جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة الفلوجة – كلية الإدارة والاقتصاد

قسم الاقتصاد

محاضـــرات في

الاقتصـــــاد القياســـــي

ECONOMETRICS

إعــداد

الأستـــاذ الدكتـــور

ناظـم عبـدالله المحمـدي

الفصل الاول: المقـــدمـــة

اولاً : تعريف الاقتصاد القياسي Definition of Econometrics

يعد الاقتصاد القياسي **(Econometrics)** أسلوب من أساليب التحليل الاقتصادي يهتم بالتقدير العددي (الكمي) للعلاقات بين المتغيرات الاقتصادية معتمداً في ذلك على النظرية الاقتصادية والرياضيات والاحصاء للوصول الى هدفه الخاص بالتقدير واختبار الفروض ومن ثم التنبؤ بالظواهر الاقتصادية.

وهذا يعني ان الاقتصاد القياسي هو مزيج او توليفة من النظرية الاقتصادية والاقتصاد الرياضي والاحصاء، وهو يستخدم الادوات الرياضية والاحصائية لتحليل الظواهر الاقتصادية، ويهدف الى اعطاء مظهر تطبيقي للنظرية الاقتصادية وذلك بالقياس الكمي للظواهر الاقتصادية والتنبؤ بها واختبار فرضياتها. الا ان ما يميز الاقتصاد القياسي عن النظرية الاقتصادية والاقتصاد الرياضي والاحصاء ان هذه الفروع العلمية الثلاثة تتكامل فيه من اجل توفير قيم عددية للمؤشرات الاقتصادية ويوفر فهماً حقيقياً وكمياً للعلاقات الاقتصادية المختلفة.
عليه يمكن تعريف الاقتصاد القياسي: بانه علم اجتماعي تستخدم فيه ادوات النظرية الاقتصادية والرياضيات والاحصاء لتحليل الظواهر الاقتصادية، وانه اصطلاح يتكون من كلمتين اصلهما اغريقي اقتصاد **Economy** وقياسي **Matrices** .

ثانياً : اهداف الاقتصاد القياسي : Goals of Econometrics

وتتمثل بما يلي :

1 - تحليل واختبار النظريات الاقتصادية المختلفة.

ان تحليل واختبار النظريات الاقتصادية يعد هدفاً رئيسياً من اهداف الاقتصاد القياسي، حيث لا يمكن إعتبار (عّد) النظرية الاقتصادية صحيحة ومقبولة مالم تجتاز اختباراً كمياً عددياً يوضح قوة النموذج ويفسر قوة ومنطقية العلاقة بين المتغيرات الاقتصادية.

2 - رسم السياسات واتخاذ القرارات.

يساهم الاقتصاد القياسي برسم السياسات واتخاذ القرارات من خلال الحصول على قيم عددية لمعلمات العلاقات الاقتصادية بين المتغيرات لتساعد المختصين في اتخاذ القرارات، من خلال توفيره لصيغ وأساليب مختلفة لتقدير المرونات والمعلمات الفنية والتكلفة الحدية والايرادات الحدية والميل الحدي للاستهلاك والادخار او الاستثمار .... وتأسيساً على ذلك فان معرفة القيم العددية لمعلمات النموذج المقدر تساعد على اجراء المقارنات واتخاذ القرار المناسب سواءً على مستوى المنشاة او الدولة.

3 - التنبؤ بقيم المتغيرات الاقتصادية في المستقبل.

يساعد الاقتصاد القياسي المختصين في وضع السياسات من خلال توفير القيم العددية لمعلمات **(parameters)** المتغيرات الاقتصادية والتنبؤ بما ستكون عليه الظاهرة الاقتصادية مستقبلاً. ان هذه التنبؤات تمكن واضعي السياسة ومتخذي القرار لتنظيم الحياة الاقتصادية واتخاذ اجراءات معينة للتأثير في متغيرات اقتصادية معينة.

ثالثا : فروع الاقتصاد القياسي

يمكن تقسيم الاقتصاد القياسي الى فرعين هما:

1- **الاقتصاد القياسي النظري** : **Theoretical econometrics**

يشمل الجهود الهادفة الى ايجاد الطرق الملائمة لقياس العلاقات الاقتصادية.

2- **الاقتصاد القياسي التطبيقي** : **Applied econometrics**

يتضمن تطبيق طرق الاقتصاد القياسي على فروع محددة من النظرية الاقتصادية.

رابعاً : علاقة الاقتصاد القياسي بالعلوم الاخرى .

للاقتصاد القياسي علاقة وثيقة بالنظرية الاقتصادية والاقتصاد الرياضي، والاحصاء الاقتصادي والاحصاء الرياضي، وان هذه الفروع تتكامل من اجل توفير قيم عددية لمعلمات المتغيرات الاقتصادية المختلفة.

ويمكن توضيح ذلك كما يلي:

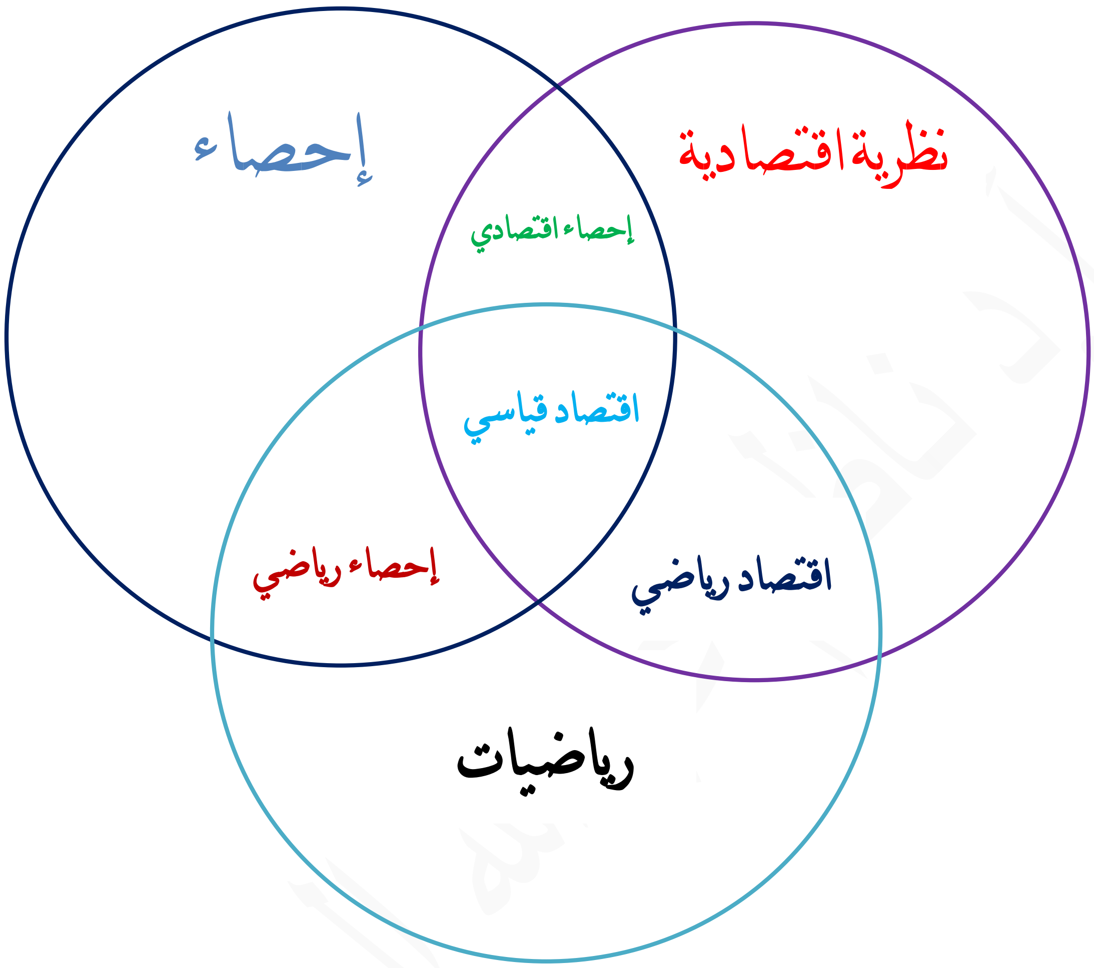
1- تقوم النظرية الاقتصادية بدراسة العلاقة بين المتغيرات الاقتصادية، فمثلاً تنص النظرية الاقتصادية الجزئية على ان زيادة سعر سلعة ما تسبب انخفاضاً في الطلب عليها، اي ان النظرية تفترض وجود علاقة عكسية بين السعر والكمية المطلوبة من السلعة، ولكنها لم تعط اي قياس عددي للعلاقة بين هذين المتغيرين فلم تبين مقدار الانخفاض في الكمية المطلوبة المصاحب لتغير معين في السعر فتصبح هذه المهمة من مهمات الاقتصاد القياسي بعد توصيفه رياضياً.

2- يهتم الاقتصاد الرياضي بإعادة صياغة العلاقة التي تم تحديدها بالاعتماد على النظرية الاقتصادية رياضياً، اي على هيئة معادلات ورموز رياضيه بدون قياس او برهنه لتلك الصياغات.

3- الاحصاء الاقتصادي ويقتصر دوره على تجميع البيانات الإحصائية الخاصة بالمتغيرات الاقتصادية، وتسجيلها وجدولتها او رسمها.

4- أما الإحصاء الرياضي فهو يجهز الباحث بأدوات تحليليه يستخدمها في دراسة العلاقات بين المتغيرات الاقتصادية، وبطريق خاصه لمعالجة أخطاء التقدير تمهيداً لاستخدامها في تحقيق أهداف القياس الاقتصادي.

لذلك يمكن النظر الى علم القياس الاقتصادي على انه نقطة التقاء ثلاث علوم رئيسيه هي الاقتصاد والرياضيات والاحصاء كما مبين في الشكل التالي:



شكل (1) الاقتصاد القياسي وعلاقته بالعلوم الاخرى

يتضح مما تقدم ان النظرية الاقتصادية تعطينا فكره عامه حول شكل العلاقة بين المتغيرات الاقتصادية ويأتي دور الاقتصاد القياسي لتحديد المقدار الكمي لتلك العلاقة بالاعتماد على الاقتصاد الرياضي الذي يحاول على تصوير العلاقة المذكورة بشكل معادله رياضيه، وعلى طرق الاحصاء الرياضي لملائمتها لطبيعة العلاقة القائمة، كل ذلك يتحقق بالاعتماد على الاحصاء الاقتصادي الذي يغذي القياس الاقتصادي بالبيانات الإحصائية اللازمة للتحليل.

خامساً: النموذج الاقتصادي: Economic Model

يعرف النموذج الاقتصادي بأنه مجموعة من العلاقات بين المتغيرات الاقتصادية لتمثيل ظاهرة معينة بهدف تحليلها أو التنبؤ بها والسيطرة عليها، قد يتكون من معادله واحده Single Equation مثل معادلة الطلب أو معادلة العرض، وقد يتكون النموذج من مجموعة من المعادلات وتسمى بالمعادلات الآنية Simultaneous Equation كنموذج السوق أو نموذج الاقتصاد العراقي.

وقد يكون الهدف من النموذج هو تقدير قيم عددية لمعلمات العلاقة بين المتغيرات الاقتصادية بغية التنبؤ أو تحليل هيكل اقتصادي أو تقييم سياسة اقتصادية. ويستخدم النموذج الاقتصادي الرموز والعلاقات الرياضية لتمثيله، وكمثال على النموذج الاقتصادي بمعادله منفردة، تفترض النظرية الاقتصادية بأن الاستهلاك **(c)** دالة في الدخل **(y)** ويعبر عن ذلك كما يلي:

$$C = f(y)$$

حيث ان:

C : الاستهلاك، ويمثل متغير الاستجابة او المتغير المعتمد او التابع Dependent variable

Y : الدخل، ويمثل المتغير التوضيحي او المتغير المستقل Independent variable

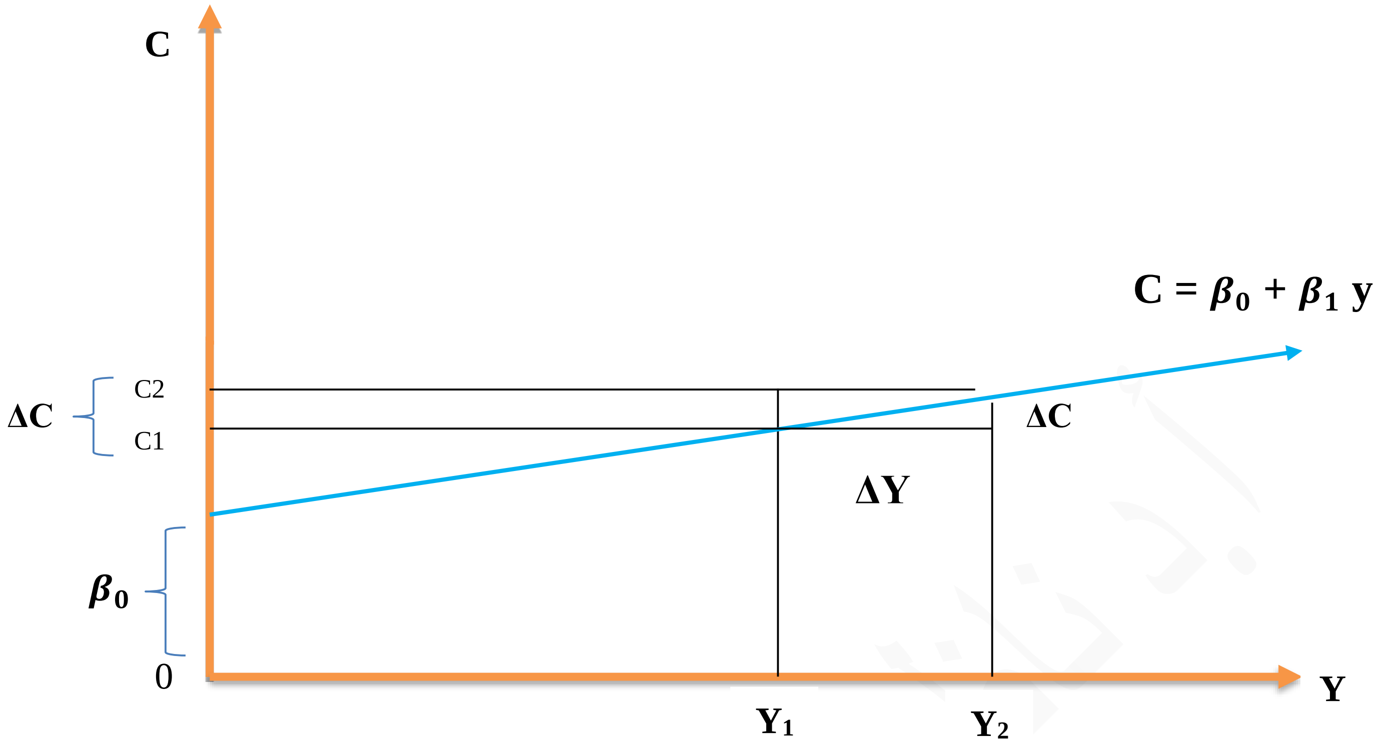
وتتحدد العلاقة بين المتغيرات الاقتصادية بعدد من العلاقات أو الصيغ، أبسطها الصيغة الخطية، فبتحويل العلاقة أعلاه الى الصيغة الخطية ستأخذ الشكل الآتي :

$$C = \beta_0 + \beta_1 y$$

حيث تمثل β_0 β_1 , المعلمات **Coefficients** والتي تعرف كما يلي :

β_0 : معلمة التقاطع وهي عبارة عن المسافة العمودية المحصورة بين نقطة الاصل ونقطة تقاطع خط الانحدار مع المحور العمودي وتمثل قيمة β_0 قيمة **C** عندما يكون قيمة **y** (الدخل) مساوية للصفر ويطلق عليها الاستهلاك الذاتي .

β_1 : ظل الزاويه التي يصنعها خط الانحدار مع مستوى الأفق وتسمى بالميل الحدي لخط الانحدار (الميل الحدي للاستهلاك) وتمثل قيمة β_1 الزيادة الحاصلة في قيمة المتغير التابع **C** نتيجة زيادة المتغير التوضيحي **y** بمقدار وحدة واحدة كما موضح في الشكل أدناه :



أي انه اذا زاد الدخل (y) بمقدار وحدة واحدة من y_1 الى y_2 $(y_2 = y_1 + 1)$ يترتب على ذلك زيادة في الاستهلاك C من c_1 الى c_2 ولنفترض بمقدار ΔC ويمكن اثبات ان $\beta_1 = \Delta C$ رياضياً ايضاً . حيث ان

$$C = \beta_0 + \beta_1 Y$$

فعندما يزداد الدخل (المتغير المستقل) y بمقدار وحدة واحدة نحصل على ان :

$$C + \Delta C = \beta_0 + \beta_1 (y + 1)$$

$$C + \Delta C = \beta_0 + \beta_1 y + \beta_1$$

وبالتعويض عن قيمة C بما يساويها نحصل على ان

$$\cancel{\beta_0} + \cancel{\beta_1} y + \Delta C = \cancel{\beta_0} + \cancel{\beta_1} y + \beta_1$$

$$\therefore \ \Delta C = \beta_1$$

وفي ضوء ذلك يمكن القول ان الميل الحدي للاستهلاك (mpc) يمثل مقدار التغير في المتغير المعتمد (c) نتيجة تغير المتغير المستقل (y) بمقدار وحدة واحدة، أي ان :

$$\beta_1 = \frac{\Delta c}{\Delta Y} = \frac{dc}{dy}$$

وحيث ان : $\Delta Y = 1$

$$\therefore \ \beta_1 = \Delta c$$

سادساً: الخصائص المرغوبة في اي نموذج اقتصادي قياسي.

وتتلخص بما يلي :

1- مطابقته للنظرية الاقتصادية بحيث يصف الظواهر الاقتصادية التي يهتم بها بشكل صحيح.
2- قدرته على توضيح المشاهدات الواقعية، بحيث يكون متناسقاً مع السلوك العملي للمتغيرات الاقتصادية التي تحدد العلاقة بين هذه المتغيرات.
3- دقته في تقدير المعلمات، حيث ان هذه التقديرات يجب ان تكون أفضل تقريب للمعلمات الحقيقية. وتتأتى هذه الدقة من اتصاف هذه التقديرات بصفات مرغوبة يحددها الاقتصاد القياسي مثل خاصية عدم التحيز، والاتساق، والكفاءة.
4- قدرة النموذج الاقتصادي على التنبؤ، بحيث يعطي تنبؤات مرضية للقيم المستقبلية للمتغيرات المعتمدة.
5- خاصية البساطة، اذ ان النموذج الاقتصادي يجب ان يبرز العلاقات الاقتصادية بأقصى حد ممكن من البساطة، فكلما قل عدد المعادلات وكان شكلها الرياضي أبسط اعتبر النموذج افضل من غيرة، شريطة ان لا يكون على حساب الصفات الاخرى . وكلما زاد عدد هذه الخصائص التي يتصف بها النموذج اعتبر أفضل للأغراض العلمية.

سابعا : انواع النماذج : هناك عدة انواع من النماذج يمكن تصنيفها كالاتي :-

1 - النماذج الاقتصادية الكلية والجزئية :-

أ- **النماذج الاقتصادية الكلية** : وهي النماذج التي تتعامل مع المتغيرات الاقتصادية التي تخص الاقتصاد الكلي اي التي تتصل بالسلوك العام والبنية العامة للاقتصاد كالدخل القومي والاستثمار العام .

ب- **النماذج الاقتصادية الجزئية** : وهي النماذج التي تتعامل مع المتغيرات الاقتصادية التي تخص الوحدات الاقتصادية الجزئية كعلاقة العرض او الطلب على سلعة معينة .

2- النماذج الاقتصادية الساكنة والمتحركة.

أ- **النماذج الاقتصادية الساكنة:** وهي النماذج التي لا يكون الزمن أحد متغيراتها او مؤشرا في تغير قيم احد المتغيرات الداخل فيها، اي بدون فترة ارتداد زمني ، وهذا يعني ان لكل متغير قيمة معينة في السنة التي يقع فيها ، فمثلاً تكون دالة الطلب الساكنة كالاتي :

$$D_t = f\,(p_t)$$

ب- **النماذج الاقتصادية المتحركة:** وهي النماذج التي يكون الزمن أحد متغيراتها او مؤثراً في أحد متغيراتها. اي ان هذه النماذج توضح كيفية تأثير الزمن في المتغيرات الاقتصادية.

وتعد هذه النماذج أكثر واقعية، فمثلا تأخذ دالة العرض المتحركة الشكل الاتي:

$$S_t = f\ (p_{t-1})$$

اي ان العرض في السنة الحالية (t) يعتمد على سعر السلعة في السنة السابقة (t−1) ويسمى المتغير الحركي بمتغير مرتد زمنيا Lagged variable مثل (p_{t-1}) .

ثامناً : مكونات النموذج : وتتمثل بما يلي

1- **معادلات النموذج**: يتكون النموذج الاقتصادي من مجموعة من المعادلات تسمى بالمعادلات الهيكلية لأنها توضح الهيكل الاساس للنموذج المراد بناؤه، ويختلف عدد المعادلات من نموذج لأخر تبعا لنوع النموذج والهدف من بنائه، وتتقسم المعادلات الهيكلية الى:

أ- **المعادلات السلوكية**: هي المعادلات التي تعبر عن العلاقات الدالية بين المتغيرات الاقتصادية ويمكن التعبير عنها بدالة ذات متغير مستقل واحد او عدة متغيرات مستقلة، مثل العلاقة بين الكمية المطلوبة من سلعة ما (D) والسعر (P) والدخل (Y) والتي يعبر عنها بالشكل الاتي:

$$D = B_0 + B_1P + B_2Y + u$$

ب- **المعادلات التعريفية او المتطابقات**: هي المعادلات التي تعبر عن علاقة اقتصادية ناتجة عن تعاريف متفق عليها، اوهي العلاقة التي تحدد قيمة المتغير التابع من خلال تحديد تعريف له في صورة علاقة مساواة. مثل متطابقة الدخل القومي (y) والتي تبين ان الدخل القومي (y) يساوي مجموع الاستهلاك (c) والاستثمار (I) والانفاق الحكومي (G) اي ان

$$\boldsymbol{Y = C + I + G}$$

2- **متغيرات النموذج** :

تتكون معادلات النموذج من عدد من المتغيرات يمكن تصنيفها الى عدة انواع هي :

أ- **المتغيرات الداخلية**: **endogenous variables**

وهي المتغيرات التي تؤثر في النموذج وتتأثر به وتتحدد قيمتها من داخل النموذج عن طريق المعاملات وقيم المتغيرات الخارجية، وتسمى هذه المتغيرات ايضا بالمتغيرات المعتمدة او التابعة.

ب- المتغيرات الخارجية: Exogenous variables

وهي المتغيرات التي تؤثر بالنموذج ولا تتأثر به وتتحدد قيمتها بعوامل خارجة عن النموذج، وفي بعض الاحيان تتحدد قيمتها عن طريق نموذج أخر مختلف عن النموذج الاصلي وتسمى بالمتغيرات المستقلة أو التوضيحية.

ت- المتغيرات المرتدة زمنيا: lagged variables

وهي المتغيرات التي تنتمي الى فترة زمنية سابقة او التي تؤخذ قيمتها من الفترة السابقة مثل (y_{t-1}) والتي تمثل الدخل في السنة السابقة، (p_{t-1}) والتي تمثل الاسعار في السنة السابقة .

تاسعاً : منهجية البحث في الاقتصاد القياسي :

يهتم **الاقتصاد القياسي (التطبيقي)** بقياس معاملات النموذج المستخدم في التقدير والتنبؤ بقيم المتغيرات الاقتصادية وهذا يتطلب اتباع منهجية معينة في البحث، لان العلاقة بين المتغيرات الاقتصادية سببية Causal اي بمعنى ان التغيير في بعض المتغيرات يحدث اثراً في المتغيرات الاخرى، ويمكن تحديد اربع مراحل للبحث العلمي في الاقتصاد السياسي وهي:

1- مرحلة التوصيف Specification Stage

تعتبر مرحلة التوصيف وصياغة النموذج من اهم مراحل بناء النموذج وأصعبها وذلك من خلال ما تتطلبه من تحديد للمتغيرات التي يجب ان يشتمل عليها النموذج او التي يجب استبعادها منه، وفي هذه المرحلة يتم الاعتماد على النظرية الاقتصادية في تحديد نوع واتجاه العلاقة بين المتغيرات الاقتصادية، كما يتم الاعتماد على الاقتصاد الرياضي لتحويل العلاقة المذكورة الى معادلات رياضية باستخدام الرموز الرياضية، مثل العلاقة بين الكمية المطلوبة من سلعة ما (D) والسعر (P) والدخل(Y) والتي تصاغ كالاتي:

$$\boldsymbol{D = \beta_0 + \beta_1 p + \beta_2 y + u}$$

فمن نظرية الطلب نتوقع الحصول على اشارة سالبة للمعامل $\boldsymbol{\beta_1}$ وذلك لوجود علاقة عكسية بين الكمية المطلوبة من سلعة ما وسعرها حسب النظرية الاقتصادية، واشارة موجبة للمعامل $\boldsymbol{\beta_2}$ لوجود علاقة طردية بين الكمية المطلوبة ودخل المستهلك .

2- مرحلة التقدير Estimation Stage

في هذه المرحلة يتم جمع البيانات المتعلقة بالظاهرة الاقتصادية او المشكلة قيد الدراسة، ومن ثم يتم تقدير معالم العلاقة التي تم وصفها وصياغتها رياضياً في المرحلة الاولى، اي تقدير قيم رقمية للمعالم $\boldsymbol{\beta_0}$ و $\boldsymbol{\beta_1}$ و $\boldsymbol{\beta_2}$ في دالة الطلب اعلاه، وفي هذه المرحلة يتم تقييم المعالم المقدرة من النواحي الاقتصادية والاحصائية والقياسية .

فمن الناحية الاقتصادية تجري عملية مقارنة بين قيم واشارات معالم النموذج التي تم تقديرها مع القيم والاشارات المتوقعة لهذه المعالم في ضوء النظرية الاقتصادية.

ومن الناحية الاحصائية يتم حساب الانحرافات الكلية والجزئية في المتغيرات التي يتكون منها النموذج واختبار معنوية المعالم من خلال اختيار (t) ومعامل التحديد (R^2) واختبار (F) .

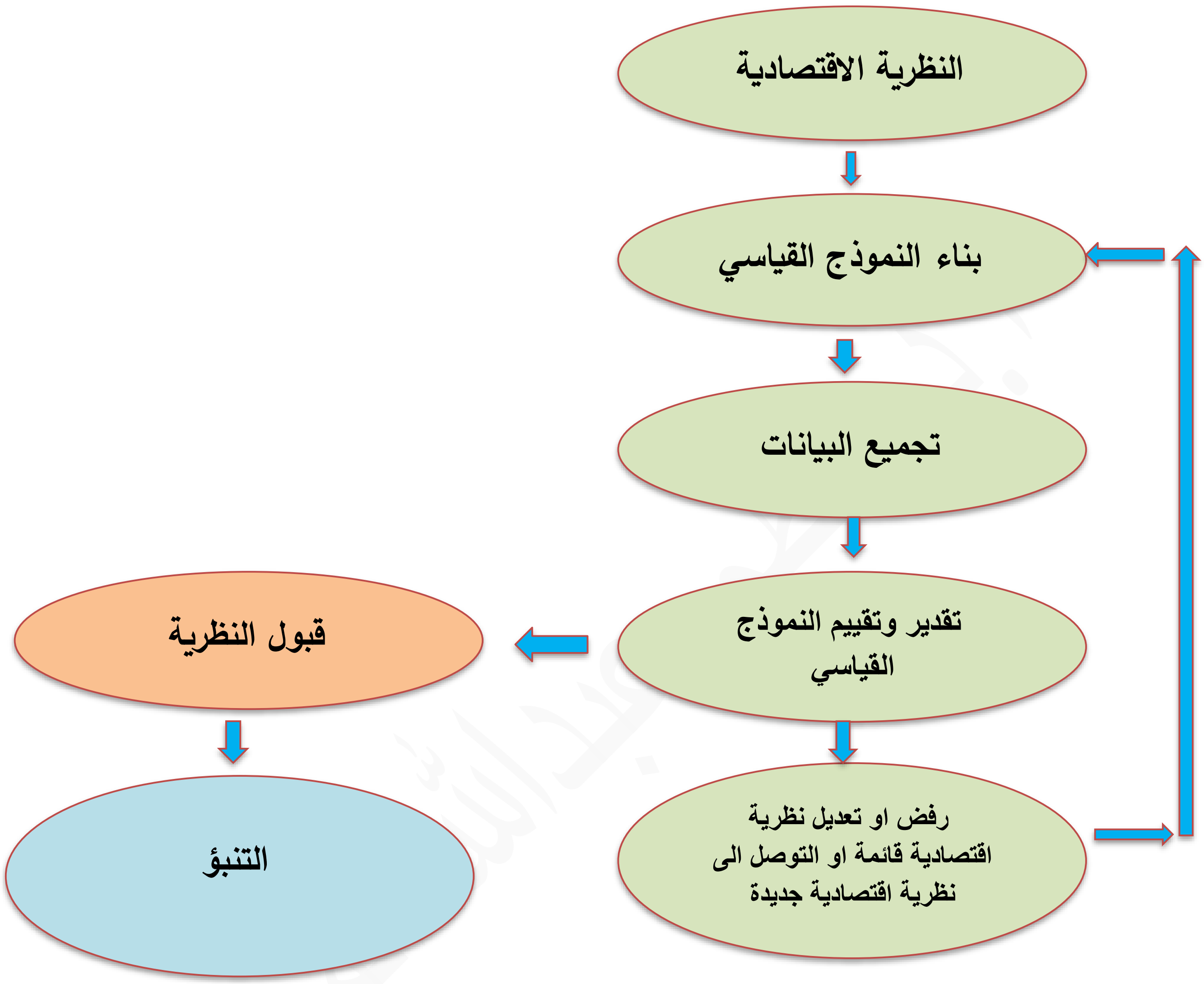
اما من الناحية القياسية فيتم اختبار مدى انسجام وتحقق الفروض الخاصة بالمتغير العشوائي(u) في النموذج القياسي المقترح حيث ان وجود الاختلافات يعني وجود المشاكل مثل مشكلة الارتباط الذاتي، التعدد الخطي وعدم ثبات تجانس التباين.

3- مرحلة الاختبار Testing Stage

في هذه المرحلة يتم اختبار قوة ومعنوية النموذج المقدر باعتماد طرق احصائية معينة للتأكد من صلاحية النموذج وقدرته على التنبؤ، وقد يواجه الباحث بعض المشاكل منها مشكلة تباين (تغاير) حد الخطأ او الارتباط الذاتي او الازدواج الخطي وغيرها من المشاكل والتي على الباحث ان يعالجها قبل البدء بعملية التقييم والتحليل.

4- مرحلة التنبؤ Prediction Stage

لأهمية وضرورة التنبؤ بالمستقبل والتعرف عليه مسبقاً وعلى مختلف المستويات الكلية والجزئية وفي مختلف المجالات الاقتصادية والاجتماعية ولمختلف المدد(الفترات) القصيرة والمتوسطة والطويلة، عليه يتم في هذه المرحلة إعداد تقديرات مستقبلية للمتغيرات المدروسة كحجم الطلب على السلعة (D) في المثال اعلاه، ويمكن توضيح منهجية البحث العلمي في الاقتصاد القياسي بالشكل الاتي :



الشكل (2) منهجية البحث العلمي في الاقتصاد القياسي

الفصل الثاني:. الانحدار الخطي البسيط Simple Liner Regression

ان دراسة العلاقة بين المتغيرات الاقتصادية تتطلب تحديد المتغيرات المؤثرة في تلك العلاقة، ومن ابسط واسهل انواع العلاقات في التقدير والتحليل الاحصائي والاقتصادي العلاقة بين متغيرين مثلY , X والتي تأخذ الشكل الاتي :

$$Y = f\,(x) \ldots\ldots\ldots (1-2)$$

حيث ان : Y : المتغير التابع dependent variable x : المتغير المستقل independent variable

ويمكن الاستعانة بالنظرية الاقتصادية والاقتصاد الرياضي والاحصاء لتحديد شكل هذه العلاقة فيما اذا كانت خطية او غير خطية، وصياغة العلاقة واختيار المتغيرات، ويمكن تحديد شكل العلاقة هذه بعدة اشكال او (صيغ) أبسطها واكثرها شيوعاً هي الصيغة الخطية. وتسمى العلاقة الخطية بين متغيرين بالانحدار الخطي البسيط Simple Linear Regression ، فالعلاقة الخطية بين X ولتكن دخل الاسرة و Y ولتكن الانفاق على سلعة معينة يمكن ان تكتب بالصيغة الرياضية الاتية :

$$y_i = \beta_0 + \beta_1\, x_i \ldots\ldots\ldots (2-2)$$

حيث ان β_1 و β_0 عبارة عن معلمات مجهولة القيم (ثوابت) وان

β_0 : تقاطع خط الانحدار مع المحور العمودي وهي عبارة عن القيمة التي تتخذها Y عندما تكون قيمة X مساوية للصفر.

β_1 : الميل الحدي للاستهلاك وهي عبارة عن مقدار الزيادة الحاصلة في قيمة المتغير التابع Y نتيجة زيادة المتغير المستقل X بمقدار وحدة واحدة .

غير ان العلاقة (2-2) اعلاه لا يمكن ان تشرح (تفسر) العلاقة بين المتغيرين بشكل دقيق، فهناك اسباب مهمة تجعل هذه المعادلة غير معبرة عن العلاقة بين X و Y تعبيرا كاملا، فقد يكون هناك انحراف بين العلاقة الحقيقية والمعادلة الاحصائية التي تمثلها نتيجة اخطاء في القياسات أو في اختيار المتغير المستقل، مما يتطلب اضافة متغير جديد يسمى بالحد العشوائي (Random Variable) او حد الخطأ او الاضطراب Error Term او Disturbance Term ويرمز له عادة بالرمز (u). لذلك فانه الصيغة المعبرة عن حقيقة العلاقة اعلاه يجب ان تضم حد الخطأ العشوائي لتأخذ الشكل الاتي:

$$y_i = \beta_0 + \beta_1\, x_i + u_i \quad \ldots\ldots\ldots (3-2)$$

ويمكن ارجاع الانحرافات بين القيمة الحقيقية والقيمة التقديرية لها الى الاسباب (العوامل) الاتية:

اسباب او مبررات اضافة المتغير العشوائي ui

1- صعوبة ادخال كافة المتغيرات المؤثرة في الظاهرة (حذف بعض المتغيرات من الدالة) .
2- السلوك العشوائي للجنس البشري.
3- عدم كفاية صياغة الشكل الرياضي للنموذج، او الصياغة الناقصة للنموذج الرياضي، فقد يجعل العلاقة خطية في حين يجب ان تكون غير خطية.
4- أخطاء التجميع Aggregation Error
5- أخطاء القياس Measurement Error

ان الاخطاء الاربعة الاولى تؤدي الى حصولنا على شكل خاطئ للمعادلة وتعرف عادة بالأخطاء في المعادلات Error In Equations اما المصدر الخامس للخطأ فيسمى بخطأ القياس للمشاهدة ذاتها. وهذا يتطلب الإلمام بخواص الخطأ العشوائي (ui) في العلاقة الخطية المدروسة.

الفرضيات الخاصة بالمتغير العشوائي (ui)

1- ui متغير عشوائي حقيقي. اي إن كل قيمة من قيم ui وفي اي فترة زمنية تعتمد على عامل الصدفة، وقد تكون هذه القيم سالبة او موجبة او مساوية للصفر.
2- E(ui)=0 (متوسط قيم ui في اية فترة مساوي للصفر) وهذا يعني ان كل قيمة من قيم المتغير المستقل سوف يأخذ قيماً مقابلة ل ui والاخيرة هذه تكون اكبر او اصغر او مساوية للصفر، وحاصل جمع هذه القيم يكون مساويا للصفر واستنادا لهذا الفرض فان :

$$\boldsymbol{E\,(y_i) = \beta_0 + \beta_1\, x_i}$$

3-

$$\boldsymbol{var\,(u_i) = E\,[\,u_i - E\,(u_i)\,]^2 = E\,(u_i)^2 = \sigma_u^2}$$

لان

$$\boldsymbol{E(u_i) = 0}$$

وهذا يعني بان تباين قيم $(\boldsymbol{u_i})$ حول متوسطها يكون ثابتاً في كل فترة زمنية بالنسبة لجميع قيم المتغير المستقل Xi ، واذا كان تباين الحد العشوائي (الخطأ) غير ثابت عندئذ تظهر مشكلة تسمى "عدم تجانس التباين" Heteroscedasticity .

4- ان المتغير العشوائي (u_i) يتوزع توزيعاً طبيعياً حول القيمة المتوقعة او حول الوسط الحسابي المساوي للصفر ويكون متماثلاً عند كل قيمة من قيم المتغير المستقل xi ، والفروض الاربعة السابقة يمكن اختصارها بالتعبير الاحصائي الاتي :

$$ui \sim N\left(0, \sigma_u^2\right)$$

اي ان الخطأ العشوائي يتوزع طبيعياً (N) لكن بوسط حسابي صفر وتباين ثابت قيمته σ_u^2

5- القيم المختلفة للمتغير العشوائي (ui) تكون مستقلة عن بعضها البعض، بعبارة اخرى التباين المشترك ل (ui) مع اي (uj) يكون مساوي للصفر، اي ان قيمة العنصر العشوائي في اي فترة لا تعتمد على قيمته في فترة اخرى، اي ان

$$COV\left(u_i u_j\right) = E\left(u_i u_j\right) = 0 \qquad (i.j = 1.2.\dots\dots.n.i \neq j)$$

واذا حدث وجود ارتباط بينهما تظهر لدينا مشكلة تسمى مشكلة الارتباط الذاتي **Autocorrelation**

6- ان قيم المتغير العشوائي (ui) غير مرتبطة باي من المتغيرات المستقلة (Xi)، اي انعدام التباين المشترك $(COV.)$ بين (ui) و (xi) اي ان

$$COV\left(x_i u_i\right) = E\left(u_i x_j\right) = 0$$

ويمكن اثبات ذلك كما يلي :

$$COV\left(x_i u_i\right) = E\left(\left[x_i - E\left(x_i\right)\right]\left[u_i - E\left(u_i\right)\right]\right)$$

$$= E\left(\left[x_i - E\left(x_i\right)\right] u_i\right) \qquad E\, u_i = 0 \text{ لان}$$

$$= E\left(x_i u_i\right) - E\left(x_i\right) E\left(u_i\right)$$

$$= E\left(x_i u_i\right) = x_i E\left(u_i\right) = 0$$

x_i ثابتة لكل العينات

7- انعدام العلاقة بين المتغيرات المستقلة (X_i). فاذا كان هناك اكثر من متغير مستقل (توضيحي) فان هذه المتغيرات يجب ان لا يكون بينها ارتباط قوي لكي يسهل التعرف على اثر كل منها على المتغير المعتمد بشكل منفصل، وفي حالة وجود علاقة قوية بين المتغيرات المستقلة تظهر مشكلة تسمى مشكلة الارتباط الخطي المتعدد Multi Collinearity .

8- ان تكون العلاقة المراد تقديرها قد تم تشخيصها. اي ان يكون النموذج المدروس ذو شكل رياضي مميز ولا يحتوي على نفس المتغيرات التي تحويها علاقة اخرى في نفس المجال.

توزيع المتغير المعتمد (Yi) :

من الضروري التعرف على توزيع المتغير المعتمد **(yi)** وقيم معاملات هذا التوزيع اي المتوسط والتباين لفائدتها في المراحل اللاحقة من دراستنا هذه .

ان توزيع المتغير المعتمد هو توزيع طبيعي(معتدل) Normal Dist تكون قيمة المتوسط فيه (التوقع) كما يلي:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

$$E\,(y_i) = \bar{Y} = \beta_0 + \beta_1 x_i$$

اما التباين variance فيعطى بالشكل الاتي :

$$var(y_i) = E\,[\,y_i - E\,(y_i)]^2 = E\,(u_i)^2 = \sigma_u^2$$

ويمكن اختصار خواص المتغير المعتمد (y_i) بالشكل الاتي :

$$y_i \sim N\,(\beta_0 + \beta_1 x_i\,, \sigma_u^2)$$

تقديرات المربعات الصغرى الاعتيادية

The Ordinary Least Square Estimations (OLS)

بالرجوع الى العلاقة الخطية بين دخل الاسرة X وانفاقها على سلعة معينة Y وفق الصيغة

$$y_i = \beta_0 + \beta_1 x_i + u_i \quad \dots\dots\dots (4-2)$$

يتبين لنا بان تأثير الدخل في الانفاق على السلعة موضوع البحث يتحدد من خلال العلاقة المنتظمة $(\beta_0 + \beta_1 x_i)$ اما تأثير العوامل الاخرى فأنها متجسدة في حد الخطأ (u_i) . ولمعرفة العلاقة الحقيقية بين دخل الاسرة وانفاقها على السلعة فانه يتطلب حساب قيم المعالم β_0 و β_1 وهذا يتطلب سحب (اختيار) عينة من الاسر وجمع البيانات منها ومن ثم تقدير قيم المعالم ويتم التقدير بواسطة المعادلة :

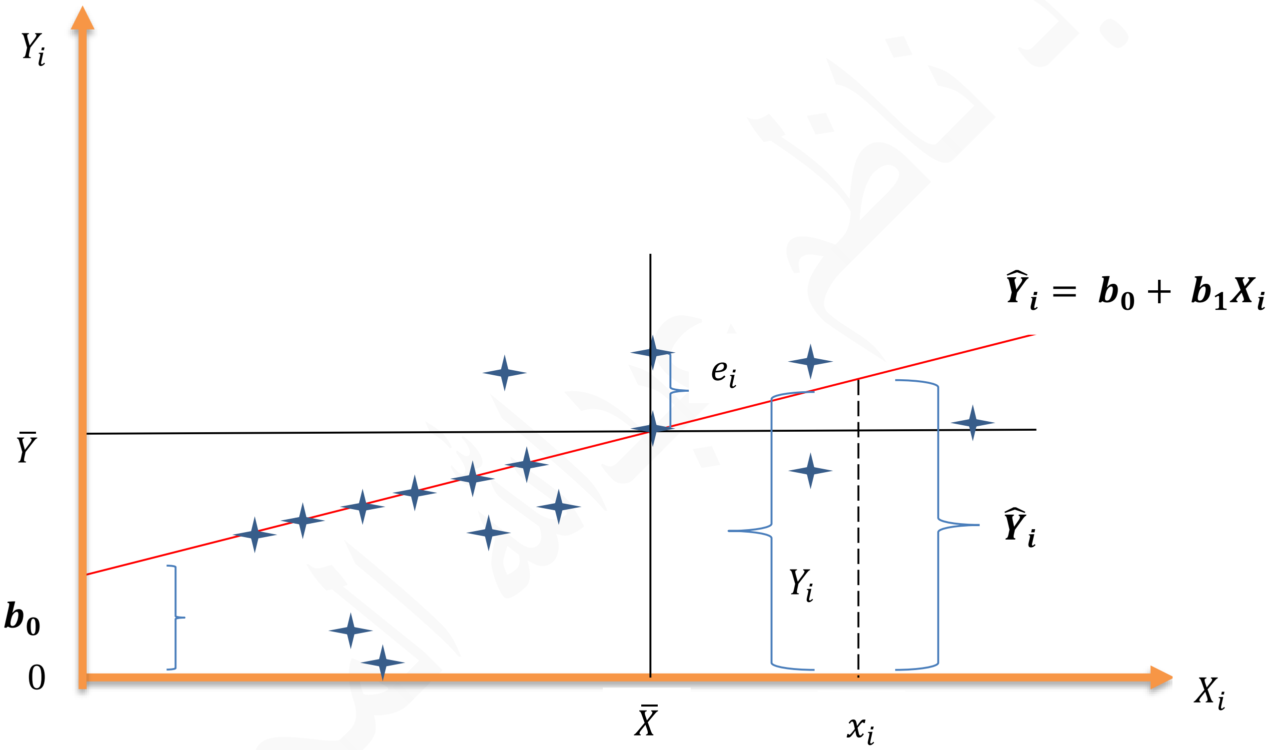
$$y_i = b_0 + b_1 x_i + e_i \quad \dots\dots\dots (5-2)$$

ولتقدير تأثير الدخل بصورة مستقلة في الانفاق فانه يتم بالعلاقة الاتية :

$$\hat{y}_i = b_0 + b_1 x_i \quad \ldots\ldots\ldots (6-2)$$

تسمى المعادلة (2−6) بمعادلة خط الانحدار المقدر ، حيث تشير العلامة (^) الى كون القيم تقديرية وليست حقيقية، وكل نقطة من نقاط $(\hat{y}_i)$ تمثل القيمة التقديرية لمتوسط انفاق جميع العوائل ذات الدخل (x_i) .

ويتبين من المعادلتين (2−5) و (2−6) بان قيم المشاهدات الفعلية (الحقيقية) (y_i) تنحرف عن القيم التقديرية $(\hat{y}_i)$ بمقدار (e_i) وكما مبين في الشكل ادناه :



ومن الشكل والعلاقات اعلاه يتبين ان $e_i = y_i - \hat{y}_i$ حيث يمكن للبواقي ei ان تكون سالبة او موجبة او مساوية للصفر حسب موضع نقطة المشاهدة من خط الانحدار المقدر . ولإيجاد افضل خط مستقيم لعينة المشاهدات يصف العلاقة الخطية تستخدم طريقة المربعات الصغرى الاعتيادية (ols) والتي تتضمن محاولة جعل مجموع مربعات انحرافات القيم الحقيقية (y_i) عن القيم التقديرية $(\hat{y}_i)$ اقل ما يمكن ، اي جعل مجموع مربعات الاخطاء العشوائية عند نهايتها الصغرى . اي ان طريقة (ols) تشترط تصغير القيمة $(\sum_{i=1}^{n} e_i^2)$ الى الحد الادنى . اي ان $\min \Rightarrow \sum_{i=1}^{n} e_i^2$

اذ ان $e_i = y_i - \hat{y}_i$ فان $\Sigma e_i^2 = \Sigma(y_i - \hat{y}_i)^2$

وبما ان معادلة الخط المستقيم الحقيقية غير المعروفة هي :

$$y_i = \beta_0 + \beta_1 X_i + e_i$$

وان معادلة الخط المستقيم التقديرية هي :

$$\hat{y}_i = b_0 + b_1 X_i$$

وبالتعويض عن $(\hat{y}_i)$ بما يساويها نحصل على :

$$\Sigma e_i^2 = \Sigma (y_i - b_0 - b_1 X_i)^2$$

وكشرط رياضي لتصغير $\sum e_i^2$ نجد المشتقات الجزئية الاولى لكل من b_0 و b_1 ومساواتها للصفر نحصل على ان :

$$\frac{\partial \sum e_i^2}{\partial\ b_0} = 2 \sum(y_i - b_0 - b_1 X_i)\,(-1) = 0$$

$$-2 \sum(y_i - b_0 - b_1 X_i)\,(-1) = 0$$

وان

$$\frac{\partial \sum e_i^2}{\partial\ b_1} = 2 \sum(y_i - b_0 - b_1 X_i)\,(-X_i) = 0$$

$$-2 \sum x_i(y_i - b_0 - b_1 X_i) = 0$$

وبالقسمة على (-2) وفك الاقواس واعادة ترتيب المعادلتين اعلاه نحصل على:

$$\left.\begin{aligned} \Sigma Y_i &= n\, b_0 + b_1\, \Sigma X_i \\ \Sigma X_i y_i &= b_0 \Sigma X_i + b_1\, \Sigma X_i^2 \end{aligned}\right\} \qquad (7-2)$$

وتسمى المعادلتين السابقتين بالمعادلتين الطبيعيتين ، ويمكن حلها انيا (او بطريقة الحذف والتعويض) نحصل على قيم لكل من b_0 و b_1 وفق الصيغ الاتية :

$$b_1 = \frac{n\sum X_i Y_i - (\sum X_i)(\sum Y_i)}{n\sum X^2{}_i - (\sum X_i)^2} \dots\dots\dots (8-2)$$

وان:

$$b_0 = \bar{Y} - b_0\,\bar{X} \quad \dots\dots\dots \; (9-2)$$

وتوجد عدة طرق اخرى لتقدير قيم b_0 و b_1 من المعادلتين الطبيعيتين في (7−2) اعلاه مثل طريقة المحددات وطريقة المصفوفات وطريقة التقدير حول نقطة المتوسط .

وسوف نوضح هذه الطرق من خلال حل المثال التالي:

مثال (1) : الجدول الاتي يمثل عدد سنوات الخدمة (X_i) ومعدل الاجر السنوي (Y_i) بالالف دينار لعينة تمثل (8) موظفين في احدى الدوائر .

المطلوب : تقدير معادلة خط الانحدار من خلال حل المعادلات الطبيعية بكافة الطرق الممكنة .

معدل الاجر Y_i	عدد سنوات الخدمة X_i	X_iY_i	X_i^2
25.6	4	102.4	16
32.7	8	261.6	64
45.4	12	544.8	144
53.9	16	862.4	256
59.0	20	1180	400
62.6	24	1502.4	576
65.0	28	1820	784
65.8	32	2105.6	1024
$\sum =$ 410	$\sum =$ 144	$\sum =$ 8379.2	$\sum =$ 3264

الحل : معادلة خط الانحدار المطلوب تقديرها تأخذ الشكل الاتي :

$$\hat{Y}_i = b_0 + b_1 X_i$$

وهذا يتطلب تقدير قيم b_0 و b_1 وهذا يلزمنا حساب $\sum X_i Y_i$ و $\sum X^2{}_i$ و $\bar{X}$ و $\bar{Y}$ بموجب الطرق الاتية:.

أولا : طريقة القيم الاصلية (الطريقة المباشرة) او طريقة الحذف والتعويض .

حيث ان :

$$b_1 = \frac{n\sum X_i Y_i - (\sum X_i)(\sum Y_i)}{n\sum X^2{}_i - (\sum X_i)^2}$$

$$= \frac{8(8379.2) - (144)\,(410)}{8(3264) - (144)^2} = \frac{7993.9}{5376}$$

$$b_1 = 1.4869$$

وان :

$$b_0 = \bar{Y} - b_1 \bar{X}$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{410}{8} = 51.25$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{144}{8} = 18$$

$$b_0 = 51.25 - (1.4869)(18)$$

$$b_0 = 24.4858$$

وعليه فان المعادلة المقدرة للعلاقة ستأخذ الشكل الاتي :

$$\hat{Y}_i = 24.4858 + 1.4869\, X_i$$

وتشير المعادلة التقديرية الى وجود علاقة طردية بين المتغير التابع (Y_i) الذي يمثل معدل الاجر السنوي للموظف والمتغير المستقل(X_i) الذي يمثل عدد سنوات الخدمة، فبزيادة خدمته الوظيفية بمقدار سنة واحدة يزداد معدل الاجر السنوي بمقدار (1.4869) دينار.

