(h-2) & \rho(1) \\ 1 & \rho(1) & \dots & \rho(h-2) & \rho(2) \\ \vdots & & \vdots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \dots & \rho(1) & \rho(h) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \dots & \rho(h-2) & \rho(1) \\ \rho(1) & 1 & \dots & \rho(h-3) & \rho(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \dots & \rho(1) & 1 \end{vmatrix}}$$

VI-6- نماذج تحليل السلسلة الزمنية:

هنا العديد من النماذج التي تصف السلسلة الزمنية بدلالة مركباتها أشهرها نوعان:

VI-6-1- النموذج التجميعي: ويوصف هذا النموذج بالعلاقة الرياضية التالية:

$$Y = T + S + C + I$$

VI-6-2- النموذج التفاعلي (الضربي): ويوصف هذا النموذج بالعلاقة الرياضية التالية:

$$Y = T \times S \times C \times I$$

حيث:

Y: تمثل قيمة الظاهرة.

T: تمثل تأثير الاتجاه العام

S: يمثل تأثير الموسمي أو الفصلي.

C: تمثل التأثير الدوري.

I: تمثل التأثير العشوائي.

ملاحظة:

يعتبر النموذج التفاعلي هو النموذج الشائع والأكثر استخدام في تحليل السلسلة الزمنية وبالتالي سنعتمد عليه في مختلف الأمثلة والتطبيقات الموالية.

إن الزمن (t) لا يعتبر متغير مؤثر بحد ذاته، بل يظهر تأثيره بطريقة غير مباشرة، حيث يعتبر انعكاسا لمحصلة مجموعة من القوى الفعلية أثرت على الظاهرة لكنها غير مشاهدة أو منظورة في العلاقة أو لا يمكن حسابها على حدة، ولهذا يظهر الزمن كمتغير مستقل يدل على محصلة القوى الحقيقية المؤثرة على الظاهرة وغير المعروفة (معروفة أحيانا) لكنها مندمجة مع المتغيرات غير المعروفة.

VI-7- كيفية الكشف عن مكونات السلسلة الزمنية:

هناك العديد من الاختبارات التي يمكننا أن نستعملها للكشف عن مكونات السلسلة الزمنية أهمها وأكثرها استعمالا:

VI-7-1- اختبار Buys-Ballot:

يعتمد هذا الاختبار على الإحصائية فيشر ويتكون من مرحلتين أساسيتين هما²⁴:

أولا: التمثيل البياني للسلسلة الزمنية وإنشاء جدول Buys-Ballot:

ويحتوي هذا الجدول على بيانات للمبيعات المحققة خلال عدة سنوات، ومتوسط المبيعات وانحرافها المعياري لكل سنة من جهة، ومن جهة أخرى على متوسط المبيعات وانحرافها المعياري لكل فصل إذا كانت البيانات فصلية، ومتوسط المبيعات وانحرافها المعياري لكل شهر إذا كانت البيانات شهرية، وأخيرا يحتوي على المتوسط العام وانحراف المعياري العام.

ثانيا: تحليل التباين واختبار فيشر (Fisher):

ليكن لدينا:

n : عدد المشاهدات.

p : عدد المشاهدات في السنة ($p = 4$ إذا كانت البيانات فصلية، $p = 12$ إذا كانت

البيانات شهرية ... الخ).

²⁴ Régis Bourbonnais et Michel Terraza "Analyse des série temporelles en économie" presses Universitaires de France 1998 p16.

• x_{ij} : قيم السلسلة الزمنية من أجل $i = 1, 2, 3, \dots, N$ و $j = 1, 2, 3, \dots, P$

N : عدد السنوات.

لنفرض أن السلسلة الزمنية تأخذ الصيغة التالية:

$$x_{ij} = m_{ij} + e_{ij}$$

حيث أن:

• $e_{ij} \rightarrow N(0, \sigma^2)$: الخطأ العشوائي مع العلم أن

m_{ij} : العناصر المكونة للسلسلة الزمنية.

والتباين الكلي يأخذ الصيغة التالية:

$$S_T = \sum_{i=1}^N \sum_{j=1}^P (x_{ij} - x_{..})^2$$

مع:

$$x_{..} = \frac{1}{N \cdot P} \sum_{i=1}^N \sum_{j=1}^P x_{ij}$$

حيث أن:

S_T : مجموع التباين الكلي مربع.

$x_{..}$: المتوسط العام للسلسلة الزمنية.

$$x_{i.} = \frac{1}{P} \sum_{j=1}^P x_{ij} \quad : i \quad \text{متوسط السنة}$$

$$x_{.j} = \frac{1}{N} \sum_{i=1}^N x_{ij} \quad : j \quad \text{متوسط الفترة}$$

الجدول -1- تحليل التباين للكشف عن التغيرات الموسمية:

العناصر	التباين	درجة الحرية	مجموع الفروق
الفترة	$v_P = \frac{S_P}{P - 1}$	$P - 1$	$S_P = N \sum_{j=1}^P (x_{.j} - x_{..})^2$
السنة	$v_A = \frac{S_A}{N - 1}$	$N - 1$	$S_A = N \sum_{i=1}^N (x_{i.} - x_{..})^2$
البواقي	$v_R = \frac{S_R}{(P - 1)(N - 1)}$	$(P - 1)(N - 1)$	$S_R = \sum_{i=1}^N \sum_{j=1}^P (x_{ij} - x_{i.} - x_{.j} + x_{..})^2$
المجموع	$v_T = \frac{S_T}{N \cdot P - 1}$	$N \cdot P - 1$	S_T

Source: Régis Bourbonnais et Michel Terraza, op-cité , p18

-1- اختبار تأثير الاتجاه العام:

لدينا الفرضيتين التاليتين:

 H_0 : السلسلة لا تحتوي على مركبة الاتجاه العام. H_1 : السلسلة الزمنية تحتوي على مركبة الاتجاه العام.

يتم حساب الإحصائية Fisher من جدول تحليل التباين أعلاه من خلال العلاقة التالية:

$$F'_{CAL} = \frac{v_A}{v_R}$$

ومقارنتها مع قيمة F'_{TAB} الجدولية والتي تستخرج من جدول Fisher (Fisher) حيث:

درجة الحرية تعطى بالشكل التالي:

$$F'_{v_3, v_2, \alpha} \Rightarrow \begin{cases} v_3 = N - 1 \\ v_2 = (N - 1)(p - 1) \end{cases}$$

 α : مستوى المعنوية.

إذا كان $F'_{TAB} < F'_{CAL}$ نرفض الفرضية العدمية H_0 وبالتالي القرار يكون كالتالي: السلسلة الزمنية تحتوي على مركبة الاتجاه العام أو تتأثر بمركبة الاتجاه العام.

2- اختبار تأثير التغيرات الموسمية:

H_0 : السلسلة الزمنية لا تحتوي على مركبة التغيرات الموسمية.

H_1 : السلسلة الزمنية تحتوي على تغيرات مركبة التغيرات الموسمية.

يتم حساب الإحصائية Fisher من جدول تحليل التباين أعلاه من خلال العلاقة التالية:

$$F_{CAL} = \frac{v_P}{v_R}$$

ومقارنتها مع قيمة F'_{TAB} الجدولية والتي تستخرج من جدول Fisher (F) حيث: درجة الحرية تعطى بالشكل التالي:

$$F_{v_3, v_2}^{\alpha} \Rightarrow \begin{cases} v_3 = p - 1 \\ v_2 = (N - 1)(p - 1) \end{cases}$$

α : مستوى المعنوية.