ثانيا : طريقة المحددات : Determinantes Method

يمكن الحصول على القيم التقديرية للمعالم b_0 و b_1 باستخدام المحددات وفق طريقة او قاعدة كرامر ، وذلك بإعادة كتابة المعادلات الطبيعية في العلاقة (2−7) في صيغة المصفوفات وعلى الشكل الاتي:

$$\sum Y_i = n\, b_0 + b_1 \sum X_i$$

$$\sum X_i Y_i = b_0 \sum X_i + b_1 \sum X_i^2$$

$$\begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X^2{}_i \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

ولتقدير b_0 و b_1 ينبغي تكوين المحددات الاتية :

$$|\Delta| = \begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum X^2{}_i \end{vmatrix} \implies \text{المحدد الاصلي}$$

$$\therefore |\Delta| = \begin{vmatrix} 8 & 144 \\ 144 & 3264 \end{vmatrix} = (8)(3264) - (144)(144)$$

$$\therefore |\Delta| = 5376$$

$$|A_0| = \begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X^2{}_i \end{vmatrix}$$

$$\therefore |A_0| = \begin{vmatrix} 410 & 144 \\ 8379.2 & 3264 \end{vmatrix}$$

$$= (410)(3264) - (144)(8379.2)$$

$$\therefore |A_0| = 131635.2$$

وان :

$$|A_1| = \begin{vmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix}$$

$$\therefore |A_1| = \begin{vmatrix} 8 & 410 \\ 144 & 8379.2 \end{vmatrix} = (8)(8379.2) - (410)(144)$$

$$\therefore |A_1| = 7993.6$$

وعليه فان :

$$b_0 = \frac{|A_0|}{|\Delta|} = \frac{131635.2}{5376} = 24.4858$$

$$b_1 = \frac{|A_1|}{|\Delta|} = \frac{7993.6}{5376} = 1.4869$$

وبذلك تكون الصيغة التقديرية للمعادلة :

$$\widehat{Y}_i = 24.4858 + 1.4869\, X_i$$

ثالثا : طريقة التقدير حول نقطة المتوسط (طريقة الانحرافات)

اسلوب التقدير وفق الطرق السابقة يعرف بالتقدير حول نقطة الاصل(التقدير باستخدام القيم الاصلية) . كما يمكن تقدير b_0 و b_1 بواسطة (باستخدام) انحرافات المتغيرين Y_i و X_i عن وسطهما الحسابي $\overline{Y}$ و $\overline{X}$ باستخدام فكرة البواقي e_i فبقسمة المعادلة الطبيعية الاتية على n نحصل على :

$$\sum Y_i = n\, b_0 + b_1 \sum X_i$$

$$\frac{\sum Y_i}{n} = \frac{n b_0}{n} + b_1 \frac{\sum X_i}{n}$$

$$\overline{Y} = b_0 + b_1 \overline{X} \dots\dots\dots\dots (10-2)$$

$$\therefore\ b_0 = \overline{Y} - b_1 \overline{X} \dots\dots\dots\dots (11-2)$$

ولإيجاد b_1 نعود الى معادلة خط المستقيم التقديري حيث ان

$$\widehat{Y}_i = b_0 + b_1 X_i \quad \dots\dots\dots\dots (12-2)$$

وبطرح المعادلة رقم (2-10) من المعادلة (2-12) نحصل على :

$$\widehat{Y}_i - \overline{Y} = b_0 + b_1 X_i - b_0 - b_1 \overline{X}$$

$$\widehat{Y}_i - \overline{Y} = b_1 X_i - b_1 \overline{X}$$

$$\therefore \widehat{Y}_i - \overline{Y} = b_1 (X_i - \overline{X})$$

$$\therefore \hat{y}_i = b_1 x_i \quad \dots\dots\dots\dots \quad (13-2)$$

اذ ان :

$$\mathbf{y_i = Y_i - \bar{Y}} \quad , \quad \mathbf{\hat{y}_i = \hat{Y}_i - \bar{Y}}$$

وان :

$$\boldsymbol{x_i} = \mathbf{X_i - \bar{X}}$$

$$\therefore\ \mathbf{e_i = y_i - \hat{y}_i}$$

$$\therefore\ \mathbf{e_i = y_i - b_1}\, \boldsymbol{x_i} \quad \ldots\ldots\ldots\ldots (\mathbf{14-2})$$

وبإدخال المجموع بعد تربيع الطرفين وتطبيق المربعات الصغرى فان

$$\sum \mathbf{e^2}_{\mathbf{i}} = \sum(\mathbf{y_i - b_1}\, \boldsymbol{x_i})^2$$

وبإيجاد المشتقة الجزئية الاولى ل $\mathbf{b_1}$ ومساواتها للصفر نحصل على ان

$$\frac{\partial \sum \mathbf{e^2}_{\mathbf{i}}}{\partial \mathbf{b_1}} = 2\sum(\mathbf{y_i - b_1\, x_i})(-\boldsymbol{x_i}) = 0$$

$$-2 \sum \mathbf{x_i}\, (\mathbf{y_i - b_1}\, \boldsymbol{x_i}) = 0$$

وبالقسمة على (-2) واعادة ترتيب المعادلة نحصل على :

$$\sum \boldsymbol{x_i}\mathbf{y}_i - \mathbf{b_1} \sum \boldsymbol{x_i^2} = 0$$

$$\sum \boldsymbol{x_i}\mathbf{y}_i = \mathbf{b_1} \sum \boldsymbol{x_i^2}$$

$$\therefore\ \mathbf{b_1} = \frac{\sum \boldsymbol{x_i\, y_i}}{\sum \boldsymbol{x_i^2}} \quad \ldots\ldots\ldots\ldots \quad (15-2)$$

اذ ان :

$\sum \boldsymbol{x_i\, y_i}$: مجموع حاصل ضرب انحرافات قيم $\boldsymbol{X}_{\mathbf{i}}$ و $\boldsymbol{Y}_{\mathbf{i}}$ عن وسطهما الحسابي $\mathbf{\bar{X}}$ و $\mathbf{\bar{Y}}$.

$\sum \boldsymbol{x_i^2}$: مجموع مربعات انحرافات قيم المتغير $\mathbf{X_i}$ عن وسطه الحسابي $\mathbf{\bar{X}}$.

وعليه فان المعادلتين $(2-11)$ و $(2-15)$ هي المعادلات الاساسية التي تستخدم في ايجاد القيم التقديرية $\mathbf{b_0}$ و $\mathbf{b_1}$ بموجب طريقة الانحرافات .

وبالرجوع الى بيانات المثال السابق رقم (1) والتعبير عنها بصيغة الانحرافات عن المتوسط يمكن الحصول على قيم $\mathbf{b_0}$ و $\mathbf{b_1}$ وكما في الجدول الاتي :

Y_i	X_i	$x_i = X_i - \bar{X}$	$y_i = Y_i - \bar{Y}$	$x_i y_i$	x_i^2
25.6	4	−14	−25.65	359.1	196
32.7	8	−10	−18.55	185.5	100
45.4	12	−6	−5.85	35.1	36
53.9	16	−2	2.65	−5.3	4
59.0	20	2	7.75	15.5	4
62.6	24	6	11.35	68.1	36
65.0	28	10	13.75	137.5	100
65.8	32	14	14.55	203.7	196
$\sum = 410$	$\sum = 144$	$\sum = 0$	$\sum = 0$	$\sum = 999.2$	$\sum = 672$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{144}{8} = 18$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{410}{8} = 51.25$$

وبموجب طريقة الانحرافات فان :

$$\mathrm{b}_1 = \frac{\sum x_i\, y_i}{\sum {x^2}_i} = \frac{999.2}{672} = 1.4869$$

وان :

$$b_0 = \bar{Y} - b_1 \bar{X} = 51.25 - (1.4869)(18)$$

$$\therefore\ b_0 = 24.4857$$

∴ الصيغة التقديرية للعلاقة هي :

$$\hat{Y}_i = 24.4857 + 1.4869\, X_i$$

وهي نفس النتيجة السابقة .

<u>**ملاحظة:**</u> في حالة إعطاء المجاميع مباشرة عن بيانات العينة فان المجاميع بدلالة الانحرافات تحسب وفق الصيغ الاتية:

$$\Sigma x^2{}_i = \Sigma X^2{}_i - \frac{(\Sigma X_i)^2}{n} = \Sigma X^2{}_i - n\,\bar{X}^2$$

$$\Sigma y^2{}_i = \Sigma Y^2{}_i - \frac{(\Sigma Y_i)^2}{n} = \Sigma Y^2{}_i - n\,\bar{Y}^2$$

$$\Sigma x_i y_i = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) = \Sigma X_i Y_i - \frac{(\Sigma X_i)(\Sigma Y_i)}{n} = \Sigma X_i Y_i - n\,\bar{X}\bar{Y}$$

رابعا:- طريقة المصفوفات Matrices Method

بالإضافة الى الطرق السابقة فانه يمكن تقدير قيم b_0 و b_1 باعتماد صيغة المصفوفات، حيث يمكن كتابة المعادلات الطبيعية على شكل مصفوفات Matrices ومتجهات Vectors وكما يلي :

$$\Sigma Y_i = nb_o + b_1 \Sigma X_i$$

$$\Sigma X_i Y_i = b_o \Sigma X_i + b_1 \Sigma X_i^2$$

$$\begin{bmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \end{bmatrix} = \begin{bmatrix} n & \Sigma x_i \\ \Sigma X_i & \Sigma X_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$X'Y = (X'X)\, b$$

$$(X'X) b = X'Y$$

بالضرب المتقدم في معكوس مصفوفة $(X'X)$ المعلومات (المعاملات) نحصل على ان:

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = (X^{\prime}X)^{-1}\, X^{\prime}Y$$

اذ ان:

$(X^{\prime}X)^{-1}$: هي معكوس مصفوفة المعلومات (المعاملات).

$X^{\prime}Y$: قيمة الحدود المطلقة للمعادلات الطبيعية.

علماً ان :

$$(X^{\prime}X)^{-1} = \frac{adj(X^{\prime}X)}{|X^{\prime}X|}$$

وبالرجوع الى بيانات المثال السابق (1) وباستخدام صيغة المصفوفات يمكن تقدير قيم معالم العلاقة بين معدل الاجر السنوي (Y_i) وعدد سنوات الخدمة (X_i) وفق الصيغة :

$$b = (X^{-}X)^{-1}\, X^{-}Y$$

اذ ان :

$$(X^{-}X) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 8 & 12 & 16 & 20 & 24 & 28 & 32 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 8 \\ 1 & 12 \\ 1 & 16 \\ 1 & 20 \\ 1 & 24 \\ 1 & 28 \\ 1 & 32 \end{bmatrix}$$

$$(X^{-}X) = \begin{bmatrix} 8 & 144 \\ 144 & 3264 \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum {X^2}_i \end{bmatrix}$$

$$(X^{-}X)^{-1} = \frac{adj(X^{-}X)}{|X^{-}X|}$$

$$|X^{-}X| = \begin{vmatrix} 8 & 144 \\ 144 & 3264 \end{vmatrix} = (8)(3264) - (144)(144) = 5376$$

$$\boldsymbol{adj}(\boldsymbol{X}^{-}\boldsymbol{X}) = \begin{bmatrix} 3264 & -144 \\ -144 & 8 \end{bmatrix}$$

$$\therefore\ (\boldsymbol{X}^{-}\boldsymbol{X})^{-\boldsymbol{1}} = \frac{\boldsymbol{1}}{\boldsymbol{5376}} \begin{bmatrix} 3264 & -144 \\ -144 & 8 \end{bmatrix}$$

$$(\boldsymbol{X}^{-}\boldsymbol{X})^{-\boldsymbol{1}} = \begin{bmatrix} 0.607142857 & -0.026785714 \\ -0.026785714 & 0.001488095 \end{bmatrix}$$

$$\boldsymbol{X}^{-}\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{1} \\ \boldsymbol{4} & \boldsymbol{8} & \boldsymbol{12} & \boldsymbol{16} & \boldsymbol{20} & \boldsymbol{24} & \boldsymbol{28} & \boldsymbol{32} \end{bmatrix} \begin{bmatrix} \boldsymbol{25.6} \\ \boldsymbol{32.7} \\ \boldsymbol{45.4} \\ \boldsymbol{53.9} \\ \boldsymbol{59.0} \\ \boldsymbol{62.6} \\ \boldsymbol{65.0} \\ \boldsymbol{65.8} \end{bmatrix}$$

$$\therefore\ \boldsymbol{X}^{-}\boldsymbol{Y} = \begin{bmatrix} \sum \boldsymbol{Y_i} \\ \sum \boldsymbol{X_i Y_i} \end{bmatrix} = \begin{bmatrix} \boldsymbol{410} \\ \boldsymbol{8379.2} \end{bmatrix}$$

$$\therefore\ \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b_0} \\ \boldsymbol{b_1} \end{bmatrix} = \widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^{-}\boldsymbol{X})^{-\boldsymbol{1}}\ \boldsymbol{X}^{-}\boldsymbol{Y}$$

$$\therefore\ \boldsymbol{b} = \begin{bmatrix} 0.607142857 & -0.026785714 \\ -0.026785714 & 0.001488095 \end{bmatrix} \begin{bmatrix} 410 \\ 8379.2 \end{bmatrix}$$

$$\therefore\ \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b_0} \\ \boldsymbol{b_1} \end{bmatrix} = \begin{bmatrix} 24.48571 \\ 1.48690 \end{bmatrix}$$

لذلك فانه معادلة خط الانحدار التقديرية ستأخذ الشكل الاتي :

$$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}} = \boldsymbol{24.4858} + \boldsymbol{1.4869}\ \boldsymbol{X_i}$$

يتضح مما سبق بان النتائج التي تم التوصل اليها لتقدير معادلة خط الانحدار متطابقة في الطرق الأربعة.

تمرين : البيانات التالية تمثل العلاقة بين الدخل (X_i) والانفاق على الرعاية الصحية (Y_i) في احدى المجتمعات خلال فترة محدودة :

$\sum X_i$	395.27	$\sum X_i^2$	9440.3
$\sum Y_i$	107.81	$\sum Y_i^2$	945.5
$\sum X_i Y_i$	2356.41	n	20

المطلوب :

1- اثبت ان معادلة خط انحدار الانفاق على الدخل هي :

$$\hat{Y}_i = 2.6513 + 0.1386\, X_i$$

2- تقدير الميل الحدي للأنفاق الصحي على الدخل .

3- تقدير المرونة الانفاقية بالنسبة للدخل عندما يكون $(\bar{X} = 20)$.

خصائص مقدرات المربعات الصغرى Properties of least squares estimators:

ان من أهم اسباب شيوع استخدام طريقة المربعات الصغرى الاعتيادية (OLS) في تقدير العلاقات الاقتصادية هو إتصاف تقديرات معاملات هذه العلاقات بخصائص مثلى، وفيما يلي تعريف بالخصائص التي تعتبرها النظرية الاحصائية التقليدية مميزات مرغوب فيها لتقديرات العينة، مع التأكيد على ان تقديرات المربعات الصغرى بالفعل تتصف بهذه الخصائص المثلى .

1- الخصائص المرغوب فيها للمقدرات

أ- خصائص المقدرات في العينات الصغيرة .

من اهم معايير (خصائص او صفات) المقدرات الجيدة في العينات الصغيرة ما يلي :

1- عدم التحيز .
2- أصغر تباين .
3- الكفاءة .

4- افضل مقدر خطي غير متحيز (BLUE) .
5- اصغر مربع لمتوسط الخطأ (MSE) .
6- الكفاية .

وفيما يلي توضيح بسيط لكل منها:

1 - مقدر غير متحيز : Unbiased estimator

يعرف تحيز **bias** اي مقدر بانه الفرق بين القيمة المتوقعة لهذا المقدر وبين قيمة المعامل الحقيقي اي ان :

$$\mathbf{Bias\ b_0 = E(b_0) - \boldsymbol{\beta}_0}$$

$$\mathbf{Bias\ b_1 = E(b_1) - \boldsymbol{\beta}_1}$$

ويكون المقدر غير متحيز اذا كان التحيز مساويا للصفر اي انه اذا كان

$$\mathbf{E\,(b_0) = \boldsymbol{\beta}_0} \qquad \textbf{and} \qquad \mathbf{E\,(b_1) = \boldsymbol{\beta}_1}$$

ويعني ذلك ان قيمة المقدر غير المتحيز تقترب من القيمة الفعلية (الحقيقية) للمعامل كلما زاد عدد العينات المستخدمة.

وعدم التحيز خاصية مرغوبة الا انها ليست مهمة بحد ذاتها ما لم تقترن بخاصية اخرى وهي صغر التباين.

2- مقدر ذو اصغر تباين (افضل مقدر) Best Estimator

يعتبر أي تقدير افضل من غيره من التقديرات اذا كان تباينه اصغر تباين بين التقديرات الاخرى التي حصلنا عليها بطرق الاقتصاد القياسي الاخرى. وبصيغة رياضية نقول ان $\boldsymbol{\hat{\beta}_i}$ افضل مقدر اذا كان :

$$E\,[\,\hat{\beta}_i - E\,(\hat{\beta}_i)]^2 \;<\; E\,[\,\beta^*{}_i - E\,(\beta^*{}_i)]^2$$

$$Var\,(\hat{\beta}_i) \;<\; Var\,(\beta^*{}_i)$$

حيث ان :

$\boldsymbol{\beta}^*{}_{\mathrm{i}}$: اي تقدير اخر للمعامل الحقيقي $\boldsymbol{\beta}_{\mathrm{i}}$ وقد يكون $\boldsymbol{\beta}^*{}_{\mathrm{i}}$ تقديرا متحيزاً، ولا تعني خاصية صغر التباين شيئا بحد ذاتها، وتزداد اهميتها باقترانها مع اصغر تحيز .

3- المقدر ذو أصغر مربع لمتوسط الخطأ: Minimum mean square error Estimator (MSE)

هذا المعيار مزيج من صفات عدم التحيز واصغر تباين وبناءً على هذا المعيار نفضل المقدر الذي يتصف بأصغر مربع لمتوسط الخطأ والذي يعرف بانه القيمة المتوقعة لمربع الفرق بين المقدر ومعامل المجتمع الاحصائي الحقيقي اي انه: $MSE = E\,(\hat{\beta}_i - \beta_i)^2$

ويمكن اثبات ان **MSE** يساوي مجموع تباين المقدر ومربع مقدار تحيز هذا المقدر اي ان :

$$MSE = var\left(\hat{\beta}_i\right) + bias^{\,2}\left(\hat{\beta}_i\right)$$

ويمكن اثبات ذلك كما يلي :

$$MSE = E\,(\hat{\beta}_i - \beta_i)^2$$

وبإضافة وطرح $E\left(\hat{\beta}_i\right)$

$$MSE = E[\{\hat{\beta}_i - E\left(\hat{\beta}_i\right)\} + \{E\left(\hat{\beta}_i\right) - \beta_i\}]^2$$

$$= E[\hat{\beta}_i - E(\hat{\beta}_i)]^2 + [E\left(\hat{\beta}_i\right) - \beta_i]^2$$
$$+ 2E\left[\{\hat{\beta}_i - E\left(\hat{\beta}_i\right)\}\{E\left(\hat{\beta}_i\right) - \beta_i\}\right]$$

$$\therefore\ MSE = Var\left(\hat{\beta}_i\right) + bias^2\left(\hat{\beta}_i\right)$$

لان الحد الثالث يساوي صفراً وكما يلي :

$$E\left\{[\hat{\beta}_i - E\left(\hat{\beta}_i\right)][E\left(\hat{\beta}_i\right) - \beta_i]\right\}$$

$$= E\{\hat{\beta}_i\, E\left(\hat{\beta}_i\right) - [E\left(\hat{\beta}_i\right)]^2 - \hat{\beta}_i\, \beta_i + \beta_i\, E\left(\hat{\beta}_i\right)\}$$

$$= [E\,(\cancel{\hat{\beta}_i})]^2 - [E\,(\cancel{\hat{\beta}_i})]^2 - \beta_i\, E\,(\cancel{\hat{\beta}_i}) + \beta_i\, E\,(\cancel{\hat{\beta}_i}) = 0$$

$$\therefore\ MSE = Var\left(\hat{\beta}_i\right) + bias^2\left(\hat{\beta}_i\right)$$

4- مقدر كفوء Efficient Estimator

يكون المقدر كفوء اذا اتصف بخاصيتي عدم التحيز واصغر تباين في نفس الوقت عند مقارنته مع بقية المقدرات الاخرى غير المتحيزة، وبصيغة رياضية يكون المقدر $\hat{\beta}_i$ كفوء اذا استوفى الشرطين التاليين :

$$1 - \qquad E\left(\hat{\beta}_i\right) = \beta_i$$

$$2 - \quad E\,[\,\hat{\beta}_i - E\,(\hat{\beta}_i)\,]^2 \; < \; E\,[\,\beta^*{}_i - E\,(\beta^*{}_i)\,]^2$$

حيث ان :

$\beta^*{}_i$: هو أي مقدر اخر غير متحيز. وبعبارة اخرى المقدر الكفوء هو أفضل مقدر في مجموعة المقدرات غير المتحيزة.

5- أفضل مقدر خطي غير متحيز Best Linear Unbiased Estimator (BLUE)

يعد اي مقدر متصفاً بالصفة **(BLUE)** اذا كان غير متحيزا وتباينه اصغر تباين قياساً الى المقدرات الاخرى غير المتحيزة وفي نفس الوقت هو عبارة عن دالة خطية في مشاهدات العينة. فاذا كانت لدينا العينة $Y_1 ,\ Y_2 ,\ldots\ldots\ldots, Y_n$ فان المقدر الخطي يتخذ الشكل الاتي :

$$\hat{\beta}_i = F\,(Y_i) = K_1Y_1 + K_2Y_2 + \cdots + K_nY_n = \sum_{i=1}^{n} K_iY_i$$

حيث ان :

K_i : *مقدار ثابت* ، $i = 1, 2, 3, \ldots\ldots\ldots n$

وتتلخص أهمية خاصية الـ **(BLUE)** للمقدرات بما يلي :

أ- التقديرات (المقدرات) الخطية تتسم بخاصية البساطة مقارنة بالتقديرات غير الخطية. بعبارة أخرى الخصائص الخطية مرغوبة لأنها تسهل عملية احتساب قيم التقديرات.

ب- عدم التحيز ليس مفيداً اذا لم يقترن بخاصية أصغر تباين.

ت- صغر التباين يفيد في اختبارات المعنوية الإحصائية وفي حساب حدود الثقة ، حيث صغر التباين يؤدي الى ان تكون المعالم المقدرة معنوية من الناحية الإحصائية وان حدود الثقة تكون اصغر في حالة صغر التباين .

ث- دقة التنبؤات المستحصلة في حالة تحقق خاصية الـ **(BLUE)** .

6- المقدر الكافي Sufficient Estimator

يكون المقدر كافي اذا كان يستخدم كل المعلومات التي تحتويها العينة عن المعاملات الحقيقة، اي انه لا يمكن لأي مقدر اخر ان يضيف اي معلومات جديدة عن المعامل الحقيقي للمجتمع الاحصائي المراد تقديره، والكفاية ليست خاصية هامة بحد ذاتها ولكنها شرط ضروري لتحقيق الكفاءة .

ب- خصائص المقدرات في العينات الكبيرة (خصائص تقاربية)

لتطبيق هذه الخصائص كمعايير لجودة مقدر ما يجب ان يكون حجم العينة متناهي الكبر اي $(n \to \infty)$ ولهذا السبب تسمى بالخصائص التقاربية لأنها تتوافر بصورة تقاربية فقط عندما يكون حجم العينة كبيرا . والخصائص التقاربية هي :

1 - عدم التحيز التقاربي Asymptotic Unbiasedness

اي مقدر يكون تقاربياً غير متحيز اذا كان تحيزه يميل الى التلاشي بزيادة حجم العينة واقترابه من اللانهاية . وبصيغة رياضية يكون المقدر تقاربياً غير متحيزا اذا كان :

$$\boldsymbol{plim} \left\{ \boldsymbol{E}(\widehat{\boldsymbol{\beta}}_i) - \boldsymbol{\beta}_i \right\} \to \mathbf{0} \quad \boldsymbol{as} \quad \boldsymbol{n} \to \infty$$

اي ان حدود احتمال **(probability limit)** التحيز هي صفر. بمعنى انه بزيادة حجم العينة يتناقص تحيز المقدر $\widehat{\boldsymbol{\beta}}_i$ حتى يصل الى الصفر .

2 - الاتساق Consistency

يكون المقدر متسقاً إذا كان تقاربياً غير متحيزاً وإذا كان توزيعه يقترب من المعامل الحقيقي للمجتمع الاحصائي باقتراب حجم العينة من اللانهاية. وبصيغة جبرية تتوافر صفة الاتساق اذا كان:

$$plim\ E\left(\widehat{\boldsymbol{\beta}}_i\right) = \boldsymbol{\beta}_i$$

اي ان حدود احتمال $\widehat{\boldsymbol{\beta}}_i$ هي قيمة المعامل الحقيقي . اي انه لكي يكون المقدر متسقا يجب ان يستوفي الشرطين التاليين:

أ- ان يكون تقاربياً غير متحيزاً.

ب-ان يتركز توزيعه اكثر فاكثر عند $\boldsymbol{\beta}_i$ باقتراب حجم العينة من اللا نهاية .

وان الشرط الضروري والكافي لكي يكون المقدر متسقا هو: $plim\ (MSE) = 0$

3 - الكفاءة التقاربية Asymptotic Efficiency

يمكن القول عن مقدر ما انه تقاربي كفوء اذا كان متسقاً وكان تباينه يقترب من الصفر اسرع من تباين غيره من المقدرات المتسقة بمعنى استخدام عينة اصغر حجماً للحصول على هذا المقدر. ولكي يكون $\widehat{\boldsymbol{\beta}}_i$ تقاربياً كفوء لابد من توفر الشروط التالية :

$$1 - \quad plim\ E\left(\widehat{\boldsymbol{\beta}}_i\right) = \boldsymbol{\beta}_i$$

$$2 - \quad Var\left(\hat{\beta}_i\right) < Var\left({\beta^*}_i\right)$$

حيث ان ${\beta^*}_i$ اي مقدر اخر متسق، ويلاحظ ان تباين كل المقدرات المتسقة يصبح مساوياً للصفر في النهاية عندما يصبح حجم العينة متناهي الكبر، الا ان تباين بعض المقدرات يكون اسرع في وصولة للصفر من غيرة من المقدرات بزيادة حجم العينة ولذلك يفضل على غيره .

2- خصائص مقدرات المربعات الصغرى Properties of (OLS)

في هذا الجزء سنحاول البرهنة على اتصاف مقدرات المربعات الصغرى بالخاصية **(BLUE)** اي افضل مقدرات خطية غير متحيزة، بشرط ان يستوفي المتغير العشوائي u_i الفروض العامة له، وهو مايطلق عليه بنظرية كاوس ماركوف **Causs-Markov Theorem** ، والمراد من هذه النظرية اثبات انه من بين كل المقدرات الخطية وغير المتحيزة تتصف مقدرات المربعات الصغرى بانها أفضل مقدرات ممكنة لان تباينها هو اصغر تباين ممكن، وسوف نتناول في ادناه هذه الخصائص بشيء من التفصيل .

1 - خاصية الخطية Linearity

أي ان تقديرات المربعات الصغرى هي دوال خطية في قيم العينة المشاهدة Y_i ، وبالرجوع الى الصيغ التي تم اشتقاقها سابقاً لاحتساب تقديرات المعاملات نجد انها تشتمل على المتغيرات X_i و Y_i وبما اننا افترضنا ان قيم X تبقى ثابتة في المعاينة المتكررة، فانه يمكننا اثبات ان تقديرات **(OLS)** تعتمد على قيم Y_i فقط . اي ان

$$b_0 = F\,(Y_i)$$

$$b_1 = F\,(Y_i)$$

ولإثبات ذلك

أ- اثبت ان $b_1 = F\,(Y_i)$

$$\therefore \quad b_1 = \frac{\sum x_i y_i}{\sum {x^2}_i}$$

اذ ان :

$$y_i = Y_i - \bar{Y}$$

فان

$$b_1 = \frac{\sum x_i(Y_i - \bar{Y})}{\sum x^2{}_i} = \frac{\sum x_i Y_i - \bar{Y}\sum x_i}{\sum x^2{}_i}$$

لكن

$$\sum x_i = \sum(X_i - \bar{X}) = 0$$

$$b_1 = \sum x_i Y_i / \sum x^2{}_i$$

$$= \sum\left[\frac{x_i}{\sum x^2{}_i}\right] Y_i$$

ولما كانت قيم X ثابته، فان $\left[\frac{x_i}{\sum x^2{}_i}\right]$ هي مقادير ثابته ويمكن ان نرمز لها بالرمز K_i وهي ثابته في كل العينات . اي ان

$$K_i = \frac{x_i}{\sum x^2{}_i}$$

$$\therefore \quad b_1 = \sum K_i Y_i \ldots\ldots\ldots\ldots\ldots \quad (18-2)$$

$$= K_1Y_1 + K_2Y_2 + \ldots\ldots\ldots\ldots + K_nY_n$$

$$\therefore \quad b_1 = \hat{\beta}_i = F\,(Y_i)$$

اي ان b_1 دالة خطية في قيم العينة Y_i . (اي ان b_1 هو مقدر خطي)

ب- اثبت ان $\quad b_0 = F\,(Y_i)$

$$\therefore \quad b_0 = \bar{Y} - b_1\bar{X}$$

لكن

$b_1 = \sum K_i Y_i$ وان $\bar{Y} = \frac{\sum Y_i}{n}$

$$\therefore \quad b_0 = \frac{\sum Y_i}{n} - \bar{X}\sum K_i Y_i$$

$$b_0 = \sum\left[\frac{1}{n} - \bar{X}K_i\right] Y_i \ \ldots\ldots\ldots\ldots (19-2)$$

ولما كانت $\bar{X}$ و K_i مقادير ثابتة فان $\left[\frac{1}{n} - \bar{X}K_i\right]$ هي مقادير ثابتة نرمز لها بالرمز w_i اي ان

$$W_i = \left[\ \frac{1}{n} - \bar{X}\,K_i\ \right]$$

$$\therefore\ \ b_O = \sum W_i\,Y_i$$

$$= W_1Y_1 + W_2Y_2 + \cdots + W_nY_n$$

$$\therefore\ \ b_O = \hat{\beta}_O = F\,(Y_i)$$

اي ان b_0 تعتمد على قيم Y_i فقط، بعبارة اخرى فان b_0 دالة خطية في قيم العينة Y_i (b_0 مقدر خطي)

خصائص الاوزان

بما ان الاوزان K_i و W_i تعتمد على قيم X الثابتة فقط فانها تعتبر ثابتة وتخضع K_i للشروط الاتية :

أ– مجموع الاوزان يساوي صفراً اي ان : $\sum_{i=1}^{n} K_i = 0$

$$K_i = \frac{x_i}{\sum x^2{}_i} \qquad\qquad \therefore\ \ \sum_{i=1}^{n} K_i = \frac{\sum x_i}{\sum x^2{}_i} = \frac{0}{\sum x^2{}_i} = 0$$

لان $\sum x_i = \sum(X_i - \bar{X}\,) = 0$

ب– مجموع حاصل ضرب K_i في قيم المتغير المستقل X_i او في انحرافاتها عن متوسطها الحسابي

(x_i) يساوي الواحد الصحيح اي ان : $\sum_{i=1}^{n} K_i\,X_i = \sum_{i=1}^{n} k_i\,x_i = 1$

$$\sum k_i x_i = \sum k_i\ \ (X_i - \bar{X}) = \sum K_i X_i - \bar{X}\sum K_i$$

$$\therefore\ \ \sum k_i = 0$$

$$\therefore\ \ \sum k_i x_i = \sum K_i X_i$$

ولإثبات انها تساوي واحد

$$\therefore\ \ k_i = \frac{x_i}{\sum x^2{}_i}$$

وبضرب الطرفين بـ (x_i) وأدخال المجموع ينتج :

$$\therefore \quad \sum k_i x_i = \cancel{\frac{\sum x^2{}_i}{\sum x^2{}_i}} = 1$$

$$\therefore \quad \sum k_i x_i = \sum K_i X_i = 1$$

ت− مجموع مربعات الاوزان $(\sum K^2{}_i)$ يساوي معكوس (مقلوب) مجموع مربعات انحرافات المتغير المستقل $(\frac{1}{\sum x^2{}_i})$ اي ان:

$$\sum_{i=1}^{n} K^2{}_i = \frac{1}{\sum_{i=1}^{n} x^2{}_i} = \frac{1}{\sum(X_i - \bar{X})^2}$$

البرهان :

$$\therefore \quad k_i = \frac{x_i}{\sum x^2{}_i} = \frac{(X_i - \bar{X})}{\sum(X_i - \bar{X})^2}$$

وعند تربيع الطرفين ثم جمعهما يكون:

$$\sum K^2{}_i = \frac{\sum(X_i - \bar{X}\cancel{)^2}}{[\sum(X_i - \bar{X})^2]^2} = \frac{1}{\sum(X_i - \bar{X})^2}$$

$$\therefore \quad \sum K^2{}_i = \frac{1}{\sum x^2{}_i}$$

2 - خاصية عدم التحيز **unbiasedness property**

اي اثبات ان $E(b_1) = \beta_1$ and $E(b_0) = \beta_0$

Ḿ اثبات ان $E(b_0) = \beta_0$

$$\therefore \quad b_0 = \sum W_i Y_i = \sum \left[\frac{1}{n} - \bar{X} K_i\right] (Y_i)$$

وبأخذ التوقع للطرفين

$$\therefore \quad E(b_0) = \sum \left[\frac{1}{n} - \bar{X} K_i\right] E(Y_i)$$

اذ ان : $E(Y_i) = \beta_0 + \beta_1 X_i$

$$\therefore \quad E(b_0) = \sum\left[\frac{1}{n} - \bar{X}\,K_i\right](\beta_0 + \beta_1 X_i)$$

$$= \frac{\cancel{n}\beta_0}{\cancel{n}} - \beta_0\,\bar{X}\sum K_i + \beta_1\,\frac{\sum X_i}{n} - \bar{X}\,\beta_1\sum K_i X_i$$

$$E(b_0) = \beta_0 - 0 + \cancel{\beta_1 \bar{X}} - \cancel{\beta_1 \bar{X}}$$

$$\therefore \quad E(b_0) = \beta_0$$

أي ان : b_0 هو تقدير غير متحيز للمعلمة الحقيقية (β_0) .

*M*ـ اثبات ان $E\,(b_1) = \beta_1$

$$\therefore \quad b_1 = \sum K_i Y_i$$

$$= \sum K_i(\beta_0 + \beta_1 X_i + u_i)$$

$$= \beta_0\sum K_i + \beta_1\sum K_i X_i + \sum K_i u_i$$

$$b_1 = \quad 0 \quad + \beta_1 + +\sum K_i u_i \ldots\ldots\ldots\ldots\ldots \quad (20-2)$$

وبأخذ التوقع للطرفين

$$\therefore \quad E(b_1) = \beta_1 + +\sum K_i\,E(u_i) \qquad \text{وبما ان} \qquad E(u_i) = 0$$

$$\therefore \quad E(b_1) = \beta_1$$

أي ان : b_1 هو تقدير غير متحيز للمعلمة الحقيقية (β_1) .

3 - خاصية أصغر تباين **The minimum variance property**

هنا نحاول اثبات Gauss-Markov theorem التي تقول بان تقديرات المربعات الصغرى (OLS) هي أفضل التقديرات بمقارنتها مع اي مقدر اخر خطي وغير متحيز يحصل عليه بطرق الاقتصاد القياسي الأخرى، وهذه الخاصية هي السبب الرئيسي وراء شيوع استخدام طريقة المربعات الصغرى.

ولغرض اثبات ان تباين $\mathbf{b_0}$ و $\mathbf{b_1}$ هو أصغر تباين ممكن فان الخطوة الاولى تتطلب ايجاد صيغة لتقدير تباين $\mathbf{b_0}$ و $\mathbf{b_1}$ كما يلي :

اشتقاق صيغة لتقدير تباين (b_0) :

$$\mathbf{b_0} = \bar{Y} - \mathbf{b_1}\,\bar{X}$$

وبالتعويض في المعادلة (2-18)

$$= \frac{\sum Y_i}{n} - \bar{X}\ \sum K_i Y_i$$

$$\therefore\ \mathbf{b_0} = \sum\left(\frac{1}{n} - \bar{X}\,K_i\right) Y_i$$

وبالتعويض عن Y_i بما يساويها نحصل على

$$\mathbf{b_0} = \sum\left(\frac{1}{n} - \bar{X}\,K_i\right)(\beta_0 + \beta_1 X_i + u_i)$$

وبعد التبسيط ينتج ان

$$\mathbf{b_0} = \beta_0 + \sum\left(\frac{1}{n} - \bar{X}\,K_i\right) u_i$$

$$\therefore\ \mathbf{b_0} - \beta_0 = \sum\left(\frac{1}{n} - \bar{X}\,K_i\right) u_i$$

اذ ان :

$$Var\,(\mathbf{b_0}) = E(\mathbf{b_0} - \beta_0)^2$$

$$\therefore\ Var\,(\mathbf{b_0}) = E\,(\mathbf{b_0} - \beta_0)^2 = E\left[\sum\left(\frac{1}{n} - \bar{X}\,K_i\right) u_i\right]^2$$

$$Var\,(\mathbf{b_0}) = \sum\left(\frac{1}{n} - \bar{X}\,K_i\right)^2 E\,(u_i)^2 + 2\sum\left(\frac{1}{n} - \bar{X}\,K_i\right)\left(\frac{1}{n} - \bar{X}\,K_j\right) E\left(u_i u_j\right)$$

$$\therefore\ E(u_i^2) = \sigma^2{}_u \qquad \text{وان} \qquad E\left(u_i u_j\right) = 0$$

$$\therefore\ Var\,(\mathbf{b_0}) = \sigma^2{}_u\left[\sum\left(\frac{1}{n} - \bar{X}\,K_i\right)^2\right]$$

وبفك الاقواس نحصل على:

$$Var\ (\mathbf{b_0}) = \sigma^2{}_u \left(\frac{\mathbf{1}}{\boldsymbol{n}} - \frac{2\overline{\boldsymbol{X}}}{n}\sum \boldsymbol{K_i} + \overline{\boldsymbol{X}}^2 \sum \boldsymbol{K_i}^2\right)$$

اذ ان $\sum \boldsymbol{K_i} = 0$ وان $\sum \boldsymbol{K_i}^2 = \frac{\mathbf{1}}{\sum \boldsymbol{x_i^2}}$

لذلك فان

$$\therefore \quad Var\ (\mathbf{b_0}) = \sigma^2{}_u \left[\frac{\mathbf{1}}{\boldsymbol{n}} + \frac{\overline{\boldsymbol{X}}^2}{\sum \boldsymbol{x_i^2}}\right] \quad \dots\dots\dots\dots \quad (21-2)$$

وبالأسلوب نفسة يمكن تحديد صيغة التباين ل $\mathbf{b_1}$ وعلى النحو التالي :

اشتقاق صيغة لتقدير تباين $(\mathbf{b_1})$:

من العلاقة (2−18) نجد ان

$$\boldsymbol{b_1} = \sum \boldsymbol{K_i Y_i}$$

وبالتعويض عن $\boldsymbol{Y_i}$ نحصل على ان

$$\boldsymbol{b_1} = \sum \boldsymbol{K_i}(\boldsymbol{\beta_0} + \boldsymbol{\beta_1 X_i} + \boldsymbol{u_i})$$

$$\therefore \quad \boldsymbol{b_1} = \boldsymbol{\beta_1} + \sum \boldsymbol{K_i u_i}$$

$$\boldsymbol{b_1} - \boldsymbol{\beta_1} = \sum \boldsymbol{K_i u_i}$$

$$\boldsymbol{Var\ (b_1) = E\ (b_1 - \beta_1)^2 = E\ [\textstyle\sum K_i u_i]^2}$$

$$= \sum_{i=1}^{n} \boldsymbol{K_i}^2\ \boldsymbol{E\ (u_i{}^2)} + 2\ \sum_{i<J}^{n} \boldsymbol{K_i K_j\ E\ (u_i u_j)}$$

$$\therefore \quad \boldsymbol{Var\ (b_1) = \sum K_i{}^2\ E\ (u_i{}^2)}$$

$$= \boldsymbol{\sigma^2{}_u \sum K_i{}^2}$$

اذ ان : $\sum \boldsymbol{K_i}^2 = \frac{\mathbf{1}}{\sum \boldsymbol{x_i}^2}$

$$\therefore \quad Var\,(b_1) = \frac{\sigma^2{}_u}{\sum x_i{}^2} \quad \ldots\ldots\ldots \quad (22-2)$$

ولإثبات ان مقدرات (OLS) هي افضل مقدرات خطية غير متحيزة اي تتصف بخاصية ال (BLUE) حسب نظرية ماركوف، فان هذا يتطلب افتراض ان (B^*) هو مقدر اخر خطي غير متحيز حصلنا عليه بإحدى طرق الاقتصاد القياسي الاخرى، ومنها نحصل على ان تباين تلك المقدرات (B^*) يكون اكبر من تباين مقدرات (OLS) اي ان $Var\,(b_0) < Var\,(b_0{}^*)$ وان $Var\,(b_1) < Var\,(b_1^*)$ وبالتالي فان الأفضلية تكون للمقدرات صاحبة التباين الاقل وهي مقدرات (OLS) والبرهان متروك للطالب.

تقدير تباين حد الخطأ العشوائي

يستعمل تباين b_0 و b_1 في اجراء اختبارات المعنوية الخاصة بتلك المقدرات، غير ان تباين هذه المقدرات يحتوي على معلمة مجهولة هي $(\sigma^2{}_u)$ وهي تباين حد الخطأ العشوائي والتي يمكن ان نحصل على تقدير غير متحيز له وفق الصيغة الاتية :

$$s^2{}_e = \hat{\sigma}^2{}_u = \frac{\sum e_i{}^2}{n-k-1}$$

والذي يرمز له ايضا بالرمز $(s^2{}_e)$ ويمكن اشتقاقه كالاتي

بما ان المعادلة التقديرية تأخذ الصيغة الاتية:

$$\hat{Y}_i = b_0 + b_1 X_i \ldots\ldots\ldots \quad (23-2)$$

وبإدخال المجموع $(\sum)$ على طرفي المعادلة

$$\sum \hat{Y}_i = n b_0 + b_1 \sum X_i$$

وبالقسمة على n

$$\frac{\sum \hat{Y}_i}{n} = b_0 + b_1 \frac{\sum X_i}{n}$$

$$\bar{\hat{Y}} = b_0 + b_1 \bar{X} \quad \ldots\ldots\ldots \quad (24-2)$$

وبطرح المعادلة (2-24) من المعادلة (2-23) نحصل على ان

$$\hat{Y}_i - \bar{\hat{Y}} = \cancel{b_0} + b_1 X_i - \cancel{b_0} - b_1 \bar{X}$$

$$\therefore \quad \hat{y}_i = b_1 X_i - b_1 \bar{X}$$

$$\therefore \quad \hat{y}_i = b_1 (X_i - \bar{X})$$

$$\therefore \quad \hat{y}_i = b_1 x_i \quad \dots\dots\dots\dots\dots \quad (25-2)$$

$$\therefore \quad e_i = y_i - \hat{y}_i$$

وبالتعويض عن $\hat{y}_i$ في العلاقة (25−2)

$$\therefore \quad e_i = y_i - b_1 x_i$$

وبتربيع طرفي العلاقة اعلاه فان

$$e_i^2 = (y_i - b_1 x_i)^2$$

وبفك الاقواس نحصل على

$$e_i^2 = y_i^2 + b_1^2 x_i^2 - 2b_1 x_i y_i$$

وبإدخال $\sum$ على طرفي المعادلة

$$\sum e_i^2 = \sum y_i^2 + b_1^2 \sum x_i^2 - 2b_1 \sum x_i y_i \quad \dots\dots\dots\dots \quad (26-2)$$

$$\therefore \quad b_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

وان حاصل ضرب الطرفين في الوسطين سيكون

$$b_1 \sum x_i^2 = \sum x_i y_i$$

وبالتعويض في العلاقة (26−2) بما يساويها نحصل على ان

$$\sum e_i^2 = \sum y_i^2 + b_1 \sum x_i y_i - 2b_1 \sum x_i y_i$$

وبعد التبسيط نحصل على ان

$$\sum e_i^2 = \sum y_i^2 - b_1 \sum x_i y_i \quad \dots\dots\dots\dots \quad (27-2)$$

اذ ان

$$\widehat{\sigma^2}_u = \frac{\sum e_i^2}{n-k-1}$$

n : حجم العينة

k : عدد المتغيرات التوضيحية (المستقلة) وفي حالة النموذج الخطي البسيط فان $k = 1$

وبالتعويض عن البسط في العلاقة (2−27) نحصل على ان

$$\widehat{\sigma^2}_u = \frac{\sum {y_i}^2 - b_1 \sum x_i y_i}{n - k - 1}$$

واذا رمزنا الى التقدير الخطي غير المتحيز لتباين الخطأ بالرمز $s^2{}_e$ فان

$$s^2{}_e = \frac{\sum {e_i}^2}{n - k - 1}$$

وهذا يعني ان s_e^2 هو افضل مقدر غير متحيز لتباين المتغير العشوائي او حد الخطأ u_i .

وفي ضوء ما سبق نستنتج ان تباين العينة $(s^2{}_e)$ يمكن تقديره بعدة صيغ يمكن استخدامها في التطبيق العملي، ويتم ذلك من خلال استخدام القيم الاصلية لكل من المتغير المعتمد والمتغير المستقل في عملية التقدير وفق الصيغة الاتية :

$$s^2{}_e = \frac{\sum {Y_i}^2 - b_0 \sum Y_i - b_1 \sum X_i Y_i}{n - k - 1} \quad \dots\dots\dots\dots \quad (28 - 2)$$

او باستخدام الانحرافات ويتم ذلك وفق الصيغ الاتية:

$$s^2{}_e = \frac{\sum {y_i}^2 - b_1^2 \sum {x_i}^2}{n - k - 1} \quad \dots\dots\dots\dots \quad (29 - 2)$$

او

$$s^2{}_e = \frac{\sum {y_i}^2 - b_1 \sum x_i y_i}{n - k - 1} \quad \dots\dots\dots\dots \quad (30 - 2)$$

وان الصيغتان (2−29) و (2−30) اعلاه متساويتان ويمكن الوصول الى الصيغة رقم (30) بمجرد التعويض عن قيمة b_1 بما يساويها .

الفصل الثالث

اختبار الفرضيات Test of Hypothesis

بعد تقدير معالم نموذج الانحدار فانه يتحتم علينا تقييم نموذج الانحدار المقدر، وذلك من خلال اجراء اختبارات المعنوية الاقتصادية والاحصائية لنتائج تقدير النموذج.

تعرف الفرضية بانها ادعاء (جملة) يضعه الباحث يمكن ان يكون صحيحاً او غير صحيح، وتثبت صحتها فقط من خلال الاختبار **(testing)**. وقبل البدء بدراسة الكيفية التي يتم على اساسها اختبار الفرضية (اختبار المعنوية الاحصائية)، لابد من اختبار المعنوية الاقتصادية للنموذج المقدر، اي أن تقديرات المعالم يجب ان تتفق ومنطقة النظرية الاقتصادية خلف الدالة المدروسة .

فمثلا العلاقة بين الكمية المطلوبة من سلعة ما وسعر تلك السلعة، وحسب منطق النظرية الاقتصادية تكون علاقة عكسية (سالبة) بين المتغير المستقل (السعر) والكمية المطلوبة (المتغير المعتمد)، وفي حالة العلاقة بين الانفاق الاستهلاكي والدخل فان منطق النظرية يشير الى ان الميل الحدي للاستهلاك **(Mpc)** يكون موجبا ولكنه اقل من الواحد الصحيح $1 > \mathbf{Mpc} > \mathbf{0}$ وان هناك حد ادنى للاستهلاك في الامد القصير، ولكن لايمكن الجزم بصحة او عدم صحة ذلك الا بعد جمع البيانات وقياس العلاقة الاقتصادية واختبارها .

ويختبر نموذج الانحدار قبل كل شيء العلاقة بين المتغير المستقل (X) والتابع (Y) وذلك للتثبت من وجودها من خلال اختبار المعنوية الاحصائية للمعلمات المقدرة b_0 و b_1 كلا على انفراد وفي هذا المجال توجد فرضيتان هي :

1- فرضية العدم Nill Hypothesis

وتنص على عدم وجود علاقة بين المتغيرين المستقل (X) والتابع (Y) اي ان :

$$H_0 : \quad b_0 = 0$$

$$b_1 = 0$$

2- الفرضية البديلة Alternative Hypothesis

وتنص على وجود علاقة بين (X) و (Y) اي ان

$$H_1: \quad b_0 \neq 0$$
$$b_1 \neq 0$$

ويستخدم لذلك اختبار t .

اختبار t : $t - test$

لأجل اختبار ما اذا كانت $b_0 = 0$ ، $b_1 = 0$ ام لا يستخدم اختبار **(t)** عند مستوى معنوية معين ودرجة حرية **(n-k-1)** والصيغ الرياضية لهذا الاختبار هي :

أ- بالنسبة الى b_1

$$t_{b1} = \frac{b_1}{S_{b1}}$$

اذ ان :

$$S_{b1} = \sqrt{S^2{}_{b1}} = \sqrt{var\,(b_1)}$$

$$\therefore \quad var\,(b_1) = S^2{}_{b1} = \frac{\boldsymbol{\sigma^2}_u}{\sum \boldsymbol{x_i}^2} = \frac{S^2{}_e}{\sum \boldsymbol{x_i}^2}$$

$$\therefore \quad S_{b1} = \sqrt{\frac{S^2{}_e}{\sum \boldsymbol{x_i}^2}}$$

وأن :

$$S^2{}_e = \frac{\sum \boldsymbol{e_i}^2}{n-k-1} = \frac{\sum \boldsymbol{y_i}^2 - b_1 \sum x_i y_i}{n-k-1}$$

اذ ان :

t : هو اختبار (t) عند مستوى معنوية معين ودرجة حرية $(n-k-1)$

(n) : عدد المشاهدات في العينة و (k) : عدد المتغيرات المستقلة .

b_1 : القيمة التقديرية لـ β_1 الحقيقية ($b_1 = \widehat{\beta_1}$)

S_{b1} : الانحراف المعياري للمعلمة المقدرة b_1 ، $S^2{}_{b1}$: تباين b_1 .

$S^2{}_e$: تباين حد الخطأ العشوائي للمقدر.

ب- بالنسبة لـ b_0 فان :

$$t_{b_0} = \frac{b_0}{S_{b_0}}$$

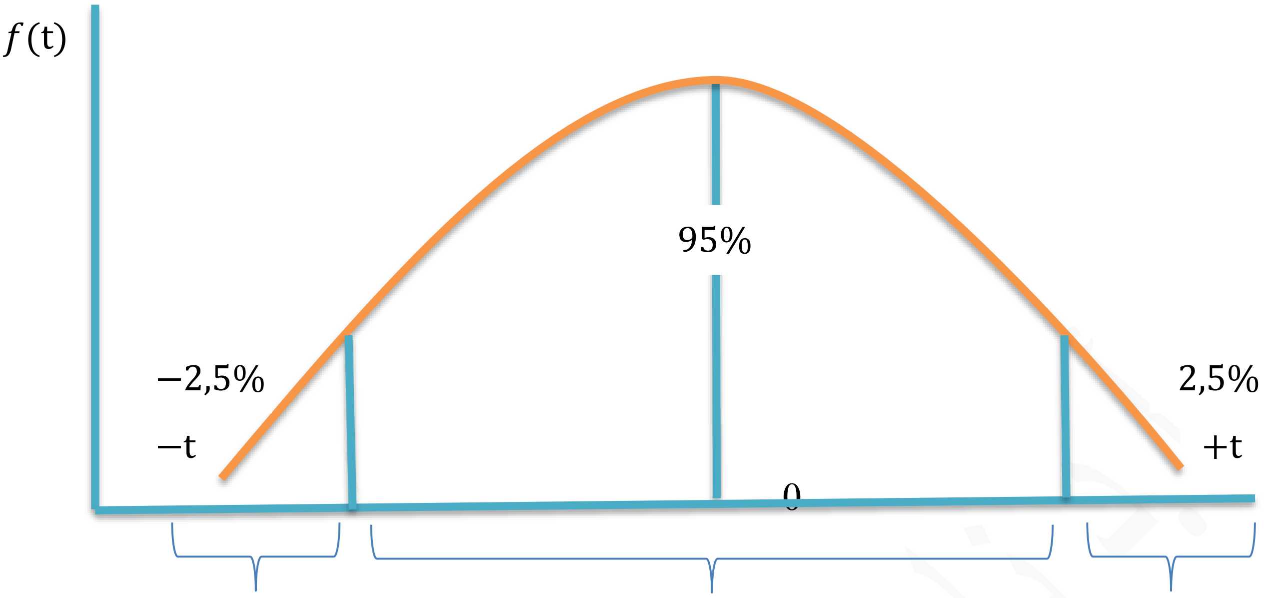
$$S_{b_0} = \sqrt{S^2{}_{b_0}} = \sqrt{var\,(b_0)}$$

$$= \sqrt{S^2{}_e\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum {x_i}^2}\right)}$$

$$S^2{}_e = \frac{\sum {e_i}^2}{n-k-1}$$

وبعد احتساب قيمة (t) تقارن مع قيمتها الجدولية المعطاة في الجداول الخاصة بها عند درجة حرية $(n-k-1)$ ومستوى المعنوية المطلوب (5% او 1%) لتحديد قبول او رفض فرضية العدم، فاذا كانت قيمة (t) المحسوبة اكبر من قيمة (t) الجدولية ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة H_1 ، بمعنى ان المعلمة المقدرة ذات معنوية احصائية، وبالعكس في حالة كون قيمة (t) المحسوبة اقل من قيمتها الجدولية حيث تقبل فرضية العدم H_0 وترفض الفرضية البديلة H_1، اي عدم معنوية المعلمة المقدرة ويمكن توضيح ذلك بالشكل ادناه :

ولاختبار معنوية المعلمات المقدرة b_1 و b_0 نعود الى بيانات المثال رقم (1) وباعتماد الصيغة الرياضية للاختبار فأننا نحتاج الى تكوين الشكل الاتي :



شكل (1) قبول او رفض الفرضية

حيث كانت نتائج التقدير كالاتي: $\hat{Y}_i = 24.48571429 + 1.486904762\, x_i$

وبالتعويض عن (X_i) بما يساويها في الجدول نحصل على قيم $(\hat{Y}_i)$ بعد ذلك نجد قيم (e_i) حيث ان $e_i = Y_i - \hat{Y}_i$ كما في الجدول الاتي :

X_i	Y_i	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$	e_i^2
4	25.6	30.43333334	−4.83333334	23.36111118
8	32.7	36.38095239	−3.68095239	13.5494105
12	45.4	42.32857143	3.07142857	9.433673461
16	53.9	48.27619048	5.62380952	31.62723352
20	59.0	54.22380953	4.77619047	22.81199541
24	62.6	60.17142858	2.42857142	5.897959142
28	65.0	66.11904763	−1.11904763	1.252267598
32	65.8	72.06666667	−6.26666667	39.27111115
$\sum = 144$	$\sum = 410$	$\sum = 410$	$\sum = 0$	$\sum = 147.204762$

اذ ان : $\hat{Y}_I = 24.48571429 + 1.486904762\ (4)$

$$\therefore\ \hat{Y}_I = 30.43333334$$

أ – لاختبار معنوية المعلمة b_1 فان

$$S^2{}_e = \frac{\sum e_i{}^2}{n-k-1} = \frac{147.204762}{8-1-1} = 24.534127$$

ويمكن استخدام الصيغ الاخرى لإيجاد $S^2{}_e$ ونحصل على نفس النتيجة .

$$S^2{}_{b_1} = \frac{S^2{}_e}{\sum x_i{}^2} = \frac{24.534127}{672} = 0.036509117$$

$$S_{b_1} = \sqrt{S^2{}_{b_1}} = \sqrt{0.036509117} = 0.191073592$$

$$t_{b1} = \frac{b_1}{S_{b1}} = \frac{1.486904762}{0.191073592} = 7.78184335$$

وبما ان قيمة (t) المحسوبة والبالغة (7.78184335) هي اكبر من قيمة (t) الجدولية عند مستوى معنوية (5%) ودرجة حرية (6) والبالغة (2.45) عليه ترفض فرضية العدم $(H_0 : b_1 = 0)$ وتقبل الفرضية البديلة $(H_1 : b_1 \neq 0)$ وهذا يعني معنوية المعلمة المقدرة b_1 من الناحية الاحصائية .

ب- لاختبار معنوية المعلمة b_0 فان

$$S^2{}_e = 24.534127$$

$$S^2{}_{b_0} = S^2{}_e\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum x_i{}^2}\right] = 24.534127\left[\frac{1}{8} + \frac{(18)^2}{672}\right] = 14.89571996$$

$$S_{b_0} = \sqrt{S^2{}_{b_0}} = \sqrt{14.89571996} = 3.859497372$$

$$t_{b_0} = \frac{b_0}{S_{b_0}} = \frac{24.48571429}{3.859497372} = 6.3442754$$

وبما ان قيمة (t) المحسوبة والبالغة (6.34427) هي اكبر من قيمة (t) الجدولية عند مستوى معنوية (5%) ودرجة حرية (6) والبالغة (2.45) عليه ترفض فرضية العدم $(H_0 : b_O = 0)$ ونقبل الفرضية البديلة $(H_1 : b_O \neq 0)$ وهذا يعني معنوية المعلمة المقدرة (b_O) .

وكحالة عامة ترفض فرضية العدم H_0 اذا كانت $S_{\hat{B}_i} < \frac{\hat{B}}{2}$

حدود الثقة لمعاملات الانحدار :

نعني بحدود الثقة (فترات الثقة) لمعاملات الانحدار تقدير المجال الذي تقع ضمنه القيمة الحقيقية للمعلمة اي معلمة المجتمع. والمقصود بحدي الثقة الحد الادنى Lower Limit والذي يرمز له بالرمز (L) والحد الاعلى Upper Limit والذي يرمز له بالرمز (U) . ونعني بذلك تحديد المجال او المدى الذي تتراوح فيه قيمة (β) بين هذين الحدين .

والصيغة الرياضية لتقدير حدود الثقة هي :

(الانحراف المعياري للمعلمة) $(t_{n-k-1}, \frac{\alpha}{2})$ ± المعلمة المقدرة = معلمة المجتمع

حيث ان : $(n - k - 1)$: تمثل درجة الحرية .

(α) : مستوى المعنوية ، $(1 - \alpha)$: تمثل معامل الثقة .

وتتراوح قيمة معامل الثقة بين 90% و 100%، كما ان مستوى المعنوية هو احتمال تكميلي لمعامل الثقة، اي ان حاصل جمع معامل الثقة ومستوى المعنوية يساوي واحد، فاذا كان معامل الثقة يساوي 95% فان مستوى المعنوية يساوي 5%.

وعلى هذا الاساس فان حدود الثقة للحد الثابت B_O هي :

$$\beta_O = b_0 \pm \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{bo}$$

ومنه فان الحد الادنى (L) يحسب كالاتي :

$$L = b_0 - \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{bo}$$

والحد الاعلى يحسب كالاتي :

$$U = b_0 + \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{bo}$$

وبنفس الاسلوب يمكن وضع حدود الثقة للميل الحدي β_1 وكالاتي :

$$\beta_1 = b_1 + \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{b1}$$

وان :

$$L = b_1 - \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{b1}$$

$$U = b_1 + \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{b1}$$

وكحالة عامة فان :

$$p_r \{L \leq B_i \leq u\} = 1 - \propto$$

مثال:- ولحساب حدود الثقة لمعاملات الانحدار b_o و b_1 لبيانات المثال رقم (1) حيث ان $\left(\left(t_{n-k-1}, \frac{\propto}{2}\right) = (t_6\,,\ 0.025) = 2.45\right)$ لذلك فان :

حدود الثقة للمعلمة β_o تحسب كما يلي :

$$\beta_O = b_0 \pm \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{bo}$$

$$\beta_O = 24.4857 \pm (2.45)(3.85949)$$

$$\beta_O = 24.4857 \pm 9.45576$$

$$\therefore L = 15.02994$$

$$\therefore U = 33.94148$$

$$\therefore Pr = \{15.02994 \leq \beta_o \leq 33.94148\} = 0.95$$

وهذا يعني ان هناك احتمال 95% ان تقع القيمة الحقيقية لمعلمة المجتمع β_o بين الحدين الاعلى (U) والادنى (L) ، وان هناك احتمال 5% ان تقع β_o خارج هذين الحدين .

حدود الثقة للمعلمة β_1 :

$$\beta_1 = b_1 + \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{b1}$$

$$\beta_1 = 1.486904 \pm (2.45)\ (0.191073)$$

$$\beta_1 = 1.486904 \pm 0.46813$$

$$\therefore L = 1.01877$$

$$\therefore U = 1.95503$$

$$\therefore Pr = \{1.01877 \leq \beta_1 \leq 1.95503\} = 0.95$$

معامل التحديد: coefficient of determination (R^2)

هو مقياس يوضح نسبة التغير في المتغير التابع (Y) التي سببها التغير في المتغير المستقل (X) .

أي نسبة الانحرافات الموضحة من قبل خط الانحدار الى الانحرافات الكلية .

ويمكن حساب هذا المعامل حسب الصيغ الاتية :

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

$$\therefore R^2 = \frac{\sum \hat{y}_i{}^2}{\sum y_i{}^2}$$

$$\therefore R^2 = \frac{b_1 \sum x_i y_i}{\sum y_i{}^2}$$

$$\therefore R^2 = r^2$$

$$\therefore R^2 = \frac{b_1{}^2 \sum x_i{}^2}{\sum y_i{}^2}$$

$$\therefore R^2 = 1 - \frac{\sum e_i{}^2}{\sum y_i{}^2}$$

ويمكن توضيح الصيغة الاخيرة من خلال الاشتقاق التالي :

الانحرافات غير الموضحة + الانحرافات الموضحة من قبل خط الانحدار = الانحرافات الكلية

أي ان: $\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$

أي ان : $\sum y_i{}^2 = \sum \hat{y}_i{}^2 + \sum e_i{}^2$

وبقسمة الطرفين على مجموع مربعات الانحرافات الكلية نحصل على ان :

$$\frac{\sum y_i{}^2}{\sum y_i{}^2} = \frac{\sum \hat{y}_i{}^2}{\sum y_i{}^2} + \frac{\sum e_i{}^2}{\sum y_i{}^2}$$

$$1 = R^2 + \frac{\sum e_i{}^2}{\sum y_i{}^2}$$

$$\therefore R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2}$$

وبما ان المقدار $1 \geq \frac{\sum e_i^2}{\sum y_i^2} \geq 0$ لذلك يمكن القول ان قيمة معامل التحديد R^2 تنحصر بين الصفر والواحد اي ان $1 \geq R^2 \geq 0$ حيث ان $R^2 = 1$ عندما تقع جميع نقاط الانتشار على خط الانحدار المقدر اي ان $Y_i = \hat{Y}_i$ وهنا تكون العلاقة تامة .

وان $R^2 = 0$ (او تقترب منه) عندما يكون خط انحدار العينة خطا افقيا اي ان $\hat{Y}_i = \bar{Y}_i$ وهذا يعني انه لا توجد علاقة بين المتغير التابع والمتغير المستقل .

ولمعرفة مقدار ما يفسره المتغير المستقل (X) من التغير في المتغير التابع (Y) في المثال السابق (1)، وبالاعتماد على جميع الصيغ الخاصة بحساب R^2 فان :

$\hat{y}_i = \hat{Y}_i - \bar{Y}$	$\hat{y}_i^2$	$y_i^2 = (Y_i - \bar{Y})^2$
−20.81666666	433.3336108	657.9225
−14.86904761	221.0885768	344.1025
−8.92142857	79.59188773	34.2225
−2.97380952	8.843543061	7.0225
2.973809953	8.843543121	60.0625
8.92142858	79.59188791	128.8225
14.86904763	221.0885774	189.0625
20.81666667	433.3336113	211.7025
$\Sigma\hat{y}_i = 0$	**1485.715238**	**1632.92**

$$\therefore \boldsymbol{R^2} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{\mathbf{1485.715238}}{\mathbf{1632.92}} = \mathbf{0.90985} \;\; = \mathbf{90.98}\ \%$$

وهذا يعني ان المتغير المستقل (X_i) (عدد سنوات الخدمة) تفسر (توضح) حوالي 90.98% من التغير الحاصل في المتغير التابع (Y_i) (معدل الاجر السنوي)، وان النسبة الباقية وهي 9.02% تمثل تأثير المتغيرات الاخرى غير المضمنة في المعادلة (العلاقة) . كما ان :

$$\therefore R^2 = \frac{b_1 \sum x_i y_i}{\sum y_i^2} = \frac{(1.486904)(999.2)}{1632.92} = 90.98\ \%$$

$$\therefore R^2 = \frac{b_1^{\,2} \sum x_i^2}{\sum y_i^2} = \frac{(1.486904)^2\,(672)}{1632.92} = 90.98\ \%$$

$$\therefore R^2 = 1 - \frac{\sum e_i{}^2}{\sum y_i{}^2} = 1 - \frac{147.204762}{1632.92}$$

$$\therefore R^2 = 1 - 0.090148177 = 90.98\ \%$$

اختبار F : F-Test

لاختبار مدى معنوية العلاقة المفترضة بين المتغير المعتمد والمتغير المستقل يستخدم اختبار **F** ويعتمد هو الاخر على نوعين من الفرضيات هما :

1- فرضية العدم : وتنص على عدم معنوية العلاقة بين المتغير التابع والمتغير المستقل : اي ان

$$H_O : \beta_1 = 0$$

2-الفرضية البديلة : وتنص على وجود علاقة جوهرية (معنوية) من الناحية الاحصائية بين المتغير المعتمد والمتغير المستقل : اي ان

$$H_O : \beta_1 \neq 0$$

والصيغة الرياضية لهذا الاختبار هي :

$$F = \frac{\sum \hat{y}_i{}^2/k}{\sum e_i{}^2/n - k - 1}$$

اي أن اختبار F هو عبارة عن نسبة الانحرافات الموضحة من قبل خط الانحدار مقسومة على عدد المتغيرات المستقلة (K) الى الانحرافات غير الموضحة مقسومة على درجات الحرية التي تساوي عدد المشاهدات (n) مطروحا منها (K) ناقصاً واحد.

وبعد احتساب قيمة (F) تقارن مع قيمة (F) الجدولية المعطاة في الجداول الخاصة بها عند مستوى المعنوية المطلوب (5% او 1%) ودرجة حرية (k , n-k-1) للبسط والمقام لتحديد قبول او رفض فرضية العدم فاذا كانت قيمة (F) المحسوبة اكبر من قيمة (F) الجدولية ترفض فرضية العدم ونقبل الفرضية البديلة (H_1) اي معنوية العلاقة المقدرة وبالعكس في حالة كون (F) المحتسبة اقل من الجدولية حيث تقبل فرضية العدم (H_O) اي عدم معنوية معادلة خط الانحدار المقدرة .

ويمكن احتساب قيمة (F) بالاعتماد على الصيغ الاخرى الآتية :

$$F = \frac{b_1 \ \sum x_i y_i \ / \ k}{\sum e_i{}^2 \ / \ (n-k-1)}$$

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

$$F = \frac{b_1^2 \ \sum x_i^2 \ / \ k}{\sum e_i{}^2 \ / \ (\,n-k-1)}$$

$$F = \ t_{b_1}^2$$

العلاقة بين اختبار F واختبار t

يمكن توضيح العلاقة بين احصائتي F و t في نموذج الانحدار الخطي البسيط (متغيرين) كالاتي :

$$F = \frac{b_1^2 \ \sum x_i^2 \ / \ k}{\sum e_i{}^2 \ / \ \ n-k-1}$$

وحيث ان k : عدد المتغيرات المستقلة = 1

$$F = \frac{b_1^2 \ \sum x_i^2 \ / \ k}{\sum e_i{}^2 \ / \ n-2} = \frac{b_1^2 \ \sum x_i^2}{\sigma_u{}^2}$$

$$F = \frac{b_1^2}{\sigma_u{}^2 / \ \sum x_i^2}$$

$$\therefore \ \ var\,(b_1\,) = \ S^2{}_{b_1} \ = \frac{\sigma_u{}^2}{\sum x_i^2}$$

$$\therefore \ \ F = \frac{b_1^2}{S^2{}_{b_1}} = \left[\frac{b_1}{S_{b_1}}\right]^2 = \ t^2{}_{b_1}$$

$$\therefore \ \ \mathbf{F = \ t_{b_1}^2}$$

ولاختبار معنوية (جوهرية) العلاقة بين المتغير التابع (Y_i) (معدل الاجر السنوي) والمتغير المستقل (X_i) (عدد سنوات الخدمة) لبيانات المثال رقم (1) وبالاعتماد على الصيغ الرياضية الخاصة باختبار F نحصل على ان :

$$F=\frac{\sum \hat{y}_i{}^2/k}{\sum e_i{}^2/n-k-1}=\frac{1485.715238/1}{147.204762/8-1-1}=\frac{1485.715238}{24.534127}=60.5570$$

وبما ان قيمة F المحسوبة والبالغة (60.557) اكبر من قيمة (F) الجدولية عند مستوى معنوية (5%) ودرجة حرية (6,1) للبسط والمقام والبالغة (5.99) عليه ترفض فرضية العدم (H_O) ونقبل الفرضية البديلة (H_1) اي معنوية (جوهرية) العلاقة المقدرة .

$$\mathrm{F}=\frac{\mathrm{b}_1\ \sum x_i \mathrm{y}_i\ /\ \mathrm{k}}{\sum \mathrm{e}_\mathrm{i}{}^2\ /\ \mathrm{n}-\mathrm{k}-1}=\frac{(1.486904762)(999.2)}{24.534127}=60.5570$$

$$\mathrm{F}=\frac{\mathrm{R}^2/\mathrm{k}}{(1-\mathrm{R}^2)/\ (\mathrm{n}-\mathrm{k}-1)}=\frac{0.909851822/1}{(1-0.909851822)/8-1-1}$$
$$=\frac{0.909851822}{0.015024696}=60.5570$$

$$\mathrm{F}=\frac{\mathrm{b}_1^2\ \sum x_i^2\ /\ \mathrm{k}}{\sum \mathrm{e}_\mathrm{i}{}^2\ /\ (\ \mathrm{n}-\mathrm{k}-1)}=\frac{(1.486904762)^2\ (672)\ /1}{24.534127}=60.557$$

$$\mathrm{F}=\ \mathrm{t}_{\mathrm{b}_1}^2\ =(7.781843354)^2=60.557$$

معامل الارتباط البسيط :- Simple correlation coefficient (r)

يقصد بالارتباط وجود علاقة بين ظاهرتين (متغيرين) او اكثر ، ويسمى المقياس الذي تقاس به درجة الارتباط بمعامل الارتباط الذي يرمز له بالرمز (r) .

ويعرف معامل الارتباط البسيط بانه القيمة العددية للعلاقة بين متغيرين وتكون خالية من وحدات القياس وقيمته تكون محصورة بين (1+) و (1−) اي ان $(-1\le r\le +1)$ حيث تمثل الاشارة (موجبة او سالبة) نوعية العلاقة (طردية او عكسية)، اما القيمة العددية فتمثل قوة العلاقة بين المتغيرين او الظاهرتين.

ويحسب معامل الارتباط حسب الصيغ الاتية:

$$r=\left[\sqrt{\frac{\sum x_i{}^2}{\sum y_i{}^2}}\right]b_1$$

$$r = \frac{\sum x_i y_i}{\sqrt{(\sum x_i{}^2)(\sum y_i{}^2)}}$$

$$r = \pm\sqrt{R^2} \quad \rightarrow \quad r = \left(b_1 \text{ اشارة}\right)\sqrt{R^2}$$

$$r = \frac{cov\ (X_i, Y_i)}{S_x S_y} = \frac{\sum x_i y_i}{n S_x S_y}$$

اذ ان :

$$S_x = \sqrt{\frac{\sum x_i{}^2}{n}} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

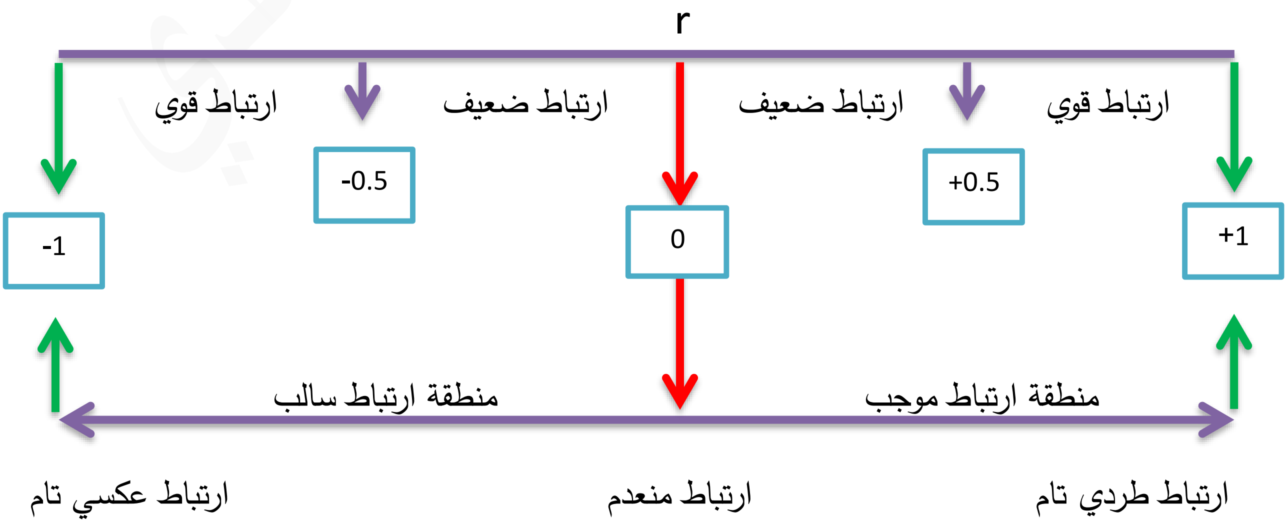
$$S_y = \sqrt{\frac{\sum y_i{}^2}{n}} = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n}}$$

وفي حالة استخدام القيم الاصلية (الحقيقية) للمشاهدات فان معامل الارتباط يحسب وفق الصيغ الاتية :

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

$$r = \frac{\sum X_i Y_i - n\, \bar{X}\, \bar{Y}}{\sqrt{(\sum X_i^2 - n\, \bar{X}^2)}\sqrt{[\sum Y_i^2 - n\, \bar{Y}^2]}}$$

والشكل ادناه توضيح لقيم معامل الارتباط الممكنة وتفسرها .



وبالرجوع الى بيانات المثال السابق رقم (1) وباعتماد الصيغ الرياضية اعلاه يمكن حساب معامل الارتباط (r) كالاتي:

$$r = \left[\sqrt{\frac{\sum x_i^2}{\sum y_i^2}}\right] b_1 = \left[\sqrt{\frac{672}{1632.92}}\right](1.4869) = 0.95386$$

اي ان العلاقة بين المتغيرين (Y, X) (عدد سنوات الخدمة ومعدل الاجر السنوي) طردية لان الاشارة موجبة وقوية (لان القيمة العددية اكبر من 0.5) . كما ان:

$$r = \frac{\sum x_i y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}} = \frac{999.2}{\sqrt{(672)(1632.92)}} = \frac{999.2}{1047.5314978} = 0.9538$$

وان :

$$r = \pm\sqrt{R^2} = \sqrt{0.9098518} = 0.9538$$

لان إشارة المعلمة (b_1) موجبة .

وان :

$$r = \frac{\sum x_i y_i}{n S_x S_y}$$

$$S_x = \sqrt{\frac{\sum x_i^2}{n}} = \sqrt{\frac{672}{8}} = 9.16515139$$

$$S_y = \sqrt{\frac{\sum y_i^2}{n}} = \sqrt{\frac{1632.92}{8}} = 14.2868821$$

$$r = \frac{999.2}{(8)(9.16515139)(14.2868821)}$$

$$= \frac{999.2}{1047.531498} = 0.9538$$

جدول تحليل التباين :- Analysis of Variance Table (ANOVA TABLE)

يهدف جدول تحليل التباين الى توضيح تأثير المتغير المستقل في المتغير التابع (معرفة مدى معنوية العلاقة المفترضة بين المتغير المعتمد والمتغير المستقل) وتزداد اهمية هذا الجدول عند دراسة الانحدار المتعدد حيث يستفاد منه في معرفة تأثير كل متغير من المتغيرات المستقلة في المتغير التابع وبالتالي تحديد المتغيرات المؤثرة في النموذج . ويمكن بناء جدول تحليل التباين كالاتي : $\sum e_i^2/n-k-1$

مصدر التباين Source of variance S.O.V	مجموع المربعات Sum of square S.S.	درجات الحرية Degrees of Freedom d.f	متوسط المربعات Mean sum square M.S.S	F F-test
الانحرافات الموضحة من قبل خط الانحدار	$\sum\hat{y}_i^2 = b_1\sum x_i y_i$ $= R^2\sum y_i^2$ S.S.R	K	$\sum\hat{y}_i^2$ / K OR MSSR $=\frac{b_1 \sum x_i y_i}{K}$	$F=\frac{\sum\hat{y}_i^2/k}{\sum e_i^2/n-k-1}$
الانحرافات غير الموضحة (المتبقية)	$\sum e_i^2 = (1-R^2)\sum y_i^2$ $=\sum y_i^2 - b_1\sum x_i y_i$ S.S.E	$n-k-1$	$\sum e_i^2/n-k-1$ Or MssE $=\frac{\sum y_i^2 - b_1\sum x_i y_i}{n-k-1}$	
مجموع الانحرافات (الانحرافات الكلية)	$\sum y_i^2$ S.S.T	$n-1$		

ويمكن بناء جدول تحليل التباين للمثال رقم (1) السابق الذي يمثل العلاقة بين معدل الاجر السنوي وعدد سنوات الخدمة كالاتي:

مصدر التباين S.O.V	مجموع المربعات Sum of square	درجات الحرية d.f	متوسط المربعات M.S.S	F-test
الانحرافات الموضحة من قبل خط الانحدار X_i	1485.715238	1	$\frac{1485.715238}{1}$	$F = \frac{1485.715238/1}{147.204762/6}$
الانحرافات غير الموضحة (المتبقية)	147.204762	$8-1-1 = 6$	$\frac{147.204762}{6}$	$= 60.55708$
مجموع الانحرافات (الانحرافات الكلية)	1632.92	$8-1=7$		

وبمقارنة قيمة F المحسوبة والبالغة (60.557) مع قيمتها الجدولية لدرجة الحرية (6,1) ولمستوى معنوية (5%) والبالغة (5.99) نجد ان قيمة F المحسوبة هي اكبر من قيمة F الجدولية ومنه نستنتج ان العلاقة الخطية المفترضة بين معدل الاجر السنوي وعدد سنوات الخدمة معنوية اي ان المتغير (X_i) يمارس تأثيره في المتغير المعتمد .

مفهوم التنبؤ: Prediction or Forecasting

أحد الاهداف الرئيسة للاقتصاد القياسي هو استخدام النموذج المقدر للتنبؤ بقيمة المتغير التابع استنادا الى قيم المتغيرات المستقلة، من اجل التعرف على مسار الظاهرة موضوع البحث في المستقبل. حيث يعرف التنبؤ بانه تحليل بيانات الماضي وتطبيق نتائجها على المستقبل من خلال استخدام نموذج رياضي (اي تقدير لما سوف تكون عليه الحالة في المستقبل بناءً على توفر بيانات سابقة).

اي ان $(\widehat{Y}_i)$ يمكن ان تستخدم للتنبؤ بقيمة (Y_i) الجديدة ولتكن (Y_{t+1}) بالاعتماد على قيمة (X_i) الجديدة ولتكن (X_{t+1}) لذلك فان المعادلة التنبؤية للفترة $(t+1)$ ستأخذ الشكل الاتي :

$$Y_{t+1} = b_o + b_1 X_{t+1}$$

توجد العديد من التعاريف للتنبؤ فيما يلي بعض منها:

- **يشير التنبؤ** الى تقدير قيمة المتغير التابع بمعلومية القيم الفعلية او المتوقعة للمتغير المستقل.
- **التنبؤ العلمي:** هو تقدير كمي للقيم المتوقعة للمتغيرات التابعة في المستقبل القريب بناءاً على ما هو متاح لدينا من معلومات عن الماضي والحاضر.

أنواع التنبؤ العلمي: انطلاقا من عدة معايير يمكن ان نميز بين أنواع كثيرة للتنبؤ نذكر منها ما يلي:

1. **صيغة التنبؤ**: ويتم التفريق هنا بين تنبؤ النقطة وتنبؤ الفترة.

- **فتنبؤ النقطة**: هو التنبؤ بقيمة واحدة للمتغير التابع في كل فترة مقبلة. او التنبؤ بالقيمة المتوقعة الشرطية لـ (y) عن مستوى معين لـ (X) وليكن (X_0) وهي نقطة تقع على خط انحدار المجتمع نفسه. ويسمى هذا النوع من التنبؤ بالتنبؤ المتوسط (mean prediction) مثلا نرغب بالتنبؤ بـ $E(y/x_o=100)$.

- **تنبؤ الفترة** (Inter prediction): هو التنبؤ بمدى او مجال معين تقع ضمنه قيمة المتغير التابع (y) وباحتمال معين. بمعنى إعطاء قيمة دنيا وقيمة عليا للقيمة المتنبأ بها؟

2. **فترة التنبؤ:** بناءا على هذا المعيار يمكن ان نميز بين نوعين من التنبؤ، تنبؤ قبل التحقق وتنبؤ بعد التحقق، وكلا النوعين يتنبأن بالقيم التي تلي فترة تقدير النموذج الا ان التنبؤ بعد التحقق يتوقع قيماً للمتغير التابع في فترة متاح عنها بيانات فعلية عن المتغير التابع (y) وهذا يتيح فرصة للتأكد من مدى صحة التوقعات من خلال مقارنتها بالبيانات الفعلية المتاحة. اما فيما يتعلق بالتنبؤ قبل التحقق فهو يحدد قيم للمتغير التابع في فترات مستقبلية لا تتاح عنها بيانات خاصة بالمتغير التابع.

3. **درجة التأكد**: وفقا لهذا هناك نوعين من التنبؤ هما:

- **التنبؤ المشروط**: نحصل على هذا النوع من التنبؤ عندما تكون قيم المتغيرات التفسيرية (المستقلة) التي سيتم على أساسها توقع قيم المتغير التابع لا تكون معروفة على وجه التأكيد وانما يتعين توقعها او تخمينها. ومن ثم فأن دقة التنبؤ بقيمة المتغير التابع تكون مشروطة بمدى دقة القيم المفترضة للمتغير التفسيري (المستقل).
- **التنبؤ غير المشروط**: هو التنبؤ بقيم المتغير التابع بناءاً على معلومات مؤكدة متاحة عن المتغيرات المستقلة (التفسيرية). ومن ثم فأن كل أنواع التنبؤ بعد التحقق تعتبر تنبؤ غير مشروط.

4. **درجة الشمول**: يقصد بها عدد المعادلات المكونة للنموذج وعلى هذا الأساس قد يتم التنبؤ باستخدام نموذج مكون من معادلة واحدة او استخدام نموذج مكون من عدة معادلات.

5. **أسلوب التنبؤ:** يوجد اسلوبين للتنبؤ هما:

- **التنبؤ القياســي**: باســتخدام نماذج الاقتصــاد القياســي (النماذج الســببية) الذي يعتمد على نماذج الانحدار التي تربط بين متغيرات تابعة وأخرى مســـتقلة ويســـتخدم للتنبؤ في الآجال الطويلة. وتتطلب هذه النماذج ما يلي:

- تحديد النظرية الاقتصادية الخاصة بموضوع البحث.
- صياغة النموذج.
- جمع البيانات الخاصة بمتغيرات النموذج.
- تقدير النموذج.
- اختبار ملائمة النموذج.
- استخدام النموذج في التنبؤ.

ـ **التنبؤ باستخدام السلاسل الزمنية**: يعتمد على القيم الماضـية لمتغير ما للتنبؤ بقيمته المسـتقبلية، ويستخدم للتنبؤ للآجال القصيرة.

6. **دقة التنبؤ**: توجد عدة معايير لقياس مقدرة النموذج على التنبؤ نذكر منها:

- **اختبار معنوية الفروق:** يسـتخدم هذا المعيار في حالة التنبؤ بعد التأكد، من خلال مقارنة القيمة المقدرة للمتغير مع القيمة الفعلية له، ويستخدم لاختبار الفرضية التالية:

$$H_0: y_F = ya$$

$$H_1: y_F \neq ya$$

حيث ان:

y_F: القيمة المقدرة (المتنبأ بها) للمتغير التابع (**y**)

ya: القيمة الفعلية (الحقيقية) للمتغير التابع (**y**)

ويستخدم لذلك اختبار (t) وفق الصيغة التالية:

$$t = \frac{ya - yF}{\sigma_y}$$

حيث ان σy: الانحراف المعياري لقيم (**y**)

- **معامل عدم التساوي لثايل** : **Thil's Inequality coefficient**

ويحسب وفق الصيغة التالية:

$$T=\sqrt{\frac{\sum_{i=1}^{n}(\mathbf{dfi}-\mathbf{dai})^2}{\sum_{i=1}^{n}\mathbf{dai}^2}}$$

حيث ان:

T: معامل ثايل.

$\mathbf{d_{fi}}$: التغير في القيم المتوقعة (المتنبأ بها) للمتغير التابع.

$\mathbf{d_{ai}}$: التغير في القيم الفعلية للمتغير التابع **(y)**.

إذا كانت قيمة معامل ثايل تساوي الصفر او تقترب منه دلّ ذلك على مقدرة النموذج الكبيرة على التنبؤ، اما إذا كانت (**T**) مساوية للواحد الصحيح فهذا يعني ان المتغير التابع سوف يكون ثابتا عبر الزمن، اما إذا كانت قيمة (**T**) أكبر من الواحد الصحيح دلّ ذلك على انخفاض مقدرة النموذج على التنبؤ.

- **اختبار جاو للتنبؤ** : **Chow Forecast test**

يستخدم (اختبار جاو) لاختبار مقدرة النموذج على التنبؤ لمجموعة فرعية من المشاهدات، وهو يعرض نتائج اختبار (Chow Breaking Test) وغالباً ما تكون نتيجة الاختبارين متساوية او متقاربة وأحيانا تختلف. وتعطي نتائج الاختبار قيمة لإحصائية (F) ونسبة الاحتمال المقرونة بدلالتها الإحصائية عند استخدام مستوى دلالة (0.05)، فاذا كان مستوى الدلالة الإحصائية أكبر من (0.05) دلّ ذلك على مقدرة النموذج على التنبؤ.

مثال:- بالرجوع الى بيانات المثال رقم (1) السابق وللتنبؤ بقيمة Y_{t+1} عندما $(X_{t+1}=36)$ فان

$$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}} = 24.4857 + 1.4869\, X_i$$

ومنها فان :

$$Y_{t+1} = 24.4857 + (1.4869)(36)$$
$$= 78.01428$$

اي ان معدل الاجر السنوي لموظف له من الخدمة (36) سنة يساوي 78.01428 ألف دينار.

ملاحظة:- ولكون قيمة (X_{t+1}) واقعة خارج نطاق العينة اي خارج مدى قيم X_i في العينة لذلك يطلق على التنبؤ بقيمة Y_{t+1} بالاستقطاب الخارجي (Extrapolation) اما اذا كانت قيمة (X_{t+1}) تقع داخل نطاق العينة (ضمن مدى قيم X_i في العينة) فيدعى التنبؤ بالاستقطاب الداخلي (Interpolation) .

ولتقدير حدود الثقة للمعلمة Y_{t+1} فان

$$Y_{t+1} = \hat{Y}_{t+1} \pm \left(t\frac{\propto}{2}\right)\left(S_{\hat{Y}_{t+1}}\right)$$

حيث ان

$$S^2{}_{\hat{Y}_{t+1}} = S^2{}_e\left[1+\frac{1}{n}+\frac{(X_{t+1}-\bar{X})^2}{\sum x_i{}^2}\right]$$

ومن بيانات المثال السابق فان

$$S^2{}_{\hat{Y}_{t+1}} = 24.534127\left[1+\frac{1}{8}+\frac{(36-18)^2}{672}\right]$$

$$= 24.534127\ [1+0.125+0.482142857]$$

$$= 39.42984696$$

$$\therefore \quad S^2{}_{\hat{Y}_{t+1}} = \sqrt{39.42984696} = 6.2793189$$

$$\therefore \quad Y_{t+1} = \hat{Y}_{t+1} \mp \left(t(n-k-1),\frac{\propto}{2}\right)\left(S_{\hat{Y}_{t+1}}\right)$$

$$\therefore \quad Y_{t+1} = 78.0142857 \mp (2.45)(6.2793189)$$

$$\therefore \quad Y_{t+1} = 78.0142857 \mp 15.3843315$$

ولذلك فان :

$$62.6299542 \leq Y_{t+1} \leq 93.3986172$$

أي ان :

$$p_r\{62.6299542 \leq Y_{t+1} \leq 93.3986172\} = 0.95$$

تحليل دوال الطلب :

تعرف العلاقة التي تربط بين كمية طلب المستهلك من سلعة معينة والعوامل المؤثرة على ذلك الطلب بدالة الطلب ويمكن التعبير عن تلك العلاقة بالشكل التالي:

$$Y_{iJ} = f\left(p_1, p_2, \dots\dots\dots, p_n, X_{iJ}\right) \dots\dots\dots 1$$

حيث ان Y_{iJ} : طلب المستهلك (J) على السلعة (i) .

p_i : سعر السلعة (i)

X_J : دخل المستهلك (J)

وتستخدم مرونة الطلب الدخلية لقياس مدى استجابة الطلب للتغير الحاصل في الدخل، وتعرف هذه المرونة بانها النسبة المئوية للتغير في كمية الطلب على السلعة عند تغير الدخل بنسبة (1%) ويعبر عنها كالاتي :

$$\eta_i = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} = \frac{\Delta Y}{\Delta X} * \frac{X}{Y} \quad \dots\dots\dots \quad 2$$

حيث ان :

η_i : مرونة الطلب الدخلية للسلعة (i)

Δ : ترمز الى مقدار التغير اي ان

$$\Delta Y = Y_2 - Y_1 \quad , \quad \Delta X = X_2 - X_1$$

وفي ضوء المرونة الدخلية يمكن تصنيف السلع الى ثلاثة انواع كالاتي :

كمالية اذا كانت $\eta_i > 1$

ضرورية اذا كانت $0 < \eta_i < 1$

متدنية اذا كانت $\eta_i < 0$

يتضح من التصنيف اعلاه بان اسلع الكمالية تتصف بان الطلب عليها يرتفع بنسبة أكبر من نسبة ارتفاع الدخل. في حين ان الطلب على السلع الضرورية يرتفع بنسبة اقل من نسبة ارتفاع الدخل وذلك نتيجة لتوجه المستهلك نجو السلع الكمالية، اما بالنسبة للسلع المتدنية فان الطلب عليها ينخفض بارتفاع الدخل وذلك لتحول المستهلك الى سلع اخرى اكثر جودة من السلع المتدنية.

وتجدر الاشارة هنا الى ان مرونة الطلب الدخلية المعطاة في الصيغة (2) تعبر عن التغير في الطلب عند تغير الدخل في نقطة معينة، ويمكن الحصول على نفس المؤشر بتفاضل دالة الطلب في (1) وكالاتي :

$$\eta_i = \frac{dY_{iJ}}{dX_J} * \frac{X_J}{Y_{iJ}}$$

علما ان :

$$\frac{dY_{iJ}}{dX_J} = \frac{\Delta Y}{\Delta X} = MPC$$

حيث ان :

η_i : مرونة الطلب الدخلية .

MPC : الميل الحدي للاستهلاك . Marginal propensity to consume

ويعرف بانه نسبة التغير الحاصل في الطلب الى التغير الحاصل في الدخل. عليه يمكن القول بان مرونة الطلب الدخلية تساوي الميل الحدي للاستهلاك مضروبا في نسبة الدخل الى الطلب. وبما ان الميل الحدي للاستهلاك يتغير بتغير الطلب وهذا بدوره يؤدي الى تغير المرونة الدخلية عند كل مستوى من مستويات الدخل، لذا يفضل اعادة كتابة الصيغة (3) بشكلها العام لتأخذ بنظر الاعتبار كافة التغيرات التي يمكن ان تحدث في مختلف المستويات وكالاتي :

$$\eta_i = \frac{dy_{iJ}}{dx_J} * \frac{\bar{X}}{\bar{Y}} \quad \dots\dots\dots \quad 4$$

اضافة الى مؤشر **المرونة الدخلية** اعلاه، هناك مؤشر اخر يمكن الحصول عليه من خلال دالة الطلب رقم (1) **ويعرف بمرونة الطلب السعرية** والتي تقيس مدى استجابة الطلب للتغير في السعر، علما انه اذا كان السعر هو سعر السلعة ذاتها عندئذ تعرف تلك المرونة بمرونة الطلب السعرية المباشرة او الذاتية

والتي تعني النسبة المئوية للتغير في كمية الطلب على السلعة (own or Direct price Elasticity) عند تغير سعر السلعة ذاتها بنسبة (1%) ويعبر عنه رياضيا كالاتي :

$$\eta_{ii} = \frac{dY_{iJ}}{dP_i} * \frac{P_i}{Y_{iJ}} \quad \dots\dots\dots \quad 5$$

حيث ان η_{ii} : مرونة الطلب السعرية المباشرة .

وتقاس مدى استجابة كمية الطلب على السلعة للتغير في سعر سلعة اخرى بمرونة الطلب السعرية الغير مباشرة او التبادلية (indirect or cross price elasticity) . وتحسب بموجب الصيغة الاتية :

$$\eta_{ir} = \frac{dY_{iJ}}{dP_r} * \frac{P_r}{Y_{iJ}} \quad \dots\dots\dots \quad 6$$

حيث ان η_{ii} : مرونة الطلب السعرية غير المباشرة .

وتجدر الاشارة هنا الى ان المرونة الدخلية والمرونات السعرية بشقيها التبادلية والذاتية ماهي الا عبارة عن مؤشرات خالية من الوحدات القياسية وعليه يمكن استخدامها لأغراض المقارنة حتى وان كانت تخص سلع ذات وحدات قياسية مختلفة او بلدان ذات عملات محلية متباينة .

ويعتمد تحليل دالة الطلب في (1) اعالاه على ثلاث انواع رئيسية من البيانات، النوع الاول يمثل بيانات مقطعية (Cross-Section Data) مصدرها الاساس هو بحث ميزانية الاسرة. اما النوع الثاني فيمثل بيانات السوق التي تكون عادة بشكل سلاسل زمنية (Time-Series Data) . اما النوع الثالث فيجمع ما بين البيانات المقطعية وبيانات السلاسل الزمنية (pooling Data) ويمكن استخدام النوع الثاني والثالث من البيانات لتحليل تغيرات الاسعار بشكل مباشر في حين يساعد النوع الاول من البيانات وفي ظل افتراضات معينة على تقدير المرونات السعرية ومرونات الاحلال والادخار. وبما ان بحوث ميزانية الاسرة تجري في فترة قصيرة نسبياً فانه يفترض عدم حصول تغيرات في اسعار السلع خلال فترة البحث ويترتب على مثل هذا الافتراض تحويل دالة الطلب في العلاقة (1) الى الصيغ التالية والتي تعرف بدالة انجل (Engle function).

$$Y_{iJ} = f\,(X_J) \quad \dots\dots\dots \quad 7$$

هذه العلاقة تمثل نقطة البداية في تحليل طلب المستهلك حيث انها تعبر عن الطلب كدالة في الدخل في ظل افتراض ثبات الاسعار. وقد اطلقت عليها التسمية اعلاه نسبة الى العالم الالماني ارنست انجل (Ernst Engel) وتأخذ العلاقة رقم (7) صيغا مختلفة قد تكون خطية او غير خطية وذلك حسب سلوك المستهلك تجاه السلعة عند تغير دخله، الامر الذي يستوجب تحديد الصيغة الملائمة للسلعة اولا ومن ثم تحليل الطلب على تلك السلعة ثانيا .

ومن اهم صيغ دوال انجل هو ما يعرف بالصيغة الخطية Linear Form التالية :

$$Y_{iJ} = \beta_0 + \beta_1 X_J + U_{ij} \quad \dots\dots\dots \quad 8$$

بموجب هذه الصيغة الخطية يكون الميل الحدي للاستهلاك ثابتا حيث ان المشتقة الاولى لها تساوي :

$$\frac{dY_{iJ}}{dX_J} = \beta_1 = MPC$$

وبالتعويض في الصيغة رقم (4) نجد ان مرونة الطلب الدخلية بموجب هذه الصيغة الخطية تساوي :

$$\eta_i = \frac{dY_{iJ}}{dP_J} * \frac{\bar{X}}{\bar{Y}} = \beta_1 \frac{\bar{X}}{\bar{Y}}$$

ومن صيغ دوال انجل غير الخطية هي الصيغة التالية :

$$Y_{iJ} = \beta_0 X_J^{B_1} + e^{u_{ij}} \quad \dots\dots\dots \quad 9$$

والتي يمكن تحويلها الى صيغة خطية وذلك بأخذ اللوغاريتم الى طرفيها وعندها تعرف بالصيغة اللوغاريتمية المزدوجة Double – logarithmic form كالاتي :

$$ln\, Y_{iJ} = ln\, \beta_0 + \beta_1\, ln\, X_J + U_{ij} \quad \dots\dots\dots \quad 10$$

حيث ان (ln) تعني اللوغاريتم الطبيعي للأساس (e) .

ويحسب الميل الحدي للاستهلاك من الصيغة اعلاه كالاتي :

$$mpc = \frac{dY_{iJ}}{dX_J} = \beta_1 \frac{\bar{X}}{\bar{Y}}$$

وبالتعويض مرة اخرى في الصيغة رقم (4) نحصل على مرونة الطلب الدخلية وكالاتي :

$$\eta_i = \frac{dY_{iJ}}{dX_J} * \frac{\bar{X}}{\bar{Y}}$$

$$= \beta_1 \frac{\bar{Y}}{\bar{X}} * \frac{\bar{X}}{\bar{Y}} = \beta_1$$

وتجدر الاشارة هنا الى ان الدالة رقم (7) اعلاه يمكن ان تأخذ عدة صيغ رياضية اضافة الى الصيغتين اعلاه، وفي ما يلي بعض هذه الصيغ مع بيان الميل الحدي للاستهلاك (mpc) والمرونة الدخليه (η_i) مصنفة حسب نوع الدالة .

نوع الدالة	الصيغة	mpc	η_i
1- الخطية	$Y_{iJ} = \beta_0 + \beta_1 X_J + U_{ij}$	β_1	$\beta_1 \frac{\bar{X}}{\bar{Y}}$
2- نصف اللوغاريتمية	$Y_{iJ} = \beta_0 + \beta_1 ln\, X_J + U_{ij}$	$\frac{\beta_1}{\bar{X}}$	$\frac{\beta_1}{\bar{Y}}$
3-اللوغاريمية المزدوجة	$lnY_{iJ} = \beta_0 + \beta_1 ln\, X_J + U_{ij}$	$\beta_1 \frac{\bar{Y}}{\bar{X}}$	β_1
4-الدالة العكسية	$Y = \beta_0 - \frac{\beta_1}{X_J} + U_{ij}$	$\frac{\beta_1}{\bar{X}^2}$	$\frac{\beta_1}{\bar{X}\bar{Y}}$
5-اللوغاريتمية العكسية	$lnY_{iJ} = \beta_0 - \frac{\beta_1}{X_J} + U_{ij}$	$\beta_1 \frac{\bar{Y}}{\bar{X}^2}$	$\frac{\beta_1}{\bar{X}}$
6-النسبية نصف اللوغاريتمية	$\frac{Y_{iJ}}{X_J} = \beta_0 + \beta_1 ln\, X_J + U_{ij}$	$\beta_1 \frac{\bar{Y}}{\bar{X}}$	$1 + \beta_1 \frac{\bar{X}}{\bar{Y}}$
7-الجذرية	$Y_{iJ} = \beta_0 + \beta_1 \sqrt{X_J} + U_{ij}$	$\frac{\beta_1}{2\sqrt{\bar{X}}}$	$\frac{\beta_1 \sqrt{\bar{X}}}{2\,\bar{Y}}$

تعد المرونة الدخلية والحاصل عليها من الصيغ المختلفة لدوال أنجل مؤشر أساسي في المجال التخطيطي، حيث يمكن توظيف هذا المؤشر لأغراض التنبؤ بالطلب في ضوء التغير المتوقع في الدخل وبالتالي وضع خطط الانتاج وخطة التجارة الخارجية وتحديد سياسات الاسعار والدخول.

وتشتق صيغة التنبؤ هذه من تعريف مرونة الطلب الدخلية ذاتها. وبالرجوع الى الصيغة رقم (2) وباعتبار أن (Y_{iJ}) و (Y_J) تعود لفترة أساس مناسبة تتوفر عنها البيانات المطلوبة، عندئذ يستخرج طلب المستهلك المقدر في الفترة (t) كالاتي :

$$\eta_i = \frac{\Delta Y}{\Delta X} * \frac{X_O}{Y_{io}}$$

حيث ان :

X_O : متوسط دخل الفرد في سنة الاساس.

Y_{io} : متوسط الانفاق على المجموعة السلعية (i) في سنة الاساس .

$$\therefore\ \Delta Y X_O = \eta_i\, \Delta X Y_{io}$$

$$\Delta Y = \frac{\eta_i\, \Delta X Y_{io}}{X_O}$$

$$\therefore\ Y_{it} - Y_{io} = \frac{\eta_i\, \Delta X Y_{io}}{X_O}$$

$$Y_{it} = Y_{io} + \eta_i \left(\frac{\Delta X}{X_O}\right) Y_{io}$$

$$\therefore\ Y_{it} = Y_{io}\left[1 + \eta_i \frac{\Delta X}{X_O}\right] \ldots\ldots\ldots\ldots\ 11$$

أي أن الطلب المتوقع للفرد على المجموعة السلعية (i) في الفترة (t) يساوي طلبة في فترة الاساس لتلك السلعة مضافا اليه التغير في الطلب نتيجة التغير في الدخل .

مثال 1 : بلغت مرونة الطلب الدخلية لسلعة معينة (1.30)، وقد بلغ متوسط استهلاك الفرد من السلعة (14) كغم في سنة الاساس (1980) فما هو مقدار الطلب الكلي على السلعة في سنة (2000)، علما بان دخل الفرد المتوقع سوف يرتفع بنسبة (20%) خلال الفترة (1980–2000) وان عدد سكان العراق المتوقع عام (2000) سوف يبلغ (28) مليون نسمة ؟

مثال2 : في ضوء المجاميع الاتية التي تمثل العلاقة بين الكمية المطلوبة من سلعة معينة (Y_i) وسعرها (X_i) مع بقاء العوامل الاخرى ثابتة :

$$n = 8 \ , \sum Y_i = 30 \ , \sum X_i = 24 \ , \sum X_i Y_i = 86 \ , \sum {X_i}^2 = 96 \ , \sum {Y_i}^2 = 140$$

المطلوب :

1- تقدير معلمات العلاقة بطريقة القيم الاصلية (التقدير حول نقطة الاصل)، المحددات، طريقة الانحرافات (التقدير حول نقطة المتوسط)، طريقة المصفوفات .

2- اختبار معنوية المعالم المقدرة عند مستوى معنوية (5%) اذا علمت بان قيمة (t) الجدولية عند مستوى معنوية 5% ودرجة حرية (6) هي 2.31 .

3- كون حدود ثقة 95% لكل من β_0 و β_1 .

4- وضح مقدار ما يفسره المتغير المستقل من التغير في المتغير التابع .