إذا كان $F'_{TAB} < F'_{CAL}$ نرفض الفرضية العدمية H_0 وبالتالي القرار يكون كالتالي: السلسلة الزمنية تحتوي على مركبة التغيرات الموسمية أو تتأثر بالتغيرات الموسمية.

VI-8- إستقرارية السلسلة الزمنية:

الاستقرارية هي خاصية مهمة جدا عند دراسة وتحليل السلسلة الزمنية، هذه الخاصية ضرورية لتطبيق نظريات معينة على السببية.

VI-8-1- السلسلة المستقرة:

نقول بأن السلسلة $\{y_t\}$ بحيث $t = 1, 2, \dots, T$ مستقرة إذا وفقط إذا كان:

1- التذبذب حول المتوسط الحسابي ثابت عبر الزمن كما توضح العلاقة التالية:

$$E(y_t) = \mu \quad \forall t$$

2- التباين ثابت عبر الزمن كما توضح العلاقة التالية:

$$\text{VAR}(y_t) = \sigma_y^2 \quad \forall t$$

3- التباين المشترك بين إي قيمتين يعتمد على الفجوة الزمنية يعتمد على الفجوة الزمنية بين القيمتين وليس على القيمة الفعلية للزمن كما توضح العلاقة التالية

$$\text{COV}(y_t, y_{t+k}) = E(y_t - \mu)(y_{t+k} - \mu) = \gamma_k$$

بالنسبة للسلسلة $\{\varepsilon_t\}$:

$$\text{COV}(\varepsilon_t, \varepsilon_{t+k}) = 0, \text{VAR}(\varepsilon_t) = \sigma_\varepsilon^2, E(y_t) = \mu$$

مستقرة، وتسمى بالتشويش الأبيض (الضوضاء البيضاء)

السلسلة المستقرة يجب تكون خالية من مركبة الاتجاه العام ومن المركبة الموسمية.

VI-8-2- السلسلة غير المستقرة:

بينت العديد من الدراسات التطبيقية أن أغلب السلاسل الزمنية والتي تخص أهم المتغيرات الاقتصادية أنها غير مستقرة عند المستوى، ويرجع عدم استقرار هذه السلاسل إلى أنها يمكن أن تكون من النوع (Trend stationary) TS، أو من النوع (Difference stationary) DS.

أولاً: النموذج TS

تعتبر هذه النماذج غير مستقرة وتبرز عدم استقرارية تحديدية (Deterministic) وتكتب على الشكل التالي:

$$y_t = f(t) + \varepsilon_t$$

حيث:

$f(t)$: دالة كثيرة حدود بدلالة الزمن (خطية أو غير خطية)

ε_t : تشويش أبيض.

كما أن أكثر هذه النماذج انتشارا يأخذ شكل كثير حدود من الدرجة الأولى ويمكن كتابته على

الشكل التالي:

$$y_t = a_0 + a_1 t + \varepsilon_t$$

هذا النموذج غير مستقر لأن متوسطه $E(y_t) = y_0 + a_1 t$ مرتبط بالزمن ويمكن أن نجعله

مستقرا وذلك بتقدير معلمه a_0 و a_1 بطريقة المربعات الصغرى العادية (OLS) وذلك بطرح القيمة

$$y_t - \hat{a}_0 + \hat{a}_1 t \text{ أي } y_t - \hat{a}_0 + \hat{a}_1 t$$

ثانيا: النموذج DS:

تعتبر هذه النماذج هي الأخرى غير مستقرة وتبرز عدم استقرار عشوائية (Stochastic)

وتكتب على الشكل التالي: $y_t = y_{t-1} + a_1 + \varepsilon_t$ ويمكننا جعلها مستقرة باستعمال الفروقات

$$\Delta y_t^d = a_1 + \varepsilon_t \text{ أي:}$$

حيث:

a_1 : ثابت حقيقي.

d : درجة الفروقات.

وغالبا ما تستعمل الفروقات من الدرجة الأولى في هذه النماذج $d=1$ وتكتب على الشكل:

$$\Delta y_t = a_1 + \varepsilon_t \text{ وتأخذ هذه النماذج شكلين:}$$

-1 إذا كان $a_1 = 0$ يسمى نموذج DS بدون مشتقة ويكتب على الشكل $y_t = y_{t-1} + \varepsilon_t$

$$y_1 = y_0 + \varepsilon_1$$

$$y_2 = y_1 + \varepsilon_2 = y_0 + \varepsilon_1 + \varepsilon_2$$

... ..

$$y_t = y_0 + \sum_{i=1}^t \varepsilon_t$$

وبما أن ε_t هو تشويش أبيض فإن النموذج يسمى نموذج السير العشوائي (Random walk model).

2- إذا كان $a_1 \neq 0$ يسمى نموذج DS بنموذج السير العشوائي بالمشتقة ويكتب على الشكل $y_t = y_{t-1} + a_1 + \varepsilon_t$ ويمكن التفرقة بين النموذجين باستعمال اختبار جذر الوحدة. في هذه الحالة لدينا:

$$y_1 = y_0 + a_1 + \varepsilon_1$$

$$y_2 = y_1 + a_1 + \varepsilon_2 = y_0 + 2a_1 + \varepsilon_1 + \varepsilon_2$$

.....

$$y_t = y_0 + a_1 t + \sum_{i=1}^t \varepsilon_i$$

حيث: ε_i موزعة بشكل مستقل ومتطابق

النموذج DS بالمشتقة غير مستقر وذلك لأن توقعه الرياضي وتباينه يعتمد على الزمن:

$$E(y_t) = y_0 + a_1 t$$

$$VAR(y_t) = VAR\left(\sum_{i=1}^t \varepsilon_i\right) = \sum_{i=1}^t VAR(\varepsilon_i) = \sum_{i=1}^t \sigma_\varepsilon^2 = t\sigma_\varepsilon^2$$

أي عندما $t \rightarrow \infty$ فإن $E(y_t) \rightarrow \infty$ و $VAR(y_t) \rightarrow \infty$

VI-9- اختبارات الاستقرار (اختبارات جذر الوحدة):

هناك العديد من الاختبارات التي يمكننا استعمالها لاختبار استقرار السلاسل الزمنية تسمى باختبارات جذر الوحدة ومن أهمها نجد:

VI-9-1- إختبار ديكي فولر البسيط (Dickey-Fuller simple) 1979:

يمكن من خلال هذا الاختبار معرفة ما إذا كانت السلسلة الزمنية مستقرة أم لا، كما أنه يجعل من الممكن تحديد الطريقة الصحيحة لإضفاء طابع الاستقرار على السلسلة.

فرضيات هذا النموذج هي:

H_0 : السلسلة غير مستقرة، وتتوافق مع أشكال عدم الاستقرار التالية:

$$[1] Y_t = \phi Y_{t-1} + \varepsilon_t$$

$$[2] Y_t = \phi Y_{t-1} + c + \varepsilon_t$$

$$[3] Y_t = \phi Y_{t-1} + bt + c + \varepsilon_t$$

مع $\phi = 1$ و ε_t هو حد الخطأ العشوائي ويفترض أن يكون وسطه الحسابي معدوم وتباينه

ثابت أي: $\varepsilon_t \rightarrow N(0, \sigma_\varepsilon^2)$

$$|\phi| < 1 : H_1$$

يمكننا أيضا كتابة الفرضيات بالشكل التالي:

H_0 : السلسلة غير مستقرة، وتتوافق مع أشكال عدم الاستقرار التالية:

$$[1] \Delta Y_t = (\phi - 1) Y_{t-1} + \varepsilon_t$$

$$[2] \Delta Y_t = (\phi - 1) Y_{t-1} + c + \varepsilon_t$$

$$[3] \Delta Y_t = (\phi - 1) Y_{t-1} + bt + c + \varepsilon_t$$

مع $(\phi - 1) = 0$ و ε_t هو حد الخطأ العشوائي ويفترض أن يكون وسطه الحسابي معدوم

وتباينه ثابت أي: $\varepsilon_t \rightarrow N(0, \sigma_\varepsilon^2)$

$$|\phi| < 1 : H_1$$

إذا كانت H_0 صحيحة، إحصائية الاختبار للمقدر Φ معطاة بالعلاقة التالية:

$$t_{\hat{\Phi}} = \frac{\hat{\Phi} - 1}{\hat{\sigma}_{\hat{\Phi}}}$$

نبدأ بدراسة النموذج العام [3] وننظر إذا كان b يختلف جوهريا عن 0 أم لا. إذا كان $b \neq 0$ لا يختلف جوهريا عن الصفر نمر لدراسة النموذج [2] ونحاول معرفة إذا كان c يختلف جوهريا عن 0 أم لا، إذا كان c لا يختلف جوهريا عن 0 نقو بدراسة النموذج [1].

ملاحظة:

إذا كانت H_0 صحيحة، قيم إحصائية t (Student) للثابت ولالاتجاه يجب مقارنتهما بالقيم الموجودة في جدول Dickey-Fuller (بالنسبة لعينة أكبر من 500 مشاهدة) القيم الحرجة هي 2.78 ل 5 بالنسبة للاتجاه للنموذج [3]، 2.52 بالنسبة للثابت للنموذج [2]، و -1.95 بالنسبة للمعلمة Φ . إذا كانت H_0 صحيحة فإن السلسلة غير مستقرة، ومقدر Φ لا يتبع قانون التوزيع الطبيعي.

قواعد القرار تكون كالتالي:

- إذا كانت $t_{\Phi_1} > t_{DF}$ مع t_{DF} تمثل القيمة الحرجة التي المعطاة من خلال جدول DF ، فإننا نقبل H_1 وبالتالي معامل المتغير المفسر يختلف اختلافا جوهريا عن 0. إذا كان b يختلف اختلافا جوهريا عن 0 بالنسبة للنموذج [3] الاختبار يتوقف هما ولا نقوم بدراسة النماذج الأخرى. ونفس الشيء إذا كنا في النموذج [2] والثابت يختلف اختلافا جوهريا عن 0 فإن الاختبار يتوقف عند النموذج [2].

- إذا كانت $|t_{\Phi}| > t_{DF}$ نقبل الفرضية H_0 وبالتالي السلسلة غير مستقرة.

VI-9-2- نموذج ديكي فولر المطور (Dickey-Fuller Augmenté):

في نموذج ديكي فولر الذي قمنا بدراسته سابقا الخطأ العشوائي ε_t يفترض أن يكون عبارة عن تشويش أبيض، وبالتالي عدم وجود ارتباط ذاتي للخطأ العشوائي، لكن هذا الافتراض لا يصبح قائما في اختار ديكي فولر المطور.

فرضيات اختبار ديكي فولر المطور تعرض على النحو التالي:

H_0 : السلسلة غير مستقرة، وتتوافق مع أشكال عدم الاستقرار التالية:

$$[1] \Delta Y_t = \rho Y_{t-1} - \sum_{k=2}^p \gamma_k \Delta Y_{t-k+1} + \eta_t$$

$$[2] \Delta Y_t = \rho Y_{t-1} - \sum_{k=2}^p \gamma_k \Delta Y_{t-k+1} + c + \eta_t$$

$$[3] \Delta Y_t = \rho Y_{t-1} - \sum_{k=2}^p \gamma_k \Delta Y_{t-k+1} + bt + c + \eta_t$$

مع : $\phi = 1, \rho = 0$ و ε_t هو حد الخطأ العشوائي ويفترض أن يكون وسطه الحسابي أي :

$$\varepsilon_t \rightarrow N(0, \sigma_\eta^2)$$

$$|\phi| < 1 : H_1$$

- بالنسبة للنموذج [1]

لدينا: $Y_t = \phi Y_{t-1}$ مع $Y_t = \phi Y_{t-1} \rightarrow AR(p-1)$ لا يعتبر تشويش أبيض

$$\varepsilon_t = \sum_{i=1}^{p-1} \theta_i \varepsilon_{t-i} + \eta_t, \quad \eta_t \rightarrow N(0, \sigma_\eta^2)$$

يمكننا كتابة هذه المعادلات باستخدام عامل التأخير B حيث $BY_t = Y_{t-1}$

$$(1 - \phi B)Y_t = \varepsilon_t \text{ فنحصل على: } B^p Y_t = Y_{t-p}$$

مع

$$\varepsilon_t = \theta_1 B \varepsilon_t + \theta_2 B^2 \varepsilon_t + \dots + \theta_{p-1} B^{p-1} \varepsilon_t + \eta_t \Leftrightarrow \left(1 - \sum_{i=1}^{p-1} \theta_i B^i\right) \varepsilon_t = \eta_t$$

وبالتالي يمكننا كتابة:

$$\left(1 - \sum_{i=1}^{p-1} \theta_i B^i\right) (1 - \phi B) Y_t = \eta_t$$

$$\Leftrightarrow \left(1 - \sum_{i=1}^{p-1} \theta_i B^i\right) (Y_t - \phi Y_{t-1}) = \eta_t$$

$$\Leftrightarrow (Y_t - \phi Y_{t-1}) - \sum_{i=1}^{p-1} \theta_i B^i (Y_t - \phi Y_{t-1}) = \eta_t$$

$$\Leftrightarrow (Y_t - \phi Y_{t-1}) - \sum_{i=1}^{p-1} \theta_i Y_{t-i} + \phi \sum_{i=1}^{p-1} \theta_i Y_{t-1-i} = \eta_t$$

$$\Leftrightarrow Y_t - \phi Y_{t-1} - \theta_1 Y_{t-1} - \theta_2 Y_{t-2} - \dots - \theta_{p-1} Y_{t-(p-1)} + \phi \theta_1 Y_{t-2}$$

$$+ \phi \theta_2 Y_{t-3} + \dots + \phi \theta_{p-1} Y_{t-p} = \eta_t$$

$$\Leftrightarrow Y_t = (\phi - \theta_1) Y_{t-1} + (\theta_2 - \phi \theta_1) Y_{t-2} + \dots + (\theta_{p-1} - \phi \theta_p) Y_{p-1}$$

$$- \phi \theta_{p-1} Y_{t-p} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = [(\phi - 1)(1 - \theta_2 - \dots - \theta_{p-1})] Y_{t-1} - \sum_{i=1}^{p-1} \gamma_k \Delta y_{t-k} + \eta_t$$

بما أن لدينا:

$$\Leftrightarrow Y_t = (\phi - \theta_1) Y_{t-1} + (\theta_2 - \phi \theta_1) Y_{t-2} + \dots + (\theta_{p-1} - \phi \theta_p) Y_{p-1}$$

$$- \phi \theta_{p-1} Y_{t-p} + \eta_t$$

$$\Leftrightarrow Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_{p-1} Y_{t-p+1} + \alpha_p Y_{t-p} + \eta_t$$