5- اوجد معامل الارتباط بين المتغيرين X_i و Y_i مفسرا النتيجة .

6- اختبر معنوية المعادلة المقدرة اذا علمت بان قيمة F الجدولية هي $F(1,6,0.05) = 5.99$

7- كون جدول تحليل التباين $(ANOVA)$.

8- تنبأ بقيمة (Y_{t+1}) عندما $X_{t+1} = 10$ وحساب حدود الثقة لها .

9- حساب مرونة الطلب السعرية؟

الحل :-

1-تقدير معلمات العلاقة بطريقة القيم الاصلية:

العلاقة المطلوب تقديرها هي :

$$Y_i = b_0 + b_1 X_i + e_i$$

حيث ان :

$$b_1 = \frac{n \sum X y_i - (\sum X_i)(\sum Y_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

$$b_1 = \frac{8\,(86) - (24)(32)}{8(96) - (24)^2} = \frac{-80}{192} = -0.416$$

$$b_0 = \bar{Y} - b_1 \bar{X}$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{32}{8} = 4$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{24}{8} = 3$$

$$\therefore \quad b_0 = 4 - (-0.416)(3)$$

$$b_0 = 5.248$$

$$\therefore \hat{Y}_i = 5.248 - 0.416\, X_i$$

تشير المعادلة التقديرية الى وجود علاقة عكسية بين المتغير المستقل X_i الذي يمثل السعر والمتغير التابع Y_i الذي يمثل الكمية المطلوبة فكل زيادة في السعر X_i بمقدار وحدة واحدة سوف تؤدي الى انخفاض الكمية المطلوبة بمقدار (0.416) وحدة .

طريقة المحددات :

$$\begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$\begin{bmatrix} 32 \\ 86 \end{bmatrix} = \begin{bmatrix} 8 & 24 \\ 24 & 96 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$|A| = \begin{vmatrix} 8 & 24 \\ 24 & 96 \end{vmatrix} = 192$$

$$\begin{vmatrix} 32 & 24 \\ 86 & 96 \end{vmatrix} = 1008 |A_0| =$$

$$|A_1| = \begin{vmatrix} 8 & 32 \\ 24 & 86 \end{vmatrix} = -80$$

$$\therefore \quad b_0 = \frac{|A_0|}{|A|} = \frac{1008}{192} = 5.25$$

$$\therefore \quad b_1 = \frac{|A_1|}{|A|} = \frac{-80}{192} = -0.416$$

والصيغة التقديرية تاخذ الشكل التالي

$$\therefore \quad \hat{Y}_i = 5.25 - 0.416 X_i$$

التقدير بطريقة الانحرافات (التقدير حول نقطة المتوسط) :-

من بيانات السؤال يمكن استخراج الانحرافات كما يلي :

$$\sum x_i y_i = \sum X_i Y_i - n\,\bar{X}\,\bar{Y} = 86 - (8)(3)(4) = -10$$

$$\sum x_i^2 = \sum X_i^2 - n\,(\bar{X})^2 = 96 - (8)(3)^2 = 24$$

$$\sum y_i^2 = \sum Y_i^2 - n\,(\bar{Y})^2 = 140 - 8\,(4)^2 = 12$$

$$\therefore\ b_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{-10}{24} = -0.416$$

$$b_0 = \bar{Y} - b_1\,\bar{X} = 4 - (-0.416)(3)$$

$$\therefore\ b_0 = 5.248$$

لذلك فان المعادلة التقديرية ستأخذ الشكل الاتي :

$$\therefore\ \hat{Y}_i = 5.25 - 0.416 X_i$$

التقدير بطريقة المصفوفات :- بموجب طريقة المصفوفات فان

$$\begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$\hat{B} = \underline{b} = (X'X)^{-1}X'Y$$

$$(X^{-}X) = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum X_i^2 \end{bmatrix} = \begin{bmatrix} 8 & 24 \\ 24 & 96 \end{bmatrix}$$

المحدد $$|X'X| = \begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum X_i^2 \end{vmatrix} = \begin{vmatrix} 8 & 24 \\ 24 & 96 \end{vmatrix} = 192$$

$$(X'X)^{-1} = \frac{adj(X'X)}{|X^{-}X|}$$

$$\therefore adj\ \ (X^{-}X) = \begin{bmatrix} 96 & -24 \\ -24 & 8 \end{bmatrix}$$

$$\therefore\ (X^{\acute{}}X)^{-1} = \frac{1}{192}\begin{bmatrix} 96 & -24 \\ -24 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5 & -0.125 \\ -0.125 & 0.04166 \end{bmatrix}$$

$$X^{\acute{}}Y = \begin{bmatrix} \sum \mathbf{Y_i} \\ \sum \mathbf{X_i Y_i} \end{bmatrix} = \begin{bmatrix} 32 \\ 86 \end{bmatrix}$$

$$\hat{B} = \underline{b} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = (X^{-}X)^{-1}X^{-}Y$$

$$= \begin{bmatrix} 0.5 & -0.125 \\ -0.125 & 0.04166 \end{bmatrix}\begin{bmatrix} 32 \\ 86 \end{bmatrix}$$

$$\begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 5.25 \\ -0.416 \end{bmatrix}$$

$$\therefore\ \hat{Y}_i = 5.25 - 0.416\, X_i$$

ويلاحظ الحصول على نفس الصيغة التقديرية باستخدام الطرق المختلفة.

2-اختبار معنوية المعالم المقدرة :-

وهذا يتطلب تقدير تباين العينة للخطأ العشوائي (S_e^2) والذي يستخرج وفق الصيغ التالية:

$$S_e^2 = \frac{\sum e_i^2}{n-k-1}$$

$$S_e^2 = \frac{\sum Y_i^2 - b_0 \sum Y_i - b_1 \sum X_i Y_i}{n-k-1}$$

$$= \frac{140 - (5.248)(32) - (-0.416)(86)}{8-1-1} = 1.306$$

ونفس التقدير اعلاه لتباين حد الخطأ يمكن الحصول عليه باستخدام الانحرافات أي أن :

$$S_e^2 = \frac{\sum y_i^2 - b_1^2 \sum x_i^2}{n-k-1} = \frac{\sum y_i^2 - b_1 \sum x_i y_i}{n-k-1} = 1.306$$

تباين الحد الثابت $(\boldsymbol{b_0})$

$$var\ (\boldsymbol{b_0}) = S_{b0}^2 = S_e^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum x_i^2}\right]$$

$$= 1.306 \left[\frac{1}{8} + \frac{(3)^2}{24}\right] = 0.653$$

$$\therefore\ S_{bo} = \sqrt{S_{bo}^2} = \sqrt{0.653} = 0.8081$$

$$t_{b0} = \frac{b_0}{S_{bo}} = \frac{5.248}{0.8081} = 6.494$$

وبما ان قيمة t المحسوبة والبالغة (6.494) اكبر من قيمة t الجدولية عند مستوى معنوية (5%) ودرجة حرية (6) والبالغة (2.31) عليه ترفض فرضية العدم (H_O) ونقبل الفرضية البديلة اي معنوية المعلمة المقدرة (bo) .

تباين الميل الحدي $(\boldsymbol{b_1})$

$$\overset{\wedge}{var}\ (\boldsymbol{b_1}) = S_{b1}^2 = \frac{S_e^2}{\sum x_i^2} = \frac{\mathbf{1.306}}{\mathbf{24}} = \mathbf{0.0544}$$

$$\therefore\ S_{b1} = \sqrt{S_{b1}^2} = \sqrt{0.0544} = 0.233$$

$$t_{b1} = \frac{b_1}{S_{b1}} = \frac{-0.416}{0.233} = -1.785$$

وبما ان قيمة t المحتسبة والبالغة (1.785) أقل من قيمة t الجدولية عند مستوى معنوية (0.05) ودرجة حرية (6) والبالغة (2.31) عليه نقبل فرضية العدم (H_0) اي عدم معنوية المعلمة المقدرة b_1 .

3- حدود الثقة للمعالم المقدرة :-

بالنسبة الى المعلمة $\boldsymbol{b_0}$ فان :

$$\boldsymbol{B_0} = \boldsymbol{b_0} \pm \left(\boldsymbol{t_{n-k-1}}, \frac{\propto}{\mathbf{2}}\right) \boldsymbol{S_{b0}}$$

$$= \mathbf{5.248} \pm (\mathbf{2.31})(\mathbf{0.8081})$$

$$\boldsymbol{B_0 = 5.248 \pm 1.8667}$$

$$\boldsymbol{3.3813 \leq B_0 \leq 7.1147}$$

$$\therefore \ \mathbf{pr}\{\boldsymbol{3.3813 \leq B_0 \leq 7.1147}\} = \boldsymbol{0.95}$$

أي أن هناك احتمال 95% ان تقع القيمة الحقيقية لمعلمة المجتمع $\boldsymbol{B_0}$ بين الحدين الاعلى (7.1147) والادنى (3.3813) وهناك احتمال 5% ان تقع خارج هذين الحدين .

بالنسبة الى المعلمة $\boldsymbol{b_1}$ فان :

$$\boldsymbol{B_1 = b_1 \pm \left(t_{n-k-1}, \frac{\propto}{2}\right) S_{b1}}$$

$$\boldsymbol{= -0.416 \pm (2.31)(0.233)}$$

$$\boldsymbol{B_1 = -0.416 \pm 0.53823}$$

$$\boldsymbol{-0.95423 \leq B_1 \leq 0.12223}$$

$$\therefore \ \mathbf{pr}\{\boldsymbol{-0.95423 \leq B_1 \leq 0.12223}\} = \boldsymbol{0.95}$$

اي هناك احتمال 95% ان تكون قيمة $\boldsymbol{B_1}$ محصورة (تقع) بين الحدين الاعلى (0.12223) والادنى (−0.954) .

4−معامل التحديد $\boldsymbol{R^2}$:

$$R^2 = \frac{\boldsymbol{b_1} \sum x_i y_i}{\sum y_i^2} = \frac{(-0.416)(-10)}{12} = \frac{4.16}{12} = 0.34666$$

$$\therefore \ \mathrm{R}^2 = 0.34666 \quad = \quad \% \, 34.666$$

هذا يعني ان المتغير المستقل X_i والذي يمثل السعر يفسر (يوضح) حوالي 34.66 % من التغيرات التي تحصل في المتغير المعتمد Y_i والذي يمثل الكمية المطلوبة من السلعة وان النسبة المتبقية والبالغة 65.334 % تمثل تأثير متغيرات اخرى لم تضمن في المعادلة كالدخل، اسعار السلع البديلة.....الخ .

5− معامل الارتباط بين $\boldsymbol{X_i}$ و $\boldsymbol{Y_i}$:

$$r = \left(\sqrt{\frac{\sum x_i^2}{\sum y_i^2}}\right) b_1 = \sqrt{\frac{24}{12}} \, (-0.416)$$

$$r = \sqrt{2} \, (-0.416) = -0.5883128$$

اي أن العلاقة عكسية ومتوسطة القوة بين X_i و Y_i .

6- اختبار F

$$F = \frac{b_1 \sum x_i y_i / \mathrm{k}}{\sum e_i^2 / \ \mathrm{n-k-1}} = \frac{(-0.416)(-10)/1}{7.84/\,8-1-1}$$

$$= \frac{4.16}{1.306666}\ 3.183673$$

بما ان قيمة (F) المحسوبة والبالغة (3.183) هي اقل من قيمة (F) الجدولية عند مستوى معنوية 5% ودرجة حرية $(6,1)$ للبسط والمقام والبالغة (5.99)عليه نقبل فرضية العدم (H_0) والتي تنص على عدم معنوية العلاقة المقدرة .

7-تكوين جدول تحليل التباين ANOVA

مصدر التباين source of variance s.o.v	مجموع المربعات Sum of square	درجات الحرية d.f	متوسط المربعات Mean sum square M.S.S	$F-test$
الانحرافات الموضحة من قبل خط الانحدار X_i	4.16 SSR	1	4.16	$F = \frac{4.16}{1.306666}$ $= 3.1836$
الانحرافات غير الموضحة (المتبقية)	7.84 SSE	$8-1-1$	$S^2{}_e = 1.306666$	
مجموع الانحرافات (الانحرافات الكلية)	12 SST	$8-1=7$		

8- التنبؤ :-

$$\hat{Y}_i = 5.248 - 0.416\, X_i$$

$$Y_{t+1} = 5.248 - 0.416\,(10)$$

$$= 5.248 - 4.16 = 1.088$$

9- مرونة الطلب السعرية :

$$\eta_i = b_1 * \frac{\bar{X}}{\bar{\bar{Y}}}$$

$$= (-0.416)\left(\frac{3}{4}\right) = -0.312$$

$$\therefore\ \eta_i < 0$$

لذلك فان السلعة متدنية (رديئة) .

وفيما يلي الصيغة التقديرية الكاملة للعلاقة بين الكمية المطلوبة من السلعة وسعرها .

$$\hat{Y}_i = 5.248 - 0.416\, X_i \qquad R^2 = 0.3466$$

$$S.E \quad (0.8081)\ \ (0.233) \qquad F = 3.1836$$

$$S_e^2 = 1.3066$$

تمارين واجب:

1. على افتراض ان باحث اقتصادي قدّر دالة الاستهلاك وحصل على النتيجة الاتية:

$$\hat{Y}_t = 21.5 + 0.84\, X_t$$

$$S.E: (9.4) \quad (0.024)$$

$n = 12 \quad , \quad R^2 = 0.992$

حيث ان الأرقام بين قوسين تشير الى الانحراف المعياري.

المطلوب:

أ – اختبر مدى معنوية المعلمات المقدرة علماً ان: $t(10.0.05) = 2.23$

ب – اختبر مدى معنوية العلاقة الخطية المفترضة بين: $Y_t\,,\ X_t$

علماً ان $F\,(1.10.0.05) = 4.10$

2. فرضاً باحث قدّر دالة الاستهلاك وحصل على النتيجة الاتية:

$$\hat{C}_t = 15 + 0.81\, Y_t$$

$$t: (3.1) \quad (18.7) \qquad\qquad n = 19$$

حيث ان الأرقام بين القوسين تشير الى قيمة (t) المحسوبة وان: C_t : الاستهلاك، Y_t : الدخل.

المطلوب:

أ- احسب الانحراف المعياري للمعالم المقدرة؟

ب- كوّن حدود الثقة ولمستوى دلالة (95%) لمعامل المتغير المستقل (Y_t) علماً ان:

$$t\,(n\text{-}2\,,\frac{\alpha}{2}) = 2.11$$

3. إذا توفرت لديك البيانات الاتية عن الكمية المطلوبة من سلعة معينة (Y_i) وسعرها (X_i)

$$\sum_{i=1}^{10}(X_i - \bar{X})^2 = 280 \quad . \quad \sum_{i=1}^{10}(Y_i - \bar{Y})^2 = 520$$

$$\sum_{i=1}^{10}(X_i - \bar{X})(Y_i - \bar{Y}) = -360 \quad , \quad \sum X_i = 40 \quad , \quad \sum Y_i = 100$$

المطلوب:

أ - قدّر معلمات العلاقة: $Y_i = b_o + b_1 X_I + ei$ مع التفسير الاقتصادي لها ؟

ب - كوّن جدول تحليل التباين، واختبر معنوية العلاقة المقدرة إذا علمت ان:

F (1,8,0.05) = 5.3 وان F (2,10,0.05) = 4.1

ج – اوجد معامل التحديد (R^2) ، ومعامل الارتباط (r) مبيناً العلاقة بينهما ومفسراً النتيجة ؟

4. استناداً للمعطيات الإحصائية التالية عن العلاقة بين الكمية المعروضة من سلعة معينة (Y_i) وسعرها (X_i):

$R^2 = 0.65$, $F_{5\%} = 8.81$, $t_{5\%} = 2.45$

$$\sum x_i^2 = 34 , \sum x_i y_i = 28 , \bar{Y} = 13 , \bar{X} = 8$$

المطلوب:

أ – تقدير معلمات العلاقة : $\hat{Y}_i = b_o + b_1 X_i$ مفسراً النتيجة اقتصادياً ؟

ب – أكمل جدول تحليل التباين الاتي:

S . O . V	S .S	d.f	M . S . S	F - test
الانحرافات الموضحة			23.044	
الانحرافات غير الموضحة				
الانحرافات الكلية		7		

ج- اختبر معنوية العلاقة الخطية المقدرة؟

ء- اختبار معنوية المعلمات المقدرة؟

هـ - قدّر فترة الثقة 95% للمعلمة (b_1) فقط؟

5. اذا توفرت لديك البيانات التالية عن المتغيرين (X_i) , (Y_i) :

$$n = 8 , \sum Y_i = 104 , \sum X_i = 64 , \sum x_i y_i = 28$$

$$\sum (X_i - \bar{X})^2 = 34 , \sum (Y_i - \bar{Y})^2 = 36$$

المطلوب:

أ- تقدير معالم العلاقة الخطية بين (X_i) , (Y_i) مفسراً النتيجة ؟

ب- حساب R^2 ، $\bar{R}^2$ ، مفسراً النتيجة ؟

ت- اختبار معنوية المعلمة (b_1) وحساب حدود الثقة لها؟

ث- تكوين جدول تحليل التباين (ANOVA)؟

تمارين واجب المجموعة الثانية:

1. البيانات التالية تمثل استيرادات العراق (Y_i) والناتج المحلي الإجمالي (X_i) مقاسة بـ (بليون دينار) للمدة (1976 - 1983).

Y_i	1.2	1.3	1.5	1.7	2.2	2.3	2.9	1.9
X_i	5.2	5.9	7.0	11.2	15.6	11.2	12.6	12.5

المطلوب:

Æ تقدير معلمات العلاقة الخطية بين الاستيرادات والناتج المحلي الإجمالي $\hat{Y}_i = b_o + b_1 X_I$

مفسراً النتيجة اقتصادياً ؟

ب- إيجاد حدود الثقة لمعاملات العلاقة المقدرة؟

ت- حساب معامل التحديد ومعامل الارتباط البسيط موضحاً العلاقة بينهما؟

ث- تكوين جدول تحليل التباين (ANOVA).

2. إذا توفرت لديك البيانات التالية عن عينة بحجم (5) مشاهدات

$$\sum X_i = 45\,,\quad \sum X_iY_i = 665\,,\ \sum Y_i = 70$$

$$\sum X_i^2 = 425\,,\quad \sum Y_i^2 = 1054$$

المطلوب:

أ – تقدير معلمات العلاقة الخطية بين (Y_i) , (X_i) ؟

ب – اختبار معنوية المعلمة (b_1) وحساب حدود الثقة (95%) لها فقط ؟

ت - كوّن جدول تحليل التباين ((ANOVA)) ثم اختبر معنوية النموذج مبيناً الفرضية المطلوب اختبارها؟

ث- حقق صحة العلاقة التالية:

$$F = t_{b1}^{2}.$$

تمارين واجب: مثال: (3)

بوبت الاسر الحضرية المشمولة في بحث ميزانية الاسرة لعام (1979) وفق الفئات المبينة في الجدول التالي المتضمن متوسط دخل الفرد ومتوسط انفاقه الشهري على مجموعة المواد الغذائية في قطاع الحضر العراقي.

فئة دخل الفرد دينار/ شهر	إنفاق الفرد على المواد الغذائية (دينار/ شهر) Y_{ij}	دخل الفرد (دينار/ شهر) X_j
4 فأقل	6.10	10.20
6-	5.48	10.32
8-	6.09	10.50
10-	6.83	12.40
12-	7.41	13.84
14-	7.59	14.81
16-	8.15	15.86
18-	9.00	18.31
20-	9.15	18.47
25-	9.76	20.65
30-	10.40	22.96
35-	11.43	26.22
40-	11.74	27.88
45-	12.67	29.81
50-	14.07	34.75
اكثر من 50	15.63	48.40
كافة الفئات (المعدل)	9.46 Y_{ij}	20.96 X_j

المطلوب: تقدير الميل الحدي للاستهلاك ومرونة الطلب الدخلية باستخدام الصيغ التالية لدوال انجل:

$$1.\; Y_{ij} = B_0 + B_1 X_j + u_{ij}$$

$$2.\; Y_{ij} = B_0 + B_1 lnX_j + u_{ij}$$

$$3.\; lnY_{ij} = lnB_0 + B_1 lnX_j + u_{ij}$$

ثم قارن بين النتائج التي تحصل عليها وبيّن أي دالة أكثر ملائمة لنمط إنفاق المستهلك.

الحل: العمليات الحسابية التالية تم الحصول عليها من الجدول أعلاه:

$$\Sigma X_j = 335.38 \;, \Sigma X_j^2 = 8670.4962 \;, \bar{X} = 20.96$$

$$\Sigma Y_{ij} = 151.5 \;, \Sigma Y_{ij}^2 = 1569.3194 \;, \bar{Y} = 9.46$$

$$\Sigma X_j Y_{ij} = 3635.8734 \;, \Sigma lnY_{ij} lnX_j = 105.2052$$

$$\Sigma lnY_{ij} = 35.14 \;, \Sigma lnX_j = 46.92 \;, \quad n = 16$$

$$\Sigma Y_{ij} lnX_j = 464.94 \;, \Sigma (lnX_j)^2 = 140.798$$

$$\Sigma (lnY_{ij})^2 = 78.6482$$

اولاً: تقدير المؤشرات الخاصة بالصيغة الخطية:

$$b_1 = \frac{n\,\Sigma X_j Y_{ij} - (\Sigma Y_{ij})(\Sigma X_j)}{n\,\Sigma X_j^2 - (\Sigma X_j)^2} = 0.28$$

$$b_o = \bar{Y} - b_1 \bar{X} = 3.59$$

$$\hat{Y}_{ij} = 3.59 + 0.28\, X_j$$

$$MPC = b_1 = 0.28$$

$$\eta = b_1 \frac{\bar{X}}{\bar{Y}} = 0.62$$

$$R^2 = \frac{b_1\, \Sigma x_j y_{ij}}{\Sigma y_{ij}^2}$$

اذ ان:

$$\sum x_j y_{ij} = \sum X_j Y_{ij} - n\bar{X}\bar{Y} = 463.3678$$

$$\sum y_{ij}^2 = \sum Y_{ij}^2 - n\,\bar{Y}^2 = 137.4538$$

$$R^2 = 0.94$$

وان:

$$F = \frac{R^2/K}{1 - R^2/n - K - 1}$$

ثانيا: تقدير المؤشرات الخاصة بالصيغة نصف اللوغاريتمية:

$$b_1 = \frac{n\sum Y_{ij} lnX_j - (\sum lnX_j)(\sum Y_{ij})}{n\sum (lnX_j)^2 - (\sum lnX_j)^2} = 6.45$$

$$b_o = \bar{Y} - b_1 ln\bar{X} = -9.45$$

$$\hat{Y}_{ij} = -9.45 + 6.45\, ln\, X_j$$

$$\mathrm{MPC} = \frac{b_1}{\bar{X}} = 0.31$$

$$\eta = \frac{b_1}{\bar{Y}} = 0.68$$

$$R^2 = \frac{b_1 \sum y_{ij} lnx_J}{n\sum y_{ij}^2}$$

اذ ان : $\sum y_{ij} lnx_j = \sum Y_{ij} lnX_j - n\,\bar{Y}\, ln\bar{X} = 20.6268$

$$R^2 = 0.97$$

$$F = \frac{R^2/K}{1 - R^2/n - K - 1} = 453.27$$

ثالثا: تقدير المؤشرات الخاصة بالصيغة اللوغارتمية المزدوجة:

$$b_1 = \frac{n\sum lnY_{ij}lnX_j - (\sum lnX_j)(\sum lnY_{ij})}{n\sum (lnX_j)^2 - (\sum lnX_j)^2} = 0.67$$

$$b_o = ln\overline{Y} - b_1 ln\overline{X} = 0.23$$

$$ln\hat{Y}_{ij} = 0.23 + 0.67\, ln\, X_j$$

ويمكن وضع الصيغة التقديرية أعلاه بشكلها غير الخطي كالاتي:

$$Anti - ln(b_o) = Anti - ln(0.23) = 1.26$$

$$\hat{Y}_{ij} = 1.26X_j^{0.67}$$

$$MPC = b_1 \frac{\overline{Y}}{\overline{X}} = 0.30$$

$$\eta = b_1 = 0.67$$

$$R^2 = \frac{b_1 \sum lnx_j lny_{ij}}{\sum (lny_{ij})^2}$$

اذ ان:

$$\sum lny_{ij} lnx_j = \sum lnY_{ij} lnX_j - n\, ln\overline{Y}\, ln\overline{X} = 2.15715$$

وان:

$$\sum (lny_{ij})^2 = \sum (Y_{ij})^2 - n(ln\overline{Y})^2 = 1.472$$

$$R^2 = 0.98$$

$$F = \frac{R^2/K}{1 - R^2/n - K - 1} = 690.14$$

وبعد تقدير معالم الصيغ المختلفة لدوال أنجل واحتساب بعض المؤشرات الإحصائية والاقتصادية الخاصة بكل صيغة، يمكن المقارنة بين النتائج واختيار الدالة الأكثر ملائمة لسلوك نمط انفاق المستهلك على مجموعة من المواد الغذائية، ويتم ذلك وفق نوعين من المعايير هي المعايير الاقتصادية والمعايير الإحصائية، وتتلخص المعايير الاقتصادية بمدى اتفاق نتائج الصيغة المقدرة مع ما تمليه النظرية الاقتصادية، اذ ينبغي ان تكون قيمة المرونة متوافقة مع طبيعة السلعة من حيث كونها كمالية او ضرورية او متدنية، إضافة الى ذلك يستوجب توافق قيم الميل الحدي للاستهلاك ومستوى الاشباع.

أما فيما يتعلق بالمعايير الإحصائية فأهمها معامل التحديد (R^2) والذي بموجبه يمكن معرفة قدرة الدالة التفسيرية ومدى منطقية وتحقق الافتراضات الخاصة بكل دالة حيث انه كلما كان هذا المؤشر ذا قيمة عالية دلّ ذلك على ان الدالة أكثر قدرة للتعبير عن العلاقة ما بين الطلب والدخل.

وتجدر الإشارة هنا الى ان قيم معامل التحديد هذا ينبغي ان تحسب جميعها على أساس تجانس قيم المتغير المعتمد ولجميع الدوال، اذ لا تصح المقارنة حينما يكون معامل التحديد محسوب على أساس ان المتغير المعتمد هو (Y_{ij}) بالنسبة للدالة الخطية ونصف اللوغارتمية في حين يأخذ المتغير المعتمد القيمة (lnY_{ij}) في حالة الدالة اللوغاريتمية المزدوجة. ويمكن تحقيق الاتساق ما بين معامل التحديد (R^2) للدوال المتخلفة عن طريق إيجاد معامل التحديد المصحح ويتم ذلك وفق الصيغة الاتية:

$$\overline{R}^2 = 1 - \frac{(1 - R^2)(n-1)}{n - K - 1)}$$

وبالتطبيق على هذا المثال ولكافة الصيغ المقدرة نجد ان قيمة معامل التحديد المصحح في حالة الصيغة الخطية مساوياً الى $\overline{R}^2 = 0.93$ وفي حالة الصيغة نصف اللوغاريتمية فان $\overline{R}^2 = 0.96$ وفي حالة اللوغاريتمية المزدوجة فان $\overline{R}^2 = 0.97$. اما فيما يتعلق بالافتراضات الخاصة بالدوال المختلفة والتي تدور حول شكل الدالة المفترضة فيكن التعرف عليها من خلال اختبار (F) القرة لكافة الصيغ.

الجدول ادناه يبين المؤشرات النهائية والتي يمكن على أساسها المفاضلة بين الصيغ الثلاث المقدرة لمجموعة من المواد الغذائية.

ت	الصيغة التقديرية	MPC	µ	$\overline{R}^2$	F المحسوبة
1	$\hat{Y}_{ij} = 3.594 + 0.28X_j$	0.28	0.62	0.93	219.63
2	$\hat{Y}_{ij} = -9.45 + 6.45lnX_j$	0.31	0.68	0.96	453.27
3	$ln\hat{Y}_{ij} = 0.23 + 6.45lnX_j$	0.30	0.76	0.97	690.14

ملاحظة: علماً ان: $F(1\ ,14,0.05) = 4.60$

الفصل الرابع

الانحدار الخطي المتعدد Multiple Linear Regression

المقدمة :

ان الانحدار البسيط يركز على دراسة العلاقة بين متغيرين أحداهما المتغير المستقل (X) والاخر المتغير التابع (Y). غير ان واقع الحياة الاقتصادية والاجتماعية مبني بشكل عام على تأثر اي ظاهرة بأكثر من متغير مستقل. فعلى سبيل المثال دالة الانتاج والتي توصف بانها دراسة للعلاقة بين عناصر الانتاج (العمل ورأس المال) والناتج اي ان : $Q_i = f\ (L_i\,.K_i\ )$

وكذلك دالة الطلب التي تحدد العلاقة بين الكميات المطلوبة من سلعة معينة والاسعار الخاصة بهذه السلعة واسعار السلع البديلة وعلى دخل المستهلك فضلا عن المتغيرات الاخرى. لذلك لا بد من توسيع نموذج الانحدار الخطي البسيط السابق ليشتمل على انحدار المتغير التابع (y) على العديد من المتغيرات المستقلة $(X_1.X_2.\ldots.X_k)$ ويسمى هذا بنموذج الانحدار الخطي المتعدد Multiple Linear Regression .

النموذج الخطي المتعدد :-

يستند النموذج الخطي المتعدد على افتراض وجود علاقة خطية بين متغير تابع (y_i) وعدد من المتغيرات المستقلة $(X_1.X_2.\ldots.X_k)$ وحد عشوائي ui . ويعبر عن هذه العلاقة بالنسبة لـ n من المشاهدات و k من المتغيرات المستقلة بالشكل التالي :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + U_i \quad \ldots\ldots\ldots (1)$$

وفي واقع الامر فان هذه المعادلة هي واحدة من جملة معادلات يبلغ عددها n

اذ ان

$(i = 1.2.3.\ldots\ldots n)$ تكون نظام المعادلات الاتي :

$$Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_k X_{1k} + U_1$$

$$Y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_k X_{2k} + U_2$$

$$. \quad . \quad . \quad . \quad . \quad . \quad .$$

$$. \quad . \quad . \quad . \quad . \quad . \quad .$$

$$Y_n = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_k X_{nk} + U_n$$

هذه المعادلات تتضمن $(k+1)$ من المعالم المطلوب تقديرها علما بان الحد الاول منها (β_0) يمثل الحد الثابت اي ان $(X_{i0} = 1)$. لذلك يتطلب اللجوء الى المصفوفات والمتجهات لتقدير تلك المعالم. عليه يمكن صياغة هذه المعادلات في صورة المصفوفات كالاتي:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & \cdots & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{n2} & \cdots & X_{nk} \end{bmatrix} . \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix}$$

وبشكل اكثر اختصارا كالاتي:

$$Y = XB + U \quad \ldots\ldots\ldots \quad (2)$$

اذ ان:

Y : متجه عمودي ابعاده (n*1) يحتوي مشاهدات المتغير التابع.

X :مصفوفة ابعادها (n*k+1) تحتوي مشاهدات المتغيرات المستقلة ويحتوي عمودها الاول على قيم الواحد الصحيح ليتضمن الحد الثابت (β_0) .

B : متجه عمودي من الرتبة (k+1*1) يحتوي على المعالم المطلوب تقديرها.

U : متجه عمودي من الرتبة (n*1) يحتوي على الاخطاء العشوائية.

ولغرض تقدير المعالم المجهولة في العلاقة (2) فانه يستوجب تحقق الفروض الاساسية الخاصة بالمتغير العشوائي ui التالية :

$$\boldsymbol{U_i \sim N\,(0\,.\sigma^2\, I_n)}$$

والذي يعني ان ui يتوزع توزيعاً طبيعياً (N) متعدد المتغيرات بوسط حسابي (توقعه) يساوي صفر ومصفوفة تباين وتباين مشترك عددية هي $(\boldsymbol{\sigma^2 I_n})$.

فرضيات النموذج الخطي المتعدد :-

عند استخدام طريقة (ols) في تقدير نموذج الانحدار الخطي المتعدد. فانه يجب تحقق الافتراضات الاتية

1- القيمة المتوقعة لمتجه حد الخطأ تساوي صفرا اي ان $\mathrm{E(ui)} = 0$

$$\mathrm{E(ui)} = E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} E(u_1) \\ E(u_2) \\ \vdots \\ E(u_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

2- تباين العناصر العشوائية ثابت. والتباين المشترك بينهما يساوي صفراً . اي ان :

$$var\ (u) = E(uu^-) = \ \sigma^2 In$$

أي ان :

$$var\ (u) = E(uu^-) = E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} [u_1\ u_2\ \ \dots\dots\ u_n]$$

$$= E\begin{bmatrix} u_1 & u_1 \dots\dots & u_1 u_n \\ u_2 & u_2 \dots\dots & u_2 u_n \\ \vdots & \vdots & \vdots \\ u_n u_1 & u_n u_2 \cdots\cdots & u_n^2 \end{bmatrix}$$

$$E\ = \begin{bmatrix} E(u_1^2) & E(u_1 u_2) \dots\dots & E(u_1 u_n) \\ E(u_2 u_1) & E(u_2^2) \dots\dots & E(u_2 u_n) \\ \vdots & \vdots & \vdots \\ E(u_n u_1) & E(u_n u_2) \cdots\cdots & E(u_n^2) \end{bmatrix}$$

$$var(u) = \begin{bmatrix} var(u_1) & cov(u_1 u_2) & \dots\ \dots & cov(u_1 u_n) \\ cov(u_2 u_1) & var(u_2) & \dots\ \dots & cov(u_2 u_n) \\ \vdots & \vdots & & \vdots \\ cov(u_n u_1) & cov(u_n u_2) & \dots\ \dots & var(u_n) \end{bmatrix}$$

$$\therefore\ var(u_i) = E\left(u_i^2\right) = \ \sigma^2$$

$$cov(u_i u_j) = E(u_i u_j) = 0 \qquad \forall \ i \neq j$$

$$\therefore var(u) = \begin{bmatrix} \sigma^2 & 0 & \dots\dots & 0 \\ 0 & \sigma^2 & \dots\dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots\dots & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{bmatrix}$$

حيث ان : $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \cdots = \sigma_n^2 = \sigma^2$

$$\therefore var(u) = \sigma^2 I_n$$

وتسمى المصفوفة اعلاه بمصفوفة التباين والتباين المشترك لحد الخطأ العشوائي u .Variance Covariance Matrix . حيث تشكل العناصر القطرية في المصفوفة تباين قيم u . بينما العناصر غير القطرية (اعلى واسفل القطر) تكون مساوية للصفر لانعدام التباين المشترك والترابط بين قيم ui . وان (I_n) هي مصفوفة الوحدة (Identity Matrix) من الرتبة (nxn) .

3- ليس هناك علاقة خطية تامة (تعدد خطي تام) بين المتغيرات المستقلة. كما وان عدد المشاهدات (n) يجب ان يكون أكبر من عدد المعالم المطلوب تقديرها اي ان

$$Rank(x) = R(X) = k + 1 \qquad < n$$

حيث ان $R(X)$ تمثل رتبة المصفوفة (X) وتساوي عدد المتغيرات المستقلة (k) زائد (1) الحد الثابت . وهي اصغر من عدد المشاهدات (n) . وهذه الفرضية ضرورية جدا لضمان ايجاد معكوس المصفوفة $(X\acute{X})$ وفي حالة عدم تحقق هذا الفرض تظهر مشكلة الارتباط الخطي المتعدد وبالتالي لا يمكن الحصول على مقدرات المربعات الصغرى الاعتيادية (OLS) .

طرق تقدير معلمات النموذج :-

عند تحقق الفروض اعلاه . يمكن استخدام طريقة المربعات الصغرى الاعتيادية (OLS)لتقدير معالم النموذج رقم (2) اعلاه. حيث يمكن اعادته بالشكل الاتي:

$$Y = X\beta + u$$

$$\therefore u = Y - X\beta$$

$$u^{-}u = (Y - X\beta\acute{)}\,(Y - X\beta) = (\acute{Y} - \acute{\beta}\,\acute{X})(Y - X\beta)$$

$$\therefore u^{-}u = \acute{Y}\mathrm{Y} - \acute{Y}\mathrm{X}\beta - \acute{\beta}\,\acute{X}Y + \acute{\beta}\acute{X}X\beta$$

ويلاحظ من الصيغة اعلاه بان الحد الثاني والثالث ذات قيمة محدودة ولا علاقة لهما بالمصفوفات وعليه يمكن جمعها

$$\therefore\ \acute{u}u = \acute{Y}Y - 2\acute{\beta}\acute{X}Y + \acute{\beta}\acute{X}X\beta$$

$$\therefore\ \frac{\partial \acute{u}u}{\partial \acute{B}} = -2\acute{X}Y + 2\acute{X}Xb = 0$$

وبقسمة طرفي المعادلة على (2) والتبسيط ينتج ان

$$\acute{X}Xb = \acute{X}Y$$

$$\therefore\ b = (\acute{X}X)^{-1}\,\acute{X}Y \qquad \dots\dots\dots (3)$$

$$or\quad b(ols) = (\acute{X}X)^{-1}\acute{X}Y$$

التقدير اعلاه يعرف بالتقدير حول نقطة الاصل (باستخدام القيم الاصلية) .

التقديرات اعلاه غير متحيزة وتمتلك مصفوفة تباين وتباين مشترك يمكن اشتقاقها كالاتي:

أ- خاصية عدم التحيز :-

بالرجوع الى النموذج رقم (2) اعلاه والعلاقة في (3) حيث ان

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

وان :

$$Y = X\beta + u$$

وبالتعويض في اعلاه نحصل على ان

$$b = (\acute{X}X)^{-1}\,\acute{X}\,(X\beta + u)$$

$$= (\acute{X}X)^{-1}\,\acute{X}X\beta + (\acute{X}X)^{-1}\,\acute{X}u$$

$$b = \beta + (\acute{X}X)^{-1}\,\acute{X}u \quad \dots\dots\dots \quad (4)$$

وبإدخال التوقع على الطرفين وحيث ان $E(u) = 0$ فان

$$E(b) = \beta + (\acute{X}X)^{-1}\,\acute{X}\,E(u)$$

$$\therefore E(b) = \beta$$

وهذا يعني ان b(ols) هي عبارة عن مقدمات غير متحيزة للمعلمة الحقيقية β.

ب- ولاشتقاق مصفوفة التباين والتباين المشترك لمقدرات نموذج الانحدار الخطي المتعدد يتم وفق الاتي:

بإعادة كتابة الصيغة رقم (4) كالاتي:

$$b - \beta = (\acute{X}X)^{-1}\,\acute{X}u$$

$$\therefore var - cov(b) = E\left[\,(b-\beta)(b-\acute{\beta})\right]$$

وبالتعويض في الصيغة أعلاه فان

$$var - cov(b) = \mathrm{E}\left\{\left[(\acute{X}X)^{-1}\acute{X}u\right]\left[(\acute{X}X)^{-1}\acute{X}u\,\right]^{-}\right\}$$

$$= E\left[(\acute{X}X)^{-1}\acute{X}u\acute{u}X(\acute{X}X)^{-1}\,\right]$$

$$= (\acute{X}X)^{-1}\acute{X}\,E(u\acute{u})X(\acute{X}X)^{-1}$$

$$E(u\acute{u}) = \sigma_u^2\, I_n$$

$$\therefore var - cov(b) = \sigma_u^2\, I_n(\acute{X}X)^{-1}\,\acute{X}X\,(\acute{X}X)^{-1}$$

$$\therefore var - cov(b) = \sigma_u^2\,(\acute{X}X)^{-1} \quad \dots\dots\dots \quad (5)$$

وهذا يعني ان قيمة تباين اي عنصر من عناصر المتجه (b) هو عبارة عن حاصل ضرب قيمة σ_4^2 بما يقابلها من العناصر الواقعة على القطر الرئيسي للمصفوفة $(\acute{X}X)^{-1}$. اما التباين المشترك بين اي اثنين من عناصر المتجه (b) فهو عبارة عن حاصل ضرب σ_u^2 بالعنصر المقابل لهما والواقع خارج القطر الرئيسي للمصفوفة $(\acute{X}X)^{-1}$. علما ان الصيغة التقديرية لتباين الخطأ العشوائي (تباين العينة) وهي S_e^2 اي ان $S_e^2 = \sigma_u^{\wedge 2}$

فيمكن حسابها وفق الصيغة الاتية:

$$S_e^2 = \frac{\sum e_i^2}{n-k-1} = \frac{\acute{e}e}{n-k-1}$$

$$\therefore\ \boldsymbol{S_e^2(n-k-1) = \acute{e}e = (Y-\hat{Y})\acute{}\,(Y-\hat{Y})}$$

اذ ان $\hat{Y} = Xb$

وبعد التبسيط نحصل على الصيغة النهائية الاتية:

$$S_e^2 = \frac{\acute{Y}Y - \acute{b}\acute{X}Y}{n-k-1} \quad \dots\dots\dots \quad (6)$$

حيث ان:

n: حجم العينة تحت البحث.

K: عدد المتغيرات المستقلة.

وعلى هذا الاساس فان الصيغة التقديرية لمصفوفة التباين والتباين المشترك لمتجه المعالم (b) في العلاقة (5) وتأخذ الشكل الاتي:

$$var - cov(b) = \ S_e^2\ (\acute{X}X)^{-1}$$

التقدير حول نقطة المتوسط (طريقة الانحرافات) :-

من الممكن تسهيل العمليات الحسابية في تقدير معاملات الانحدار المتعدد باستخدام اسلوب الانحرافات او ما يسمى بالمتوسطات. اي التعامل مع الانحرافات لكل من المتغير المعتمد والمتغيرات المستقلة عن اوساطها الحسابية. ولتوضيح ذلك نأخذ نموذج يحتوي على متغيرين مستقلين X_2 . X_1 اي ان

$$Y_i = b_0 + b_1 X_{i1} + \ b_2 X_{i2} + e_i$$

وبأخذ المتوسط لهذه المعادلة where $\bar{e} = 0$

$$\bar{Y} = b_0 + b_1 \bar{X}_1 + \ b_2 \bar{X}_2 + \bar{e}$$

$$Y_i - \bar{Y} = b_1 (X_{i1} - \bar{X}_1) + \ b_2 (X_{i2} - \bar{X}_2) + \ e_i$$

$$y_i = \ b_1 x_{i1} + b_2 x_{i2} + e_i \ \dots\dots\dots \ (7)$$

اي ان y_i . x_{i1} . x_{i2} : تمثل انحرافات المتغير المعتمد والمتغيرات المستقلة عن اوساطها الحسابية. بشكل عام ول (k) من المتغيرات المستقلة وحجم عينة (n) يمكن التعبير عن النموذج في العلاقة (7) بصيغة المصفوفات بالشكل الاتي:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

وبشكل اكثر اختصارا $y = xb + e$

حيث ان

y : متجه عمودي ابعاده $(n*1)$ لانحرافات قيم المتغير المعتمد عن وسطه الحسابي.

x : مصفوفة لانحرافات المتغيرات المستقلة من الرتبة (n*k) . حيث انها لا تتضمن العمود الاول الذي يمثل الحد الثابت. حيث يمكن استخراج الحد الثابت b_0 من خارج المصفوفة وفق الصيغة الاتية:

$$b_0 = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2$$
$$= \bar{Y} - (b_1\bar{X}_1 + b_2\bar{X}_2)$$

b: متجه عمودي لمعالم النموذج من الرتبة $(K * 1)$ ويلاحظ ان هذا المتجه قد حذف منه العنصر الاول والمتمثل بالحد الثابت (b_0) .

e: متجه عمودي يحتوي على الاخطاء العشوائية (البواقي) من الرتبة $(\text{n} * 1)$.

ويمكن التوصل الى صيغة لتقدير معالم نموذج الانحدار المتعدد بأسلوب الانحرافات باتباع الخطوات الاتية:

يمكن اعادة كتابة العلاقة (7) على النحو الاتي:

$$e_i = y_i - b_1x_{i1} - b_2x_{i2}$$

ولما كانت أفضل طريقة للحصول على أصغر قيمة ممكنة للانحرافات تتم بواسطة تربيعها وجعل مجموع مربعاتها أصغر ما يمكن. وبأخذ المشتقة الجزئية لها بالنسبة لكل من b_2 و b_1 ومساواتها للصفر وتبسيطها نحصل على المعادلات الطبيعية الاتية :

$$\Sigma e_i^2 = \Sigma(y_i - b_1x_{i1} - b_2x_{i2})^2$$

$$\frac{\partial \sum e_i^2}{\partial b_1} = 2\,\Sigma\,(y_i - b_1x_{i1} - b_2x_{i2})(-\,x_{i1}) = 0$$

$$-2\Sigma\, x_{i1}\,(y_i - b_1x_{i1} - b_2x_{i2}) = 0$$

وبالقسمة على (-2) وفك القوس والتبسيط نحصل على ان

$$\Sigma x_{i1}\, y_i = b_1\Sigma x_{i1}^2 + b_2\Sigma x_{i1}\, x_{i2} \dots\dots\dots \quad (8)$$

$$\frac{\partial \sum e_i^2}{\partial b_2} = 2\Sigma(y_i - b_1x_{i1} - b_2x_{i2})(-x_{i2}) = 0$$

$$-2\,\Sigma x_{i2}\,(y_i - b_1x_{i1} - b_2x_{i2}) = 0$$

وبالقسمة على (-2) وفك القوس والتبسيط نحصل على ان

$$\Sigma x_{i2}\, y_i = b_1\Sigma x_{i1}\, x_{i2} + b_2\Sigma x_{i2}^2 \dots\dots\dots \quad (9)$$

ويمكن كتابة المعادلتين الطبيعيتين في العلاقات (8) و (9) بهيئة المصفوفات على الشكل الاتي:

$$\begin{bmatrix} \Sigma x_{i1}\, y_i \\ \Sigma x_{i2}\, y_i \end{bmatrix} = \begin{bmatrix} \Sigma x_{i1}^2 & \Sigma x_{i1}\, x_{i2} \\ \Sigma x_{i1}\, x_{i2} & \Sigma x_{i2}^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

والذي يمكن التعبير عنه بالشكل الاتي:

$$\acute{x}\mathrm{y} = (\acute{x}x)\underline{b}$$

ولذلك فان تقدير المعالم لنموذج الانحدار الخطي المتعدد باستخدام المصفوفات وبطريقة الانحرافات سيأخذ الصيغة الاتية:

$$\boldsymbol{\underline{b} = (\acute{x}x)^{-1}\, \acute{x}\mathrm{y} \dots\dots\dots \quad (10)}$$

وبعد احتساب المتجه $(\acute{x}\mathrm{y})$ ومحدد المصفوفة $(\acute{x}x)$ اي $|\acute{x}x|$ الذي ينبغي ان لا يساوي صفرا. نجد معكوس المصفوفة حيث ان:

$$(\acute{x}x)^{-1} = \frac{adj(\acute{x}x)}{|\acute{x}x|}$$

ثم تطبيق القانون في الصيغة (10) اعلاه. اما b_0 فيمكن حسابه من الصيغة الاتية:

$$b_0 = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2$$

هذا ويمكن حساب (استخراج) القيم بالانحرافات من مجاميع البيانات الاصلية وفق الصيغ الاتية:

$$\sum y_i^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

$$\sum x_1^2 = \sum X_1^2 - \frac{(\sum X_i)^2}{n}$$

$$\sum x_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n}$$

$$\sum x_1 y = \sum X_1 Y - \frac{(\sum X_1)(\sum Y)}{n}$$

$$\sum x_2 y = \sum X_2 Y - \frac{(\sum X_2)(\sum Y)}{n}$$

$$\sum x_1 x_2 = \sum X_1 X_2 - \frac{(\sum X_1)(\sum X_2)}{n}$$

<u>مثال(1−4) :−</u> الجدول الاتي يتضمن البيانات الخاصة بالاستيرادات لمتغير تابع (y) والدخل القومي لمتغير مستقل اول (X_1) واسعار الاستيرادات لمتغير مستقل ثاني (X_2) في احدى الدول للفترة (1985−1993). والمطلوب استخدام طريقة الانحرافات لتقدير معالم تلك العلاقة.

السنوات	Y_i	X_1	X_2	X_1Y	X_2Y	X_1X_2	X_1^2	X_2^2	Y^2
1985	100	100	100	10000	10000	10000	10000	10000	10000
1986	106	104	99	11024	10494	10292	10816	9801	11236
1987	107	106	110	11342	11770	11660	11236	12100	11449
1988	120	111	126	13320	15120	13986	12321	15876	14400
1989	110	111	113	12210	12430	12543	12321	12769	12100
1990	116	115	103	13340	11948	11845	13225	10609	13456
1991	124	120	102	14880	12648	12240	14400	10404	15376
1992	133	124	103	16492	13699	12772	15376	10609	17689
1993	137	126	98	17262	13426	12348	15876	9604	18769
N=9	$\sum Y = 1053$	$\sum X_1 = 1017$	$\sum X_2 = 954$	$\sum X_1Y = 119870$	$\sum X_2Y = 111535$	$\sum X_1X_2$ 107690	$\sum X_1^2$ =115571	$\sum X_2^2 = 101772$	$\sum Y^2 = 124475$

العلاقة المطلوب تقديرها هي:

$$Y_i = b_0 + b_1X_1 + b_2X_2 + e_i$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{1053}{9} = 117$$

$$\bar{X}_1 = \frac{\sum X_1}{n} = \frac{1017}{9} = 113$$

$$\bar{X}_2 = \frac{\sum X_2}{n} = \frac{954}{9} = 106$$

ولإيجاد مجاميع القيم بدلالة الانحرافات فان:

$$\sum x_1 y = \sum X_1 Y - \frac{(\sum X_1)(\sum Y)}{n}$$

$$= \sum X_1 Y - n\bar{X}_1\bar{Y}$$

$$= 119870 - \frac{(1017)(1053)}{9} = 881$$

$$\sum x_2 y = \sum X_2 Y - \frac{(\sum X_2)(\sum Y)}{n}$$

$$= 111535 - \frac{(954)(1053)}{9} = -83$$

$$\sum x_1 x_2 = \sum X_1 X_2 - \frac{(\sum X_1)(\sum X_2)}{n}$$

$$= 107690 - \frac{(1017)(954)}{9} = -112$$

$$\sum x_1^2 = \sum X_1^2 - \frac{(\sum X_i)^2}{n}$$

$$= \sum X_1^2 - n\bar{X}_1^2$$

$$= 115571 - \frac{(1017)^2}{9} = 650$$

$$\sum x_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n}$$

$$= 101772 - \frac{(954)^2}{9} = 648$$

$$\sum y_i^2 = \sum Y^2 - \frac{(\sum Y)^2}{n} = 124475 - \frac{(1053)^2}{9} = 1274$$

$$\therefore \underline{b} = (\acute{x}x)^{-1}\,\acute{x}y$$

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum x_1^2 & \sum x_1x_2 \\ \sum x_1x_2 & \sum x_2^2 \end{bmatrix} \begin{bmatrix} \sum x_1y \\ \sum x_2y \end{bmatrix}$$

أي ان :

وحيث ان $$(\acute{x}x)^{-1} = \frac{adj(\acute{x}x)}{|\acute{x}x|}$$

فان $$(\acute{x}x) = \begin{bmatrix} 650 & -112 \\ -112 & 648 \end{bmatrix}$$

وان : $$|\acute{x}x| = \begin{bmatrix} 650 & -112 \\ -112 & 648 \end{bmatrix} = (650)(648) - (-112)(-112)$$

$$= 408656$$

وان :

$$(\acute{x}x)^{-1} = \frac{1}{408656}\begin{bmatrix} 648 & 112 \\ 112 & 650 \end{bmatrix}$$

$$= \begin{bmatrix} 0.001585685 & 0.000274069 \\ 0.000274069 & 0.001590579 \end{bmatrix}$$

وان :

$$\acute{x}y = \begin{bmatrix} \sum x_1y \\ \sum x_2y \end{bmatrix} = \begin{bmatrix} 881 \\ -83 \end{bmatrix}$$

$$\therefore \underline{b} = \begin{bmatrix} 0.001585685 & 0.000274069 \\ 0.000274069 & 0.001590579 \end{bmatrix} \begin{bmatrix} 881 \\ -83 \end{bmatrix}$$

$$\therefore\ \underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1.374240758 \\ 0.109436732 \end{bmatrix}$$

$$b_0 = \overline{Y} - b_1\overline{X_1} - b_2\overline{X_2}$$

$$= 117 - (1.374240758)(113) - (0.109436732)(106)$$

$$= 117 - 155.2892057 - 11.60029359 = -49.88949929$$

لذلك فان العلاقة المقدرة تأخذ الشكل

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2$$

$$\therefore\ \hat{Y}_i = -49.88949929 + 1.374240758\,X_1 + 0.10943673\,X_2$$

ويستدل من المعادلة التقديرية اعلاه الى وجود علاقة طردية بين الاستيرادات (Y) والمتغير المستقل X_1 الذي يمثل الدخل القومي فكل زيادة في X_1 بمقدار وحدة واحدة تزداد Y بمقدار (1.374) وحده مع ثبات X_2. كما تشير هذه المعادلة الى وجود علاقة طردية بين الاستيرادات (Y) والمتغير المستقل X_2 الذي يمثل السعر . فبزيادة X_2 بمقدار وحدة واحدة تزداد الاستيرادات بمقدار (0.109) وحده مع ثبات أثر X_1 وهذا مخالف لمنطق النظرية الاقتصادية والتي تشير الى وجود علاقة عكسية بين السعر والاستيرادات.