بالنسبة للنموذج AR(1)

$$Y_t = \alpha_1 Y_{t-1} + \eta_t$$

$$\Leftrightarrow Y_t - Y_{t-1} = \alpha_1 Y_{t-1} - Y_{t-1} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\alpha_1 - 1)Y_{t-1} + \eta_t$$

بالنسبة للنموذج AR(2)

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \eta_t$$

$$\Leftrightarrow Y_t - Y_{t-1} = -Y_{t-1} + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_2 Y_{t-1} - \alpha_2 Y_{t-1} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\alpha_1 + \alpha_2 - 1)Y_{t-1} + \alpha_2(Y_{t-2} - Y_{t-1}) + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\alpha_1 + \alpha_2 - 1)Y_{t-1} - \alpha_2(Y_{t-1} - Y_{t-2}) + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\alpha_1 + \alpha_2 - 1)Y_{t-1} - \alpha_2 \Delta Y_{t-1} + \eta_t$$

بالنسبة للنموذج AR(3)

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\alpha_1 + \alpha_2 + \alpha_3 - 1)Y_{t-1} - (\alpha_2 + \alpha_3)\Delta Y_{t-1} - \alpha_3 \Delta Y_{t-2} + \eta_t$$

بالنسبة للنموذج AR(P)

$$\Delta Y_t = \left(\sum_{i=1}^p \alpha_i - 1 \right) Y_{t-1} - \sum_{k=1}^{p-1} \left(\sum_{i=k+1}^p \alpha_i \right) \Delta Y_{t-k} + \eta_t$$

كنا افترضنا سابقا أن:

$$\alpha_1 = \phi - \theta_1, \alpha_2 = \theta_2 - \phi\theta_1, \alpha_p = \theta_p - \phi\theta_{p-1}$$

إذا نحصل على:

$$\Delta Y_t = (\phi - \theta_1 + \theta_2 - \phi\theta_1 + \dots + \theta_{p-1} - \phi\theta_{p-2} - \phi\theta_{p-1} - 1)Y_{t-1}$$

$$- \sum_{k=1}^{p-1} \left(\sum_{i=k+1}^p \alpha_i \right) \Delta Y_{t-k} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\phi - \theta_1 + \theta_2 - \phi\theta_1 + \dots + \theta_{p-1} - \phi\theta_{p-2} - \phi\theta_{p-1} - 1)Y_{t-1} - \sum_{k=1}^{p-1} \gamma_k \Delta Y_{t-k} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = [\phi(\theta_1 - \theta_2 - \dots - \theta_{p-1}) + \theta_1 + \theta_2 + \dots + \theta_{p-1} - 1]Y_{t-1} - \sum_{k=1}^{p-1} \gamma_k \Delta Y_{t-k} + \eta_t$$

$$\Leftrightarrow \Delta Y_t = (\phi - 1)(\theta_1 - \theta_2 - \dots - \theta_{p-1})Y_{t-1} - \sum_{k=1}^{p-1} \gamma_k \Delta Y_{t-k} + \eta_t$$

وبافتراض أن $p = (\phi - 1)(\theta_1 - \theta_2 - \dots - \theta_{p-1})$

وبالتالي نحصل على النموذج الأول [1].

$$\Delta Y_t = \rho Y_{t-1} - \sum_{k=1}^{p-1} \gamma_k \Delta Y_{t-k} + \eta_t$$

أو أيضا:

$$\Delta Y_t = \rho Y_{t-1} - \sum_{k=2}^p \gamma_k \Delta Y_{t-k+1} + \eta_t$$

تحديد التأخر P في الاختبار:

يتم تحديد القيمة P باستخدام دالة الارتباط الجزئي لسلسلة الفروق ΔY_t ، بمجرد تحديد القيمة P تتبع نفس الخطوات كما في اختبار ديكي فولر البسيط، نبدأ بدراسة دلالة b في النموذج [3]، وقاعدة القرار تبقى نفسها كما في نموذج ديكي فولر البسيط. إحصائية الاختبار للمقدر ϕ هي:

$$t_{\hat{\phi}} = \frac{\hat{\phi} - 1}{\hat{\sigma}_{\hat{\phi}}}$$

VI-9-3- اختبار فيليبس بيرون (Phillips-Perron):

يأخذ هذا الاختبار بعين الاعتبار الارتباط الذاتي للأخطاء بالإضافة إلى عدم تجانسها ويستند على نفس النماذج المدروسة في اختبار ديكي فولر، غير أنه يقترح تصحيح غير معلمي للإحصائية.

يتم إجراء هذا الاختبار بالمرور بأربع مراحل هي:

1- التقدير بالاعتماد على طريقة المربعات الصغرى الاعتيادية للنماذج الثلاثة لاختبار ديكي فولر البسيط وحساب البواقي $\hat{\varepsilon}_t$.

2- تحديد ما يسمى بالتباين قصير المدى والمعرف بالعلاقة التالية:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^2$$

3- تقدير معامل التصحيح S_t^2 والذي يسمى التباين على المدى البعيد والمعرف بالعلاقة التالية:

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

مع:

$$b \approx 4 \left(\frac{n}{100}\right)^{2/9}$$

4- حساب الإحصائية فيليبس بيرون من خلال العلاقة التالية:

$$t_{\hat{\phi}}^{pp} = \sqrt{k} \times \frac{\hat{\phi} - 1}{\hat{\sigma}_{\hat{\phi}}} + \frac{n(k-1)\hat{\sigma}_{\hat{\phi}}}{\sqrt{k}}$$

مع:

$$k = \frac{\hat{\sigma}^2}{S_t^2}$$

فيليبس وبيرون سنة 1988 بينا أن هذا التصحيح غير المعلمي الذي تم إجراؤه على الإحصائية $t_{\hat{\phi}}$ لا يعدل التوزيع المقارب للإحصائية الذي يبقى متطابقاً مع ذلك الذي تمت ملاحظته في حالة اختبار ديكي فولر البسيط. ونتيجة لذلك، تبقى القيم الحرجة التي وضعها جدولي وديكي وفولر صالحة أيضاً لاختبار Phillips-Perron.

VI-10- النماذج الخطية للسلسلة الزمنية:

الهدف الرئيسي لهذه نماذج هو السماح بالتنبؤ بالتطور المستقبلي للظواهر الاقتصادية. ويستند تطويرها في مجال الاقتصاد القياسي على هذا المبدأ. لكن هناك اهتمام آخر أكثر أهمية في مجال البحث العلمي هو فهم الأهمية النظرية لهذه النماذج المختلفة. من الواضح أن هذا التفسير يعتمد على طبيعة الظاهرة المدروسة، وعلى النماذج المتاحة للباحث لمحاستها.

تتكون تشكيلة النماذج العشوائية من نماذج الانحدار الذاتي (AR)، ونماذج المتوسطات المتحركة (MA)، والنماذج المختلطة من نماذج الانحدار الذاتي ونماذج المتوسطات المتحركة (ARMA) بالإضافة إلى النماذج الممتدة (ARIMA, SARIMA)، ومن شروط استعمال هذه النماذج يجب أن تكون السلسلة الزمنية مستقرة.