اختبار فرضيات النموذج الخطي المتعدد :-

يهدف هذا الجزء الى اجراء الاختبارات الاتية لنموذج الانحدار الخطي المتعدد .

1 - اختبار معنوية المعالم (اختبار t) :-

يستخدم اختبار (t) لتقييم معنوية تأثير المتغيرات المستقلة بصورة منفردة $(X_1 . X_1 \ldots\ldots\ldots . X_k)$ في المتغير التابع (Y) في نموذج الانحدار المتعدد. والذي يعتمد على نوعين من الفرضيات هما :

فرضية العدم $\qquad \boldsymbol{H_O : \beta_1 = \beta_2 = \beta_3 = \cdots \ldots\ldots\ldots = \beta_k = 0}$

الفرضية البديلة $\qquad \boldsymbol{H_1 : \beta_1 \neq \beta_2 \neq \beta_3 \neq \cdots \ldots\ldots\ldots \neq \beta_k \neq 0}$

وبعد احتساب قيمة (t) تقارن مع قيمتها الجدولية لتحديد قبول او رفض فرضية العدم ومن ثم تقييم معنوية معلمات النموذج المقدر. والصيغة الرياضية لهذا الاختبار تحسب كالاتي :

أ- بالنسبة الى (b_1) :- فان

$$t_{b1} = \frac{b_1}{S_{b_1}}$$

حيث ان :

$$S_{b1} = \sqrt{S_{b_1}^2} = \sqrt{var\ (b_1)}$$

وأن :

$$var - cov\ (\underline{b}) = S_e^2\ (X^{-}\mathrm{X})^{-1}$$

$$S_e^2 = \frac{e^{-}e}{n-k-1} = \frac{Y^{-}Y - b^{-}X^{-}Y}{n-k-1} = \frac{\sum y_i^2 - b_1 \sum x_1 y_i - b_2 \sum x_2 y_i}{n-k-1}$$

ب- بالنسبة الى b_2 فان:

$$t_{b_2} = \frac{b_2}{S_{b_2}}$$

حيث ان :

$$S_{b2} = \sqrt{S_{b_2}^2} = \sqrt{var(b_2)}$$

مثال:- المطلوب اختبار معنوية المعلمات b_1 و b_2 للبيانات في المثال السابق $(1-4)$. والذي يمثل العلاقة بين الاستيرادات والدخل القومي واسعار الاستيرادات.

الحل:- أ- بالنسبة الى b_1 :

$$S_e^2 = \frac{e^{-}e}{n-k-1} = \frac{Y^{-}Y - b^{-}X^{-}Y}{n-k-1} = \frac{\sum y_i^2 - b_1 \sum x_1 y - b_2 \sum x_2 y}{n-k-1}$$

$$S_e^2 = \frac{1274 - (1.37424)(881) - (0.1094367)(-83)}{9-2-1} = \frac{72.3771407}{6}$$

$$\boldsymbol{S_e^2 = 12.06285679}$$

$$\therefore \boldsymbol{var - cov(\underline{b}) = S_e^2 (X^{-}X)^{-1}}$$

$$= 12.06285679 \begin{bmatrix} 0.001585685 & 0.000274069 \\ 0.000274069 & 0.001590574 \end{bmatrix}$$

$$= \begin{bmatrix} 0.019127891 & 0.003306055 \\ 0.003306055 & 0.019186926 \end{bmatrix}$$

$$\therefore var(b_1) = S_{b_1}^2 = 0.019127891$$

$$\therefore S_{b_1} = \sqrt{S_{b_1}^2} = \sqrt{0.019127891} = 0.138303618$$

$$t_{b_1} = \frac{b_1}{S_{b_1}} = \frac{1.374240758}{0.138303618} = 9.936404975$$

وبما ان قيمة t_{b_1} المحسوبة والبالغة (9.936) اكبر من قيمتها الجدولية عند درجة حرية (6) ومستوى معنوية (5%) والبالغة (2.45) لذلك ترفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 اي ان (b_1) معنوي.

ب- بالنسبة الى b_2 :- فان

$$var(b_2) = S_{b_2}^2 = 0.019186926$$

$$\therefore S_{b_2} = \sqrt{S_{b_2}^2} = \sqrt{0.019186926} = 0.138516879$$

$$\therefore t_{b_2} = \frac{b_2}{S_{b_2}} = \frac{0.109436732}{0.138516879} = 0.7900606$$

وبما ان قيمة t_{b_2} المحسوبة والبالغة (0.790) هي اقل من قيمة (t) الجدولية عند مستوى معنوية (5%) ودرجة حرية (6) والبالغة (2.45) عليه تقبل فرضية العدم H_0 . اي عدم معنوية المعلمة المقدرة (b_2) .

2- **معامل التحديد المتعدد** : (R^2) Multiple Coefficient Of Determination

ان معامل التحديد (R^2) يعتبر مؤشر اساسي في تقييم مدى معنوية العلاقة المفترضة بين المتغير المعتمد (y) والمتغيرات المستقلة (X_k) حيث ان (k=1.2....K). بعبارة اخرى هو مقياس يوضح نسبة مساهمة المتغيرات المستقلة في تفسير التغير الحاصل في المتغير التابع.

ان معادلة الانحرافات الكلية يمكن ان تكتب على الشكل الاتي:

الانحرافات الكلية= الانحرافات الموضحة من قبل خط الانحدار + الانحرافات غير الموضحة أي

$$\acute{y}y = \acute{b}\acute{x}y + \acute{e}e$$

حيث ان:

$\acute{y}y$: الانحرافات الكلية ويرمز لها بالرمز (Tss) .

$\acute{b}\acute{x}y$: الانحرافات الموضحة من قبل خط الانحدار (Rss) .

$\acute{e}e$: الانحرافات غير الموضحة من قبل خط الانحدار (Ess) .

وبما ان معامل التحديد عبارة عن نسبة الانحرافات الموضحة من قبل خط الانحدار الى الانحرافات الكلية . فانه يمثل نسبة مجموع مربعات التغير في المتغيرات المستقلة الى مجموع المربعات الكلية: اي ان

$$R^2 = \frac{\text{Rss}}{Tss} = \frac{\acute{b}\acute{x}y}{\acute{y}y} = \frac{\acute{b}\acute{x}y}{\sum y^2}$$

كما ان:

$$R^2 = 1 - \frac{\acute{e}e}{\acute{Y}Y - n\,\bar{Y}^2}$$

او :

$$R^2 = \frac{b_1 \sum x_1\, y + \ b_2\, \sum x_2\, y}{\sum y^2}$$

ان اضافة متغيرات مستقلة جديدة الى المعادلة يؤدي الى رفع قيمة (R^2) وذلك لثبات قيمة المقام وتغير قيمة البسط بمقدار(bxy) غير ان الاستمرارية بإضافة المتغيرات المستقلة سيؤدي الى انخفاض درجات الحرية (n–k–1). مما يتطلب استخراج معامل التحديد المصحح او المعدل $\bar{R}^2$ وفق الصيغة الاتية :

$$\bar{R}^2 = 1 - \left[(1 - R^2)\ \frac{n-1}{n-k-1}\right] \ \dots\dots\dots\ (11)$$

مثال: باستخدام بيانات المثال السابق (1−4) المطلوب قياس نسبة مساهمة المتغيرات المستقلة X_1 و X_2 في التغير الحاصل في (y_i). بمعنى اخر حساب معامل التحديد R^2 او معامل التحديد المعدل المصحح $\bar{R}^2$.

الحل :

$$R^2 = \frac{\acute{b}\acute{x}y}{\sum y^2} = \frac{[1.374240758 \quad 0.109436732] \begin{bmatrix} 881 \\ -83 \end{bmatrix}}{1274}$$

$$= \frac{1210.706108 - 9.083248756}{1274} = 0.943189$$

$$\therefore \; R^2 = \%94.31$$

او يحسب وفق الصيغة (10) أعلاه.

$$R^2 = \frac{b_1 \sum x_1\, y + \; b_2 \; \sum x_2\, y}{\sum y_i^2}$$

$$= \frac{(1.3742407)(881) + (0.1094367)(-83)}{1274} = 0.943189$$

$$\therefore \; R^2 = \%94.31$$

اما معامل التحديد المصحح $\bar{R}^2$ فيحسب وفق الصيغة الاتية:

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

$$= 1 - \left[\frac{(1 - 0.943189)(9-1)}{9-2-1}\right] = 1 - 0.07574 = 0.92425 = 92.4\ \%$$

وهذا يعني ان المتغيرين المستقلين x_2 و X_1 يفسران حوالي 92.4% من التغير الحاصل في المتغير التابع (y) وان النسبة الباقية والبالغة 7.6% تمثل تأثير متغيرات اخرى غير مضمنة في المعادلة (الدالة).

3- **إختبار** – F :   F-statistics-Test

يهدف هذا الاختبار ايضا الى معرفة مدى معنوية العلاقة الخطية المفترضة بين المتغيرات المستقلة $(X_1.X_2.\dots\dots.X_k)$ على المتغير التابع (y). وكما هو الحال في الانحدار البسيط فانه يعتمد على نوعين من الفرضيات هي:

فرضية العدم (H_0): وتنص على عدم وجود علاقة معنوية بين المتغيرات المستقلة والمتغير التابع اي ان

$$H_0 = b_1 = b_2 = \cdots = b_k = 0$$

الفرضية البديلة H_1 : وتنص على وجود علاقة معنوية بين المتغيرات المستقلة والمتغير التابع

أي ان :

$$H_1: b_1 \neq b_2 \neq \cdots \neq b_k \neq 0$$

والصيغة الرياضية لهذا الاختبار هي:

$$F = \frac{\acute{b}\acute{x}y/k}{\acute{e}e\,/(n-k-1)}$$

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

ثم تقارن مع قيمتها الجدولية بدرجة حرية $(n-k-1.k)$ للبسط والمقام ولمستوى معنوية معين.

مثال: اختبر معنوية العلاقة بين الاستيرادات والدخل القومي واسعار الاستيرادات لبيانات المثال السابق (1−4) .

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{0.943189057/2}{(1-0.943189057)/(9-2-1)}$$

$$= \frac{0.471594528}{0.00946849} = 49.8067$$

وبما ان قيمة (F) المحسوبة والبالغة (49.806) اكبر من قيمة (F) الجدولية عند مستوى معنوية (5%) ودرجة حرية (6.2) للبسط و المقام والبالغة (5.14). عليه ترفض فرضية العدم (H_0). وتقبل الفرضية البديلة (H_1) التي تنص على معنوية العلاقة المقدرة . بعبارة اخرى. ان هناك على الاقل تأثير لأحد المتغيرين $X_1.X_2$ على المتغير التابع Y ؟

4- جدول تحليل التباين:- ANOVA

لبيان اثر المتغيرات المستقلة ول(k) من المتغيرات على المتغير التابع يمكن بناء جدول تحليل التباين على الشكل الاتي:

"جدول تحليل التباين (ANOVA)"

مصدر التباين s.of.v	مجموع المربعات S.S	درجات الحرية d.f	متوسط مجموع المربعات M.S.S	F– test
الانحرافات الموضحة من قبل خط الانحدار $X_1.X_2...X_K$	$\acute{b}\acute{x}y$ $(SSR) = \acute{Y}YR^2$	K	$\frac{\acute{b}\acute{x}y}{k}$	$F = \frac{b\acute{}x\acute{}y/k}{e\acute{}e/(n-k-1)}$
الانحرافات غير الموضحة Residual	$\acute{e}e = (1-R^2)\acute{Y}Y$ $= SSE$	n−k−1	$\frac{\acute{e}e}{n-k-1}$	
الانحرافات الكلية Total Variation	$\acute{y}y$ = SST	n−1		

وعلى هذا الاساس فان جدول تحليل التباين لبيانات المثال السابق$(4-1)$ للعلاقة بين Y و $X_1 . X_2$ سيأخذ الشكل الاتي:

"جدول تحليل التباين ANOVA"

s.of.v.	S.S	d.f	M.S.S	F
الانحرافات الموضحة من قبل X_1 و X_2	$\acute{b}\acute{x}y$ 1201.622859	2=k	$\acute{b}\acute{x}y/k$ 600.8114295	$F = \frac{600.8114295}{12.06285679}$ $= 49.8067$
الانحرافات غير الموضحة (البواقي)	$\acute{e}e$ 72.37714076	n−k−1 9−2−1 =6	$\acute{e}e\ /n-k-1$ 12.06285679	
الانحرافات الكلية	1274	8		

يتضح من اختبار F معنوية العلاقة المقدرة بعبارة اخرى ان هناك على الاقل تأثير لأحد المتغيرين $X_1 . X_2$على المتغير التابع Y .

ولمعرفة تأثير كل متغير مستقل في المتغير التابع بصورة منفردة. فإننا نختبر تأثير المتغير X_1 بصورة مستقلة في Y كمرحلة اولى. وفي المرحلة الثانية نختبر تأثير المتغير المستقل X_2 بصورة مستقلة في Y وبالرجوع الى بيانات المثال السابق $(4-1)$ فإننا نختبر ما يلي:

<u>أولا</u> :- تأثير متغير الدخل القومي (X_1) في الاستيرادات (Y) :

لغرض اختبار التأثير المستقل للدخل القومي (X_1) في دالة الاستيرادات (Y) يجب معرفة مقدار الزيادة المتحققة في قيمة مجموع مربعات الانحرافات الموضحة من قبل خط انحدار المتغير التابع (Y) على المتغير المستقل (X_2) نتيجة اضافة المتغير X_1 الى الدالة . ويتم ذلك بافتراض نموذج يتضمن المتغير X_2 اي ان:

$$Y_i = \beta_0 + \beta_2 X_2 + u_i$$

والصيغة التقديرية لهذا النموذج هي:

$$\hat{Y}_i = b_0 + b_2 X_2 + e_i$$

$$\therefore\ b_2 = \frac{\sum x_2 y}{\sum x_2^2} = \frac{-83}{648} = -0.128086419$$

ومجموع مربعات الانحرافات الناتجة من النموذج المقدر اعلاه (الناتجة عن المتغير X_2) هي

$$b_2 \Sigma x_2 y = (-0.128086419)(-83) = 10.63117284$$

اما التأثير الذي يضيفه المتغير المستقل X_1 الى الدالة فهو عبارة عن الفرق بين مجموع مربعات الانحرافات الموضحة من النموذج الكلي الذي يحتوي X_1 و X_2 ومجموع مربعات الانحرافات الناتجة من المتغير المستقل X_2 فقط اي ان:

$$b_1 \Sigma x_1 y = \acute{b}\acute{x}y - b_2 \Sigma x_2 y$$

$$= 1201.622859 - 10.63117284$$

$$= 1190.991686$$

ولاختبار مدى معنوية هذا التأثير المضاف من قبل المتغير المستقل X_1 ننظم جدول تحليل التباين الاتي:

"تحليل التأثير المستقل للمتغير X_1 في النموذج"

S.O.V.	S.S	d.f	M.S.S	F.test
الانحرافات الموضحة من قبل X_2	10.63117284 $b_2 \Sigma x_2 y$	1		
الانحرافات الموضحة من قبل X_1	1190.991686 $b_1 \Sigma x_1 y$	1	1190.991686	
الانحرافات الموضحة من قبل X_1 و X_2	1201.622859 $\acute{b}\acute{x}y$	2		$F_2 = \frac{1190.991686}{12.06285679} = 98.7321417$
الانحرافات غير الموضحة (البواقي)	72.37714076 $\acute{e}e$	6	12.06285679	
الانحرافات الكلية	1274 $\acute{y}y$	8		

وبمقارنة قيمة (F) المحتسبة والبالغة (98.73) مع مثيلتها الجدولية عند مستوى معنوية (5%) ودرجة حرية (6.1) للبسط والمقام والبالغة (5.99) ويتضح منها انها اكبر من (F) الجدولية عليه ترفض الفرضية العدم $(H_0: b_1 = 0)$. وتقبل الفرضية البديلة $(H_1: b_1 \neq 0)$. مما يدل على وجود تأثير معنوي للمتغير المستقل (X_1) الذي يمثل الدخل القومي على المتغير التابع (Y) الذي يمثل الاستيرادات.

ثانيا:- تأثير السعر (X_2) في الاستيرادات (Y) :-

لبيان أثر المتغير المستقل X_2 في الدالة نفترض نموذجا يتضمن المتغير المستقل X_1 اي ان

$$Y_i = \beta_0 + \beta_1 X_1 + u_i$$

والصيغة التقديرية لهذا النموذج هي : $\hat{Y}_i = b_0 + b_1 x_1 + e_i$

ومنه فان

$$b_1 = \frac{\sum x_1 y}{\sum x_1^2} = \frac{881}{650} = 1.355384615$$

وان مجموع مربعات الانحرافات الناجمة عن المتغير المستقل X_1 فهي

$$b_1 \Sigma x_1 y = (1.355384615)(881) = 1194.093846$$

اما التأثير الذي يضيفه المتغير المستقل X_2 الى الدالة فهو:

$$b_2 \Sigma x_2 y = \acute{b}\acute{x}\, y - b_1 \Sigma x_1 y$$

$$= 1201.622859 - 1194.093846$$

$$= 7.529013$$

ولاختبار مدى معنوية هذا التأثير المضاف من قبل المتغير المستقل (X_2) ننظم جدول تحليل التباين الاتي:

"تحليل تأثير المستقل للمتغير (X_2) في النموذج"

s.o.v.	S.S	d.f	M.S.S	F.test
الانحرافات الموضحة من قبل X_1	$b_1\Sigma x_1 y$ 1194.093846	1		
الانحرافات الموضحة من قبل X_2	$b_2\Sigma x_2 y$ 7.529013	1	7.529013	
الموضحة من قبل X_1 و X_2	$\acute{b}\acute{x}\ y$ 1201.622859	2		$F_2 = \dfrac{7.529013}{12.06285679}$ $= 0.6241484$
الانحرافات غير الموضحة (البواقي)	$e^{-}e$ 72.37714076	6	12.06285679	
الانحرافات الكلية	$\Sigma y^2 = \acute{y}y$ 1274	8		

وبمقارنة قيمة (F_2) المحتسبة والبالغة (0.624) مع القيمة الجدولية المقابلة لها عند مستوى معنوية (5%) ودرجة حرية (6.1) والبالغة (5.99) والتي يتضح بأنها اقل من الجدولية . عليه تقبل فرضية العدم $(H_0 = b_2 = 0)$ مما يدل على ان المتغير المستقل (X_2) غير معنوي . اي لا يمارس تأثيرا معنويا على المتغير التابع (Y) .

ومما سبق نستنتج ان المتغير المستقل X_1 الذي يمثل الدخل القومي يمارس تأثيرا معنويا على المتغير التابع (Y) الذي يمثل الاستيرادات . في حين ان المتغير المستقل X_2 والذي يمثل اسعار الاستيرادات لا يمارس تأثيرا معنويا على (Y) ويجب استبعاده من النموذج (العلاقة) المدروس واعتماد النموذج الذي يحتوي على المتغير X_1 ذو التأثير المعنوي في المتغير (Y) اي ان النموذج سيأخذ الشكل الاتي:

$$Y_i = \beta_0 + \beta_1 X_1 + u_i$$

والذي يتم تقديره وتقييمه وفق أسلوب الانحدار الخطي البسيط وكالاتي:

$$b_1 = \frac{\sum x_1 y}{\sum x_1^2} = \frac{881}{650} = 1.355384615$$

$$b_0 = \bar{Y} - b_1 \bar{X}_1 = 117 - (1.3553846)(133)$$

$$= -36.1584615$$

ومنه فان

$$\hat{Y}_i = -36.1584615 + 1.3553846\, X_1$$

وتعني ان زيادة المتغير المستقل X_1 والذي يمثل الدخل القومي بمقدار وحدة واحدة سوف تؤدي الى زيادة في المتغير التابع (Y) بمقدار (1.355) وحدة .

$$R^2 = \frac{b_1 \sum x_1 y}{\sum y^2} = \frac{(1.3553846\ )(881)}{1274} = 0.937279$$

$$\therefore\ R^2 = \%93.7279$$

وان

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{(0.937279)/1}{(1-0.937279)/(9-1-1)}$$

$$= 104.60592$$

ولاختبار معنوية المعلمات b_1 و b_0 نحتاج الى البيانات الاتية:

Y_i	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$	e_i^2
100	99.38	0.62	0.3844
106	104.8015385	1.1984615	1.436309967
107	107.5123077	-0.5123077	0.262459179
120	114.2892308	5.7107692	32.61288486
110	114.2892308	-4.2982308	18.39750086
116	119.7107692	-3.7107692	13.76980806
124	126.4876923	-2.4876923	6.188612979
133	131.9092308	1.0907692	1.18977448
137	134.62	2.38	5.6644
1053	**1053**	**0**	**79.90615335**

$$S_e^2 = \frac{\sum e_i^2}{n-k-1} = \frac{79.90615335}{9-1-1} = 11.41516476$$

ولاختبار معنوية b_1 :- فان

$$S_{b_1}^2 = \frac{S_e^2}{\sum x_i^2} = \frac{11.41516476}{650} = 0.017561791$$

$$S_{b_1} = \sqrt{S_{b_1}^2} = \sqrt{0.017561791} = 0.132520911$$

$$t_{b_1} = \frac{b_1}{S_{b_1}} = \frac{1.355384615}{0.132520911} = 10.2277037$$

ولاختبار b_0 : فان

$$S_{b_0}^2 = S_e^2\left[\frac{1}{n} + \frac{\bar{X}_1^2}{\sum x_1^2}\right] = 11.41516476\left[\frac{1}{9} + \frac{(113)^2}{650}\right]$$

$$= 225.5148729$$

$$S_{b_0} = \sqrt{S_{b_0}^2} = \sqrt{225.5148729} = 15.01715262$$

$$\therefore\ t_{b_0} = \frac{b_0}{S_{b_0}} = \frac{-36.1584615}{15.01715262} = -2.407810749$$

وعليه فان الصيغة التقديرية للنموذج المدروس ستأخذ الشكل الاتي:

$$\hat{Y}_i = b_0 + b_1 X_i$$

$$\hat{Y}_i = -36.1584615 + 1.355384615\ X_i$$

$$t \qquad (-2.407) \qquad (10.227)$$

$$R^2 = \%\ 93.72 \qquad .F = 104.605 \qquad S_e^2 = 11.415$$

(تمارين)

1-4 : لدراسة نموذج الانحدار الخطي المتعدد للعلاقة بين الكمية المطلوبة من سلعة معينة (Y) وسعر السلعة (X_1) والدخل الفردي (X_2) .

$$Y_i = b_0 + b_1X_1 + b_2X_2 + e_i$$

وحصلت على البيانات الاتية :

$$(\acute{X}X)^{-1} = \begin{bmatrix} 1.5 & -1.2 & 1.8 \\ & 1 & -1.7 \\ & & 2 \end{bmatrix} . \quad \acute{X}Y = \begin{bmatrix} 42 \\ 30 \\ 10 \end{bmatrix} \ . \ S_e^2 = 10$$

والمطلوب:– أ– قدر معالم هذه الدالة $(b_2 \,.\, b_1 \,.\, b_0)$ مع التفسير الاقتصادي لها ؟

ب– احسب مصفوفة التباين والتباين المشترك لمعلمات الدالة أي $var - cov(\underline{b})$

ج– أوجد تباين المعلمة b_2 ؟

2– لدراسة دالة الانفاق على المواد الغذائية اخذت (6) أسر واعتمد لوغاريتم الانفاق على المواد الغذائية كمتغير معتمد (Y_i) وعلاقته بلوغاريتم سعر المواد الغذائية (X_{1i}) ولوغاريتم الدخل المتاح (X_{2i}) كمتغيرات مستقلة . حيث كانت النتائج كالاتي:

$$(\acute{X}X)^{-1} = \begin{bmatrix} 0.86 & -0.15 & -0.07 \\ & 0.32 & -0.45 \\ & & 0.55 \end{bmatrix} . \acute{X}Y = \begin{bmatrix} 9 \\ 12 \\ 16 \end{bmatrix}$$

والمطلوب:- قدر معلمات العلاقة بين (Y_i) والمتغيرين (X_{1i}) و (X_{2i}) مفسرا النتيجة اقتصاديا؟

3- ماذا تعني النتائج في الدالة المقدرة الاتية: -

$$\ln Y_t = 1.24 + 0.25 \ln L_t + 0.78 \ln k_t$$

حيث ان Y_t : الإنتاج . L_t :العمل. k_t : راس المال.

4- لنموذج الانحدار الخطي المتعدد الاتي:-

$$Y = X\beta + u$$

$$n \sim n\ (0.\sigma^2 \ln)\ .\ E(u_i u_j) = 0 \qquad \forall_i \neq J$$

المطلوب:-

أ- اثبت ان قيمة المعالم (b) المقدر بأسلوب (OLS) غير متحيز اي ان $E(b) = \beta$ ؟

ب- اشتق صيغة لتقدير مصفوفة التباين والتباين المشترك لموجه المعالم $(\underline{b})$ اي

$$var - cov(\underline{b}).$$

5- النتائج الاتية تم الحصول عليها من عينة تمثل (13) عائلة حيث (y) يمثل الانفاق الاسبوعي على الشؤون المنزلية . (X_1) يمثل الدخل العائلي . (X_2) يمثل عدد الاطفال في الاسرة.

$$\hat{Y}_i = 6.26 + 0.45X_1 - 0.38X_2$$

وان $\sum x_1^2 = 74$. $\sum x_2\, y = -28$. $\sum x_1\, y = 38$

$$\Sigma x_2^2 = 60 \quad .\Sigma x_1 x_2 = -12 \quad .\Sigma\, e_i^2 = 12.27$$

$$t(0.05) = 2.228$$

المطلوب:-

1- حساب مقدار ما تفسره المتغيرات X_1 وX_2 من التغير الحاصل في Y ؟

2- اختبر معنوية المعلمة b_1 ؟

3- قدر حدود الثقة 95% للمعلمة b_1 ؟

4- كم يكون الانفاق على الشؤون المنزلية عندما $X_1 = 300$. $X_2 = 4$ ؟

5- اوجد الانحراف المعياري ل (b_2) ؟

مثال: 4-2 : البيانات التالية تمثل متوسط انفاق الفرد العراقي Y_t ومتوسط دخله القابل للتصرف X_{1t} اضافة الى متوسط انفاقه للسنة السابقة $(X_{2t} = Y_{t-1})$ خلال الفترة (1964-1980) والبيانات مقاسة بالدينار وبالأسعار الثابتة .

المطلوب:1- تقدير معالم دالة الاستهلاك التالية $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$

ثم بيان اهم المؤشرات اللازمة لاعتماد الصيغة التقديرية لدالة الاستهلاك؟

2- اختبار مدى تأثير كل من متوسط دخل الفرد القابل للتصرف ومتوسط انفاقه للسنوات السابقة على متوسط انفاقه الحالي وعلى التوالي. مستخدماً مستوى دلالة قدره (5%) ؟

السنة	Y_t	X_{1t}	X_{2t}
1964	75.3	96.0	72.1
1965	85.0	103.4	75.3
1966	87.97	106.4	85.0
1967	82.0	105.7	87.97
1968	85.9	107.4	82.0
1969	81.4	101.8	85.9
1970	81.5	97.3	81.4
1971	84.9	95.2	81.5
1972	75.9	99.1	84.9
1973	57.5	94.2	75.9
1974	70.0	121.8	57.5
1975	127.5	151.7	70.0
1976	139.4	160.8	127.5
1977	148.0	162.7	139.4
1978	173.6	191.7	148.0
1979	174.6	237.95	173.6
1980	185.8	212.4	174.6

$$\Sigma X_{2t} = 1702.57 \quad . \quad \Sigma X_{1t}^2 = 330182.7025 \qquad n = 17$$

$$. \ \Sigma Y_t = 1816.2 \quad . \quad \Sigma X_t = 2245.55$$

$$\sum X_{2t}^2 = 192593.0409 \quad . \quad \sum Y_t^2 = 221916.2707$$

$$\sum X_{1t} Y_t = 269159.988 \quad . \sum X_{2t} Y_t = 204717.03$$

$$\sum X_{1t} X_{2t} = 249413.499$$

ومنها يمكن الحصول على قيم الانحرافات مباشرة على الشكل الاتي:

$$\sum x_{1t}^2 = \sum X_{1t}^2 - n\,\overline{X}_1^2 = 33565.36118$$

$$\sum x_{2t}^2 = \sum X_{2t}^2 - n\,\overline{X}_2^2 = 22078.65238$$

$$\sum y_t^2 = \sum Y_t^2 - n\,\overline{Y}^2 = 27867.05249$$

$$\sum x_{1t} y_t = \sum X_{1t} Y_t - n\,\overline{X}_1 \overline{Y} = 29246.74691$$

$$\sum x_{2t} y_t = \sum X_{2t} Y_t - n\,\overline{X}_2 \overline{Y} = 22815.45271$$

$$\sum x_{1t} x_{2t} = \sum X_{1t} X_{2t} - n\,\overline{X}_1 \overline{X}_2 = 24519.02468$$

اذ ان

$$\underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = (\acute{x}x)^{-1}\,\acute{x}y$$

وان

$$(\acute{x}x) = \begin{bmatrix} \sum x_{1t}^2 & \sum x_{1t}x_{2t} \\ \sum x_{1t}x_{2t} & \sum x_{2t}^2 \end{bmatrix} = \begin{bmatrix} 33565.36118 & 24519.02468 \\ 24519.02468 & 22078.65238 \end{bmatrix}$$

$$\acute{x}y = \begin{bmatrix} \sum x_{1t} y_t \\ \sum x_{2t} y_t \end{bmatrix} = \begin{bmatrix} 29246.74691 \\ 22815.45271 \end{bmatrix}$$

$$(\acute{x}x)^{-1} = \frac{adj(\acute{x}x)}{|\acute{x}x|} = \begin{bmatrix} 0.000157822 & -0.000175266 \\ -0.000175266 & 0.000239931 \end{bmatrix}$$

$$|\acute{x}x| = 741077941.5 - 601182571.3 = 139895370.2$$

$$\therefore \underline{b} = \begin{bmatrix} 0.000157822 & -0.000175266 \\ -0.000175266 & 0.000239931 \end{bmatrix} \begin{bmatrix} 29246.74691 \\ 22815.45271 \end{bmatrix}$$

$$\therefore \underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0.617007 \\ 0.348174 \end{bmatrix}$$

وان $$b_0 = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2 = -9.531802$$

لذلك فان الصيغة التقديرية لدالة الاستهلاك تأخذ الشكل

$$\hat{Y}_t = -9.531 + 0.617\, X_{1t} + 0.348\, X_{2t}$$

لتقييم معالم الصيغة التقديرية أعلاه . يستوجب حساب المؤشرات التالية:

$$S_e^2 = \frac{\acute{y}y - \acute{b}\acute{x}y}{n-k-1} = \frac{\sum y_t^2 - b_1\sum x_{1t}y_t - b_2\sum x_{2t}y_t}{n-k-1}$$

وبعد التعويض نحصل على ان

$$S_e^2 = 134.4308664$$

$$\therefore var - cov(\underline{b}) = S_e^2(\acute{x}x)^{-1}$$

$$= 134.4308664 \begin{bmatrix} 0.000157822 & -0.000175266 \\ -0.000175266 & 0.000239931 \end{bmatrix}$$

$$\therefore var - cov(\underline{b}) = \begin{bmatrix} 0.021216 & -0.023561 \\ -0.023561 & 0.032254 \end{bmatrix}$$

ومنه فان

$$\widehat{var}(b_1) = 0.021216 \quad : \quad \widehat{var}(b_2) = 0.032254$$

$$\widehat{cov}(b_1.b_2) = -0.023561$$

ولاختبار مدى دقة (معنوية) الميل الحدي للاستهلاك أي (b_1) فان الفرضية هي:

$$H_0 \ : \ \beta_1 = 0$$

$$H_1 \ : \ \beta_1 \neq 0$$

وان
$$t_{b_1} = \frac{b_1}{S_{b_1}} = \frac{0.617}{\sqrt{0.021216}} = 4.2376$$

وبمقارنة قيمة (t) المحسوبة والبالغة (4.23) مع نظيرتها الجدولية لدرجة حرية (14) ومستوى دلالة 5% حيث ان t(14;0.025)=2.145 وحيث ان قيمة (t) المحسوبة اكبر من قيمة (t) الجدولية لذلك ترفض فرضية العدم H_0 ونأخذ بالفرض البديل (H_1) والتي تعني بان متوسط الدخل الفردي يمارس تأثيرا على متوسط انفاق الفرد (Y_t) . ولاختبار مدى معنوية وتأثير نمط انفاق الفرد في السنة السابقة على متوسط انفاقه الحالي (b_2) نضع الفرضية الاتية:

$$H_0 : \ \beta_2 = 0$$

$$H_1 : \ \beta_2 \neq 0$$

$$t_{b_2} = \frac{b_2}{S_{b_2}} = \frac{0.348}{\sqrt{0.0323}} = 1.93656$$

وبما ان قيمة (t) المحسوبة هي اقل من قيمة (t) الجدولية (1.936<2.145) لذلك تقبل فرضية العدم (H_0) بمعنى ان نمط انفاق الفرد في السنة السابقة (t-1) لا يمارس تأثيرا على انفاقه في السنة الحالية (t) . ولاختبار مدى معنوية العلاقة الخطية المفترضة لتقدير معالم دالة الاستهلاك وبيان اثر كل من (X_{1t}) و (X_{2t})على متوسط انفاق الفرد نستخدم اختبار (F) والذي بدوره يتطلب حساب المؤشرات الاتية:

$$\acute{y}y = \Sigma y_t^2 = 27867.05249$$

$$R^2 = \frac{\acute{b}\acute{x}y}{\acute{y}y} = \frac{[0.617 \quad 0.348]\begin{bmatrix}29246.74691\\22815.45271\end{bmatrix}}{27867.05249}$$

$$\therefore\ R^2 = 0.93246 = \%93.246$$

$$\acute{b}\acute{x}y = R^2 \Sigma y_t^2 = 25985.02036$$

$$\acute{e}e = \Sigma e_t^2 = (1 - R^2)\Sigma y_t^2 = 1882.032127$$

لذلك فان جدول تحليل التباين سيأخذ الشكل الاتي

"جدول تحليل التباين ANOVA"

s.o.v	s.s.	d.f	M.s.s	F.test
الانحرافات الموضحة من قبل X_{1t} و X_{2t}	25985.02036	2	12992.51018	$F = \frac{12992.51018}{134.4308662} = 96.648$
الانحرافات غير الموضحة (البواقي)	1882.032127	14	134.4308662	
الانحرافات الكلية	27867.05248	16		

وبمقارنة قيمة (F) المحسوبة مع نظيرتها الجدولية لدرجة حرية (2) و (14) ومستوى دلالة (0.05) .

$$F(k.n-k-1.x) = F(2.14.0.05) = 3.74$$

اذ ان $96.648 > 3.74$ لذلك ترفض فرضية العدم وتقبل الفرضية البديلة (H_1) حيث ان

$$H_1 = \beta_1 \neq \beta_2 \neq 0$$

مما يعني معنوية العلاقة الخطية المفترضة لتقدير معالم دالة الاستهلاك. وان هناك على الأقل تأثير من احد المتغيرين المستقلين (X_{1t}) و (X_{2t}) على المتغير المعتمد (Y_t).

**ولغرض الوقوف على تأثير كل من متوسط الدخل الفردي القابل للتصرف ومتوسط انفاق الفرد للسنة السابقة على متوسط انفاق الفرد الحالي. نتابع الاختبار فنضع نموذجا يتضمن متوسط انفاق الفرد للسنة السابقة (X_{2t}). أي ان

$$Y_t = b_0 + b_2 X_{2t}$$

$$b_2 = \frac{\sum x_{2t} y_t}{\sum x_{2t}^2} = \frac{22815.45271}{22078.65238} = 1.0333716$$

ومنه نحسب الانحرافات الموضحة بالشكل التالي:

$$b_2 \Sigma x_{2t} y_t = 23576.84126$$

وبالتالي يمكن حساب التأثير الذي يضيفه المتغير (X_{1t}) متوسط دخل الفرد

$$b_1 \Sigma x_{1t} y_t = \acute{b}\acute{x}y - b_2 \Sigma x_{2t} y_t = R^2 \; \Sigma y_t^2 - b_2 \Sigma x_{2t} y_t = 2408.1791$$

اما اختبار مدى معنوية التأثير الذي يضيفه المتغير المستقل (X_{1t}) في العلاقة الخطية المدروسة فيمكن بيانه من خلال جدول تحليل التباين الاتي:

"جدول تحليل التباين ANOVA"

s.o.v.	S.S	d.f	MS.S	F.test
الانحرافات الموضحة من قبل X_{2t}	23576.84126	1	23576.84126	
الانحرافات الموضحة من قبل X_{1t}	2408.1791	1	2408.1791	$F_1 = \frac{2408.1791}{134.4308662} = 17.194$
الموضحة من قبل X_{1t} و X_{2t}	25985.02036	2	12992.51018	
الانحرافات غير الموضحة (البواقي)	1882.032127	14	134.4308662	
الانحرافات الكلية	27867.05248	16		

وبمقارنة قيمة (F_1) العملية مع نظيرتها الجدولية حيث ان

$$F(1.14.0.05) = 4.60$$

وحيث ان $17.194 > 4.60$ وهذا يعني ان المتغير المستقل (X_{1t}) له تأثير معنوي على متوسط انفاق الفرد .

**ولاختبار مدى معنوية المتغير المستقل (X_{2t}) أي متوسط انفاق الفرد للسنة السابقة في تقدير دالة الاستهلاك. ففي مثل هذه الحالة يجب وضع نموذجا يتضمن متوسط دخل الفرد القابل للتصرف (X_{1t}) وكالاتي:

$$\hat{Y}_t = a_0 + b_1 X_{1t}$$

$$b_1 = \frac{\sum X_{1t}\, y_t}{\sum x_{1t}^2} = \frac{29246.74691}{33565.36118} = 0.871337172$$

والانحرافات الموضحة لهذا النموذج البسيط هي

$$b_1 \Sigma x_{1t}\, y_t = 25483.77965$$

اما التأثير الذي يضيفه المتغير المستقل (X_{2t}) فيحسب كالاتي:

$$b_2 \Sigma x_{2t}\, y_t = \acute{b}\acute{x}y - b_1 \Sigma x_{1t}\, y_t = R^2 \Sigma y_t^2 - b_1 \Sigma x_{1t}\, y_t = 501.24262$$

ولاختبار مدى معنوية هذا التأثير المضاف من قبل المتغير (X_{2t}) تكون جدول تحليل التباين كالاتي

"جدول تحليل التباين ANOVA"

s.o.v.	S.S	d.f	M.S.S	F.test
الانحرافات الموضحة من قبل X_{1t}	25483.77965	1	25483.77965	$F_2 = \frac{501.24262}{134.4308662} = 3.729$
الانحرافات الموضحة من قبل X_{2t}	501.24262	1	501.24262	
الانحرافات الموضحة من قبل X_{1t} و X_{2t}	25985.02036	2	12992.51018	
الانحرافات غير الموضحة (البواقي)	1882.032127	14	134.4308662	
الانحرافات الكلية	27867.05248	16		

وبمقارنة قيمة(F_2) العملية مع قيمتها الجدولية لدرجة حرية (1، 14) وعند مستوى معنوية 5% والبالغة (4.60) وحيث ان $4.60 > 3.729$ لذلك تقبل فرضية العدم (H_0) مما يعني ان المتغير المستقل

(X_{2t}) لا يمارس تأثيرا على المتغير المعتمد (Y_t) الذي يمثل متوسط انفاق الفرد العراقي بالتالي يمكن حذفه من النموذج .

تمرين واجب :-

باحث اقتصادي قدر معالم العلاقة الخطية بين المتغير المعتمد (Y_i) والمتغيرات المستقلة (X_1) و (X_2) مستخدماً عينة ذات حجم (20) مشاهدة وحصل على النتائج الاتية :

$$\hat{Y}_i = 6.96 + 0.335X_1 + 1.466X_2$$

وكانت نتائج العمليات الحسابية حول نقطة المتوسط (بالانحرافات) كالاتي :

$$\Sigma x_1^2 = 64106.67 \quad ; \quad \Sigma x_2^2 = 297 \quad ; \quad \Sigma y_i^2 = 12799.83$$

$$\Sigma y_i x_1 = 22791.34 \quad ; \quad \Sigma y_i x_2 = 725.4$$

المطلوب : اختبار مدى معنوية تأثير المتغير (X_1) على (Y_i) وكذلك تأثير (X_2) على (Y_i) ؟

الفصل الخامس

مشكلة الارتباط الذاتي The Autocorrelation Problem

المقدمة: -

من جملة الافتراضات الأساسية التي يقوم عليها النموذج الخطي ، افتراض انعدام الارتباط بين قيم المتغير العشوائي u في السنة (t) وقيمته في السنوات السابقة (u_{t-1}، u_{t-2}،....) او اللاحقة u_{t+1}, u_{t+2} ... الخ أي ان قيم u تكون مستقلة عن بعضها البعض، ويعبر عن ذلك بمساواة التباين المشترك للأخطاء المتتالية بالصفر أي ان :

$$cov(u_t u_{t-s}) = 0 \qquad (t = 1,2,3,\dots n)$$

وتعني بالوقت نفسه أيضا عدم تأثير الظاهرة الاقتصادية المتحققة في السنة (t) على تلك التي ستتحقق في السنة $(t+1)$ ، غير ان الواقع الاقتصادي يشير الى عكس ذلك، اذ يوجد تأثير للظاهرة الاقتصادية المتحققة في السنة (t) على تلك التي ستتحقق في السنة $(t+1)$, $(t+2)$ او في السنوات $(t-1),(t-2)$....الخ

ولذلك تظهر مشكلة الارتباط الذاتي في اغلب الدراسات التي تأخذ شكل السلاسل الزمنية Time Series Data وكذلك في البحوث التي تعتمد على بيانات مقطعية Cross-Section Data وبخاصة البيانات المقطعية التي تأخذ شكل أوساط مجاميع (Grouping of observation) .

وفي حالة اعتماد الأخطاء العشوائية على بعضها البعض ينتفي الافتراض الخاص بانعدام الارتباط بينها فتظهر مشكلة تدعى مشكلة الارتباط الذاتي The Autocorrelation Problem أي ان :

$$cov(u_t . u_{t-s}) \neq 0$$

أولا- أسباب (مسببات الارتباط الذاتي

يظهر الارتباط الذاتي للأسباب التي يمكن أن نوجزها بما يلي :-

1 -الاثار الممتدة لبيانات السلاسل الزمنية :-

ان بعض العوامل العشوائية الطارئة وغير المتكررة قد ينتج عنها ترابط في قيم العنصر العشوائي u_t . لأكثر من فترة زمنية واحدة ، فالحروب والفيضانات والزلازل تمتد أثارها وانعكاساتها على فعالية الاقتصاد لعدة سنوات متتالية ، مما يتسبب في حصول ارتباط ذاتي بين قيم u_i المتلاحقة .

2- حذف بعض المتغيرات المستقلة من العلاقة المدروسة لسبب أو لأخر:

مثل عدم توفر البيانات المناسبة عنها أو لغرض تبسيط هيكل النموذج وقد يكون من بين هذه المتغيرات المحذوفة متغير او اكثر مترابطة ذاتيا ، الأمر الذي يؤدي الى جعل العنصر العشوائي يتضمن المتغيرات المرتبطة ومن ثم فان u_t لا يعكس الخطأ العشوائي في النموذج فحسب انما يعكس أيضا المتغيرات المحذوفة .

3- معالجة البيانات :-

تجري على البيانات المنشورة أحيانا عمليات تشذيب وقد يتم تقدير قيم بعض المشاهدات اعتمادا على قيم مشاهدات أخرى. وحيث ان عمليات التشذيب والتقدير تعتمد في العادة على أخذ معدلات قيم المشاهدات المتتالية، مما يحدث علاقة بين أخطاء تلك المشاهدات وبالتالي التأثير على طبيعة توزيعها.

4- الصياغة غير الدقيقة للنموذج :-

بمعنى ان شكل العلاقة الدالية المستخدمة لا يتطابق مع الشكل الحقيقي لها ، فاذا افترضنا أن العلاقة خطية بين المتغيرين Y و X في حين ان العلاقة الحقيقية غير خطية فانه يمكن ان ينتج عن ذلك ترابط ذاتي في حد الخطأ العشوائي u .

ثانيا - تحليل الارتباط الذاتي :-

لتحليل الفكرة الأساسية للارتباط الذاتي يمكننا أخذ النموذج الخطي البسيط الاتي :

$$Y_t = \beta_0 + \beta_1 X_t + u_t \qquad \dots\dots\dots \quad (1)$$

اذ ان :

$$E(u_t) = 0 \qquad , \quad var(u_t) = E(u_t^2) = \sigma_u^2$$

ولكنه لا يستوفي فرض انعدام الارتباط الذاتي بين الأخطاء العشوائية أي ان توزيع الأخطاء العشوائية في النموذج (1) يتبع الارتباط الذاتي من الدرجة الأولى First Order Autoregressive Scheme أي ان

$$u_t = \mathcal{P} u_{t-1} + e_t \qquad \dots\dots\dots \quad (2)$$

حيث ان $\mathcal{P}$ او (rho) : معامل الارتباط الذاتي البسيط بين الأخطاء العشوائية وتنحصر قيمته بين $(+1$ و $-1)$ أي ان $-1 \leq \mathcal{P} \leq +1$. e_t : متغير عشوائي يتوزع توزيعا طبيعياً بوسط حسابي مقداره صفر $E(e_t) = 0$ وتباين ثابت مقداره σ_e^2 حيث ان $var(e_t) = E(e_t^2) = \sigma_e^2$ وانه مستقل عن e_{t-1} أي ان

$$cov(e_t e_{t-1}) = E(e_t e_{t-1}) = 0$$

وبعبارة عامة فان:

$$e_t \sim N\,(0\,,\sigma_e^2)$$

ومن المعادلة (2) يمكن الحصول على الاتي :

$$u_{t-1} = \mathcal{P}u_{t-2} + e_{t-1} \quad \dots\dots\dots\dots (3)$$

وبتعويضها في (2) نحصل على ان :

$$u_t = \mathcal{P}(\mathcal{P}u_{t-2} + e_{t-1}) + e_t$$

$$\therefore\ u_t = \mathcal{P}^2 u_{t-2} + \mathcal{P}e_{t-1} + e_t$$

وكذلك يكون لدينا :

$$u_{t-2} = \mathcal{P}u_{t-3} + e_{t-2}$$

لذلك فان :

$$u_t = \mathcal{P}^2(\mathcal{P}u_{t-3} + e_{t-2}) + \mathcal{P}e_{t-1} + e_t$$

$$\therefore\ u_t = \mathcal{P}^3 u_{t-3} + \mathcal{P}^2 e_{t-2} + \mathcal{P}e_{t-1} + e_t$$

وبالتعويض المتسلسل (المتتالي) للأخطاء الناتجة نحصل على ان

$$u_t = e_t + \mathcal{P}e_{t-1} + \mathcal{P}^2 e_{t-2} + \mathcal{P}^3 e_{t-3} + \cdots \dots\dots \quad (4)$$

ولـ r من الفترات فان العلاقة (4) ستأخذ الشكل الاتي:

$$u_t = \sum_{r=0}^{\infty} \mathcal{P}^r\ e_{t-r}$$

وبأخذ التباين للصيغة رقم (4) أعلاه حيث ان

$$var(u_t) = E(u_t^2) \qquad ; \qquad E(e_t^2) = \sigma_e^2$$

وبعد تربيع الطرفين واخذ التوقع والتبسيط نحصل على ان

$$\boldsymbol{\sigma_u^2 = \frac{\sigma_e^2}{1 - \mathcal{P}^2}} \quad \dots\dots\dots \quad \mathbf{(5)}$$

أما التباين المشترك cov بين الأخطاء العشوائية u_t و u_{t-1} فهو

$$E(u_t\, u_{t-1}) = \frac{\mathcal{P}\sigma_e^2}{1-\mathcal{P}^2} \qquad \dots\dots\dots.\ (6)$$

وبمقارنة العلاقة رقم (6) مع العلاقة رقم (5) السابقة نجد ان

$$E(u_t\, u_{t-1}) = \mathcal{P}\sigma_u^2 \qquad \dots\dots\dots\ (7)$$

والعلاقة رقم (7) أعلاه يمكن ان توضع بشكل عام ولـ s من الارتباط المتسلسل كالاتي:

$$E(u_t\, u_{t-s}) = \mathcal{P}^s\sigma_u^2 \qquad \neq 0$$

$$S\ = 0,1,2,3, \dots n-1$$

ففي حالة كون قيمة S=0 نحصل على ان :

$$E(u_t\, u_{t-s}) = E(u_t\, u_t) = (Eu_t^2) = \sigma_u^2$$

وفي حالة s=1 فان

$$E(u_t\, u_{t-1}) = \mathcal{P}\sigma_u^2$$

وفي حالة s=2 فان

$$\begin{array}{lcl} E(u_t\, u_{t-1}) = \mathcal{P}^2 & & \sigma_u^2 \\ \quad\vdots & & \vdots \\ \quad\vdots & & \vdots \\ E\left(u_t\, u_{t-(u-1)}\right) = \mathcal{P}^{n-1} & & \sigma_u^2 \end{array}$$

وهكذا عندما $S\ = n-1$

أي ان التباين المشترك لا يساوي صفراً .

وبجمع هذه الحدود في مصفوفة تباين الأخطاء في حالة النموذج الخطي العام فان :

$$E(u\acute{u}) = \begin{bmatrix} E(u_1^2) & E(u_1u_2) & E(u_1u_3)\dots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & E(u_2u_3)\dots & E(u_2u_n) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ E(u_nu_1) & E(u_nu_2) & E(u_nu_3)\dots & E(u_n^2) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_u^2 & \mathcal{P}\sigma_u^2 & \mathcal{P}^2\sigma_u^2 & \mathcal{P}^{n-1}\,\sigma_u^2 \\ \mathcal{P}\sigma_u^2 & \sigma_u^2 & \mathcal{P}\sigma_u^2 & \mathcal{P}^{n-2}\,\sigma_u^2 \\ \vdots & \vdots & \vdots & \vdots \\ \mathcal{P}^{n-1}\,\sigma_u^2 & \mathcal{P}^{n-2}\,\sigma_u^2 & \mathcal{P}^{n-3}\,\sigma_u^2 & \sigma_u^2 \end{bmatrix}$$

$$\therefore\ \ E(u\acute{u}) = \sigma_u^2 \begin{bmatrix} 1 & \mathcal{P} & \mathcal{P}^2 & \mathcal{P}^{n-1} \\ \mathcal{P} & 1 & \mathcal{P} & \mathcal{P}^{n-2} \\ \vdots & \vdots & \vdots & \vdots \\ \mathcal{P}^{n-1} & \mathcal{P}^{n-2} & \mathcal{P}^{n-3} & 1 \end{bmatrix}$$

$$= \sigma_u^2\ \boldsymbol{\Omega}$$

اذ أن : $\boldsymbol{\Omega}$: مصفوفة الارتباطات الذاتية .

ثالثا- طرق تقدير معامل الارتباط الذاتي $(\mathcal{P})$:

هناك عدة طرق لتقدير معامل الارتباط الذاتي من الدرجة الأولى $(\mathcal{P})$ نذكر منها ما يلي:

أ ' طريقة اختبار إحصاءة ديربن واتسن $(D-W)$:

يتم تقدير معامل الارتباط الذاتي $(\mathcal{P})$ من خلال طريقة اختبار ديربن واتسن للكشف عن وجود الارتباط الذاتي حيث ان:

$$D-W = 2\left(1-\hat{\mathcal{P}}\right)$$

وبعد إعادة ترتيب المعادلة نحصل على ان :

$$\frac{D-W}{2} = 1-\hat{\mathcal{P}}$$

$$\therefore\ \hat{\mathcal{P}} = 1-\frac{D-W}{2} \quad \dots\dots\dots \quad (8)$$

والتي يمكن استخدامها لتقدير قيمة معامل الارتباط الذاتي $\hat{\mathcal{P}}$.

ب –طريقة ثايل - ناجر **Theil-Nagar**

يتم تقدير قيمة معامل الارتباط الذاتي $(\mathcal{P})$ باستخدام الصيغة الاتية:

$$\hat{\mathcal{P}} = \frac{n^2\left[1 - \frac{D-W}{2}\right] + (k+1)^2}{n^2 - (k+1)^2}$$

حيث ان:

K : عدد المتغيرات المستقلة.

K+1 : عدد معلمات الانحدار المقدرة .

ج-طريقة كوكران – أوركات -: **Cochrane – Orcutt**

يتم تقدير قيمة معامل الارتباط الذاتي$(\hat{\mathcal{P}})$ من خلال تقدير قيم المتغير العشوائي وفق الصيغة الاتية:

$$\hat{\mathcal{P}} = \frac{\sum \hat{u_i}\widehat{u_{i-1}}}{\sum(\widehat{u_{i-1}})^2}$$

او باستخدام طريقة انحدار قيمة المتغير العشوائي $\hat{u_i}$ على $\widehat{u_{i-1}}$ وكما يلي:

$$u_i = \hat{\mathcal{P}}\ \widehat{u_{i-1}}$$

د-طريقة ديربن -: **Durbin Method**

يتم تقدير معامل الارتباط الذاتي بموجب طريقة ديربن على مرحلتين تنطوي المرحلة الأولى على تقدير $(\hat{\mathcal{P}})$ وكالاتي:

$$(y_t - \mathcal{P}y_{t-1}) = \beta_0(1-\mathcal{P}) + \beta_1(x_t - \mathcal{P}x_{t-1}) + u_t$$
$$y_t = \beta_0(1-\mathcal{P}) + \mathcal{P}y_{t-1} + \beta_1 x_t - \beta_1\mathcal{P}x_{t-1} + u_t$$
$$y_t = \beta_0 + \hat{\mathcal{P}}y_{t-1} + \beta_1 x_t - yx_{t-1} + u_t$$

اذ ان : $\beta_0 = \beta_0(1-\mathcal{P}) \quad : \quad r = \beta_1\mathcal{P}$

ويتضح من الصيغة (المعادلة) أعلاه بأن هناك ثلاث متغيرات مستقلة هي y_{t-1} ، x_t ، x_{t-1} والميل الحدي للمتغير y_{t-1} يعطينا تقدير لمعامل الارتباط الذاتي من الدرجة الأولى $\hat{\mathcal{P}}$. حيث يمكن تقديرها بطريقة (OLS) اما المرحلة الثانية فتتضمن تقدير معالم النموذج الاتي :

$$(y_t - \hat{\mathcal{P}}y_{t-1}) = \beta_0 + \beta_1(x_t - \hat{\mathcal{P}}x_{t-1}) + u_t \quad \ldots\ldots (9)$$

اذ ان : $\beta_0 = \beta_0(1-\hat{\mathcal{P}})$

رابعا - النتائج المترتبة على وجود الارتباط الذاتي:-

تبقى مقدرات المربعات الصغرى الاعتيادية (OLS) تتصف بالخاصية الخطية وعدم التحيز . الا انها تفقد خاصية أفضل او أصغر تباين (حيث ان التباين بوجود الارتباط الذاتي يفوق التباين بغياب الارتباط الذاتي). كما يؤثر الارتباط الذاتي على نتائج تحليل الانحدار فتعطي الاختبارات (t , F) نتائج أقل دقة من تلك في حالة عدم وجوده، حيث تستند تلك الاختبارات على التباين والخطأ المعياري، فضلاً عن عدم دقة التنبؤات المستحصلة بطريقة (OLS)، مما يتطلب استخدام طرق أخرى للتقدير كطريقة المربعات الصغرى العامة (الشاملة) (GLS) . *Generalized Least Squares Method*

لأنها تعطي افضل تقدير خطي غير متحيز في حالة وجود الارتباط الذاتي شريطة ان تكوم قيمة($\mathcal{P}$) معلومة.

خامسا - اختبار وجود الارتباط الذاتي :-

هناك عدد من الاختبارات الخاصة بالارتباط الذاتي، الا ان اكثرها شيوعا هو اختبار ديربن واتسن *(Durbin – Watson Test)* الذي يرمز له بالرمز *(D–W)* والذي يمكن اعتماده في العينات الصغيرة ويستند اختبار *(D–W)* على وجود فرضيتين أساسيتين هما :

1- فرضية العدم : التي تنص على انعدام الارتباط الذاتي $H_0 : \mathcal{P} = 0$

2- الفرضية البديلة : والتي تنص على وجود الارتباط الذاتي $H_1 : \mathcal{P} \neq 0$

ويستخدم لذلك الاختبار *(D–W)* الذي بدوره يعتمد على الأخطاء العشوائية الناتجة من النموذج الخطي العام الاتي :

$$Y = X\beta + u$$

حيث يفترض ان الفروق للأخطاء العشوائية تأخذ الصيغة التالية:

$$u_t = \mathcal{P}u_{t-1} + e_t$$

في الواقع التطبيقي يتم حساب الأخطاء العشوائية للنموذج أعلاه كالاتي :

$$e_t = Y_t - \hat{Y}_t$$

وبالتالي تحسب قيمة *(D – W)* بموجب الصيغة التالية:

$$D - W = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \quad \dots\dots\dots \quad (10)$$

ان قيمة إحصاءه ديربن واتسن $(D-W)$ مجدولة بقيمتين احدهما تسمى بالحد الأدنى Lower Limit ويرمز لها بالرمز (dl) والأخرى وتسمى بالحد الأعلى Upper Limit ويرمز لها بالرمز (du) حسب درجات الحرية n , $\acute{k}$ ولمستوى معنوية معين .

حيث ان : n : عدد المشاهدات في العينة . $\acute{k}$: العدد الكلي للمتغيرات المستقلة.

ويتم الاختبار على أساس مقارنة قيمة $(D-W)$ المحتسبة بقيم (dl) و(du) المجدولة لاتخاذ القرار الاحصائي المطلوب وعلى الشكل الاتي:

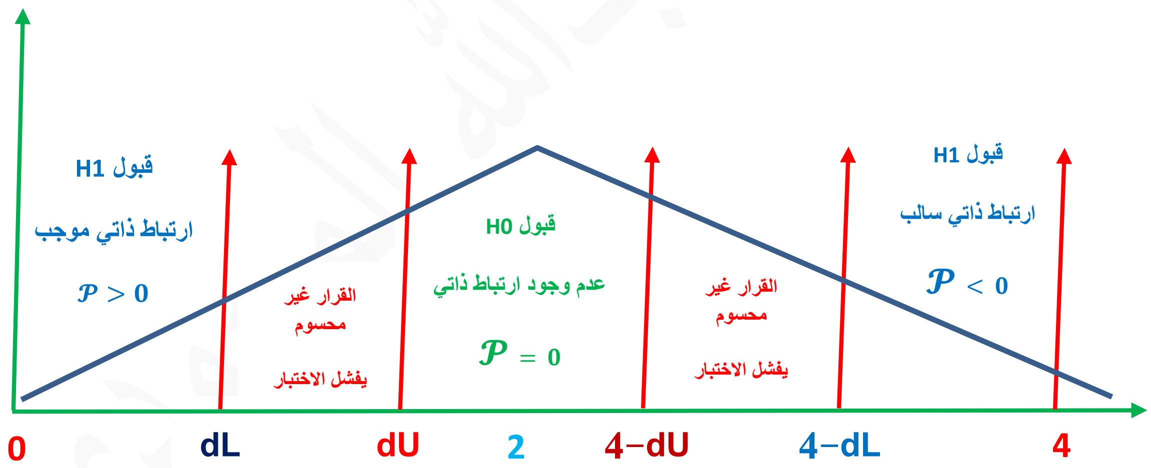
1- اذا كانت $0 < D\text{-}W < dl$ نرفض H_0 ونقبل H_1 أي ان هناك مشكلة ارتباط ذاتي موجب .

2- عندما $dL < D\text{-}W < du$ أو $4\text{-}du < D\text{-}W < 4\text{-}dL$ يكون الاختبار غير محسوم وتترك الحرية للباحث بقبول او رفض فرضية العدم اذ قد يكون السبب في وجود المشكلة خطأ في صيغة النموذج وليس بسبب ارتباط الأخطاء (سوء توصيف النموذج) .

3- عندما $du < D\text{-}W < 4\text{-}du$ نقبل H_0 أي انعدام وجود مشكلة الارتباط الذاتي .

4- عندما $4\text{-}dL < D\text{-}W < 4$ نرفض H_0 ونقبل H_1 بمعنى ان هناك مشكلة ارتباط ذاتي سالب .

5- يمكن توضيح الاحتمالات الواردة في أعلاه بالشكل الاتي :



ويمكن تلخيص مدى اختبار (D–W) في الجدول الاتي :

ت	الحالة (قيمة D,W)	النتيجة
1.	4–dl < D–W < 4	ترفض H_0 يوجد ارتباط ذاتي سالب
2.	4–du < D–W < 4–dl	الاختبار فاشل لا يمكن الجزم بشيء
3.	du < D–W < 4–du	نقبل H_0 انعدام وجود الارتباط الذاتي
4.	dl < D–W < du	الاختبار فاشل لا يمكن الجزم بشيء
5.	0 < D–W < dl	ترفض H_0 يوجد ارتباط ذاتي موجب

سادسا: معالجة مشكلة الارتباط الذاتي :-

تتوقف الطريقة التي تعالج فيها مشكلة الارتباط الذاتي من الدرجة الأولى على سبب حدوث المشكلة، أي ان :

1- عندما يكون السبب هو اهمال متغير او متغيرات مستقلة من النموذج يتعين إضافة ذلك المتغير او المتغيرات الى النموذج .

2- عندما يكون سبب المشكلة هو الصياغة غير الدقيقة للنموذج فان المعالجة تتوقف على إعادة صياغة النموذج المراد دراسته من واقع العلاقة .

3- اما اذا كان سبب المشكلة هو وجود علاقة فعلية بين قيم حد الخطأ او المتغير العشوائي فيصبح معالجتها بتحويل المتغيرات المستقلة بالشكل الذي يضمن التخلص من الارتباط الذاتي .

فاذا كان الارتباط الذاتي من الدرجة الأولى يتعين تحويل المتغيرات المعتمدة والمستقلة في النموذج المراد دراسته بإحدى الطرق الاتية :

أ- طريقة التحويل.

ب- طريقة التكرار.

ج- طريقة الفرق العام.

د- طريقة الفرق الأول.

وسوف نوضح كيفية معالجة الارتباط الذاتي من الدرجة الأولى باستخدام طريقة التكرار فقط.

طريقة التكرار :- Iterative Method

بموجب هذه الطريقة نتبع الخطوات الاتية :-

1- نبدأ باستخدام طريقة المربعات الصغرى الاعتيادية (OLS) لتقدير معلمات النموذج ، ولغرض الايضاح نفترض نموذجاً يتضمن متغير مستقل واحد وكالاتي:

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

حيث يتم الحصول على تقدير للحد الثابت (β_0) والميل الحدي للنموذج (β_1) باتباع طريقة (OLS) وكالاتي :

$$\hat{Y}_t = b_1 + b_1 X_t$$

وباستخدام النموذجين أعلاه يتم حساب البواقي (e_t) (الأخطاء) الناتجة من الفرق بين القيمة المشاهدة والقيمة التقديرية للمتغير المعتمد أي ان :

$$e_t = Y_t - \hat{Y}_t$$

2- حساب قيمة إحصاءه ديربن واتسن (D−W) بموجب الصيغة الاتية :

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

ثم تقارن قيمة (D−W) مع القيم الجدولية لها du , dl لتحديد قبول او رفض فرضية العدم H_0 . وفي حالة رفض H_0 وقبول الفرضية البديلة H_1 مما يعني وجود مشكلة الارتباط الذاتي وبالتالي يستوجب تقدير قيمتة .

3-حساب (تقدير) قيمة معامل الارتباط الذاتي $(\hat{\mathcal{P}})$ وفق الصيغة الاتية :

$$\hat{\mathcal{P}} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2} \quad \dots\dots\dots \quad (11)$$

t = 2,3,4,……n

4- بعد حساب قيمة $(\hat{\mathcal{P}})$ يتم تحويل بيانات كل من المتغير المعتمد والمتغيرات المستقلة بالشكل التالي:

$$Y_t - \hat{\mathcal{P}} Y_{t-1} = b_0\,(1 - \hat{\mathcal{P}}) + b_1(X_t - \hat{\mathcal{P}}\, X_{t-1}) + (1 - \hat{\mathcal{P}})u_t$$

أي ان :

$$Y_t^* = b_0^* + b_1\, X_t^* + u_t^* \quad \dots\dots\dots \quad (12)$$

حيث ان :

$$Y_t^* = Y_t - \hat{\mathcal{P}} Y_{t-1} \qquad : \qquad X_t^* = X_t - \hat{\mathcal{P}}\, X_{t-1}$$

$$b_0^* = b_0\,(1-\hat{\mathcal{P}}) \qquad : \qquad u_t^* = (1-\hat{\mathcal{P}})u_t$$

5- يعاد تقدير معالم النموذج في العلاقة (12) من جديد في ضوء البيانات المحولة للمتغيرين بطريقة (OLS) ، أي تقدير خط انحدار (Y_t^*) على المتغير المستقل (X_t^*) حيث نحصل على:

$$\hat{Y}_t^* = \hat{b}_0^* + \hat{b}_1\, X_t^*$$

ومنه يمكن حساب البواقي الجديدة (e_t^*) أي ان :

$$e_t^* = Y_t^* - \hat{Y}_t^*$$

6- حساب قيمة (D−W) بموجب الصيغة الاتية:

$$\mathrm{D-W} = \frac{\sum_{t=2}^{n}(e_t^* - e_{t-1}^*)^2}{\sum_{t=}^{n} e_t^{2*}}$$

ثم تقارن قيمة (D−W) مع القيمة الجدولية لقبول او رفض فرضية العدم، ففي حالة قبول H_0 يعني انعدام الارتباط الذاتي، والتوقف عند هذا الحد (أي قبول التقديرات الناتجة) .

اما في حالة قبول H_1 عندها تجري عملية تنقية البيانات مرة ثانية باتباع الخطوات السابقة نفسها، أي تكرار ما قمنا به وبالأسلوب ذاته لرؤية مدى تناقص الارتباط الذاتي، ويمكن الاستمرار في عملية التصحيح (التنقية) والتقدير الى ان تتقارب القيم التقديرية لكل من b_0 و b_1 للنموذج المدروس بين مرحلة وأخرى .

العلاقة بين $\hat{\mathcal{P}}$ و (D-W) :-

تربط بين إحصاءه (D−W) و $(\hat{\mathcal{P}})$ تقدير معلمة الارتباط الذاتي البسيط بين الأخطاء العشوائية للعلاقة الاتية :

$$\mathrm{D-W} = 2(1-\hat{\mathcal{P}})$$

ويمكن اشتقاق العلاقة أعلاه رياضياً كالاتي :

$$\mathrm{D-W} = \frac{\sum_{t=2}^{n}(e_t^* - e_{t-1}^*)^2}{\sum_{t=}^{n} e_t^{2*}}$$

وبفك الاقواس في البسط نحصل على :

$$D - W = \frac{\sum_{t=2}^{n} e_t^2 + \sum_{t=2}^{n} e_{t-1}^2 - 2\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}$$

وعندما يكون حجم العينة كبير جداً $(n \to \infty)$ فاننا نحصل على الاتي :

$$\sum_{t=1}^{n} e_t^2 \simeq \sum_{t=2}^{n} e_t^2 \simeq \sum_{t=2}^{n} e_{t-1}^2$$

وعليه فان :

$$D - W = \frac{2\sum_{t=2}^{n} e_t^2 - 2\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}$$

$$\therefore \quad D - W = 2\left(1 - \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}\right)$$

وبالتعويض عن المقام :

$$D - W = 2\left(1 - \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}\right)$$

$$\therefore \quad \hat{\mathcal{P}} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}$$

ولذلك فان :

$$\therefore D - W = 2\left(1 - \hat{\mathcal{P}}\right)$$

او بصيغة أخرى فانه:

$$\hat{\mathcal{P}} = 1 - \frac{D - W}{2}$$

وبما ان قيمة $\hat{\mathcal{P}}$ تنحصر بين *(-1 , +1)* لذلك فان قيمة *(D-W)* تتراوح بين *(4 , 0)*

فاذا كانت $\hat{\mathcal{P}} = 0$ (انعدام الارتباط الذاتي كلياً)

فان D-W = 2(1-0) = 2

وعندما $\hat{\mathcal{P}} = 1$ فان :

(وجود ارتباط ذاتي موجب) . D-W=2(1-1)=0

وعندما $\hat{\mathcal{P}} = -1$ فان:

(وجود ارتباط ذاتي سالب).　　D-W=2(1-(-1))=4

مثال(1) : الجدول الاتي يتضمن البيانات الخاصة بعدد حوادث السرقات (Y_t) وعدد مكاتب الشرطة (X_t) في احدى الدول للمدة (1990–1999) ، وباستخدام طريقة المربعات الصغرى (OLS) تم تقدير العلاقة وكانت كالاتي :

$$\hat{Y}_t = 1346.28941 - 12.10030471\, X_t$$

$$t \qquad (5.319) \qquad\qquad (3.105)$$

$$R^2 = 0.546 \qquad , \qquad F = 9.641$$

والمطلوب : معرفة ما اذا كان النموذج الخاص بالعلاقة أعلاه يعاني من مشكلة الارتباط الذاتي مستخدما إحصاءه D–W عند مستوى معنوية 5% اذا علمت ان du = 1.36 , dL = 1.08 .

السنة	عدد حوادث السرقات Y_t	عدد مكاتب الشرطة X_t	$\hat{Y}_t$	$e_t = Y_t - \hat{Y}_t$	e_{t-1}
1990	580	60	620.2711274	−40.2711274	-------
1991	890	59	632.3714321	257.628567	−40.2711274
1992	430	77	414.5659473	15.434052	257.628567
1993	690	52	717.0735651	−27.073565	15.434052
1994	310	87	293.562900	16.4371	−27.073565
1995	750	50	741.2741745	8.7258255	16.4371
1996	460	80	378.2650332	81.734966	8.7258255
1997	630	52	717.0735651	−87.07356	81.734966
1998	800	53	704.9732604	95.026739	−87.07356
1999	215	67	535.5689944	−320.56899	95.026739
n=10	**5755**	**637**	**5755**	**0**	

$e_t - e_{t-1}$	$(e_t - e_{t-1})^2$	e_t^2
---------	---------	1621.763702
297.8996953	88744.22846	66372.479
-242.194515	58658.18319	238.2099827
-42.507617	1806.897571	732.9779272
43.5106651	1893.177977	270.1782564
-7.7112745	59.46375441	76.14003066
73.0091413	5330.334713	6680.604798
-168.808531	28496.32044	7581.805739
182.100304	33160.52097	9030.081239
- 415.59573	172719.8141	102764.4802
	390868.9412	**195368.7208**

$$D - W = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{390868.9412}{195368.7208}$$

$$\therefore D - W = 2.000673$$

وبما انه $\quad du < D\text{-}W < 4\text{-}du$

أي ان $\quad 1.36 < 2.000673 < 2.64$

وعليه فان النموذج لا يعاني من مشكلة الارتباط الذاتي ذلك لان قيمة D–W المحتسبة تقع في منطقة القبول أي قبول فرضية العدم (H_0) التي تنص بعدم وجود ارتباط ذاتي بين القيم المتتالية للمتغير العشوائي.

مثال(2) :- عينة عشوائية بحجم n=15 فيها المتغير Y_t يرتبط خطياً بالمتغير المستقل X_t ، والمطلوب الاختبار لوجود الارتباط الذاتي ، فان وجد قدر قيمته واستبعد اثره مستخدما طريقة التكرار.

Y_t	33	34	38	43	46	46	45	37	40	38	40	43	44	54	55
X_t	10	10	11	12	12	13	13	13	13	13	14	14	15	16	16

الحل :- لاختبار وجود الارتباط الذاتي نستخدم اختبار ديربن واتسون (D–W) ، وهذا يتطلب تقدير معالم خط الانحدار (Y_t) على (X_t)، ثم إيجاد القيم التقديرية $(\widehat{Y}_t)$ وذلك بتطبيق طريقة (OLS) مباشرة على بيانات المتغيرين في الجدول أعلاه: أي ان

$$b_{(ols)} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = (\acute{X}X)^{-1}\,\acute{X}Y = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

وفي ادناه العمليات الحسابية اللازمة لذلك :

$$n = 15\,, \qquad \sum X_i = 195 \;\; , \sum X_i^2 = 2583 \;\; , \;\; \sum Y_i = 636 \;\; , \qquad \sum X_i Y_i = 8400$$

$$\therefore \; \underline{b} = \begin{bmatrix} 15 & 195 \\ 195 & 2583 \end{bmatrix}^{-1} \begin{bmatrix} 636 \\ 8400 \end{bmatrix}$$

$$= \frac{1}{720} \begin{bmatrix} 2583 & -195 \\ -195 & 15 \end{bmatrix} \begin{bmatrix} 636 \\ 8400 \end{bmatrix}$$

$$\therefore \; \underline{b}_{(ols)} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 6.65 \\ 2.75 \end{bmatrix}$$

$$\therefore \; \widehat{Y}_t = 6.56 + 2.75 X_t$$

وباستخدام الصيغة التقديرية أعلاه ، تم الحصول على القيم التقديرية للمتغير المعتمد $(\widehat{Y}_t)$ وبالتالي اجراء كافة العمليات الحسابية اللازمة لاحتساب اختبار ديربن واتسن (D–W) كما في الجدول ادناه :

Y_t	$\hat{Y}_t$	$e_t = Y_t - \hat{Y}_t$	e_{t-1}	$e_t - e_{t-1}$
33	34.15	−1.15	------	------
34	34.15	−0.15	−1.15	1.00
38	36.90	1.10	−0.15	1.25
43	39.65	3.35	1.10	2.25
46	39.65	6.35	3.35	3.00
46	42.40	3.60	6.35	2.75
45	42.40	2.60	3.60	−1.00
37	42.40	−5.40	2.60	−8.00
40	42.40	−2.40	−5.40	3.00
38	42.40	−4.40	−2.40	−2.00
40	45.15	−5.15	−4.40	−0.75
43	45.15	−2.15	−5.15	3.00
44	47.90	−3.90	−2.15	−1.75
54	50.65	3.35	−3.90	7.25
55	50.65	4.35	3.35	1.00

ومن الجدول أعلاه يمكن الحصول على العمليات الحسابية التالية:

$$\sum_{t=1}^{15} e_t^2 = 204.6 \quad , \quad \sum_{t=2}^{15}(e_t - e_{t-1})^2 = 168.375$$

$$D - W = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{168.375}{204.6} = 0.8229$$

ومن جداول (D-W) لمستوى دلالة (5%) ودرجة حرية (15,1) نجد بأن

dL=1.08 , du= 1.37 لذلك فان 0.0 < 0.8229 < 1.08

لذلك ترفض فرضية العدم (H_0) ، $(\mathcal{P}=0)$ أي ان هناك ارتباط ذاتي موجب وبالتالي يستوجب تقديره بموجب الصيغة الاتية :

$$\hat{\mathcal{P}} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}$$

علماً بان

$$\sum_{t=2}^{15} e_t e_{t-1} = 110.29$$

$$\sum_{t=2}^{15} e_{t-1}^2 = 185.6775$$

$$\therefore \hat{\mathcal{P}} = \frac{110.29}{185.6775} = 0.59$$

او

$$\hat{\mathcal{P}} = 1 - \frac{D-W}{2} = 1 - \frac{0.82}{2} = 1 - 0.41 = 0.59$$

طريقة التكرار :- بعد ثبوت وجود الارتباط الذاتي ، وتقدير قيمته $(\hat{\mathcal{P}} = 0.59)$ يمكن استخدام طريقة التكرار لتنقية بيانات العينة من اثر وجوده ، حيث يتطلب ذلك العمليات الحسابية التالية :

$$Y_t^* = Y_t - \hat{\mathcal{P}} Y_{t-1} \quad : \quad X_t^* = X_t - \hat{\mathcal{P}} X_{t-1}$$

Y_t	Y_{t-1}	Y_t^*	X_t	X_{t-1}	X_t^*
33	---	---	10	---	---
34	33	41.53	10	10	4.1
38	34	17.94	11	10	5.1
43	38	20.58	12	11	5.51
46	43	20.63	12	12	4.92
46	46	18.86	13	12	5.92
45	46	17.86	13	13	5.33
37	45	10.45	13	13	5.33
40	37	18.17	13	13	5.33
38	40	14.4	13	13	5.33
40	38	17.58	14	13	6.33
43	40	19.4	14	14	5.74
44	43	18.63	15	14	6.74
54	44	28.04	16	15	7.15
55	54	23.14	16	16	6.56

من بيانات الجدول أعلاه يمكن الحصول على العمليات الحسابية التالية:

$$n = 14 \quad , \quad \Sigma X_t^* = 79.39 \quad , \quad \Sigma Y_t^* = 260.21$$

$$\sum X_t^{*2} = 458.6687 \quad , \quad \sum X_t^* Y_t^* = 1508.592$$

وبتطبيق طريقة المربعات الصغرى (OLS) على بيانات العينة المنقاة ، يمكن تقدير معالم العلاقة الخطية بالشكل الاتي :

$$\boldsymbol{b}^*_{(ols)} = \begin{bmatrix} \boldsymbol{b}_0^* \\ \boldsymbol{b}_1^* \end{bmatrix} = (\boldsymbol{X}^{*-}\boldsymbol{X}^*)^{-1}\, \boldsymbol{X}^{*-}\boldsymbol{Y}^*$$

$$= \begin{bmatrix} n & \sum X_t^* \\ \sum X_t^* & \sum X_t^{*2} \end{bmatrix}^{-1} \begin{bmatrix} \sum Y_t^* \\ \sum X_t^* Y_t^* \end{bmatrix}$$

$$\therefore \begin{bmatrix} \hat{b}_0^* \\ \hat{b}_1^* \end{bmatrix} = \begin{bmatrix} 14 & 79.39 \\ 79.39 & 454.6687 \end{bmatrix}^{-1} \begin{bmatrix} 260.21 \\ 1508.592 \end{bmatrix} = \begin{bmatrix} 0.5009 \\ 3.18928 \end{bmatrix}$$

حيث ان الحد الثابت في الصيغة التقديرية يجب تعديله بالشكل الاتي :

$$\hat{Y}_t^* = \hat{b}_0^* + \hat{b}_1^* X_t^*$$

$$\hat{b}_0 = \frac{\hat{b}_0^*}{1-\hat{\mathcal{P}}} = \frac{0.50}{0.41} = 1.22$$

عليه فان الصيغة التقديرية بعد التكرار الأول تكتب على الشكل الاتي:

$$\hat{Y}_t^* = 1.22 + 3.19\, X_t^*$$

ونستمر في عملية التكرار هذه مستخدمين معالم العلاقة الجديدة أعلاه في إيجاد الانحرافات وبالتالي الاختبار لوجود الارتباط الذاتي ، فاذا كانت نتيجة الاختبار تؤكد وجود الارتباط الذاتي ، استوجب استبعاد أثره مرة ثانية من بيانات العينة ، اما اذا أظهرت نتيجة الاختبار عدم وجوده عندئذ تعتمد معالم العلاقة المقدرة الأخيرة .

والجدول التالي يوضح العمليات الحسابية اللازمة لعملية التكرار الثانية (لاختبار وجود الارتباط الذاتي).

Y_t^*	$\widehat{Y}_t^*$	e_t^*	e_{t-1}^*	$e_t^* - e_{t-1}^*$
14.53	14.299	0.231	---	---
17.94	17.489	0.451	0.231	0.220
20.58	18.7969	1.7831	0.451	1.3321
20.36	16.9148	3.7152	1.7831	1.9321
18.86	20.1048	−1.2448	3.7152	−4.960
17.86	18.2227	−0.3627	−1.2448	0.8821
10.45	18.2227	−7.7727	−0.3627	−7.41
18.17	18.2227	−0.0527	−7.7727	7.72
14.40	18.2227	−3.8227	−0.0527	−3.77
17.58	21.4127	−3.8327	−3.8227	−0.01
19.40	19.5306	−0.1306	−3.8327	3.7021
18.63	22.7206	−4.0906	−0.1306	−3.96
28.04	24.0285	4.0115	−4.0906	8.021
23.14	22.1464	0.9936	4.0115	−3.0179

ومن الجدول أعلاه يمكن الحصول على العمليات الحسابية التالية:

$$\sum(e_t^* - e_{t-1}^*)^2 = 263.79398$$

$$\sum e_t^{*2} = 142.4697$$

$$\therefore D - W = \frac{\sum_{t=2}^{n}(e_t^* - e_{t-1}^*)^2}{\sum_{t=1}^{n} e_t^{*2}} = \frac{263.79398}{142.4697} = 1.852$$

وباستخدام حجم عينة n=14 وعدد متغيرات مستقلة k=1 وعند مستوى دلالة 5% نجد ان du=1.350 و dl=1.045

وعليه فان $du < 1.852 < 4-du$

أي ان $1.350 < 1.852 < 2.650$

وهذا يعني عدم وجود ارتباط ذاتي في بيانات العينة تحت البحث ، وبذلك يجب التوقف عند هذا الحد. واعتماد الصيغة التقديرية الأخيرة أعلاه .

تمارين الفصل الخامس

الأول : للنموذج الخطي العام التالي:

$$Y = X\beta + u$$

حيث ان :

$$U \sim N\,(0, \sigma_u^2 \boldsymbol{\Omega})$$

$$u_t = \mathcal{P} u_{t-1} + e_t$$

$$e_t \sim N\,(0\sigma_e^2) \qquad , \qquad E(e_t e_s) = 0 \;\; \forall \;\; t \neq s$$

في ظل الفروض أعلاه اشتق عناصر المصفوفة **(Ω)**.

الثاني : باحث اقتصادي استخدم أسلوب (OLS) لتقدير دالة الاستيرادات التالية وكانت النتائج كالاتي :

$$\hat{Y}_t = -2461 + 0.28\, X_t$$

$$S.E \qquad (250) \qquad (0.01) \qquad r^2 = 0.98$$

ومن الانحرافات $e_{t=Y_t - \hat{Y}_t}$ وجد أن

$$\sum_{t=2}^{20} (e_t - e_{t-1})^2 = 537192 \qquad : \qquad \sum_{t=1}^{20} e_t^2 = 573069$$

المطلوب :- إحسب احصاءة (D-W) واختبر لوجود مشكلة الارتباط الذاتي مستخدما مستوى دلالة (5%) علما ان dL=1.20 , du=1.41

الثالث : البيانات التالية تمثل الفرق بين القيمة الحقيقية والقيمة التقديرية للمتغير المعتمد (Y_t) في عينة عشوائية ذات حجم (15) مشاهدة.

ت	$e_{t=Y_t-\hat{Y}_t}$	ت	$e_{t=Y_t-\hat{Y}_t}$	ت	$e_{t=Y_t-\hat{Y}_t}$
1	−0.15	8	−5.40		
2	−0.15	9	−2.40		
3	1.10	10	−4.40		
4	3.35	11	−5.15	15	−2.15
5	6.35	12	−3.90		
6	3.60	13	3.35		
7	2.60	14	4.35		

المطلوب : أحسب معامل ديربن واتسن (D−W) ثم اختبر لوجود الارتباط الذاتي ($\mathcal{P}$) مستخدما مستوى دلالة (5%) فان وجد قدر قيمته وبين نوعه ، علما ان هناك متغير مستقل واحد في النموذج المدروس أي ان K=1

الرابع : باحث قدر العلاقة بين المتغير (Y_t) والمتغير (X_t) وحصل على النتائج والمعلومات الاتية:

$$\hat{Y}_t = 6.65 + 2.75\, X_t$$

$$\mathrm{D-W} = 0.8229 \qquad : \qquad \sum_{t=2}^{n} e_t \,.\, e_{t-1} = 110.29$$

$$\sum_{t=2}^{n} e_{t-1}^2 = 185.6775$$

المطلوب:− من قيمة *(D−W)* وجد ان هناك ارتباط ذاتي موجب ، قدره ثم ضع الخطوات اللازمة لمعالجة هذه المشكلة في الجانب التطبيقي؟

الخامس : (ان مشكلة الارتباط الذاتي تنشأ عند استخدام بيانات السلاسل الزمنية ولا تنشأ عند استخدام بيانات المقاطع العرضية) ناقش هذه العبارة موضحا طبيعة المشكلة واسبابها، وماهي النتائج المترتبة على وجودها في تحليل الانحدار؟

الفصل السادس

مشكلة التعدد الخطي (الازدواج الخطي) The Multucollinearity Problem

من الافتراضات الأساسية لنموذج الانحدار الخطي المتعدد عدم وجود ارتباط خطي تام او شبه تام بين قيم مشاهدات المتغيرات المستقلة الداخلة في نموذج الانحدار المراد تقديره أي ان

$$rx_ix_j \neq 1$$

وفي حالة اعتماد قيمة احد المتغيرات المستقلة على قيمة واحد او اكثر من المتغيرات المستقلة في النموذج تحت الدراسة ينبغي الافتراض الخاص بانعدام الارتباط فتظهر مشكلة تدعى مشكلة التعدد الخطي (الازدواج الخطي) .

وبعبارة أخرى تحصل مشكلة التعدد الخطي عندما يرتبط اثنان او أكثر من المتغيرات المستقلة بعلاقة خطية قوية جدا بحيث يكون من الصعب فصل إثر كل متغير مستقل عن المتغير المعتمد. وتجدر الإشارة هنا على ان التعدد الخطي ما بين المتغيرات المستقلة لا يشكل مشكلة الا إذا تجاوزت شدته درجة معينة، حيث انه غالبا ما تكون هناك علاقة بين المتغيرات المستقلة وذلك نتيجة لتأثير المتغيرات الاقتصادية ببعضها البعض.

وتواجه مشكلة التعدد الخطي حينما تكون قيمة أحد المتغيرات المستقلة متساوية لكافة المشاهدات، أو عندما تعتمد قيمة أحد المتغيرات المستقلة على قيمة واحد او أكثر من المتغيرات المستقلة في النموذج المدروس، علماً بأن مثل هذه المشكلة تواجه الباحث سواء في ظل فرضية التجانس او عدم التجانس وسواء اخذ البيانات شكل السلاسل الزمنية أو المقطعية.

ويتضح من أعلاه بأن الفرض الخاص بالتعدد الخطي يتلخص بما يلي:

لا توجد علاقة خطية تامة او شبه تامة بين أي من المتغيرات المستقلة، إضافة الى ذلك يجب ان يكون عدد المعالم المطلوب تقديرها اقل من حجم العينة تحت البحث أي ان:

$$ranK(X) = \mathcal{P}(X) = k + 1 < n$$

أسباب حدوث ظاهرة التعدد الخطي :-

وتتلخص بما يلي :

1- عندما تتغير بعض المتغيرات المستقلة سوية، فعلى سبيل المثال، في فترة الازدهار الاقتصادي يلاحظ ان المتغيرات الاقتصادية كالدخل والاستثمار تزداد بوقت واحد، وتنخفض في وقت واحد في فترة الكساد، لذلك فعندما تستخدم هذه المتغيرات كمتغيرات تفسيرية (مستقلة) في النموذج تبرز ظاهرة التعدد الخطي، مما يجعل من الصعب او المستحيل عزل تأثيراتها الفردية على المتغير التابع.

2- عند استخدام المتغيرات المتخلفة زمنيا كمتغيرات تفسيرية في النموذج، فعلى سبيل المثال يستخدم الدخل الحالي والدخل السابق كمتغيرات تفسيرية سوية في دالة الاستهلاك (النموذج) ومن الطبيعي ان القيم المتطابقة لأي متغير اقتصادي تظهر نوع من الارتباط، ولهذا السبب تظهر مشكلة الارتباط الخطي المتعدد. فاذا كان لدينا دالة الاستهلاك الاتية:

$$C_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t$$

واذا كان :

$$X_1 = 2\, X_2 \qquad or \qquad X_1 = X_2^2$$

فان ذلك يؤدي الى وجود مشكلة التعدد الخطي .

أنواع التعدد الخطي:-

أ - غياب التعدد الخطي:- Absence Of Multicollinearity

تحدث هذه الحالة عندما لا ترتبط المتغيرات المستقلة (X_i) في نموذج الانحدار الخطي المتعدد ارتباطاً خطياً وبالتالي فأن المصفوفة $(\acute{x}x)$ تصبح مصفوفة قطرية تأخذ الشكل الاتي:

$$(\acute{x}x) = \begin{bmatrix} \sum x_1^2 & 0 & 0 & 0 \\ 0 & \sum x_2^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sum x_k^2 \end{bmatrix}$$

وان مقلوب (معكوس) المصفوفة سيكون قطريا أيضا والعناصر المعطاة في المصفوفة $(\acute{x}x)$ وتحسب المعلمات $(\hat{b}_i)$ بموجب الصيغة الاتية :

$$b_i = \frac{\sum x_i\, y_i}{\sum x_i^2} \qquad\qquad (i = 1,2, \dots, k)$$

وبالتالي لا توجد حاجة الى نموذج الانحدار المتعدد ، حيث يقدر الثابت b_0 وفق الصيغة الاتية :

$$b_0 = \bar{Y} - b_i\,\bar{X}_i \qquad\qquad (i = 1,2,\dots,k)$$

ويتضح مما سبق بان هذه الصيغ مطابقة لانحدار Y البسيط على كل متغير من المتغيرات المستقلة X_i بشكل مستقل.

ب-التعدد الخطي التام:- Exact Multicollinearity

ان حالة التعدد الخطي التام هي حالة مثالية غير ممكنة التحقق في الواقع العلمي ، ذلك لانها تتحقق فقط في حالة وجود علاقة خطية تامة بين قيم اثنين او اكثر من المتغيرات المستقلة، ويمكن حينئذ التعبير عن واحد او اكثر من المتغيرات المستقلة كتركيب خطي للمتغير او المتغيرات المستقلة الأخرى.

ولتوضيح هذه الحالة نفترض وجود نموذج انحدار خطي متعدد يحتوي على متغيرين مستقلين كالأتي:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u_i$$

والذي يمكن تقدير معلماته وفق الصيغة الاتية :

$$\underline{b} = (\acute{X}X)^{-1}\,\acute{X}Y$$

حيث محدد المصفوفة $(\acute{x}x)$ هو:

$$|\acute{x}x| = \begin{vmatrix} \sum x_1^2 & \sum x_1x_2 \\ \sum x_1x_2 & \sum x_2^2 \end{vmatrix} \neq 0$$

$$= \left(\sum x_1^2\right)\left(\sum x_2^2\right) - \left(\sum x_1x_2\right)^2 \neq 0$$

فبوجود علاقة خطية تامة بين قيم مشاهدات المتغيرين المستقلين X_1 و X_2 يمكن التعبير عنها كما يلي :

$$X_2 = \mathcal{P}X_1$$

حيث ان $\mathcal{P}$ عبارة عن قيمة ثابتة لا تساوي صفر . فالمحدد $|\acute{x}x|$ سيكون مساويا للصفر أي ان:

$$|\acute{x}x| = \begin{bmatrix} \sum x_1^2 & \mathcal{P}\sum x_1^2 \\ \mathcal{P}\sum x_1^2 & \mathcal{P}^2\sum x_1^2 \end{bmatrix} = 0$$

$$= \mathcal{P}^2\left(\sum x_1^2\right)^2 - \mathcal{P}^2\left(\sum x_1^2\right)^2 = 0$$

وبالتالي لا يمكن حساب معكوس المصفوفة $|\acute{x}x|$ وتسمى هذه المصفوفة بالمصفوفة الشاذة وعليه لا يمكن الحصول على مقدرات (OLS) وفق الصيغة $\underline{b} = (\acute{x}x)^{-1}\,\acute{x}y$ لذلك تنهار تقديرات (OLS) تحت وطأة الارتباط الخطي المتعدد التام.

ج - التعدد الخطي غير التام :- No Exact Multicollinearity

مما سبق يتضح ان حالتي التعدد الخطي التام وغياب التعدد الخطي هما حالتان لا تحدثان في الحياة العملية . وفي الواقع هناك دائما علاقة ارتباطية بدرجات متفاوتة بين المتغيرات المستقلة تقع ضمن الحالتين المتطرفتين المذكورتين (التعدد الخطي التام وغيابه). وفي حالة وجود علاقة خطية غير تامة بين قيم مشاهدات المتغيرين X_1 و X_2 ممثلة بالشكل الاتي :

$$X_2 = \mathcal{P}X_1 + v$$

حيث v متغير فيه $\sum v = \sum X_1\, v = 0$

لذلك فان محدد المصفوفة $(\acute{x}x)$ لا يساوي الصفر وهكذا يمكن الحصول على المعكوس $(\acute{x}x)^{-1}$ ومن ثم تقدير معالم النموذج المتعدد بطريقة (OLS) .

النتائج المترتبة على وجود التعدد الخطي :

يتعذر تقدير معالم النموذج عندما تكون هنالك علاقة خطية تامة بين اثنين او اكثر من المتغيرات المستقلة ويرجع السبب في ذلك الى استحالة إيجاد معكوس المصفوفة $(\acute{X}X)$ وذلك لكون محدد هذه المصفوفة سوف يكون مساويا للصفر. وبنفس الوقت فان تباين معاملات النموذج المقدرة سوف تكون مساوية الى ما لانهاية في ظل وجود العلاقة الخطية التامة بين المتغيرين المستقلين.

أي ان $var(b_i) = \infty$ وان تقديرات (b_i) تكون غير محددة (كمية غير معرفة) وتعالج المشكلة في هذه الحالة باستخدام طريقة المركبات الرئيسية (p.c.)

اما اذا كانت العلاقة الخطية غير تامة بين المتغيرات المستقلة . بعبارة أخرى ان كل من $(X_1$ و $X_2)$ يتأثران بعامل ثالث، فان مقدرات (OLS) تحتفظ بخاصية الخطية وعدم التحيز فيما لو كان النموذج قد حدد بدقة ولكنه لا يكتسب خاصية الكفاءة أي اقل تباين ممكن . وذلك نتيجة لضآلة قيمة محدد المصفوفة $(\acute{x}x)$ والذي يترتب عليه ان تكون المعالم المقدرة ذات تباين كبير جداً وذلك لان :

$$var - cov(\underline{b}) = S_e^2(\acute{X}X)^{-1} = S_e^2 \quad \frac{adj\ (\acute{X}X)}{|\acute{X}X|}$$

وبالتالي قد يستنتج خطأ بان بعض المتغيرات المستقلة غير مهمة ، اذ يظهر اختبار (t) عدم معنوية معالم تلك المتغيرات في حين انها في الواقع معنوية، ولكن النموذج يعجز عن اظهار اثر كل منها بشكل منفصل نظراً لارتباطها ببعضها البعض ومن أبرز طرق المعالجة لمثل هذه الحالة هو استخدام طريقة انحدار الحرف (Ridge Regression).

بشكل عام مشكلة التعدد الخطي تكون على نوعين: الأول يعرف بالتعدد الخطي التام (Perfect Multicollinearity) وفي مثل هذه الحالة يكون محدد مصفوفة المعلومات او مصفوفة فيشر (Fisher Matrix) مساوياً للصفر، بعبارة أخرى $(|X'X| = 0)$ ويترتب على ذلك استحالة إيجاد مقدرات معالم النموذج الخطي العام، وابرز طرق المعالجة في مثل هذه الحالة هو استخدام أسلوب المركبات الرئيسية (Principle Components).

اما النوع الثاني فيسمى بالتعدد الخطي شبه التام (Semi Multicollinearity) وفيه تكون قيمة محدد مصفوفة المعلومات صغير جداً وعندها تكون المعالم المقدرة ذات تباين كبير جداً، وقد نستنتج خطأ بان بعض المتغيرات التوضيحية غير مهمة ومن ابرز طرق المعالجة في مثل هذه الحالة هو استخدام أسلوب الانحدار الحرف (Ridge Regression) .

إختبار وجود مشكلة التعدد الخطي :

هناك عدد من الاختبارات للكشف عن وجود مشكلة التعدد الخطي ومن أهمها هي :

أ- اختبار فريش Frisch

ب-اختبار فرار وكلوبر Farrar – Glauber

ت-اختبار كلاين Klien Test

وسوف نوضح باختصار الاختبارين الأخيرين .

اختبار فرار وكلوبر : Farrar – Glauber

ويعتبر من أهم الاختبارات للكشف عن مشكلة التعدد الخطي، ويستند هذا الاختبار على احصاءة $(\mathcal{X}^2)$ (مربع كاي) حيث يتم اختبار الفرضية التالية :

$H_0 : (X_J)\ Orthogonal$ (المتغيرات مستقلة غير مرتبطة لا توجد مشكلة التعدد الخطي) متعامدة X_J

$H_1 : (X_J)\ Not\ Orthogonal$ (المتغيرات غير مستقلة توجد مشكلة التعدد الخطي) غير متعامدة X_J

والصيغة الرياضية لهذا الاختبار هي :

$$\mathcal{X}^2 = -\left[n - 1 - \frac{1}{6}\ (2k + 5)\right] Ln|R|$$

حيث ان :-

n : حجم العينة (عدد المشاهدات).

K : عدد المتغيرات المستقلة.

$Ln|R|$: اللوغاريتم الطبيعي لمحدد مصفوفة معاملات الارتباط الاتية :

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots\dots\dots & r_{1k} \\ r_{21} & 1 & r_{23} & \dots\dots\dots & r_{2k} \\ \vdots & \vdots & & \vdots & \vdots \\ r_{k1} & r_{k2} & r_{k3} & \dots\dots\dots & 1 \end{bmatrix}$$

ثم تقارن قيمة $(\mathcal{X}^2)$ المحسوبة مع قيمة $(\mathcal{X}^2)$ الجدولية بدرجة حرية مساوية الى$(k(k-1)/2)$ عند مستوى معنوية معين . فاذا كانت القيمة المحسوبة اكبر من القيمة الجدولية ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة، أي ان هناك مشكلة التعدد الخطي بين المتغيرات المستقلة والعكس صحيح.

وبعد ثبوت وجود مشكلة التعدد الخطي بموجب الاختبار أعلاه، يستوجب تحديد أي متغير من المتغيرات المستقلة مرتبط خطياً بحيث أدى الي مشكلة التعدد الخطي ، ويتم مثل هذا التشخيص باستخدام اختبار(F) حيث تستخرج القيمة المحسوبة لهذا الاختبار، وبعد تقدير معامل التحديد المتعدد ما بين المتغير المستقل (X_J) وبقية المتغيرات المستقلة الأخرى وكالاتي :

$$R_J^2.\,2.3.4\ \dots K \qquad : \qquad J = 1,2, \dots., K$$

فعلى سبيل المثال فان معامل التحديد بين أربعة متغيرات مستقلة يحسب وفق الصيغة الاتية:

$$\boldsymbol{R^2 1.234 = 1 - (1 - r_{12}^2)(1 - r_{13.2}^2)(1 - r_{14.23}^2)}$$

حيث ان $r_{13.2}$ ، $r_{14.23}$ تمثل معاملات الارتباط الجزئية والصيغة العامة للاختبار هي :

$$F_J = \frac{R_J^2.23\ \dots k/(k-1)}{(1 - R_J^2.23\ \dots k)/(n-k)}$$

لاختبار فرضية العدم التالية :

$$H_0: R_J^2.\,23.\dots k \ = 0$$

مقابل الفرضية البديلة :

$$H_1 : R_J^2.23.\dots k \neq 0$$

ثم تقارن قيمة (F_J) العملية (المحسوبة) مع القيمة الجدولية المقابلة لها بدرجة حرية مساوية الى $\left[(k-1) \text{ و } (n-k)\right]$ للبسط والمقام ولمستوى معنوية معين . فاذا كانت القيمة المحسوبة اكبر من الجدولية ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة (H_1) أي ان المتغير (X_J) يرتبط خطياً مع بقية المتغيرات المستقلة ، وبعكسه ترفض الفرضية البديلة (H_1) أي ان المتغير (X_J) لا يرتبط خطيا مع بقية المتغيرات المستقلة ولا يشكل مصدرا لمشكلة التعدد الخطي . وتكرر العملية(الاختبار) لكل متغير من المتغيرات المستقلة في النموذج حتى يتم تشخيص المتغيرات المستقلة التي ترتبط خطيا مع بقية المتغيرات المستقلة الأخرى.

ولغرض تحديد المتغيرات المستقلة المسببة لحصول مشكلة التعدد الخطي يتطلب ذلك اجراء اختبار ثالث هو اختبار (t) والذي يعتمد بدوره على قيم معاملات الارتباط الجزئية ما بين كل اثنين من المتغيرات المستقلة. فعلى سبيل المثال، فان معامل الارتباط الجزئي بين متغيرين مستقلين بثبوت متغير مستقل ثالث يحسب وفق الصيغة الاتية:

$$r_{12.3} = \frac{r_{12} - (r_{13})(r_{23})}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

اذ ان r_{12} ، r_{13} ، r_{23} تمثل معاملاً الارتباط البسيطة .

وتحسب قيمة اختبار (t_{ij}) العملية بالنسبة للمتغيرين (X_i) و (X_J) وفق الصيغة الاتية:

$$t_{ij} = \frac{r_{ij}.12\dots k\sqrt{n-k}}{\sqrt{1-r_{ij}^2.12\dots k}}$$

اذ ان :-

$(r_{ij}^2.12,\dots \mathrm{k})$: مربع معامل الارتباط الجزئي ما بين (X_i) و (X_J) باعتبار ان بقية المتغيرات المستقلة ثابتة.