VI-10-1- نموذج الانحدار الذاتي (AR) Autoregressive :

كانت أولى الدراسات الخاصة بهذا النموذج من قبل (George Udny Yule)، في مقالة [Yul27]²⁵. حيث استعمل في هذه الدراسة أول نموذج للانحدار الذاتي لنمذجة سلسلة زمنية للبقع الشمسية. ويمكن كتابة هذا النموذج من خلال علاقة خطية تربط بين المتغير التابع (y_t) مع القيم أو المشاهدات السابقة لنفس المتغير (y_{t-i}).

²⁵ U. Yule : On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. Philosophical Transactions of the Royal Society, 226, 1927.

نقول بأن التابع $(y_t) \rightarrow t \geq 0$ هو نموذج الحدار ذاتي (Auto-regressive model) من الرتبة P .AR(p).

عندما يمكن كتابته على الشكل التالي:

$$y_t = \sum_{k=1}^p \phi_k y_{t-k} + \varepsilon_t \rightarrow \{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$$

حيث:

y_t : قيمة الظاهرة في الفترة t

ϕ_k ($k = 1, 2, \dots, p$): تمثل معاملات النموذج

ε_t : تشويش أبيض (ضوضاء بيضاء)، يتبع توزيعا طبيعيا وسطه الحسابي 0 وتايته ثابت σ_ε^2 .

p : رتبة النموذج

أو بطريقة أخرى نلاحظ أنه يمكننا إعادة كتابة النموذج AR(P) مع كثير الحدود $\phi(B)$ الذي سيضعف y_t في هذه الحالة على الشكل التالي:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} = \varepsilon_t$$

$$(1 - \phi_1 B - \phi_1 B^2 - \dots - \phi_p B^p) y_t = \varepsilon_t$$

$$\phi(B) y_t = \varepsilon_t$$

حيث:

$$\phi(B) = (1 - \phi_1 B - \phi_1 B^2 - \dots - \phi_p B^p)$$

مع B يسمى معامل التأخير

VI-10-2- نموذج المتوسطات المتحركة (MA) : Moving Average

تم تقديم هذا النموذج لأول مرة من قبل Eugen Slutsky سنة 1927 في مقاله [Slu27]²⁶. وذلك من خلال منهجية تعتمد على المتوسطات المتحركة $MA(q)$ ، يمكن كتابة هذا النموذج من خلال علاقة خطية تربط بين المتغير التابع (y_t) مع القيم الخطأ العشوائي (ε_{t-i}) .

نقول بأن التابع $t \geq 0 \rightarrow (y_t)$ هو نموذج متوسط متحرك من الرتبة q عندما يمكن كتابته على الشكل التالي:

$$y_t = \sum_{k=1}^q \theta_k \varepsilon_{t-k} + \varepsilon_t \rightarrow \{\varepsilon_t\} \dots \rightarrow N(0, \sigma_\varepsilon^2)$$

حيث:

y_t : قيمة الظاهرة في الفترة t

$\theta_k (k = 1, 2, \dots, q)$: تمثل معاملات النموذج ويمكن أن تكون موجبة كما يمكن أن تكون سالبة.

ε_t : تشويش أبيض (ضوضاء بيضاء)، يتبع توزيعا طبيعيا وسطه الحسابي 0 وتايته ثابت σ_ε^2

أو بطريقة أخرى:

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

$$y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

$$y_t = \theta(B) \varepsilon_t$$

حيث:

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

مع B يسمى معامل التأخير

²⁶ E. Slutsky : **The summation of random causes as the source of cyclic processes.** Econometrica, 5:105-146, 1927.

VI-10-3- النمذج المختلطة (ARMA) Autoregressive Moving Average :

هي عبارة عن نماذج مركبة نحصل عليها بدمج نموذجين مختلفين هما نموذج الانحدار الذاتي ونموذج متوسطات متحركة، ظهرت على يد Herman Wold سنة 1938 [Wol38]²⁷. ففي العديد من الحالات يعتبر نموذج الانحدار الذاتي أو نموذج المتوسطات المتحركة مثالي، لكن هناك العديد من الحالات التي نحتاج فيها لتقدير عدد كبير من المعلمات لتتناسب مع النموذج المطروح في حالة وجود عدد قليل من المشاهدات. وهنا تميل هذه التقديرات إلى أن تكون غير دقيقة. بالإضافة إلى ذلك، إذا كان النموذج الذي يحتوي على العدد p من المعلمات مناسباً للحالة المدروسة، فليس من الجيد أن يحتوي النموذج على معلمات أكثر من العدد p . وأثبت Herman Wold أنه يمكن استخدام هذه النماذج لنمذجة أي سلسلة ثابتة طالما تم اختيار القيم p و q بشكل جيد.

نقول أن النموذج من نوع $ARMA(p, q)$ إذا كان هناك تسلسل حقيقي من (ϕ_k) (θ_k) كالتالي:

$$y_t - \sum_{k=1}^p \phi_k y_{t-k} = \varepsilon_t + \sum_{k=1}^q \theta_k \varepsilon_{t-k} \quad \dots \quad \{\varepsilon_t\} \rightarrow N(0, \sigma_\varepsilon^2)$$

يمكننا أيضًا استخدام كثيري الحدود $\phi(B)$ و $\theta(B)$ لكتابة النموذج على الشكل

$$\phi(B)y_t = \theta(B)\varepsilon_t$$

بحيث:

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

و

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

هناك بعض الخصائص المتعلقة بنماذج $ARMA(p, q)$

²⁷ H. Wold : **A study in the Analysis of Stationary Time Series**. Almqvist and Wiksell, 1938.

1- إذا كان $p = q = 0$ فإن $y_t = \varepsilon_t$ والتشويش أبيض أي التوزيع طبيعي وسطه الحسابي 0 وتباينه ثابت (σ_ε^2) .

2- إذا كان $p = 0$ و $q \neq 0$ فإن y_t يصبح نموذج المتوسطات المتحركة $MA(q)$.

3- إذا كان $p \neq 0$ و $q = 0$ فإن y_t يصبح نموذج الانحدار الذاتي $AR(p)$.

منهجية Box و Jenkins :

سنة 1970 توصل BOX-JENKINS في الولايات المتحدة الأمريكية إلى نشر عملهما المتعلق بمعالجة السلاسل الزمنية وكيفية استعمالها في مجال التنبؤ وذلك بالاعتماد على دالة الارتباط الذاتي واستخدام مبدأ المتوسطات المتحركة ومبدأ الانحدار الذاتي، هذا التحليل يخضع السلسلة الزمنية إلى العشوائية نموذج عشوائي $(S)ARIMA$.

VI-10-4- النماذج المختلطة المركبة ARIMA (Autoregressive Integrated)

(Moving Average):

تعرف هذه النماذج بالمختلطة المركبة، أو المتجانسة غير المستقرة من الدرجة d ، ويرمز لها بالرمز $ARIMA(p, d, q)$. ويكمن الاختلاف الجوهرى بينها وبين النماذج السابقة خاصة نموذج $ARMA(p, q)$ هي أن السلسلة الزمنية المدروسة تكون غير مستقرة. ولذا يجب أولاً إزالة عدم الاستقرار بالاستعمال الطريقة المناسبة. إذا كان سبب عدم الاستقرار هو الاتجاه العام فهنا نطبق طريقة الفروق من الدرجة الأولى مرة أو مرتين. أما إذا كان سبب عدم الاستقرار هو المركبة الفصلية أو الموسمية فهنا نحاول إزالة هذه المركبة.

إذا كان Δ هو معامل التمايز حيث $\Delta Y_t = y_t - y_{t-1}$ ، معامل التمايز ذو الترتيب k هو

$\Delta^k Y_t = \Delta(\Delta^{k-1} y_t)$ ، تكون y قابلة اتجاه لكثير الحدود ذو الرتبة k إذا:

$$y_t = \sum_{j=0}^k a_j t^j + \varepsilon_t$$

وعليه الفرق Δy يقبل إتجاه لكثير الحدود من الرتبة $k-1$ إذا:

في نماذج ARMA يمكننا ملاحظة أن الفرق الأول هو $\Delta Y_t = (1 - B)y_t$ ، وبشكل عام الفرق عند الترتيب d هو $\Delta^d Y_t = (1 - B)^d y_t$ ، لذلك نماذج ARIMA تعريف على الشكل التالي:

في السلسلة مستقرة y_t يمكن النموذج ARIMA (p,d,q) بالحدود الدنيا إذا فقط إذا:

$$\phi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \rightarrow \forall t \in z$$

مع القيود التالية:

- $\theta_q \neq 0$ ، $\phi_p \neq 0$
- P و q هي جذور وحدة وليست جذور مشتركة
- ε_t : يتبع توزيع طبيعي تباينه σ^2

نماذج ARIMA (p,d ,q) مناسبة لنمذجة سلسلة زمنية تشتمل على اتجاه متعدد للدرجات ، المعامل $(1 - B)^d$ يعمل على التحول من الدرجة d إلى ثابت.

طريقة تقدير معاملات نموذج ARIMA هي نفسها كما في نموذج ARMA بعد تحديد الفرق $(1 - B)^d y_t$.

VI-10-5- نموذج SARIMA (Seasonal Autoregressive Integrated Moving)

(Average):

نموذج SARIMA هو نفسه نموذج ARIMA ولكن يضاف إليه المركبة الموسمية، وبالتالي نقول أن y_t هو نموذج SARIMA (p,d,q) في s إذا فقط إذا:

$$(1 - B)^d \phi_p(B)(1 - B^s)^D \phi_p(B^s) y_t = \theta_q(B)\theta_q(B^s)\varepsilon_t \rightarrow \forall t \in z$$

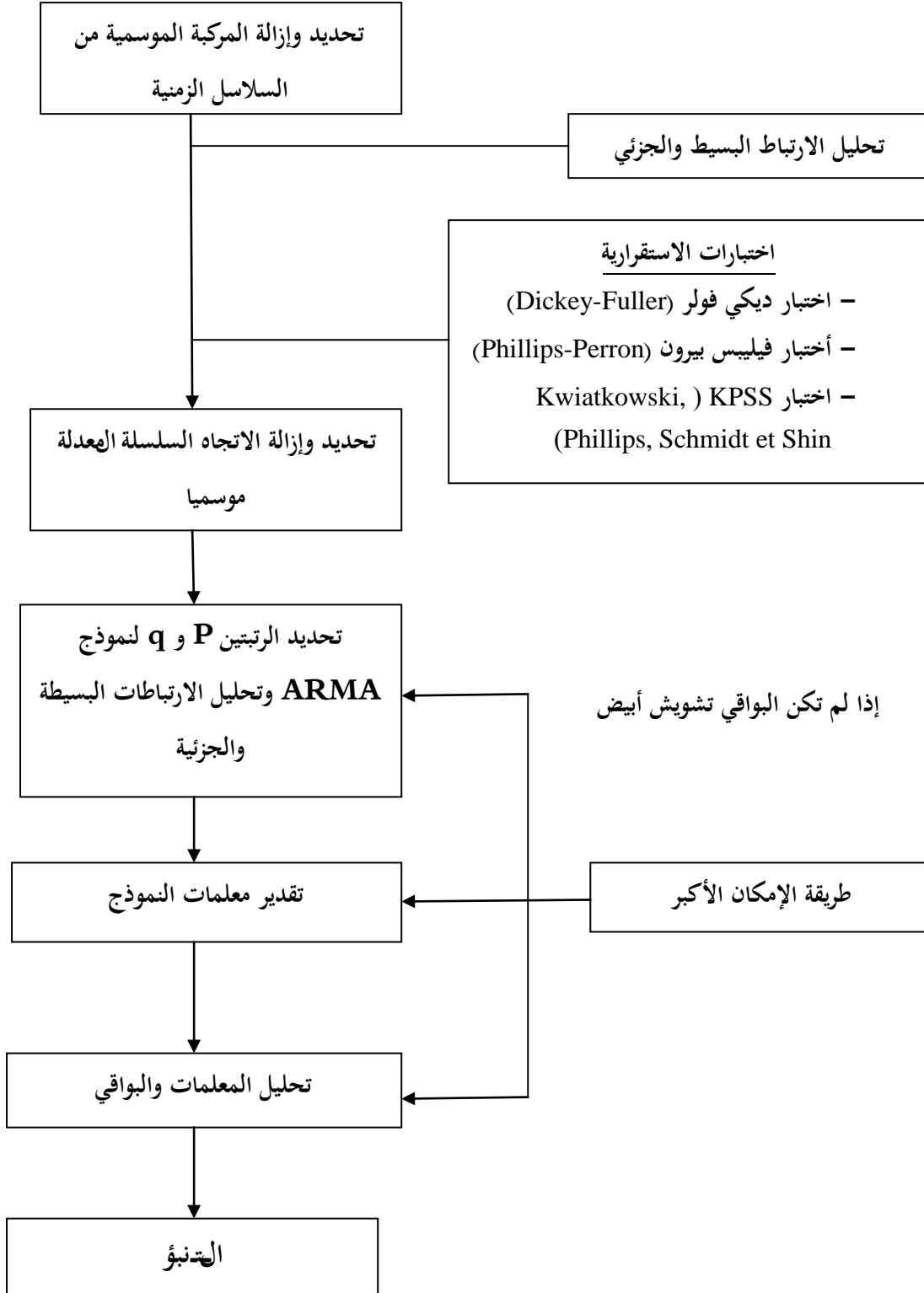
D و d : يتم اختيارها بحيث تكون السلسلة المتمايزة $(1 - B)^d$ و $(1 - B^s)^D$ تكون مستقرة.
 p و q : يتم استخراجها كما في نموذج ARMA(p,q) الاعتيادية.

P و Q: تبحث في الرتب المتعددة ل s إنطلاقاً من $(1 - B^S)$ نختار D قبل اختيار d عندما

$$1 - B^S = (1 - B)(1 + B + B^2 + \dots + B^{k-1})$$

تمكن طريقة Box و Jenkins من تحديد نموذج ARIMA الذي يمكن أن يكون مناسباً

لسلسلة زمنية وفقاً لخصائصه. ينقسم إلى عدة مراحل:



أولاً: تحديد النموذج

تتكون هذه الخطوة بشكل عام ، تحديد السلسلة التي تمت دراستها بشكل أفضل. بعد إزالة مركبتي الاتجاه العام والموسمية، وبعبارة أخرى أكثر شمولية ، فإن الهدف هو العثور على نموذج مستقر يأخذ في الاعتبار التباين بمرور الوقت مع عدم وجود ارتباط ذاتي للبواقي. كما تتضمن هذه الخطوة طرق تقدير معلمة التكامل d ، وتقدير الرتبين p و q ، واختبارات المرونة التغيرية، واختبارات عدم الاستقرار أو جذور الوحدة والتحويلات غير الخطية.