لاختبار فرضية العدم التالية : $H_0 : r_{ij}.12\dots k = 0$

مقابل الفرضية البديلة: $H_1 : r_{ij}.12\dots k \neq 0$

ثم تقارن قيمة (t_{ij}) العملية (المحسوبة) مع نظيرتها الجدولية لدرجة حرية (n–k) ومستوى دلالة معين. فاذا كانت القيمة المحسوبة اكبر من القيمة الجدولية ترفض (H_0) وتقبل (H_1) ، أي ان الارتباط الجزئي بين (X_J) و (X_i) معنوي. وباجراء هذا الاختبار لكافة المتغيرات المستقلة ، والتي تبين من اختبار (F) بانها تشكل مصدرا لمشكلة التعدد الخطي ، وبذلك تشخص بشكل نهائي المتغيرات المستقلة التي تكون سببا في حصول مشكلة التعدد الخطي.

اختبار كلاين : Klien Test

يستخدم هذا الاختبار للكشف عن وجود التعدد الخطي ، حيث يتم مقارنة معامل التحديد R^2 مع مربع معامل الارتباط بين المتغيرات المستقلة، فاذا كان معامل التحديد R^2 اكبر من مربع معامل الارتباط بين المتغيرات المستقلة فهذا يعني عدم وجود مشكلة التعدد الخطي وان كان موجودا فهذا لا يؤثر او يكون غير مؤثر أي ان $R^2 > r^2{}_{xixj}$

ومن الاختبارات الأخرى للكشف عن تعدد العلاقات الخطية هي:

1- التفاوت ومعامل تضخيم التباين: Tolerance and Variance inflation Factor

حيث ان معامل تضخيم التباين يحسب وفق الصيغة الاتية:

$$VIF = \frac{1}{(1 - r^2{}_{12})}$$

وان معامل التفاوت يحسب وفق الصيغة الاتية :

$$T_0L_J = \frac{1}{VIF_J} = \left(1 - R^2{}_J\right)$$

طرق معالجة مشكلة التعدد الخطي :-

وتتلخص بما يلي :

1- محاولة توسيع حجم العينة وذلك بإضافة بيانات كافية عن متغيرات الظاهرة المدروسة، اذ تزداد التقديرات دقة بزيادة عدد البيانات التي تعتمد عليها في عملية التقدير نظراً لوجود علاقة عكسية بين حجم العينة وقيمة التباين ، فكلما كبر حجم العينة كلما تم الحصول على معلومات إضافية تساعد على تخفيض حجم التباين .

2- حذف المتغير المستقل او المتغيرات المستقلة التي تسببت في ظهور المشكلة في النموذج، ولكن غالبا ما يستبدل هذا الحل للمشكلة بمشكلة أخرى، اذ ان حذف متغير مستقل معين له أهميته التفسيرية يوقع الباحث بمشكلة التوصيف (عدم ادخال المتغيرات المهمة في النموذج) مما يرفع احتمال تحيز المقدرات في تلك الحالة.

3- تحويل شكل الدالة باستعمال النسب او الفروقات عوضا عن المتغيرات الاصلية. فعلى سبيل المثال النموذج التالي:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u_i$$

يمكن اختيار أحد المتغيرات المستقلة وليكن مثلا X_2 كمقام (نقسم طرفي النموذج عليه) فنحصل على متغيرات جديدة كالاتي :

$$\frac{Y_i}{X_2} = \beta_0 \frac{1}{X_2} + \beta_1 \frac{X_1}{X_2} + \beta_2 + \frac{u_i}{X_2}$$

غير انه يلاحظ ان النموذج الجديد قد لا يستوفي احد فروض (OLS) الاعتيادية الا وهو ثبات التباين (أي انه لا يمتلك تباين ثابت لحدود الخطأ $(\frac{u_i}{X_2}$ حيث ان

$$E(\frac{u_i}{X_2})^2 = \frac{E\left(u_i^2\right)}{X_2^2} = \frac{\sigma_u^2}{X_2^2} \neq \sigma^2$$

مما يعني استبدال المشكلة بمشكلة أخرى.

4- استخدام أسلوب الدمج بين بيانات السلاسل الزمنية والبيانات المقطعية ، فمثلا يمكن استخدام المعلمة المستخرجة بواسطة البيانات المقطعية مع العلاقة المقدرة بواسطة السلاسل الزمنية ، كاستخدام مقدر معلمة الميل الحدي للاستهلاك لفترة معينة ولقطر معين من بيانات المقاطع العرضية مع العلاقة بين الدخل والاسعار لنفس الفترة والقطر من بيانات السلاسل الزمنية، وفضلاً عن ذلك توجد طرق أخرى يمكن استخدامها لمعالجة مشكلة التعدد الخطي نذكر منها أسلوب او طريقة ال (Ridge Regression) (طريقة او أسلوب انحدار الحرف) وطريقة المركبات الرئيسية (Principle Component) .

مثال (1) : في عينة عشوائية ذات حجم n=15 فيها المتغير (Y_i) وعلاقتة بثلاث متغيرات مستقلة (X_1 , X_2 , X_3) وتوفرت لديك المعلومات التالية:

$$r_{12} = 0.8811 \quad , \quad r_{13} = 0.9399 \quad , \quad r_{23} = 0.9866$$

والمطلوب: اختبر لوجود مشكلة التعدد الخطي:

الحل:- لغرض اجراء الاختبار فانه يتطلب احتساب قيمة محدد مصفوفة معاملات الارتباط بين المتغيرات المستقلة أي ان

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{bmatrix}$$

وعليه فان محدد مصفوفة معاملات الارتباط البسيطة هو :

$$|R| = \begin{vmatrix} 1 & 0.8811 & 0.9399 \\ 0.8811 & 1 & 0.9866 \\ 0.9399 & 0.9866 & 1 \end{vmatrix}$$

وباستخدام الطريقة العامة او الخاصة لإيجاد قيمة المحدد من الرتبة *(3X3)* علماً ان القيم بالانحرافات حيث ان :

$$|R| = 0.000969$$

$$rx_1x_2 = \frac{\sum x_1x_2}{\sqrt{\sum x_1^2}\,\sqrt{\sum x_2^2}}$$

$$\boldsymbol{\mathcal{X}^2} = -\left[n - 1 - \frac{1}{6}(2k + 5)\right] Ln\,|R|$$

$$\therefore \boldsymbol{\mathcal{X}^2} = -\left[15 - 1 - \frac{1}{6}(2(3) + 5)\right] Ln\,(0.000969)$$

$$= -\left[14 - \frac{1}{6}(11)\right](-6.939)$$

$$= -[14 - 1.833](-6.939)$$

اي ان $\boldsymbol{\mathcal{X}^2}$ المحسوبة هي $\boldsymbol{\mathcal{X}^2} = 84.427$

اما قيمة $\boldsymbol{\mathcal{X}^2}$ الجدولية بدرجة حرية 2 /k(k-1) *والتي تساوي* 2/ (2) 3=3 ولمستوى معنوية (0.05) فأنها تبلغ (7.816) وحيث ان قيمة $\boldsymbol{\mathcal{X}^2}$ المحسوبة هي اكبر من قيمة $\boldsymbol{\mathcal{X}^2}$ الجدولية أي ان$84.427 > 7.816$

وعليه ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة (H_1) أي ان هناك مشكلة التعدد الخطي بين المتغيرات المستقلة في النموذج .

مثال (2) : عينة عشوائية ذات حجم n=10 فيها المتغير (Y_i) وعلاقته بأربعة متغيرات مستقلة هي $(X_1, X_2, X_3, X_4,)$ والمطلوب اختبر لوجود مشكلة التعدد الخطي ، ثم بين أي متغير مستقل يكون مصدراً لهذه المشكلة .

Y_i	X_1	X_2	X_3	X_4
6.0	40.1	5.5	108	**63**
6.0	40.3	4.7	94	**72**
6.5	47.5	5.2	108	**86**
7.1	49.2	6.8	100	**100**
7.2	52.3	7.3	99	**107**
7.6	58.0	8.7	99	**111**
8.0	61.3	10.2	101	**114**
9.0	62.5	14.1	97	**116**
9.0	64.7	17.1	93	**119**
9.3	66.8	21.3	102	**121**

الحل :- من البيانات أعلاه يمكن الحصول على البيانات الحسابية التالية:

$$\sum X_1 = 5427 \;, \sum X_2 = 100.9 \;, \sum X_3 = 1001 \;, \sum X_4 = 1009$$

$$\sum X_1^2 = 30320.55 \;; \; \sum X_2^2 = 1303.55 \;; \; \sum X_3^2 = 100429 \;; \; \sum X_4^2 = 105573$$

$$\sum X_1X_2 = 5913.63 \quad ; \quad \sum X_1X_3 = 54173.2 \quad ; \quad \sum X_1X_4 = 56487.3$$

$$\sum X_2X_3 = 100222 \quad ; \quad \sum X_2X_4 = 10969.5 \quad ; \quad \sum X_3X_4 = 100617$$

وان :

$$n = 10 \quad ; \quad \bar{X}_1 = 54.27 \quad ; \quad \bar{X}_2 = 10.09 \quad ; \quad \bar{X}_3 = 100.1 \; ; \; \bar{X}_4 = 100.9$$

ولغرض اجراء الاختبار يستوجب احتساب قيمة محدد مصفوفة معاملات الارتباط البسيطة بين المتغيرات المستقلة والتي تأخذ الشكل الاتي :

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} \\ r_{21} & 1 & r_{23} & r_{24} \\ r_{31} & r_{32} & 1 & r_{34} \\ r_{41} & r_{42} & r_{43} & 1 \end{bmatrix}$$

ولغرض تسهيل العمليات الحسابية تم تحويل البيانات أعلاه الى الانحرافات وكالاتي:

$$\sum x_1^2 = \sum X_1^2 - n\bar{X}_1^2 = 868.221$$

$$\sum x_2^2 = \sum X_2^2 - n\bar{X}_2^2 = 285.469$$

$$\sum x_3^2 = \sum X_3^2 - n\bar{X}_3^2 = 228.9$$

$$\sum x_4^2 = \sum X_4^2 - n\bar{X}_4^2 = 3764.9$$

$$\sum x_1x_2 = \sum X_1X_2 - n\bar{X}_1\bar{X}_2 = 437.787$$

$$\sum x_1x_3 = \sum X_1X_3 - n\bar{X}_1\bar{X}_3 = -151.07$$

$$\sum x_1x_4 = \sum X_1X_4 - n\bar{X}_1\bar{X}_4 = 1728.87$$

$$\sum x_2x_3 = \sum X_2X_3 - n\bar{X}_2\bar{X}_3 = -77.89$$

$$\sum x_2x_4 = \sum X_2X_4 - n\bar{X}_2\bar{X}_4 = 788.69$$

$$\sum x_3x_4 = \sum X_3X_4 - n\bar{X}_3\bar{X}_4 = -383.9$$

ومنها تم احتساب معاملات الارتباط البسيطة بموجب الصيغة العامة الاتية:

$$r_{x_ix_J} = \frac{\sum x_ix_J}{\sqrt{\sum x_i^2}\sqrt{\sum x_J^2}}$$

وكانت النتائج كالاتي:

$$r_{x_1x_2} = r_{12} = 0.879 \quad , \quad r_{x_1x_3} = r_{13} = -0.339$$

$$r_{x_1x_4} = 0.956 \quad , \quad r_{x_2x_3} = -0.305 \quad , \quad r_{x_2x_4} = 0.761$$

$$r_{x_3x_4} = r_{34} = -0.414$$

ولذلك فان المصفوفة (R) ستأخذ الشكل الاتي:

$$R = \begin{bmatrix} 1 & 0.879 & -0.339 & 0.956 \\ 0.879 & 1 & -0.305 & 0.761 \\ -0.339 & -0.305 & 1 & -0.414 \\ 0.956 & 0.761 & -0.414 & 1 \end{bmatrix}$$

وان $|R| = 0.0098$

وان الفرضية المطلوب اختبارها هي:

H_0 : المتغيرات X_J متعامدة (مستقلة غير مرتبطة) ، لا توجد مشكلة التعدد الخطي .

H_1 : المتغيرات X_J غير متعامدة (غير مستقلة) بمعنى توجد مشكلة تعدد خطي.

وصيغة الاختبار هي:

$$\boldsymbol{\mathcal{X}}^2 = -\left[n - 1 - \frac{1}{6}(2k+5)\right] Ln\,|R|$$

$$= -\left[10 - 1 - \frac{1}{6}\,((2*4)+5)\right] Ln\,(0.0098)$$

$$= -(6.8334)(-4.6253729)$$

$$\therefore \boldsymbol{\mathcal{X}}^2 = 31.606699$$

ومن جداول $(\boldsymbol{\mathcal{X}}^2)$ ولمستوى دلالة (0.05) ودرجة حرية (6) نجد ان قيمة $\boldsymbol{\mathcal{X}}^2$ الجدولية مساوية الى (12.592) وحيث ان : $31.606699 > 12.592$

عليه ترفض فرضية العدم وتقبل الفرضية البديلة (H_1) أي ان هناك مشكلة تعدد خطي بين المتغيرات المستقلة.

ولتحديد مصدر هذا التعدد الخطي يجب اجراء اختبار (F) وفق الخطوات الاتية مبتدئين بالمتغير المستقل (X_1) حيث ان:

$$F_{X_1} = \frac{R^2_{X_1}\,.\,X_2\,,X_3\,,X_4/(k-1)}{(1 - R^2_{X_1}\,.\,X_2\,,X_3\,,X_4)/(n-k)} = \frac{0.972926/2}{(1-0.972926)/7}$$

$$\therefore F_{X_1} = 125.779$$

والفرضية المطلوب اختبارها هي:

$$H_0 = R^2_{X1}\,.\,X_2\,,X_3\,,X_4 = 0$$

$$H_1 = R^2_{X1}\,.\,X_2\,,X_3\,,X_4 \neq 0$$

ومن جداول (F) الإحصائية ولمستوى معنوية (5%) ودرجة حرية مساوية الى (7,2) نجد ان قيمة F الجدولية تساوي (4.74) وحيث ان $125.779 > 4.74$ لذلك ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة (H_1) مما يعني ان المتغير المستقل X_1 يكون مصدرا لوجود مشكلة التعدد الخطي بعبارة أخرى ان المتغير المستقل X_1 مرتبط خطياً بالمتغيرات المستقلة الأخرى .

ثم ننتقل الى اختبار المتغير المستقل الثاني (X_2) أي ان:

$$F_{X_2} = \frac{R^2_{X_2} \,.\, X_3 \,, X_3 \,, X_4/(k-1)}{(1-R^2_{X_2} \,.\, X_1 \,, X_3 \,, X_4)/(n-k)} = \frac{0.858/2}{(1-0.858)/7}$$

$$\therefore F_{X_2} = 21.148$$

وحيث ان : $$21.148 > 4.74$$

لذلك يستنتج بان المتغير المستقل (X_2) مرتبط خطيا أي انه مصدر لوجود مشكلة التعدد الخطي ثم ننتقل الى اختبار المتغير المستقل الثالث X_3 حيث ان:

$$F_{X_3} = \frac{0.261/2}{0.739/7} = 1.236$$

وحيث ان : $$1.236 < 4.74$$

لذلك تقبل فرضية العدم (H_0) أي ان المتغير X_3 غير مرتبط خطيا وبالتالي فانه لا يشكل مصدرا لمشكلة التعدد الخطي .

اما فيما يتعلق بالمتغير المستقل الرابع (X_4) فان :

$$F_{X_4} = \frac{0.9524773/2}{(1-0.9524773)/7} = 70.148549$$

وحيث ان : $$70.148549 > 4.74$$

عليه ترفض فرضية العدم (H_0) ، أي ان المتغير المستقل (X_4) مرتبط خطيا وبالتالي فانه مصدرا لمشكلة التعدد الخطي .

ولغرض تشخيص المتغيرات المستقلة المسببة بشكل نهائي في حصول مشكلة التعدد الخطي لا بد من اجراء اختبار (t) ما بين كل اثنين من المتغيرات المستقلة والتي ثبت من خلال اختبار (F) بانها مصدر لنشوء التعدد الخطي.

تمارين الفصل السادس

الأول: ماذا يقصد بمشكلة التعدد الخطي، وما هي انواعه، وضح باختصار طرق معالجة المشكلة في بيانات العينة؟

الثاني: لنموذج الانحدار الخطي المتعدد الاتي:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u_i \qquad i = 1,2, \dots , n$$

بافتراض وجود علاقة بين المتغيرين X_1 و X_2 ماهي الطريقة الأنسب لتقدير معالم النموذج أعلاه في الحالات التالية ذاكراً السبب :

1- وجود علاقة خطية غير تامة بين المتغيرين X_1 و X_2 ؟

2- وجود علاقة خطية تامة بين المتغيرين X_1 و X_2 ؟

الثالث: قام احد الباحثين بدراسة دالة الطلب على سلعة معينه Y_i فوجد ان هناك عاملين أساسيين يؤثران فيها هما سعر السلعة (X_1) وسعر سلعة أخرى منافسة (X_2) وكما مبين في الجدول ادناه:

Y_i	X_1	X_2
9	1	6
7	2	6
5	3	5
5	4	2
4	5	1

المطلوب:-

1- تقدير معلمات العلاقة الخطية بين Y_i ; X_1 ; X_2 ؟

2- الاختبار لوجود مشكلة التعدد الخطي؟

3- ماهي المعالجات الضرورية في حالة وجود المشكلة؟

الفصل السابع

مشكلة عدم تجانس التباين The Heteroscedasticity Problem

المقدمة :-

من الافتراضات الأساسية التي يقوم عليها النموذج الخطي (البسيط والعام) هي ثبات التباين لحدود الخطأ (تجانس تباين الخطأ) Homoscedasticity لجميع المشاهدات (i) والتي يعبر عنها في النموذج الخطي البسيط بالشكل الاتي :

$$E\left(u_i^2\right) = \sigma_u^2$$

$$E\left(u_i u_j\right) = 0 \quad \forall \ i \neq j$$

والتي تم وضعها على شكل مصفوفات وموجهات في حالة النموذج الخطي العام كالاتي:

$$E(u\acute{u}) = \sigma_u^2 \, \mathrm{I}_n$$

وبموجب هذه الفرضية فان $\sigma_{u_1}^2 = \sigma_{u_2}^2 = \cdots = \sigma_{u_n}^2$

وفي واقع الامر فأننا كثيراً ما نواجه حالات يتعذر فيها استيفاء الشروط أعلاه، ومن ثم فان التباين لا يكون ثابتا بل يختلف لكل مشاهدة من مشاهدات العينة وتصبح لدينا قيم مختلفة وغير ثابتة لتباينات حدود الخطأ العشوائية، وعليه فان القطر الرئيس لمصفوفة التباين والتباين المشترك الخاصة بحدود الخطأ يحتوي على قيم مختلفة وغير ثابتة أي ان:

$$E(u\acute{u}) \neq \sigma_u^2 \, \mathrm{I}_n$$

أي ان $\sigma_{u_1}^2 \neq \sigma_{u_2}^2 \neq \cdots \neq \sigma_{u_n}^2 \neq \sigma_u^2$

ويحدث هذا أساساً في الدراسات القياسية وخاصة الدراسات التي تعتمد على البيانات الإحصائية التي تأخذ شكل البيانات المقطعية Cross – Section Data كما هو الحال في بيانات بحوث ميزانية الاسرة التي تشتمل على أسراً متباينة بشكل كبير في مستويات دخولها، فتشتت مشاهدات البيانات المقطعية الخاصة بالمتغير المعتمد قد يختلف اختلافا شديدا من مستوى الى أخر من مستويات المتغير المستقل . مثال ذلك دراسة دالة الاستهلاك التي تعتمد على دخل وانفاق العوائل على مختلف السلع والخدمات ، فالعوائل ذات الدخول المرتفعة تتمتع بمرونة كبيرة في الانفاق ، أما انفاق العوائل ذات الدخول الواطئة فانه يقع عادة ضمن حدود ضيقة ، وعليه فان التباين عند قيم الدخول الكبيرة يكون أكبر من التباين عند قيم الدخول الصغيرة. وهكذا

نجد أن فرضية تجانس تباين الخطأ تصبح عديمة الجدوى في مثل هذه الحالات وخرق الافتراض هذا يؤدي الى حدوث (ظهور) مشكلة تدعى مشكلة عدم تجانس التباين .

وفي ظل وجود مشكلة عدم تجانس التباين أعلاه يكون استخدام طريقة (OLS) لتقدير معالم النموذج غير مجدي، حيث ان المعالم المقدرة تمثل هذا الأسلوب سوف لن تكون افضل تقدير خطي غير متحيز (BLUE) . بعبارة أخرى لا تمتلك المعالم المقدرة بهذا الأسلوب خاصية اقل تباين ممكن.

أسباب عدم تجانس تباين حد الخطأ :

هناك عدة أسباب لظهور هذه المشكلة ومن أهمها ما يلي :

1- سلوكية وتصرف الافراد التي تقل الأخطاء فيها بمرور الزمن، وعليه فان تباين حد الخطأ (σ_i^2) يتناقص هو الاخر خلال الفترة الزمنية .

2- يزداد تباين حد الخطأ (σ_i^2) مع زيادة مستوى الدخل وذلك لتباين وتعدد إختيارات الناس في سلوكهم مثال ذلك تباين الانفاق على المواد الغذائية بين الأسر يمكن ان يزيد بزيادة دخل الأسرة .

3- مع تحسن أساليب جمع البيانات يقل تباين حد الخطأ(σ_i^2) ، ذلك لان جمع البيانات الدقيقة والواقعية تقلل من الأخطاء ، مثال ذلك ان الأخطاء التي ترد في مستندات المؤسسات الحكومية والتي تستخدم الحاسب الألي في جمع وتحليل البيانات تكون اقل من مثيلتها في المؤسسات التي لا تستخدم الحاسب الالي .

النتائج (الاثار) المترتبة على وجود مشكلة عدم تجانس التباين :.

وتتلخص بالاتي :

1- أن استخدام طريقة (OLS) لتقدير معالم النموذج تكون غير مجدية (غير فعالة) أي أن المقدرات التي نحصل عليها تكون غير موثوقة؟
2- اختبارات المعنوية الإحصائية وايجاد حدود الثقة تكون غير دقيقة وغير تطبيقية ؟
3- التقديرات المستحصلة لا تتسم بخاصية الكفاءة لعدم تحقق شرط صغر التباين ؟
4- التقديرات تكون خطية وغير متحيزة ؟
5- عدم دقة التنبؤات المستحصلة ؟

اكتشاف عدم تجانس تباين حد الخطأ :

هناك عدة طرق (اختبارات) يمكن بواسطتها اختبار فيما اذا كان النموذج يعاني من مشكلة عدم تجانس التباين أم لا، منها :

أولا : اختبار كولد فيلد وكوانت Gold feld And Quandt Test

يعد من الاختبارات المهمة لغرض الكشف عن مشكلة عدم تجانس تباين الخطأ، ويتم استخدامه في حالة العينات كبيرة الحجم، وتتلخص خطوات اجراء هذا الاختبار بما يلي:

1- ترتيب البيانات الخاصة بالمتغير المستقل X_i من أصغر قيمة الى أكبر قيمة (وفقاً لقيمة وإشارة المتغير المستقل) .

2- حذف المشاهدات الوسطية من بيانات العينة ويفضل حذف $\left(\frac{1}{5}\right)$ من المشاهدات الكلية .

3- تقسيم المشاهدات الباقية الى عينتين جزئيتين متساويتين تتضمن العينة الأولى على قيم X_i الصغيرة وتتضمن العينة الثانية على قيم X_i الكبيرة.

4- يتم تقدير معلمات العلاقة الخطية بين المتغير التابع (Y_i) والمتغير المستقل (X_i) لكل عينة جزئية على انفراد.

5- يتم احتساب تباين الخطأ للعينة الجزئية الأولى (S_{e1}^2) ، وللعينة الجزئية الثانية (S_{e2}^2) . بموجب الصيغة الاتية:

أ- بالنسبة للعينة الجزئية الأولى فان:

$$S_{e1}^2 = \frac{\sum e_{i1}^2}{n_1 - 2}$$

اذ ان :

$(\sum e_{i1}^2)$: مجموع مربعات البواقي للعينة الأولى وان $e_i = Y_i - \hat{Y}_i$

(n_1): حجم العينة الجزئية الأولى

(2): عدد معلمات النموذج الخطي البسيط (b_1 , b_0)

ب- بالنسبة للعينة الجزئية الثانية فان:

$$S_{e2}^2 = \frac{\sum e_{i2}^2}{n_2 - 2}$$

6- إحتساب إحصاءة (اختبار) F وفق الصيغة الاتية :

$$F = \frac{S_{e2}^2}{S_{e1}^2}$$

ثم تقارن قيمة F المحسوبة مع قيمة F^* الجدولية عند مستوى معنوية معين (1% او 5%) ودرجة حرية قدرها $(n_1 - 2\ ,\ n_2 - 2)$ للبسط والمقام ، فاذا كانت قيمة F المحسوبة اكبر من قيمة F^* الجدولية ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة (H_1) والتي تنص على وجود مشكلة عدم تجانس تباين الخطأ أي ان :

$$H_1: \sigma_1^2 \neq \sigma_2^2 \neq \cdots \neq \sigma_n^2$$

وبالعكس اذا كانت قيمة F المحسوبة أصغر من قيمة F^* الجدولية عندئذ تقبل فرضية العدم (H_0) وترفض الفرضية البديلة (H_1) مما يعني عدم وجود مشكلة عدم تجانس التباين (أي تجانس تباين حد الخطأ) حيث أن :

$$H_o: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2$$

مثال:- في ضوء البيانات الإحصائية الاتية التي تبين الانفاق الاستهلاكي (Y) والدخل القابل للتصرف (X) في اقتصاد احدى الدول للمدة 1983-1994؟

السنة	Y	X	السنة	Y	X
1983	26.1	38.3	1989	50.1	77.2
1984	29.3	43.5	1990	54.5	86.1
1985	35.6	53.5	1991	60.1	94.6
1986	39.4	60.8	1992	64.9	102.4
1987	42.7	66.4	1993	69.2	109.9
1988	46.3	71.2	1994	73.1	115.6

المطلوب:- الاختبار لوجود مشكلة عدم تجانس التباين مستخدماً اختبار كولد فيلد وكوانت Gold feld And Quandt Test.

الحل :- البيانات الخاصة بالمتغير المستقل (X_i) مرتبة في السؤال من أصغر قيمة الى أكبر قيمة ثم نحذف $\left(\frac{1}{5}\right)$ من المشاهدات الوسطية ، أي ان $2.5 = 12\left(\frac{1}{5}\right)$ مشاهدة ، لذلك تحذف المشاهدتين الخاصتين بعامي 1988 و 1989 لكي تقسم البيانات الى مجموعتين جزئيتين متساويتين تضم كلا منها (5) مشاهدات وكالاتي:

المجموعة الجزئية الأولى: - لتقدير معالم النموذج الخطي البسيط لبيانات هذه العينة وحساب الأخطاء العشوائية فان ذلك يتطلب حساب البيانات الاتية:

حيث ان $\bar{X} = 52.5$, $\bar{Y} = 34.62$

السنة	Y_i	X_i	$y_i = Y_i - \bar{Y}$	$x_i = X_i - \bar{X}$	$x_i y_i$	x_i^2	$\hat{Y}_i$
1983	26.1	38.3	−8.52	−14.2	120.984	201.64	26.246075
1984	29.3	43.5	−5.32	−9.0	47.88	81.0	29.312583
1985	35.6	53.5	0.98	1.0	0.98	1.0	35.209713
1986	39.4	60.8	4.78	8.3	39.674	68.89	39.514618
1987	42.7	66.4	8.08	13.9	112.312	193.21	42.817011
$n_1 = 5$	173.1	262.5	0	0	321.83	545.74	173.1

وان : $Y_i = b_o + b_1 X_1 + e_i$

$e_i = Y_i - \hat{Y}_i$	e_{i1}^2
−0.146075	0.0213379
−0.012583	0.0001583
0.390287	0.1523239
−0.114618	0.0131372
−0.117011	0.0136915
0	0.2006488

$$\therefore \quad b_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{321.83}{545.74}$$

$$b_1 = 0.589713$$

وان :

$$b_0 = \bar{Y} - b_1\bar{X} = 34.62 - (0.589713)(52.5)$$

$$b_0 = 3.660068$$

لذلك فان المعادلة التقديرية لبيانات العينة الجزئية الأولى هي :

$$\hat{Y}_i = 3.660068 + 0.589713\ X_i$$

ولذلك فان :

$$S_{e1}^2 = \frac{\sum e_{i1}^2}{n_1 - 2} = \frac{0.2006488}{5-2} = 0.0668829$$

المجموعة الجزئية الثانية:- لتقدير معالم النموذج الخطي البسيط لهذه العينة وحساب الأخطاء العشوائية فان ذلك يتطلب اجراء العمليات الحسابية التالية:

السنة	Y_i	X_i	$y_i = Y_i - \bar{Y}$	$x_i = X_i - \bar{X}$	$x_i y_i$	x_i^2	y_i^2
1990	54.5	86.1	−9.86	−15.62	154.0132	243.9844	**97.2196**
1991	60.1	94.6	−4.26	−7.12	30.3312	50.6944	**18.1476**
1992	64.9	102.4	0.54	0.68	0.3672	0.6424	**0.2916**
1993	69.2	109.9	4.84	8.18	39.5912	66.9124	**23.4256**
1994	73.1	115.6	8.74	13.88	121.3112	192.6544	**76.3876**
$n_2 = 5$	321.8	508.6	0	0	354.614	554.708	215.472

وان :

$$\bar{Y} = 64.36 \quad ; \quad \bar{X} = 101.72$$

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{354.614}{554.708} = 0.6230557$$

$$b_0 = \bar{Y} - b_1 \bar{X} = 64.36 - (0.6230557)(101.72)$$

$$\therefore b_0 = 0.982775$$

لذلك فان المعادلة التقديرية للعينة الثانية هي :

$$\hat{Y}_i = 0.982775 + 0.6230557\, X_i$$

وان :

$$\sum \hat{y}_i^2 = b_1 \sum x_i y_i = (0.6230557)(354.614) = 215.33677$$

اذ ان :

$$\sum e_i^2 = \sum y_i^2 - \sum \hat{y}_i^2 = 215.472 - 215.33677 = 0.13523$$

او تحسب $\sum e_i^2$ وفق الصيغة الاتية :

$$\sum e_i^2 = \sum y_i^2 - b_1 \sum x_i y_i$$

لذلك فان :

$$S_{e2}^2 = \frac{\sum e_{i2}^2}{n_2 - 2} = \frac{0.13523}{5-2} = 0.0450766$$

$$\therefore F = \frac{S_{e2}^2}{S_{e1}^2} = \frac{0.0450766}{0.0668829} = 0.673963$$

وبمقارنة قيمة F المحسوبة والبالغة (0.673) مع قيمتها الجدولية F^* والبالغة (9.28) عند مستوى معنوية (5%) ودرجة حرية (3 ، 3) للبسط والمقام يتضح بان (F) المحسوبة هي اقل من (F^*) الجدولية ، عليه تقبل فرضية العدم (H_0) والتي تنص على تجانس تباين الخطأ مما يعني عدم وجود مشكلة عدم تجانس التباين .

ثانيا :- اختبار معامل ارتباط الرتب لسبيرمان :-

Spearman's Rank Correltion Coefficient Test

يعد هذا الاختبار من أبسط وأسهل الاختبارات المستخدمة في اكتشاف مشكلة عدم تجانس التباين، ويمكن تطبيقه على العينات الكبيرة والصغيرة يعتمد هذا الاختبار على احتساب معامل ارتباط الرتب لسبيرمان بي القيم المطلقة للأخطاء $|e_i|$ (أي اهمال إشارة e_i) وقيم المتغير المستقل (X_i) موضوع الدراسة ويتطلب احتساب هذا المؤشر ما يلي :

1- تقدير معلمات النموذج باستخدام طريقة (OLS) .

2- إحتساب البواقي e_i ، حيث ان $e_i = Y_i - \hat{Y}_i$.

3- ترتيب قيم المتغير المستقل (X_i) والانحرافات المطلقة $|e_i|$ تصاعدياً أو تنازلياً ، وإعطاء كل منها رتبا معينة (تمثل سلسلة أعداد طبيعية (1,2,3,...n) وفق تسلسل القيم ثم تحسب فروقات الرتب وتربع، ومن ثم يستخرج معامل ارتباط الرتب ما بين القيم المطلقة للانحرافات وقيم المتغير المستقل المعني وفق قانون سبيرمان Spearman's لارتباط الرتب الاتي:

$$r_s = 1 - \left[\frac{6\sum_{i=1}^{n} D_i^2}{n(n^2-1)}\right]$$

اذ ان :

(n) : حجم العينة

(D_i^2) : الفرق ما بين رتب القيم المطلقة للأخطاء $(|e_i|)$ ورتب المتغير المستقل (X_i) المناظرة لها.

وكلما اقتربت قيمة r_s من الواحد الصحيح دل ذلك على وجود علاقة قوية بين e_i و X_i ومن ثم وجود مشكلة عدم تجانس التباين .

ويمكن التأكد من وجود المشكلة من خلال اختبار *(t)* على وفق الصيغة الاتية:

$$t = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$$

وبمقارنة قيمة *(t)* المحسوبة مع نظيرتها الجدولية (t^*) عند درجة حرية *(n−k−1)* ومستوى معنوية 1% أو 5% فاذا كانت المحسوبة أكبر من الجدولية ترفض فرضية العدم (H_0) وتقبل الفرضية البديلة (H_1) مما يعني وجود مشكلة عدم تجانس التباين اذ أن :

$$H_0: \sigma_1^2 = \sigma_2^2 = \ \dots \dots . \ = \sigma_n^2 \qquad (homoscedastic)$$

$$H_1: \sigma_1^2 \neq \sigma_2^2 \neq \ \dots \dots . \neq \ \sigma_n^2 \qquad (heteroscedastic)$$

اما في حالة كون قيمة *(t)* المحسوبة هي اقل من قيمة (t^*) الجدولية لذلك تقبل فرضية العدم (H_0) وترفض الفرضية البديلة (H_1) مما يعني عدم وجود مشكلة عدم تجانس التباين (تجانس تباين الخطأ).

ثالثا:- اختبار بارك كليجسر:- *Park - Glejser Test*

من الاختبارات الأخرى المستخدمة للكشف عن عدم تجانس التباين هو اختبار بارك – كليجسر ويعتمد هذا الاختبار على شكل العلاقة بين الأخطاء العشوائية (e_i) والمتغير المستقل الذي تعتمد عليه تلك الأخطاء . وتتمثل الخطوة الأولى في هذا الاختبار بتقدير معالم العلاقة الخطية ثم إيجاد قيم الأخطاء العشوائية الناتجة من الفرق بين القيم الحقيقية والقيم التقديرية للمتغير المعتمد. اما الخطوة الثانية فيتم فيها تحديد صيغ مختلفة وحسب نوع العلاقة التي يعتقد بانها موجودة ما بين القيم المطلقة للأخطاء العشوائية وقيم المتغير المستقل ومن الأمثلة على هذه الصيغ المختلفة الاتي :

$$|e_i| = \beta_0 + \beta_1 X_i$$

$$|e_i| = \beta_0 + \beta_1 \sqrt{X_i}$$

$$|e_i| = \beta_0 + \beta_1 / X_i$$

والخطوة الأخيرة في هذا الاختبار تتمثل في إجراء اختبار معنوية العلاقة المفترضة حيث يستدل منها فيما اذا كانت هناك مشكلة عدم تجانس التباين او لا، وتحديد صيغة العلاقة بين الأخطاء العشوائية والمتغير المستقل ثانيا.

وتجدر الإشارة هنا الى ان هذا الاختبار يعتمد في تحديد الصيغة الملائمة على التجربة .

رابعا:- اختبار بارتليت Bartlett Test

معالجة مشكلة عدم تجانس التباين :-

اوضحنا سابقاً ان خرق فرضية ثبات التباين لحدود الخطأ يؤدي الى وجود قيم مختلفة وغير ثابتة لتباينات حدود الخطأ العشوائية، ومن ثم فان القطر الرئيس لمصفوفة التباين والتباين المشترك الخاصة بحدود الخطأ يحتوي على قيم مختلفة وغير ثابتة أي ان :

$$E(u\acute{u}) \neq \sigma^2 I_n$$

اذ ان :

$$\sigma_1^2 \neq \sigma_2^2 \neq \ldots\ldots \neq \sigma_n^2$$

وتتمثل معالجة مشكلة عدم تجانس التباين بتحديد الأوزان ومن ثم استخدام هذه الاوزان في تحويل صيغة النموذج الى صيغة أخرى ينتج عنها قيم متساوية في القطر الرئيس للمصفوفة $\sigma^2 I_n$. ولتحقيق هذا الغرض توجد طرق مختلفة منها :

أ–عندما يزداد تباين المتغير التابع (Y_i) بشكل تناسبي مع قيمة الوسط $(\bar{X})$ فانه يمكن تحويل المتغيرات التي تعاني من مشكلة عدم تجانس التباين من خلال قسمة طرفي العلاقة الاتية:

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

على المتغير المستقل X_i وكما في الصيغة الاتية:

$$\frac{Y_i}{X_i} = \frac{\beta_0}{X_i} + \beta_1 \frac{X_i}{X_i} + \frac{u_i}{X_i}$$

مع ملاحظة ان الحد الثابت في المعادلة الرئيسية β_0 أصبح يمثل معامل الانحدار أي β_1 .

وان β_1 أصبح يمثل الحد الثابت ، مما يجعل معامل التحديد (R^2) مختلف في المعادلتين لاختلاف قيم المتغير التابع حيث كان (Y_i) في المعادلة الأولى وأصبح ($\frac{Y_i}{X_i}$) في المعادلة الثانية .

ب–عندما يزداد تباين المتغير التابع (Y_i) بشكل تناسبي مع الزيادة في قيم X_i فتتم عملية التحويل بقسمة طرفي معادلة الانحدار على $\sqrt{X_i}$ أي ان :

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

تأخذ الشكل الاتي:

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\beta_0}{\sqrt{X_i}} + \beta_1 \frac{X_i}{\sqrt{X_i}} + \frac{u_i}{\sqrt{X_i}}$$

وهنا العلاقة بين المتغيرين X_i و Y_i أصبحت علاقة بين ثلاث متغيرات هي $\frac{Y_i}{\sqrt{X_i}}$ ، $\frac{1}{\sqrt{X_i}}$ ، $\sqrt{X_i}$ علما بان قيمة الحد الثابت بعد التحويل تصبح صفراً كما وان قيمة (R^2) تصبح اكبر ، مما يجعل تقدير القيمة المتوقعة للمتغير التابع (Y) غير دقيقة.

مثال :- في بيانات المثال السابق الخاصة بالإنفاق الاستهلاكي (Y_i) والدخل القابل للتصرف (X_i) في اقتصاد احدى الدول للمدة (1983، 1994) المطلوب استخدام معامل ارتباط الرتب لسبيرمان لاختبار وجود او عدم وجود مشكلة عدم تجانس التباين .

السنة	Y_i	X_i	$x_i = X_i - \bar{X}$	$y_i = Y_i - \bar{Y}$	$x_i y_i$	x_i^2	$\hat{Y}_i$
1983	26.1	38.3	−38.325	−23.175	888.18187	1468.8056	**26.113903**
1984	29.3	43.5	−33.125	−19.975	661.67187	1097.2656	**29.256439**
1985	35.6	35.5	−23.125	−13.675	316.23437	534.76562	**35.299778**
1986	39.4	60.8	−15.825	−9.875	156.27187	250.43062	**39.711416**
1987	42.7	66.4	−10.225	−6.575	67.229375	104.55062	**43.095685**
1988	46.43	71.2	−5.425	−2.975	16.139375	29.430625	**45.996488**
1989	50.1	77.2	0.575	0.825	0.474375	0.330625	**49.622492**
1990	54.5	86.1	9.475	5.225	49.506875	89.775625	**55.001063**
1991	60.1	94.6	17.975	10.825	194.57937	323.10062	**60.137901**
1992	64.9	102.4	25.775	15.625	402.73437	664.35062	**64.851706**
1993	69.2	109.9	33.275	19.925	663.00437	1107.2256	**69.38421**
1994	73.1	115.6	38.975	23.825	928.57937	1519.0506	**72.828913**
n = 12	**591.3**	**919.5**	**0**	**0**	**4344.6067**	**7189.0822**	**591.3**

لذلك فان

$$\hat{Y}_i = b_0 + b_1 X_i$$

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{4344.6067}{7189.0822} = 0.6043$$

$$b_0 = \bar{Y} - b_1\bar{X} \qquad \Longrightarrow \qquad \bar{Y} = 49.275$$

$$\bar{X} = 76.625$$

$$\therefore b_0 = 49.275 - (0.6043)(76.625) = 2.9705$$

$$\therefore \hat{Y}_i = 2.9705 + 0.6043\, X_i$$

ومنه فأن:

$e_i = Y_i - \hat{Y}_i$	$Rank\ Of\ X_i$	$Rank\ Of\ \lvert e_i \rvert$	D_i	D_i^2
−0.013903	1	1	*0*	**0**
0.043561	2	3	−1	**1**
0.300222	3	7	−4	**16**
−0.311416	4	9	−5	**25**
−0.395685	5	10	−5	**25**
0.303512	6	8	−2	**4**
0.477508	7	11	−4	**16**
−0.501063	8	12	−4	**16**
−0.037901	9	2	7	**49**
0.048294	10	4	6	**36**
−0.18421	11	5	*6*	**36**
0.271087	12	6	6	**36**
0			**0**	**260**

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

$$= 1 - \frac{6(260)}{12\big((12^2)-1\big)} = 1 - \frac{1560}{12(144-1)}$$

$$= 1 - 0.9090909$$

$$\therefore r_s = 0.0909091$$

ولما كانت قيمة (r_s) تساوي (0.09) مما يدل على وجود علاقة ضعيفة بين e_i و X_i مما يشير الى عدم وجود مشكلة عدم تجانس التباين. وللتاكد من ذلك يتم اجراء اختبار (t) وكالاتي:

$$t = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$$

$$\therefore t = \frac{0.0909091\sqrt{12-2}}{\sqrt{1-(0.0909091)^2}} = \frac{0.2874798}{0.9958592}$$

$$t = 0.2886751$$

وبمقارنة قيمة (t) المحتسبة والبالغة (0.288) مع قيمتها الجدولية (t^*) والبالغة (2.23) عند مستوى معنوية (0.05) ودرجة حرية (10) لذلك تقبل فرضية العدم (H_0) وترفض الفرضية البديلة (H_1) والتي تنص على عدم تجانس تباين الخطأ أي قبول فرضية التجانس (عدم وجود المشكلة) .

تمارين الفصل السابع

الأول : يعطي الجدول الاتي بيانات عن كل من X_i و e_i والمطلوب استخدام معامل ارتباط الرتب لسبيرمان لاختبار مشكلة عدم تجانس التباين علما ان $t5\% = 2.31$.

التسلسل	X_i	e_i
1	12.4	1.017
2	14.4	1.260
3	14.6	0.181
4	16.0	0.202
5	11.39	0.221
6	10.0	0.602
7	16.2	0.908
8	10.4	0.110
9	13.1	0.077
10	11.3	0.038

الثاني : باستخدام عينة عشوائية مكونة من(12) عائلة توصل باحث اقتصادي الى البيانات الاتية عن المتغيرين الدخل (Y_i) والادخار (S_i) بالاف الدنانير والمطلوب:

1- تقدير دالة الادخار للعينة

2- استخدام اختبار كولد فيلد وكوانت لاختبار وجود مشكلة عدم تجانس التباين؟

التسلسل	الادخار S_i	الدخل Y_i
1	2.6	30.5
2	2.2	26.0
3	1.5	18.0
4	4.0	42.5
5	2.7	30.0
6	2.9	28.0
7	2.6	27.5
8	3.0	32.5
9	3.2	35.0
10	2.7	26.0
11	2.2	27.5
12	3.4	39.0

دالة الادخار تأخذ الشكل الاتي: $\hat{S}_i = b_0 + b_1 Y_i + e_i$

1- ترتيب البيانات الخاصة بالمتغير المستقل Y_i من اصغر قيمة الى اكبر قيمة.

2- حذف المشاهدات الوسطية (6,7) من بيانات العينة، ثم تقسم المشاهدات الباقية الى عينتين متساويتين كل منهما تحتوي على (5) مشاهدات ، ثم تقدير معلمات العلاقة الخطية لكل عينة على انفراد واحتساب (S_{e1}^2) و (S_{e2}^2) واجراء اختبار F .

الثالث : البيانات الاتية عن المتغيرين (X_i , e_i) المطلوب : استخدام معامل ارتباط الرتب لسبيرمان لاختبار وجود مشكلة عدم تجانس التباين ؟ علماً ان : $t(6 , 0.05) = 1.86$.

X_i	6	4	3	5	4	5	3	2
e_i	−6	−4	0	3	2	3	2	0

الرابع : البيانات التالية تمثل عينة من عشرة مفردات للمتغيرين المعتمد (Y_i) والمستقل (X_i) ، اختبر تجانس تباين الخطأ باستخدام اختبار ارتباط الرتب لسبيرمان علماً ان :

$\hat{Y}_i = -0.72 + 0.54\, X_i$ وان $t_{(0.05)} = 2.31$

Y_i	X_i	$\hat{Y}_i$	e_i	$Rank\|e_i\|$	$RankX_i$	D_i	D_i^2
25	42	21.96	3.04	5	1	4	16
16	47	24.66	−8.66	10	2	8	64
29	50	26.28	2.72	3	3	0	0
32	55	28.98	3.02	4	4	0	0
40	67	35.46	4.54	8	5	3	9
44	84	44.64	−0.64	2	6	−4	16
45	85	45.18	−0.18	1	7	−6	36
43	88	46.80	−3.80	7	8	−1	1
49	98	52.20	−3.20	6	9	−3	9
61	104	55.44	−5.56	9	10	−1	1
			0			**0**	**152**

$$r_s = 1 - \left[\frac{6\sum D_i^2}{n(n^2-1)}\right] = 1 - \left[\frac{6(152)}{10(100-1)}\right] = 1 - 0.9212$$

$$r_s = 0.0788$$

$$t = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}} = \frac{(0.0788)(\sqrt{8})}{\sqrt{1-(0.0788)}} = \frac{0.22288}{0.99689}$$

$$\therefore t = 0.2236$$

وبما ان القيمة الجدولية لاختبار *(t)* هي (2.31) لمستوى معنوية(5%) ودرجة حرية (8) حيث ان القيمة المستخرجة هي اقل من القيمة الجدولية وان قيمة معامل ارتباط الرتب لسبيرمان ضعيفة لذا يتم قبول فرضية العدم (H_0) أي تجانس تباين الخطأ اذ أن :

$$H_0 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 \qquad \text{(homo.)}$$

الفصل الثامن

منظومة المعادلات الانية The Simultaneous Equation System

المقدمة :

ان نموذج الانحدار الخطي الذي تم التطرق اليه في الفصول السابقة وتم مناقشته مشاكله ، ما الا عبارة عن حالة خاصة من وضع عام ، حيث افترض بموجبه ان هناك اتجاهاً واحداً للسببية بين متغير او متغيرات مستقلة الى المتغير التابع ، بمعنى ان مجموعة من المتغيرات المستقلة ($X_1 , X_2 , \dots \dots \dots X_K$) تؤثر في المتغير المعتمد (Y_i) ولا تتأثر به ، في حين ان الحالة العامة لمعظم العلاقات الاقتصادية تنطوي على الاعتماد المتبادل بين المتغيرات الداخلة في النموذج ، أي ان العلاقة السببية تكون باتجاهين فقد تكون من المتغيرات المستقلة الى المتغير التابع وكذلك من المتغير التابع الى المتغير او المتغيرات المستقلة .

أي ان هناك على الأقل عدد من المتغيرات تتحدد أنياً تؤثر وتتأثر ببعضها البعض ، وهذا التأثير المتبادل يجعل الفرض الذي يتعلق باستقلال المتغير العشوائي (u_i) عن المتغير المستقل (X_i) غير ممكن أي ان :

$$E\ (X_i\, u_i\ )\ \neq 0$$

ومن ثم فان مقدرات **(OLS)** ستكون متحيزة وغير متسقة.

فعلى سبيل المثال الاستهلاك يؤثر في الدخل ويتأثر به في الوقت ذاته، ومن ثم لا مجال للقول بأن دالة الاستهلاك تقف بمعزل عن غيرها من العلاقات التي تحدد الدخل والاستثمار. كذلك الأجور لا تتحدد بمعزل عن الأسعار ولا تتحدد الأسعار بمعزل عن الأجور وانما يتحدد الاثنان سوية ضمن منظومة متعددة المعادلات.

أن وجود تأثير ذو اتجاهين في الدالة يعني بحد ذاته ضرورة وجود معادلتين او مجموعة معادلات لوصف العلاقة بين متغيرين، فالمتغير التابع في الدالة الأولى قد يوجد ضمن مجموعة المتغيرات المستقلة

في المعادلة الثانية، وعند ذلك يؤدي دوراً مزدوجاً اذ يكون هو الأثر (تابع) في المعادلة الأولى والمؤثر (مستقل) في المعادلة الثانية. ان هذا النظام في وصف التأثير المتبادل بين المتغيرات يسمى بنظام المعادلات الأنية **Simultaneous Equation System** .

فروض منظومة المعادلات الأنية :

بشكل عام يمكن تعريف منظومة المعادلات الأنية بأنها منظومة من المعادلات التي يكون المتغير المعتمد (التابع) لواحدة او أكثر من معادلاتها متغير مستقل في معادلة او اكثر من معادلة ضمن المنظومة. وتدعى المتغيرات المعتمدة في منظومة المعادلات الانية بالمتغيرات الداخلية (Endogenous Variables) حيث يقابل كل متغير داخلي في المنظومة معادلة واحدة، وبهذا فان عدد المعادلات في منظومة المعادلات الأنية ينبغي ان يساوي عدد المتغيرات الداخلية. اما فيما يتعلق بالمتغيرات الخارجية (Exogenous Variables) فان عددها لا يتحدد بعدد المتغيرات الداخلية وانما يتوقف على طبيعة العلاقة بين مختلف معادلات المنظومة، فقد يحدث ان تكون بعض المتغيرات المستقلة (الخارجية) في المنظومة داخلية والتي بدورها تعتمد على قيم الحدود العشوائية في المنظومة، وهو امر يتناقض مع الافتراض الخاص بنموذج الانحدار العام والذي ينص على استقلالية العلاقة ما بين قيم المتغيرات المستقلة وقيم الحدود العشوائية.

ويمكن توضيح ذلك من خلال النموذج الكينزي البسيط التالي :

$$\left.\begin{aligned} Y_t &= \beta_0 + \beta_1 X_t + u_t \\ X_t &= Y_t + Z_t \end{aligned}\right\} \quad (1)$$

حيث ان : Y_t : الاستهلاك X_t : الدخل Z_t : الاستثمار

المعادلة الأولى من المنظومة أعلاه تعرف بدالة الاستهلاك ، حيث ظهر فيها متغيراً عشوائياً (u_t) كعنصر خطأ ، اما المعادلة الثانية فتعرف بأنها متطابقة او علاقة تعريفية وهي محددة ولا يوجد فيها عنصر خطأ .