ثانياً: تقدير النموذج

تعتبر طريقة الإمكان الأكبر (*vraisemblance*) هي الطريقة الشائعة لتقدير معالم النموذج، هذه الدالة مرتبطة بعينة y_1, y_2, \dots, y_t ، حيث قانون دالة الكثافة $f(y, \vartheta)$ مع $\vartheta = (\vartheta_1, \vartheta_2, \dots, \vartheta_k) \in R^+$ معرف بـ:

$$L(\vartheta) = \pi_{h=1}^t f(y, \vartheta) \dots \dots \dots (1)$$

مقدر طريقة الإمكان الأكبر (*EMV*) هو $\hat{\vartheta}$ والذي يعظم $L(\vartheta)$ بالنسبة لكل معلمة من المعلمات $(\vartheta = \vartheta_1, \vartheta_2, \dots, \vartheta_k)$ وحل نظام متكون من k معادلة.

$$\frac{\partial L(\vartheta)}{\partial \vartheta_j} = 0, (j = 1, 2, \dots, k) \dots \dots \dots (2)$$

$$L(y, \varphi, \theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{T/2}} \frac{1}{\sqrt{|\Omega\vartheta\varphi|}} \exp \left\{ \frac{1}{2\sigma^2} y^T \Omega^{-1}(\varphi, \theta) y \right\}$$

حيث:

$L(y, \varphi, \theta, \sigma^2)$: دالة الامكان الأكبر

$y = (y_1, y_2, \dots, y_d)$

σ^2 : التباين

$\Omega(\varphi, \theta) \in R^{T+T}$: هي مصفوفة (التباين-التباين المشترك) ل y

دالة لوغاريتم الإمكان الأكبر معطاة بالعلاقة التالية:

$$\ln L(y, \varphi, \theta, \sigma^2) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln(\sigma^2) - \frac{T}{2} \ln(\det(\Omega(\varphi, \theta))) - \frac{1}{2\sigma^2} y' \Omega(\varphi, \theta)^{-1} y$$

وبسبب الصعوبة حساب دالة لوغاريتم الإمكان الأكبر وبالتالي تعظيمها بالطريقة الحساب الخطي نلجأ إلى تعظيمها بواسطة خوارزمية التعظيم غير الخطية²⁸ (Newton raphson).

ثالثاً: إختبار مدى ملائمة النموذج

تمثل هذه الخطوة في التأكد من أن النموذج المختار يصف سلسلة المشاهدات، بالإضافة إلى الشروط الثلاثة للتشويش الأبيض باستخدام الاختبارات المناسبة، ومن هذه الاختبارات:

1- اختبار البواقي:

في تحليل السلاسل الزمنية، يلعب التشويش الأبيض دوراً خاصاً لأنه عملية بدون أي هيكل. الفرضية العدمية الأولى التي يتم اختبارها هي:

$$H_0: P_1 = P_2 = \dots = P_h$$

مقابل الفرضية البديلة H_1 يوجد على الأقل P_i تختلف عن 0 نلجأ إلى الإحصائيتين Q ل (Box-Pierre) أو Q ل (Ljung-Box)، والتي تعطى بالعلاقة التالية:

$$Q = \sum_{h=1}^h \hat{p}_h^2$$

$$Q = T(T+2) \sum_{h=1}^h \frac{\hat{p}_h^2}{t-j}$$

حيث:

T: عدد المشاهدات

²⁸ Didier Delignières , Analyse des séries temporelles , Septembre 2007.

h: عدد التأخرات

\hat{p}_h : الارتباط الذاتي عند h.

رابعا: اختيار النموذج

هناك العديد من المعايير التي يمكننا استخدامها لاختيار النموذج أكثرها شيوعا هو معيار المعلومات Akaike:

$$AIC = -2\text{Log}L + 2(p + q)$$

ومعيار Schwarz:

$$BIC = 2\text{Log}L + (p + q)\text{Log}T^2$$

مع $\text{Log}L$ هو لوغاريتم الامكان الأكبر للنموذج المقدر، وبالتالي نختار النموذج و p و q التي يكون قيم هذه المعايير أصغر ما يمكن.

خامسا: التنبؤ

الخطوة الأخير في عملية النمذجة هي التنبؤ بالقيمة المستقبلية للظاهرة المدروسة من خلال السلسلة الزمنية من خلال تطبيق نظرية (Wald 1948) بالاعتماد على $MA(\infty)$ ²⁹.

$$\hat{y}_{T+h} = \sum_{j=1}^{\infty} \psi_j \varepsilon_{T+h-j}$$

\hat{y}_{T+h} : المشاهدات في الفترة T+h.

T: حجم عينة المشاهدات.

h: آفاق التنبؤ.

التنبؤ \hat{y}_{T+h} يمكن أيضا كتابته على شكل توقع شرطي من خلال العلاقة التالية:

$$\hat{y}_{T+h} = E(y_{T+h}/y_{T+h-1}, y_{T+h-2}) = E(y_{T+h}/\varepsilon_{T+h-1}, \varepsilon_{T+h-2})$$

²⁹ Agnès Iagnoux, **Renforcement des Séries Chronologiques**, Université de Toulouse le Mirail. 2010-2011.

والخطأ في التنبؤ معرف بـ

$$y_{T+h} - \hat{y}_{T+h} = \sum_{j=1}^{h-1} \Psi_j \varepsilon_{T+h-j} = \varepsilon_{T+h},$$

AutoRegressive Conditional ARCH نماذج -6-10-VI (Heteroskedasticity)

تستعمل هذه النماذج في الأساس لنمذجة تقلبات السلسلة الزمنية، رأينا في نماذج (p,q) ARMA، أن تباين الخطأ العشوائي ثابت $\{\varepsilon_t\} \rightarrow N(0, \sigma_\varepsilon^2)$ لكن هذا التباين σ_ε^2 يمكن أن لا يكون ثابتا. وبالتالي نتحل إلى الحالة التي يكون فيها مشروط أي $V(\varepsilon_t/\varepsilon_{t-1}) = \sigma_\varepsilon^2$ نستعمل مجموعة نماذج ARCH لنمذجة هذا التباين الشرطي. بالإضافة إلى نموذج ARCH لدينا نماذج ARCH-M، TGARCH، EGARCH، GARCH.

أولا: نموذج ARCH(q):

يمكن التعبير عنه بالعلاقة التالية:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

حيث:

$$\alpha_0 > 0, \alpha_i \geq 0 \forall i$$

ثانيا: نموذج GARCH(p,q) (Generalized ARCH)

يمكن التعبير عنه بالعلاقة التالية:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

حيث:

ثالثاً: نموذج (Exponential GARCH) EGARCH(p, q)

يمكن التعبير عنه بالعلاقة التالية:

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i [\phi z_{t-i} + \gamma (|z_{t-i}| - E|z_{t-i}|)] + \sum_{j=1}^p \beta_j \ln \sigma_{t-j}^2$$

حيث:

$$z_{t-i} = \frac{\varepsilon_{t-i}}{\sigma_{t-i}}$$

يمكننا أن نلاحظ هنا أنه لا يوجد قيد ايجابية يؤثر على المعاملات لأنه يتم التعبير عن معادلة التباين في (log)، عكس نموذجي ARCH(q) و GARCH(p, q)، يتيح نموذج EGARCH(p, q) يسمح للتقلبات بالتفاعل بشكل مختلف اعتماداً على علامة الصدمات.

رابعاً: نموذج (Threshold GARCH) TGARCH(p, q)

يمكن التعبير عنه بالعلاقة التالية:

$$\sigma_t = \alpha_0 + \sum_{i=1}^q (\alpha_i^+ \varepsilon_{t-i}^+ - \alpha_i^- \varepsilon_{t-i}^-) + \sum_{j=1}^p \beta_j \sigma_{t-j}$$

حيث:

$$\varepsilon_t^+ \text{ هي } \max(\varepsilon_t, 0) \text{ و } \varepsilon_t^- \text{ هي } \min(\varepsilon_t, 0)$$

يتضمن النموذج مجموعة من القيود ايجابية هي:

يسمح النموذج هنا بالنظر في الآثار غير المتماثلة للصدمات على التقلبات.

$\alpha_0 >$

خامسا: نموذج ARCH-M

يمكن التعبير عنه بالعلاقة التالية:

{ σ_t^2 }

تسمى بنية ARMA بمعادلة المتوسط، في نموذج ARCH نضيف تباينا إلى معادلة المتوسط، يمكن فعل نفس الشيء مع GARCH-M و TGARCH-M وغيرها.

مجموعة نماذج ARCH أعلاه يمكن تقديرها باستعمال طريقة الإمكان الأكبر (المعقولة العظمى) لأنه ببساطة الأخطاء العشوائية في السلاسل الزمنية بالنسبة للمتغيرات المالية لا تخضع لقانون التوزيع الطبيعي. وتكتب دالة الإمكان الأكبر في هذه الحالة على الشكل التالي:

$$Ln = f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \left[\ln(\sigma_t^2) + \left(\frac{\varepsilon_t}{\sigma_t} \right)^2 \right]$$

حيث θ هو شعاع يحتوي على مقدرات معلمات لنموذج ARMA والتقلبات لـ (ARCH و GARCH وغيرها).

خاتمة:

حاولنا من خلال هذه المطبوعة تبيان العناصر الأساسية التي تعتمد عليها عملية القياس الاقتصادي فلتوجه الحالي في أبحاث الاقتصاد القياسي يظهر تغييراً في المنظور عن ذلك الذي كان سائداً حتى أوائل السبعينيات. في البداية كان الاقتصاد القياسي يهدف إلى تقدير النماذج التي طورتها النظرية الاقتصادية دون التشكيك فيها. أما في الوقت الحالي، غالباً ما تكون المشكلات التي تثيرها ممارسة الاقتصاد القياسي هي التي تقود خبراء الأعمال إلى تصميم نماذج قادرة على أخذ هذه المشكلات في الاعتبار. بطريقة أخرى علم الاقتصاد هو الذي يقود اللعبة، فهل هناك خطر يجب الخوف منه؟ وفقاً للطبيعة التجريبية الشكلية للاقتصاد، يميل المرء إلى الإجابة بشكل سلبي، فحسب Frisch (1933)، "الاختراق المتبادل للنظرية الاقتصادية الكمية والملاحظات الإحصائية هو جوهر الاقتصاد القياسي". في كلتا الحالتين، من الصعب تصور أن بناء نظرية لعلم مثل الاقتصاد، أو من المستحيل ببساطة إجراء التجارب بطريقة الفيزياء التي لا يتأثر إلى حد كبير بالملاحظات الإحصائية.

كما نأمل أن نكون قد وفقنا في تحقيق أهداف هذا العمل المتواضع والذي سنحاول تحسينه وتطويره والتعمق أكثر في النماذج الأخرى الأكثر تعقيد التي تساعد صناع القرار في تحديد السياسات الاقتصادية الناجعة في الأعمال اللاحقة إن شاء الله، ونوصي أعضائنا الطلبة والباحثين بالخوض في دراسة هذه النماذج وأستسمح القارئ الكريم العذر عما قد يجده من هفوات غير مقصودة والعصمة لله. وفوق كل ذي علم عليم.

المراجع

أولاً: باللغة العربية:

- 1 - دجلة ابراهيم مهدي، نذير عباس إبراهيم، الاقتصاد القياسي أسلوب كمي باستعمال SPSS، MINITAB، EViews، الطبعة الأولى، 2014.
- 2 - سالم عيسى بدر، عماد غصاب عبابنة، مبادئ الإحصاء الوصفي والاستدلالي، دار المسيرة للنشر والتوزيع والطباعة، الطبعة الثانية، 2010.
- 3 - شبيخي محمد، طرق الاقتصاد القياسي محاضرات وتطبيقات، دار الحامد للنشر والتوزيع، الطبعة الأولى، 2012.
- 4 - ظافر رمضان مطر، وسن رعد ذنون، مقارنة بين الاختبارات الـ Portmaunteau باستخدام المحاكاة، المجلة العراقية للعلوم الإحصائية (22) 2012.
- 5 - عبد الحميد عبد المجيد البلداوي، أساليب الإحصاء للعلوم الاقتصادية وإدارة الأعمال مع استخدام برنامج SPSS، دار وائل للنشر والتوزيع، الطبعة الأولى، الأردن، 2009.
- 6 - محمد سليمان محمد جبريل، التعدد الخطي أسبابه تأثيراته والمعالجة بانحدار الحافة وانحدار المركبات الرئيسية مع التطبيق على بيانات افتراضية، بحث لنيل درجة دكتوراه الفلسفة في الإحصاء، جامعة السودان للعلوم والتكنولوجيا، كلية الدراسات العليا، جمهورية السودان، 2014.
- 7 - محمد عبد السميع عناني، التحليل القياسي والإحصائي للعلاقات الاقتصادية، مدخل حديث باستخدام spss، الدار الجامعية، الاسكندرية، 2009.
- 8 - مكيد على، الاقتصاد القياسي دروس ومسائل محلولة، ديوان المطبوعات الجامعية، الطبعة الثانية، 2011.
- 9 - وليد إسماعيل السيفو، عيد أحمد أبو بكر، غالب عوض الرفاعي، أساسيات الأساليب الإحصائية للأعمال وتطبيقاتها في العلوم المالية والإدارية والاقتصادية، زمزم ناشرون وموزعون، الطبعة الأولى، 2010.

ثانياً: باللغة الأجنبية:

- 1- Agnès lagnoux, Renforcement des Séries chronologiques, Université de Toulouse le Mirail .2010-2011.
- 2- Didier Delignières , Analyse des séries temporelles , Septembre 2007.
- 3- E. Slutsky : The summation of random causes as the source of cyclic processes. Econometrica, 5:105-146, 1927.

- 4- Françoise seyte, **introduction a l'économétrie**, cours magistrale, université Montpellier, 2011-2012, P06.
- 5- H. Wold : **A study in the Analysis of Stationary Time Series**. Almgrist and Wiksell, 1938.
- 6- P.A.Samualson .T.C.Copmans, « Report Of Evaluative Commitee For Econometrica, Econometric », Vol 02, N 02,1954.
- 7- Régie Bourbonnais, **économétrie cours et exercices corrigés**, DUNOD, 9eme édition,2015.
- 8- Régis Bourbonnais et Michel Terraza "**Analyse des série temporelles en économie**" presses Universitaires de France 1998.
- 9- Rumana rois, muhd muzibur rahman, tapati basak, ajit majumder, **Modified breusch- godfrey test for restricted higher order autocprrelation in dynamic linear model- a distance based approach**, Researchgate, international journal of business and management, augest 2012.
- 10- U. Yule : On a method of investigating periodicities in disturbed series, with special referéce to wolfer's sunspot numbers. Philosophical Transactions of the Royal Society, 226, 1927.