اذا كان هناك استقلال بين المتغير المستقل (X_t) والمتغير العشوائي (u_t) عندها يمكن تطبيق طريقة (OLS) للحصول على تقديرات غير متحيزة لمعالم دالة الاستهلاك ، ولكن مثل هذا الشرط غير متحقق في المنظومة أعلاه ، وذلك لان المتغير (X_t) لا يمكن اعتباره متغيراً خارجياً في دالة الاستهلاك لارتباطه

بعنصر الخطأ العشوائي (u_t) في هذه الدالة ، ويتضح ذلك بشكل جلي عند التعويض عن قيمة (Y_t) بما يساويها في المعادلة الثانية من المنظومة أعلاه :

$$X_t = (\beta_0 + \beta_1 X_t + u_t) + Z_t$$

وبإعادة ترتيب الحدود نحصل على:

$$X_t - \beta_1 X_t = \beta_0 + Z_t + u_t$$

$$\therefore \quad X_t(1-\beta_1) = \beta_0 + Z_t + u_t$$

$$\therefore \quad X_t = \frac{\beta_0}{1-\beta_1} + \frac{Z_t}{1-\beta_1} + \frac{u_t}{1-\beta_1} \qquad \ldots\ldots\ldots \ (2)$$

يتبين من الصيغة رقم (2) أعلاه ان الدخل (X_t) دالة في عنصر الخطأ العشوائي (u_t) في دالة الاستهلاك أي ان $Cov\,(X_t\, u_t) \neq 0$ وذلك لان :

$$Cov\,(X_t\, u_t) = E\left[\big(X_t - E(X_t)\big)\big(u_t - E(u_t)\big)\right]$$

وبالتعويض عن القيمة المتوقعة لـ (X_t) وعلى افتراض ان $E(u_t) = 0$

$$\therefore = Cov\,(X_t\, u_t) = E\left[\left(X_t - (\frac{\beta_0}{1-\beta_1}) \ - \ (\frac{1}{1-\beta_1})Z_t\right)u_t\right]$$

وبالتعويض عن (X_t) بما يساويها في (z) فان

$$\therefore = Cov\,(X_t\, u_t) = E\left[\left(\frac{1}{1-\beta_1}\right) u_t.u_t\right] = \left(\frac{1}{1-\beta_1}\right)E\,(u_t^2) = \left(\frac{1}{1-\beta_1}\right)\boldsymbol{\sigma^2}_{\boldsymbol{u}}$$

النتيجة أعلاه غير صفرية، ويستدل منها بعدم إمكانية تطبيق طريقة المربعات الصغرى الاعتيادية (OLS) لتقدير معالم دالة الاستهلاك، أي ان تقدير معالم دالة الاستهلاك باستخدام طريقة (OLS) سوف يكون متحيزاً وغير متسقاً، لذا يستوجب اللجوء الى طرق أخرى لتقدير معالم منظومة المعادلات الانية.

الفرض أعلاه يوضح حالة بسيطة للتشابك بين متغير خارجي واحد والخطأ العشوائي في منظومة متضمنة علاقتين فقط K بشكل عام وفي حالة توفر (K) من المتغيرات الخارجية في منظومة متضمنة (G) من المعادلات، عندها يمكن استخدام أسلوب المصفوفات لبيان الفروض الواجب توفرها عند بناء

منظومة المعادلات الانية للظواهر المختلفة، ولتوضيح ذلك ندرس دعنا ندرس منظومة المعادلات الانية المعاد ترتيب حدودها بالشكل التالي:

$$\beta_{11}Y_{1t} + \beta_{12}Y_{2t} + \cdots\cdots + \beta_{1G}Y_{Gt} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + \cdots\cdots + \gamma_{1k}X_{kt} = u_{1t}$$
$$\beta_{21}Y_{1t} + \beta_{22}Y_{2t} + \cdots\cdots + \beta_{2G}Y_{Gt} + \gamma_{21}X_{1t} + \gamma_{22}X_{2t} + \cdots\cdots + \gamma_{2k}X_{kt} = u_{2t}$$
$$\vdots$$
$$\beta_{G1}Y_{1t} + \beta_{G2}Y_{2t} + \cdots\cdots + \beta_{GG}Y_{Gt} + \gamma_{G1}X_{1t} + \gamma_{G2}X_{2t} + \cdots\cdots + \gamma_{Gk}X_{kt} = u_{Gt}$$

حيث ان:

Y : المتغيرات الداخلية ، X : المتغيرات الخارجية ، u : الأخطاء العشوائية .

β, γ : معالم المنظومة الخاصة بالمتغيرات الداخلية والخارجية على التوالي.

G : عدد المتغيرات الداخلية ، k : عدد المتغيرات الخارجية .

مجموعة المعادلات رقم (3) أعلاه تمثل منظومة معادلات أنية وفيها المتغير المعتمد لواحد او اكثر من معادلاتها متغيراً مستقلاً في معادلة أخرى او اكثر من معادلة ضمن تلك المجموعة من المعادلات، بعبارة أخرى ان بعض المتغيرات المعتمدة تكون كمتغيرات تفسيرية مرتبطة بمتغيرات معتمدة أخرى.

ان منظومة المعادلات الأنية تكون على أنواع، منها منظومة المعادلات الترددية (Recursive Equations System) والتي يمكن إيجاد (تحديد) المتغيرات الداخلية فيها بالتعاقب.

وهناك نوع اخر من المنظومات الأنية تعرف بمنظومة المعادلات القطاعية – الترددية (Recursive Block Equations System) والتي فيها مجموعة المعادلات يمكن تجزئتها الى مجموعات او قطاعات من المعادلات بحيث تكون المعادلات أنياً في كل قطاع، وأن مجموعة المعادلات عبر القطاعات تأخذ صفة منظومة معادلات ترددية بحيث أن المعلومات عن المتغيرات الداخلية في القطاع الأول على سبيل المثال تساعد على تحديد المتغيرات الداخلية في القطاع الثاني وهكذا .....

فضلاً عن ذلك فان هناك حالة خاصة من منظومة المعادلات الأنية وفيها لا يظهر أي تداخل (تشابك) بين متغيراتها الداخلية والخارجية، ولكن لا يزال هناك تداخل بين معادلاتها السلوكية وذلك نتيجة لترابط

الأخطاء العشوائية بين المعادلات المختلفة. هذا النوع يعرف بمنظومة معادلات الانحدار غير المرتبط ظاهرياً (Seemingly Unrelated Regression Equations) (SURE) وفيها مجموعة المعادلات قد تكون مرتبطة احصائياً حت ولو انها لا تبدو كذلك هيكلياً.

وبالرجوع الى منظومة المعادلات الأنية رقم (3) حيث يمكن إعادة كتابته باستخدام المصفوفات والموجهات وكالاتي:

$$\beta Y_t + \Gamma Y_t = u_t \quad \cdots\cdots\cdots \quad (4)$$

الصيغة رقم (4) أعلاه تعرف بالشكل الهيكلي او النموذج الهيكلي (structural model) وكل معادلة هيكلية فيه تعبر عن احد المتغيرات الداخلية بدلالة المتغيرات الخارجية والداخلية وكذلك المتغيرات الداخلية المرتدة زمنياً ان وجدت ، علماً ان كل من Y_t و X_t و u_t تمثل موجهات في حين β و Γ تمثل مصفوفات وبالشكل التالي :

$$Y_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{Gt} \end{bmatrix} \quad , \quad X_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Kt} \end{bmatrix} \quad , \quad u_t = \begin{bmatrix} u_{1t} \\ u_{1t} \\ \vdots \\ u_{Gt} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots\cdots\cdots & \beta_{1G} \\ \beta_{21} & \beta_{22} & \cdots\cdots\cdots & \beta_{2G} \\ \vdots & \vdots & & \vdots \\ \beta_{2G} & \beta_{2G} & \cdots\cdots\cdots & \beta_{2G} \end{bmatrix} \quad , \quad \Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \cdots\cdots\cdots & \Gamma_{1K} \\ \Gamma_{21} & \Gamma_{22} & \cdots\cdots\cdots & \Gamma_{2K} \\ \vdots & \vdots & & \vdots \\ \Gamma_{G1} & \Gamma_{G2} & \cdots\cdots\cdots & \Gamma_{GK} \end{bmatrix}$$

وان موّجة خطأ المنظومة (u_t) يخضع للفرض التالي :

$$u_t \sim N\,(0\,,\phi)$$

وبمصفوفة تباين وتباين مشترك للأخطاء العشوائية كالاتي:

$$Var - Cov\,(u_t) = E\,(u_t u_t^-) = \phi = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots\cdots\cdots & \sigma_{1G} \\ \sigma_{21} & \sigma_{22} & \cdots\cdots\cdots & \sigma_{2G} \\ \vdots & \vdots & & \vdots \\ \sigma_{G1} & \sigma_{G2} & \cdots\cdots\cdots & \sigma_{GG} \end{bmatrix}$$

حيث ان قطر المصفوفة يمثل التباين للأخطاء العشوائية في مختلف معادلات المنظومة، اما العناصر خارج نطاق قطر المصفوفة فتمثل التباين المشترك بين أخطاء كل معادلتين من معادلات المنظومة الانية .

التشخيص والاختزال : Reduce and Identification

تعتبر مشكلة التشخيص (Identification) من المشاكل الأساسية لبناء النماذج القياسية، ويقصد بالتشخيص اختيار كل معادلة من معادلات المنظومة من حيث صياغتها بشكل نهائي وبدون اهمال أي متغير أساسي او ثانوي فيها، وبالتالي بيان الأسلوب الملائم لتقدير معالم المنظومة تحت البحث. ولغرض توضيح الفكرة الأساسية للتشخيص دعنا ندرس منظومة العرض والطلب التالية:

$$D = A_O + A_1 P + u_1$$

$$S = \beta_O + \beta_1 P + \beta_2 W + u_1$$

$$D = S$$

حيث ان : D : الكمية المطلوبة S : الكمية المعروضة P : السعر w : الرقم القياسي لحالة الجو وان (D ، S ، P) تعرف بالمتغيرات الداخلية وان (W) يمثل متغير خارجي .

كل معادلة هيكلية في منظومة المعادلات أعلاه تعبر عن احد المتغيرات الداخلية بدلالة المتغيرات الخارجية والداخلية وكذلك المتغيرات الداخلية المرتدة زمنياً (Lagged Endogenous Variables) ومعالم النموذج الهيكلي تسمى بالمعالم الهيكلية (Structural Parameters) . وبالرجوع الى النموذج الهيكلي أعلاه حيث يمكن اشتقاق نموذجاً اخر، وذلك بالتعبير عن كل متغير داخلي فيه بدلالة المتغيرات الخارجية والداخلية المرتدة زمنياً ان وجدت وكالاتي:

وبما ان $D = S$

$$A_O + A_1 P + u_1 = \beta_O + \beta_1 P + \beta_2 W + u_2$$

$$\therefore\ P(A_1 - \beta_1) = \beta_O - A_O + \beta_2 W + u_2 - u_1$$

ومنه فأن

$$P = \frac{\beta_O - A_O}{A_1 - \beta_1} + \frac{\beta_2}{A_1 - \beta_1} W + \frac{u_2 - u_1}{A_1 - \beta_1} \ldots\ldots\ldots \quad (5)$$

وبتعويض ذلك في دالة العرض او الطلب نحصل على:

$$D = A_O + A_1 \left[\frac{\beta_O - A_O}{A_1 - \beta_1} + \frac{\beta_2}{A_1 - \beta_1} W + \frac{u_2 - u_1}{A_1 - \beta_1}\right] + u_1$$

$$D = S = \frac{A_1\beta_O - A_O\beta_1}{A_1 - \beta_1} + \frac{A_1\beta_2}{A_1 - \beta_1} W + \frac{A_1 u_2 - u_1\beta_1}{A_1 - \beta_1} \ldots\ldots\ldots \quad (6)$$

الصيغتين رقم (5) و (6) أعلاه والمشتقة من النموذج الهيكلي تسمى بالنموذج المختزل (Reduce Form Model) وظهرت فيه المتغيرات الداخلية كدالة في المتغيرات الخارجية إضافة الى عنصر الخطأ العشوائي ، وبمقارنة معالم النموذج الهيكلي بمعالم النموذج المختزل يتضح ان معالم النموذج المختزل ما هي الا دوال لمعالم النموذج الهيكلي ، وكذلك الحال بالنسبة لعنصر الخطأ العشوائي ، علماً بأن معالم النموذج الهيكلي تقيس الأثر المباشر للمتغير المستقل (داخلياً كان او خارجياً) على المتغير الداخلي ، في حين أن معالم النموذج المختزل تقيس الأثر الكلي أي الأثر المباشر والأثر غير المباشر للتغير في المتغيرات المستقلة (خارجية كانت او داخلية مرتدة زمنياً) على كل متغير داخلي .

يتضح من الصيغتين أعلاه أن معالم النموذج المختزل تأخذ في نظر الاعتبار التشابك (الترابط) المتبادل بين المتغيرات الداخلية ، فعلى سبيل المثال زيادة وحدة واحدة في المتغير (W) في النموذج المختزل يؤدي الى زيادة في السعر قدرها $\left(\frac{\beta_2}{A_1-\beta_1}\right)$ وزيادة في العرض او الطلب مقدارها $\left(\frac{A_1\beta_2}{A_1-\beta_1}\right)$ ومثل هذه الخاصية لها فائدة كبيرة في التنبؤ وتحليل السياسات ، وذلك لان ما يهم واضع السياسة هو الأثر الكلي وليس فقط الأثر المباشر لتغير المتغيرات الخارجية على المتغيرات الداخلية .

وبإعادة كتابة النموذج المختزل بالشكل التالي :

$$P = \prod\nolimits_{11} + \Pi_{11} W + V_1$$

$$D = S = \Pi_{21} + \Pi_{22} W + V_2 \qquad \ldots\ldots\ldots \quad (7)$$

اذ أن:

$$\Pi_{11}=\frac{\beta_O - A_O}{A_1-\beta_1} \quad , \quad \Pi_{12}=\frac{\beta_2}{A_1-\beta_1} \quad , \quad V_1=\frac{u_2-u_1}{A_1-\beta_1}$$

$$\Pi_{21}=\frac{A_1\beta_O - A_O\beta_1}{A_1-\beta_1} \quad , \quad \Pi_{22}=\frac{A_1\beta_2}{A_1-\beta_1} \quad , \quad V_2=\frac{A_1u_2-u_1\beta_1}{A_1-\beta_1}$$

حيث يمكن استخدام طريقة (OLS) بشكل غير مباشر لتقدير معالم منظومة العرض والطلب أعلاه، وذلك في حالة تحقيق الشروط اللازمة لتشخيص كل معادلة من معادلاته، ويقصد بشرط التشخيص هو إمكانية الحصول على تقدير وحيد لمعالم النموذج الهيكلي من تقديرات معالم النموذج المختزل، وعندها يكون النموذج تحت البحث مشخصاً تماماً (Just Identify) . اما اذا لم يكن النموذج الهيكلي مشخصاً تماماً (under Identify) فأن ذلك يعني عدم إمكانية الحصول على تقديرات وحيدة للمعالم الهيكلية من تقديرات معالم النموذج المختزل، وفي الحالة التي يكون فيها إمكانية للحصول على قيم من النموذج المختزل لكل معلمة من معالم النموذج الهيكلي، عندها سوف لن يكون هناك تقديراً وحيداً لهذه المعالم وبالتالي يكون النموذج المدروس فوق التشخيص (Over Identify) .

الأن وفي حالة مثالنا السابق ، نلاحظ من نموذج الاختزال رقم (7) يمكن الوصول الى تقديرات وحيدة لمعالم دالة الطلب (A_1 ، A_0) في النموذج الهيكلي وكالاتي :

$$A_1=\frac{\Pi_{22}}{\Pi_{12}} \quad , \quad A_o=\Pi_{12}-\frac{\Pi_{22}}{\Pi_{12}}\ \Pi_{11}$$

وهذا يعني ان دالة الطلب مشخصة ، اما معالم دالة العرض (β_O ، β_1 ، β_2) فلا يمكن تقديرها من خلال معالم النموذج المختزل ، وعليه فهي دالة غير مشخصة وبالتالي منظومة العرض والطلب ككل غير مشخصة ويجب إعادة النظر ببنائها .

وبالرجوع الى النظرية الاقتصادية وبالذات الى مفهوم دالة الطلب نجد ان الطلب ما هو الا دالة في السعر والدخل وليس فقط في السعر كما ورد في المعادلة الأولى من المنظومة، عليه يستوجب إعادة صياغة منظومة العرض والطلب بالشكل التالي:

$$D = A_O + A_1 P + A_2 Y + u_1$$
$$S = \beta_O + \beta_1 P + \beta_2 W + u_2 \qquad \ldots\ldots\ldots\ldots \quad (8)$$
$$D = S$$

حيث أن : Y : الدخل ، وبقية المتغيرات تأخذ نفس التعريف السابق .

النموذج الهيكلي أعلاه متكون من ثلاث معادلات بثلاث متغيرات داخلية $(D$ ، S ، $D)$ ويحتوي على متغيرين خارجيين هما الدخل (Y) وحالة الجو (W) ولغرض تشخيص كل معادلة من معادلاته لابد من إيجاد الصيغة المختزلة له ، أي التعبير عن كل متغير داخلي بدلالة المتغيرات الخارجية وذلك يتم بالإحلال المتتابع وكالاتي :

$$A_O + A_1 P + A_2 Y + u_1 = \beta_O + \beta_1 P + \beta_2 W + u_2$$
$$P(\beta_1 - A_1) = A_O - \beta_O + A_2 Y - \beta_2 W + u_1 - u_2$$
$$\therefore \quad P = \frac{A_O - \beta_O}{\beta_1 - A_1} + \frac{A_2}{\beta_1 - A_1} Y - \frac{\beta_2}{\beta_1 - A_1} W + \frac{u_1 - u_2}{\beta_1 - A_1} \quad \ldots\ldots\ldots\ldots \quad (9)$$

وبالتعويض في معادلة العرض أو الطلب نحصل على:

$$D = A_O + A_1 P + A_2 Y + u_1$$
$$D = A_O + A_1 \left[\frac{A_O - \beta_O}{\beta_1 - A_1} + \frac{A_2}{\beta_1 - A_1} Y - \frac{\beta_2}{\beta_1 - A_1} W + \frac{u_1 - u_2}{\beta_1 - A_1}\right] + A_2 Y + u_1$$
$$\therefore \quad D = \frac{A_O\beta_1 - A_1\beta_O}{\beta_1 - A_1} + \frac{A_2\beta_1}{\beta_1 - A_1} Y - \frac{A_1\beta_2}{\beta_1 - A_1} W + \frac{u_1\beta_1 - u_2 A_1}{\beta_1 - A_1} = S \ldots\ldots (10)$$

الصيغتان (9) و (10) تمثلان الشكل المختزل للنموذج الهيكلي والذي يمكن إعادة كتابته بالشكل التالي:

$$D = S = \prod_{10} + \prod_{11} \mathrm{Y} + \prod_{12} \mathrm{W} + V_1$$
$$P = \prod_{20} + \prod_{21} \mathrm{Y} + \prod_{22} \mathrm{W} + V_2 \qquad \ldots\ldots\ldots\ldots \quad (11)$$

اذ ان:

$$\prod_{10} = \frac{A_O\beta_1 - A_1\beta_O}{\beta_1 - A_1} \quad , \quad \prod_{11} = \frac{A_2\beta_1}{\beta_1 - A_1} \quad , \quad \prod_{12} = \frac{-A_1\beta_2}{\beta_1 - A_1}$$

$$V_1 = \frac{u_1\beta_1 - u_2A_1}{\beta_1 - A_1} \quad , \quad \prod_{20} = \frac{A_O - \beta_O}{\beta_1 - A_1} \quad , \quad \prod_{21} = \frac{A_2}{\beta_1 - A_1}$$

$$\prod_{22} = \frac{-\beta_O}{\beta_1 - A_1} \quad , \quad V_2 = \frac{u_1 - u_2}{\beta_1 - A_1}$$

وبملاحظة النموذج المختزل رقم (11) أعلاه نجد ان هناك إمكانية لتقدير معالم دالة الطلب في النموذج الهيكلي وعلى النحو التالي:

$$A_O = \prod_{20}\left[\frac{\prod_{10}}{\prod_{20}} - \frac{\prod_{12}}{\prod_{22}}\right] \quad , \quad A_1 = \frac{\prod_{12}}{\prod_{22}}$$

$$A_2 = \left[\frac{\prod_{11}}{\prod_{21}} - \frac{\prod_{12}}{\prod_{22}}\right]\prod_{12} = \prod_{11} - \frac{\prod_{12}\prod_{21}}{\prod_{21}}$$

وكذلك الحال بالنسبة لمعالم دالة العرض في النموذج الهيكلي، حيث يمكن تقديرها من معالم النموذج المختزل وكالاتي:

$$\beta_O = \prod_{20}\left[\frac{\prod_{10}}{\prod_{20}} - \frac{\prod_{11}}{\prod_{21}}\right] \quad , \quad \beta_1 = \frac{\prod_{11}}{\prod_{21}}$$

$$\beta_2 = \prod_{22}\left[\frac{\prod_{12}}{\prod_{22}} - \frac{\prod_{11}}{\prod_{21}}\right] = \prod_{12} - \frac{\prod_{22}\prod_{11}}{\prod_{21}}$$

وبما ان هناك ستة علاقات مستنبطة من النموذج المختزل تمثل ستة معالم في النموذج الهيكلي ، أي ان هناك تقديرات وحيدة لمعالم المنظومة الهيكلية وبالتالي فان كل من دالة الطلب والعرض مشخصة والمنظومة ككل مشخصة تماماً ، حيث يمكن تطبيق طريقة (OLS) بشكل غير مباشر لتقدير معالم النموذج الهيكلي وذلك عن طريق تقدير معالم النموذج المختزل اولاً ثم استخدام العلاقات السابقة الستة أعلاه لتقدير المعالم الهيكلية (β_2 ، β_1 ، β_O ، A_2 ، A_1 ، A_O) وهذا الأسلوب في التقدير يعرف بطريقة المربعات الصغرى الغير مباشر (ILS) ، (Indirect Least Square) .

الشروط الأساسية للتشخيص :

من مناقشة مشكلة التشخيص اتضح لنا بأن تشخيص أي معادلة هيكلية يتطلب إيجاد الشكل المختزل لها أولاً، ولكن في الواقع ليس من الضروري أن نختزل المعادلة الهيكلية في كل مرة لغرض تشخيصها، فهناك قاعدة عامة يمكن استخدامها لهذا الغرض دون اللجوء الى الاختزال وتتمثل هذه القاعدة بتحقيق الشرطين التاليين:

1- **شرط الترتيب** **Order Condition**

2- **شرط الرتبة** **Rank Condition**

الشرط الأول ضروري ولكنة غير كافي لتشخيص أي معادلة هيكلية من معادلات المنظومة، لذا وتأكيداً للاختبار الأول يستوجب اجتياز المعادلة شرط الاختبار الثاني.

1- الشرط الأول (شرط الترتيب) :

بشكل عام تكون المعادلة مشخصة بموجب الشرط الأول عندما يكون عدد المتغيرات المستبعدة منها ولكنها داخلة في المعادلات الأخرى للنموذج الهيكلي مساوياً لعدد المعادلات في المنظومة مطروحاً منه واحد، فاذا كان عدد معادلات المنظومة الهيكلية (G) وعدد المتغيرات سواء كانت خارجية او داخلية مرتدة زمنياً في النموذج الهيكلي مساوياً الى (K)، وعدد المتغيرات في المعادلة محل الاختبار مساويه الى (M) فان شرط الترتيب لتشخيص هذه المعادلة يأخذ الصيغة العامة التالية:

$$K - M \geq G - 1$$

فاذا كانت $K - M = G - 1$ تكون المعادلة مشخصة تماماً . (Exactly Or Just Identified)

اما اذا كانت $K - M > G - 1$ تكون المعادلة فوق التشخيص (Over Identified) .

واخيراً في حالة كون $K - M < G - 1$ عندها تكون المعادلة غير مشخصة او تحت التشخيص (under Identified) .

2- الشرط الثاني (شرط الرتبة) :

لغرض تأكيد المرحلة الأولى من الاختبار لابد من اجراء الاختبار الثاني والمتمثل بشرط الرتبة والذي يلخص بالشكل التالي:

ترتيب كافة المعالم الهيكلية بدلالة جميع المتغيرات في المنظومة في جدول ، ثم تؤخذ المعالم المقابلة للمعالم المفقودة في المعادلة المطلوب اختبارها وتوضع بشكل مصفوفة ، بعدها نجد قيمة محدد هذه المصفوفة والتي تكون من الرتبة $(G-1)$ ، فاذا كانت قيمة محددها لا يساوي صفر تكون المعادلة مشخصة ، وبخلافة اذا كانت قيمة المحدد مساوياً للصفر عندها توصف المعادلة موضع البحث بانها غير مشخصة (تحت التشخيص) . اما اذا حدث وان كانت المصفوفة المستخرجة من المعالم الهيكلية غير مربعة ، عندها يستوجب تجزئتها الى كافة المصفوفات الجزئية الممكنة وذات الرتبة $(G-1)$ ، فاذا كانت على الأقل واحدة من قيم هذه المحددات لا يساوي صفر تكون المعادلة مشخصة ، اما اذا كانت كافة قيم المحددات الرتبة $(G-1)$ مساوية للصفر فان المعادلة سوف تكون غير مشخصة

وبالرجوع الى المثال السابق (النموذج رقم 8) والذي تم اختزاله والمتعلق بمنظومة العرض والطلب والمتكون من معادلتين ومتطابقة، ولغرض تشخيص كل معادلة من معادلاته يستوجب إعادة كتابته بالشكل التالي:

$$-D + A_O + A_1 P + A_2 Y + u_1 = 0$$

$$-S + \beta_O + \beta_1 P + \beta_2 W + u_2 = 0$$

$$-D + S = 0$$

ثم نرتب كافة المعالم للمنظومة أعلاه بدلالة كافة المتغيرات سواء كانت متغيرات داخلية او خارجية او داخلية مرتدة زمنياً وكالاتي:

المعادلة	المتغيـــــــــــــرات					
	I	D	S	P	Y	W
1	A_O	-1	0	A_1	A_2	0
2	β_O	0	-1	β_1	0	β_2
3	0	-1	1	0	0	0

اولاً: شرط الترتيب .: **Order Condition**

$$K - M \geq G - 1$$

عدد المتغيرات في المعادلة المطلوب تشخيصها $= M$ ، $G = 3$ ، $K = 5$

1- بالنسبة للمعادلة الأولى فان:

$$5 - 3 = 3 - 1$$

$2 = 2$ لذلك فان المعادلة مشخصة تماماً .

2- بالنسبة للمعادلة الثانية فان :

$$5 - 3 = 3 - 1$$

$2 = 2$ لذلك فان المعادلة مشخصة تماماً .

ثانياً : شرط الرتبة .: **Rank Condition**

1- بالنسبة للمعادلة الأولى فان :

$$Z_1 = \begin{bmatrix} -1 & \beta_2 \\ 1 & 0 \end{bmatrix} \Rightarrow \therefore |Z_1| = \begin{vmatrix} -1 & \beta_2 \\ 1 & 0 \end{vmatrix} = -\beta_2$$

فاذا كانت قيمة (β_2) لا تساوي صفر فان المعادلة الأولى تكون مشخصة تماماً .

2- بالنسبة للمعادلة الثانية فان :

$$Z_2 = \begin{bmatrix} -1 & A_2 \\ -1 & 0 \end{bmatrix} \Rightarrow \therefore |Z_2| = \begin{vmatrix} -1 & A_2 \\ -1 & 0 \end{vmatrix} = A_2$$

فاذا كانت قيمة (A_2) لا تساوي صفر فان المعادلة الثانية تكون مشخصة تماماً .

وتجدر الإشارة الى ان المنظومة الهيكلية تكون غير مشخصة اذا كانت واحدة او اكثر من معادلاتها غير مشخصة، وهذا النوع من النماذج الهيكلية لا يمكن تقدير معالمة باي أسلوب من أساليب القياس الاقتصادي، اما اذا كانت كافة معادلات المنظومة مشخصة تماماً عندها تكون المنظومة ككل مشخصة

تماماً، وبالتالي يمكن تقدير معالمها الهيكلية بموجب طريقة المربعات الصغرى غير المباشرة (ILS) (Indirect Least Square) . او طريقة المربعات الصغرى ذات المرحلتين (2SLS) (Tow Stage Least Square) او طريقة المتغيرات المساعدة (IV) (Instrumental Variable) او طريقة الإمكان الأعظم ذات المعلومات المحدودة (LIML) (Limited in Formation Maximum Likelihood)

واخيراً اذا كانت معادلات المنظومة فوق التشخيص فان معالمها الهيكلية يمكن ان تقدر بإحدى الأساليب (الطرق) القياسية التالية:

أ- طريقة المربعات الصغرى ذات الثلاث مراحل (3SLS) (Three Stage Least Square)

ب- طريقة الإمكان الأعظم ذات المعلومات الكاملة (التامة) (FIML) (Full in formation Maximum Likelihood)

وتجدر الإشارة هنا الى ان الطريقتين الأخيرتين تسمى بطريقة المنظومة لكون التقدير بموجبها يتم أنياً وعلى مستوى كافة معادلات المنظومة ، في حين الطرق الأربعة الأولى تعرف بطرق أحادية المعادلة وذلك لإمكانية تطبيقها لتقدير معالم كل معادلة من معادلات المنظومة على انفراد .

مثال (1) .: شخص كل معادلة من معادلات المنظومة التالية :

$$Y_1 = 3Y_2 - 2X_1 + X_2 + u_1$$

$$Y_2 = Y_3 + X_3 + u_2$$

$$Y_3 = Y_1 - Y_2 - 2X_3 + u_3$$

علماً ان Y_1 ، Y_2 ، Y_3 متغيرات داخلية وان X_1 ، X_2 ، X_2 متغيرات خارجية .

الحل : نُعيد كتابة النموذج الهيكلي بعد دمج المتغيرات الداخلية والمتغيرات الخارجية وكالاتي :

$$-Y_1 + 3Y_2 - 2X_1 + X_2 + u_1 = 0$$

$$-Y_2 + Y_3 + X_3 + u_2 = 0$$

$$-Y_3 + Y_1 - Y_2 - 2X_3 + u_3 = 0$$

وبإهمال الأخطاء العشوائية وإعادة كتابة المعالم الهيكلية بدلالة كافة المتغيرات في المنظومة للحصول على الجدول الاتي:

المعادلة	المتغيـــــــــــــرات					
	Y_1	Y_2	Y_3	X_1	X_2	X_3
1	-1	3	0	-2	1	0
2	0	-1	1	0	0	1
3	1	-1	-1	0	0	-2

اولاً : اختبار شرط الترتيب .

الصيغة العامة للاختبار هي $K - M \geq G - 1$

$$G = 3 \qquad K = 6$$

1- بالنسبة للمعادلة الأولى فان:

$$6 - 4 = 3 - 1$$

$2 = 2$ لذلك فان المعادلة مشخصة تماماً .

2- بالنسبة للمعادلة الثانية فان :

$$6 - 3 = 3 - 1$$

$3 \neq 2$ المعادلة فوق التشخيص .

3- بالنسبة للمعادلة الثانية فان :

$$6 - 4 = 3 - 1$$

$2 = 2$ لذلك فان المعادلة مشخصة تماماً .

ثانياً : شرط الرتبة .:

1- بالنسبة للمعادلة الأولى فان :

$$|Z_1| = \begin{vmatrix} 1 & 1 \\ -1 & -2 \end{vmatrix} = -1$$

وحيث ان قيمة المحدد لا تساوي صفر فان المعادلة الأولى تكون مشخصة تماماً .

2- بالنسبة للمعادلة الثانية فان :

$$|Z_2| = \begin{vmatrix} -1 & -2 & 1 \\ 1 & 0 & 0 \end{vmatrix}$$

وبما ان المصفوفة غير مربعة لذلك يجب تجزئتها الى عدد من المصفوفات الفرعية ثم نجد قيم محدداتها وعلى النحو الاتي :

$$|Z_{21}| = \begin{vmatrix} 1 & -2 \\ 1 & 0 \end{vmatrix} = 2$$

$$|Z_{22}| = \begin{vmatrix} -1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

$$|Z_{23}| = \begin{vmatrix} -2 & 1 \\ 0 & 0 \end{vmatrix} = 0$$

وبما ان هناك على الأقل واحد من قيم هذه المحددات الفرعية للمصفوفة المجزأة (Z_2) لا يساوي صفراً لذلك فالمعادلة الثانية مشخصة تماماً .

1- بالنسبة للمعادلة الثالثة فان :

$$|Z_3| = \begin{vmatrix} -2 & 1 \\ 0 & 0 \end{vmatrix} = 0$$

وبما ان قيمة محدد المصفوفة (Z_3) يساوي صفراً ، لذلك فأن المعادلة الثالثة غير مشخصة .

مثال (2) .: شخص كل معادلة من معادلات المنظومة التالية :

$$C_t = A_O + A_1 Y_t + A_2\, C_{t-1} + u_{1t}$$

$$I_t = \beta_O + \beta_1 R_t + \beta_2\, I_{t-1} + u_{2t}$$

$$R_t = \gamma_O + \gamma_1 Y_t + \gamma_2\, M_t + u_{3t}$$

$$Y_t = C_t + I_t + G_t$$

علماً ان C_t ، Y_t ، I_t ، R_t : الاستهلاك ، الدخل القابل للتصرف ، الاستثمار ، سعر الفائدة ، على التوالي وتمثل متغيرات داخلية . وان M_t ، G_t : السعر الجاري والنفقات الحكومية على التوالي ، وهي متغيرات خارجية . وان C_{t-1} ، I_{t-1} : متغيرات داخلية مرتدة زمنياً بسنة واحدة .

الحل : نُعيد كتابة النموذج الهيكلي اعلاه بعد دمج المتغيرات الداخلية مع الخارجية وكذلك مع المتغيرات المرتدة زمنياً .

$$-C_t + A_O + A_1 Y_t + A_2\, C_{t-1} + u_{1t} = 0$$

$$-I_t + \beta_O + \beta_1 R_t + \beta_2\, I_{t-1} + u_{2t} = 0$$

$$-R_t + \gamma_O + \gamma_1 Y_t + \gamma_2\, M_t + u_{3t} = 0$$

$$-Y_t + C_t + I_t + G_t = 0$$

وبإهمال الأخطاء العشوائية وإعادة كتابة المعالم بدلالة كافة المتغيرات نحصل على الجدول التالي :

المعادلة	المتغيـــــــــرات								
	C_t	I_t	R_t	Y_t	1	C_{t-1}	I_{t-1}	M_t	G_t
1	-1	0	0	A_1	A_O	A_2	0	0	0
2	0	-1	β_1	0	β_O	0	β_2	0	0
3	0	0	-1	γ_1	γ_O	0	0	γ_2	0
4	1	1	0	-1	0	0	0	0	1

اولاً : اختبار شرط الترتيب .

الصيغة العامة للاختبار هي $K - M \geq G - 1$

$$G = 4 \quad K = 8$$

1- بالنسبة للمعادلة الأولى فان :

$$8 - 3 > 4 - 1$$

$\therefore \quad 5 > 3$ لذلك فان المعادلة فوق التشخيص .

2- بالنسبة للمعادلة الثانية فان :

$$8 - 3 > 4 - 1$$

$\therefore \quad 5 > 3$ لذلك فان المعادلة فوق التشخيص .

3- بالنسبة للمعادلة الثالثة فان :

$$8 - 3 > 4 - 1$$

$\therefore \quad 5 > 3$ لذلك فان المعادلة فوق التشخيص .

ثانياً : شرط الرتبة .:

1- بالنسبة للمعادلة الأولى فان :

$$Z_1 = \begin{vmatrix} -1 & \beta_1 & \beta_2 & 0 & 0 \\ 0 & -1 & 0 & \gamma_2 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{vmatrix}$$

المصفوفة أعلاه غير مربعة لذا يستوجب تجزئتها الى مصفوفات مربعة من الرتبة (G_{-1}) وكالاتي :

$$\begin{bmatrix} -1 & \beta_1 & \beta_2 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad , \quad \begin{bmatrix} -1 & \beta_1 & 0 \\ 0 & -1 & \gamma_2 \\ 1 & 0 & 0 \end{bmatrix}$$

وهكذا لكافة المصفوفات الممكنة ثم إيجاد قيم محدداتها، فان وجد على الأقل واحد من قيم محدداتها الجزئية لا يساوي صفر كانت المعادلة مشخصة وبعكسه تكون المعادلة غير مشخصة .

2- بالنسبة للمعادلة الثانية فان :

$$Z_2 = \begin{vmatrix} -1 & A_1 & A_2 & 0 & 0 \\ 0 & \gamma_1 & 0 & \gamma_2 & 0 \\ 1 & -1 & 0 & 0 & 1 \end{vmatrix}$$

وبنفس الأسلوب السابق، يجب تجزئة المصفوفة أعلاه ثم إيجاد قيم محدداتها الجزئية .

3- بالنسبة للمعادلة الثالثة فان :

$$Z_3 = \begin{vmatrix} -1 & 0 & A_2 & 0 & 0 \\ 0 & -1 & 0 & \beta_2 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{vmatrix}$$

وكذلك الحال هنا ، بما ان المصفوفة غير مربعة عليه يجب تجزئتها الى كافة المصفوفات الجزئية من الرتبة (G_{-1}) ثم إيجاد قيم محدداتها الجزئية ، فان وجد على الأقل واحد من قيم محدداتها الجزئية لا يساوي صفر كانت المعادلة مشخصة وبعكسه تكون غير مشخصة . وتجدر الإشارة هنا ان المتطابقات والمعادلات التعريفية والتوازنية لا يجري لها اختبار التشخيص وذلك لعدم وجود معالم فيها.

الفصــل التاسـع

نماذج المتغيرات المرتدة زمنياً: The Lagged Variables Models

يرتبط المتغير المستقل (X) مع المتغير المعتمد (Y) في نماذج الانحدار الخطية ضمن السلسة الزمنية المحددة، ولكن هنالك سلوك اقتصادي يمكن ان يؤثر على سير السلسة الزمنية ويتأثر هذا السلوك من خلال السنوات السابقة للحالة قيد الدراسة، حيث ان السلوك الاقتصادي يكون متحركاً وربما يحدث الأثر بشكل جوهري بعد سببه مثال على ذلك انفاق الاسر على الاستهلاك لا يعتمد فقط على مستوى الدخل الحالي بل يعتمد على مستوى الدخل في السنوات السابقة، وهذا يعني ان حالة استجابة المتغير التابع للمتغيرات المستقلة ستكون في حالة تباطؤ عبر الزمن. وتعرف تلك النماذج
(بنماذج المتغيرات المرتدة زمنياً).

تعريف النموذج المرتد زمنياً:

يعرّف النموذج المرتد زمنياً بانه: النموذج الذي يحتوي على متغيرات خارجية او داخلية مرتدة زمنياً وتكون من بين المتغيرات الموجودة في النموذج والتي يكون تأثير المتغير المســـتقل فيها على المتغير المعتمد يتوزع عبر عدد معين من قيم المتغير المســـتقل المرتد زمنياً والى (S) من فترات الارتداد الزمنية والتي يمكن ان تكون محددة Finite او غير محدودة Infinite.

وبشكل عام يمكن كتابة الصيغة الرياضية لنموذج الارتداد الزمني كالاتي:

$$y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots \; + \; \beta_k X_{t-k} + u_t \; \ldots (1)$$

ويمثل النموذج (1) أعلاه العلاقة بين الانفاق الاســتهلاكي(y_t) الذي يمثل المتغير المعتمد، والمتغير المستقل (X_t) خلال فترات زمنية سابقة $(t_{-1}\,, t_{-2}\,, \ldots .\, t_{-k})$. وتمثل المعلمة (β_0) التأثير القصــير الاجل لأنها تعطي التغير في متوســط قيمة المتغير المعتمد (y) عند تغير المتغير المســتقل (X) بمقدار وحدة واحدة لنفس الفترة الزمنية. وان $(\beta_0 + \beta_1)$ تعطي التغير في متوسـط قيمة (y) للفترة القادمة. وتشـــير المعلمات $(\beta_2 + \beta_1)$ الى التأخر الزمني الذي يقيس متوســـط التأثير على المتغير التابع (y_t) عندما يتغير (X_t) بوحدة واحدة خلال الفترات الزمنية الســابقة، وان (u_t) هو المتغير العشوائي.

وان: $\sum_{i=0}^{k} \beta_i = \beta_0 + \beta_1 + \beta_2 \, \ldots + \beta_k = \beta$

تمثل معاملات الارتداد طويلة الأمد او الاجل (الأثر طويل الاجل).

وإذا عرفنا:

$$\beta_i^* = \frac{\beta_i}{\sum_{i=1}^{k} \beta_i}$$

سنحصل على (β_i) المعيارية.

هذا وتعد المتغيرات المرتدة زمنياً متغيرات مستقلة مهمة جداً في معظم العلاقات الاقتصادية، ذلك لان السلوك الاقتصادي لكل فترة زمنية محددة يتأثر الى حد ما بنمط السلوك السائد في الفترة الزمنية السابقة، كما وتعد نماذج الارتداد الزمني من النماذج المهمة جداً في عملية اتخاذ القرار.

أسباب وجود الارتداد الزمني: -

يظهر الارتداد الزمني في نماذج القياس الاقتصادي لعدة أسباب أهمها ما يلي:

أ- العادات والتقاليد الاستهلاكية للمجتمعات (الأسباب النفسية).

ب- تغيرات الأسعار والأجور وتصاميم واشكال السلع (الأسباب الفنية).

ت- التشريعات الحكومية وتأثيرها في القرار الاقتصادي (أسباب مؤسسية وتشريعية).

لذلك يعتبر الارتداد الزمني مهم جداً في تحليل النشاط الاقتصادي لكل فترة زمنية سواء كان في الاجل القصير او الاجل الطويل. فمثلاً نقول ان الميل الحدي للاستهلاك (mpc) في الاجل القصير اقل منه في الاجل الطويل بوجود الارتداد الزمني في العوامل المؤثرة على الاستهلاك، وكذلك الحال بالنسبة للقيمة المطلقة لمرونة الدخل.

تقدير نماذج الارتداد الزمنية: Estimated of Lagged Models

1-نموذج كويك. Koyck Model

اقترح هذا العالم طريقة جيدة ومناسبة لتقدير نماذج التوزيع المرتد، حيث افترض بأن اوزان المتغيرات المستقلة المرتدة زمنياً تتناقص تدريجياً بمتوالية هندسية وفق الصيغة الاتية:

$$\beta_k = \beta_0 \lambda^k \quad , \quad 0 < \lambda < 1 \quad , (k = 0,1,2 \dots) \qquad \dots . (2)$$

حيث تمثل (λ) سرعة الاستجابة او معامل التباطؤ. وان $(1 - \lambda)$ تمثل سرعة التوافق او التكيف. ويوضح شرط المتوالية الهندسية أعلاه ان المعلمات (β_k) تتناقص بصورة مستمرة، اذ ان $(\lambda < 1)$ فكلما بَعُد الزمن كلما قل تأثير المتغير المرتد على المتغير التابع.

ويتضـــح من العلاقة (2) أعلاه ان كل معامل من معاملات (β) المتعاقبة هي اقل من الســـابقة لها وكما يلي:

$$\beta_1 = \beta_0 \lambda^1$$

$$\beta_2 = \beta_0 \lambda^2$$

$$\beta_3 = \beta_0 \lambda^3$$

.

.

.

$$\beta_k = \beta_0 \lambda^k$$

أي ان قيمة معاملات الارتداد (β_k) تعتمد على قيمة المعامل (β_0) وعلى قيمة (λ)، وتبدأ قيم (β_k) بالتناقص الى ان تصل الى الصفر تقريباً والمثال التالي يوضح ذلك:

λ	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
0.7	β_0	0.7 β_0	0.49 β_0	0.34 β_0	0.24 β_0	0.17 β_0	0.12 β_0	0.08 β_0	0.06 β_0	0.04 β_0	0.03 β_0
0.3	B_0	$0.03\ B_0$	$0.09\ B_0$	$0.03\ B_0$	$0.01\ B_0$	$0.002\ B_0$	$0.001\ B_0$	$0.0002\ B_0$	$0.0001\ B_0$	$0.00001\ B_0$	0.00

من ميزات نموذج (كويك Koyck) ان قيمة (λ) هي غير سالبة وان مجموع (B) والتي تعطي الارتداد طويل الامد هي لانهائية أي ان:

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 \left(\frac{1}{1-\lambda}\right)$$

اذ ان:

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 \sum_{k=0}^{\infty} \lambda^k$$

$$= \beta_0 (1 + \lambda + \lambda^2 + \lambda^3 + \cdots)$$

$$= \beta_0 \left(\frac{1}{1-\lambda}\right)$$

حسب قانون المجموع لمتوالية هندسية لا نهائية = $\frac{\text{الحد الاول}}{1-\text{الاساس}}$

وبتعويض العلاقة (2) في نموذج الارتداد الزمني (1) أعلاه يمكن كتابته بالشكل التالي:

$$y_t = \alpha + B_0\, X_t + B_0\, \lambda\, X_{t-1} + B_0\, \lambda^2\, X_{t-2} + \cdots \quad + u_t \ldots (3)$$

ان النموذج في (3) أعلاه ليس من الســـهولة تقديره طالما يحتوي على عدد لانهائي من المعالم، حيث ان طرق التقدير المعروفة لا يمكن تطبيقها على هكذا نماذج، لذلك اقترح العالم (كويك Koyck) طريقة ارتد بها بفترة زمنية واحدة للنموذج في (3) أعلاه (جعل النموذج (3) مرتد لفترة زمنية واحدة) لنحصل على:

$$y_{t-1} = \alpha + B_0\, X_{t-1} + \lambda\, B_0\, X_{t-2} + \lambda^2 B_0\, X_{t-3} + \cdots \quad + u_{t-1} \ldots (4)$$

وبضرب طرفي النموذج في (4) في λ نحصل على:

$$\lambda\, y_{t-1} = \lambda\, \alpha + \lambda\, B_0\, X_{t-1} + \lambda^2 B_0\, X_{t-2} + \lambda^3 B_0\, X_{t-3} + \cdots \quad + \lambda\, u_{t-1} \ldots (5)$$

وبطرح المعادلة (5) من المعادلة (3) نحصل على:

$$y_t - \lambda\, y_{t-1} = \alpha - \lambda\, \alpha + \; B_0\, X_t + u_t - \lambda\, u_{t-1} \ldots (6)$$

$$y_t - \lambda\, y_{t-1} = \alpha(1 - \lambda\,) + \; B_0\, X_t + u_t - \lambda\, u_{t-1}$$

وبالمناقلة والترتيب نحصل على ان:

$$y_t = \alpha(1 - \lambda\,) + B_0\, X_t + \lambda\, y_{t-1} + V_t \quad \ldots \quad (7)$$

حيث ان $V_t = u_t - \lambda u_{t-1}$ هو حد الخطأ الجديد

وتســمى هذه الصــيغة في (7) أعلاه بتحويلة (مبدلة) (كويك Koyck Transformation) والتي تســتخدم في نقل (تحويل) النموذج الأصــلي والذي يحتوي على عدد لانهائي من المتغيرات المســتقلة المرتدة زمنياً الى نموذج يحتوي على متغيرين مســتقلين فقط هما $(y_{t-1}\ ,\ X_t)$ كما تناقصـــت عدد المعلمات المطلوب تقديرها من العدد اللانهائي في النموذج الأصـــلي الى ثلاث معلمات فقط في نموذج كويك المحول هي $(\lambda\,, B_0\ \,,\alpha)$. وبتطبيق طريقة (OLS) يمكن تقدير

مـــعـــلـــمـــات مـــعـــادلـــة كـــويـــك لـــلـــحـــصـــول عـــلـــى الـــمـــقـــدرات $(\hat{\lambda}, \hat{B}_0 , \hat{\alpha})$ ثم تجري عملية استعادة تقديرات المعلمات الاصلية (B_K) بتطبيق الصيغة:

$$\beta_k = \beta_0 \lambda^k \qquad , (k = 0,1,2, \ldots\ldots)$$

مثال: فلو كانت نتائج تقدير العلاقة (7) أعلاه كالاتي:

$$\hat{y}_t = 64 + 0.3X_t + 0.6y_{t-1}$$

عندئذ فان:

$$\hat{\alpha}(1 - \hat{\lambda}) = 64 \quad , \quad \hat{\lambda} = 0.6 \quad \text{و ان} \quad \hat{B}_0 = 0.3$$

لذلك فان:

$$1 - \hat{\lambda} = 1 - 0.6 = 0.4$$

وان:

$$\hat{\alpha} = \frac{64}{1-\hat{\lambda}} = \frac{64}{0.4} = 160$$

وعند تعويض هذه المقدرات في النموذج (3) أعلاه والمعاد كتابته:

$$y_t = \alpha + B_0 X_t + B_0 \lambda X_{t-1} + B_0 \lambda^2 X_{t-2} \ \ldots + u_t$$

نحصل على ان:

$$\hat{y}_t = 160 + 0.3 X_t + 0.18 X_{t-1} + 0.108 X_{t-2} \ \ldots$$

واستناداً لكل ما سبق يمكن إثارة الملاحظات الاتية حول نموذج كويك:

- بدأنا بنموذج متغيرات مرتدة زمنياً موزعة بصورة لانهائية، غير انه تحول الى نموذج انحدار ذاتي لـ y_t على X_t ، y_{t-1}
- ان المتغير التابع المرتد زمنياً يظهر ضمن المتغيرات المستقلة مما يخرق الفرض اللازم للحصول على مقدرات (OLS) والمتضمن ثبات قيم المتغيرات المستقلة في المعاينات المتكررة.
- ان حد الخطأ العشوائي الخاص بالنموذج الأصلي هو (u_t) بينما حد الخطأ الخاص بنموذج كويك Koyck المحول هو (V_t) فاذا كانت (u_t) لا ترتبط ذاتياً فحسب فروض (OLS) فان (V_t) لا تحتفظ بتلك الخاصية.

ونتيجة العشـــوائية (y_{t-1}) وعدم اســتقلال (V_t)عن قيمها الســـابقة لذلك تنتهك فروض (OLS) الاعتيادية مما يقود الى مقدرات متحيزة وغير متسقة لنموذج كويك المتخلف زمنياً.

مثال: البيانات التالية تمثل مصاريف الاستهلاك y_t والدخل X_t وان مصاريف الاستهلاك مرتدة زمنياً لفترة واحدة (y_{t-1}).

y_t	13.5	14.1	14.2	14.3	14.4	14.5	14.8	15.1
X_t	15.3	15.8	16.1	16.2	16.3	16.4	16.7	17.8

المطلوب: تقدير العلاقة (7) أعلاه الخاصة بنموذج كويك:

الحل: من البيانات السابقة نحصل على ان:

$$(\acute{X}X)=\begin{bmatrix} 8 & 130 & 99.8 \\ & 2135.76 & 1660.49 \\ & & 1423.84 \end{bmatrix}=\begin{bmatrix} n & \Sigma X_t & \Sigma y_{t-1} \\ \Sigma X_t & \Sigma X_t^2 & \Sigma X_t y_{t-1} \\ \Sigma y_{t-1} & \Sigma X_t y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}$$

$$\acute{X}Y\begin{bmatrix} \Sigma y_t \\ \Sigma X_t y_t \\ \Sigma y_t y_{t-1} \end{bmatrix}=\begin{bmatrix} 114.9 \\ 1878.07 \\ 1459.93 \end{bmatrix}$$

$$\mathrm{b}=(\acute{X}X)^{-1}\,\acute{X}Y$$

لذلك فان المعادلة التقديرية ستأخذ الشكل التالي:

$$\hat{y}_t = 7.7766 + 0.3494\,X_t + 0.0728\,y_{t-1}$$

$$R^2 = 0.995$$

B_0 : تمثل الميل الحدي للاستهلاك في الاجل القصير وهو: $\mathrm{B}_0 = MPC = 0.3494$

$$\hat{\lambda} = 0.0728$$

$$1-\hat{\lambda} = 1-0.0728 = 0.9272$$

$$\hat{\alpha}(1-\hat{\lambda}) = 7.7766$$

$$\therefore \quad \hat{\alpha} = \frac{7.7766}{0.9272} = 8.387$$

وهذا يوضح ان مصاريف الاستهلاك تسير بشكل منتظم مع الدخل.

وعند تعويض هذه المقدرات في العلاقة (3) نحصل على النموذج الأصلي بالشكل الاتي:

$$\hat{y}_t = 8.387 + 0.3494\, X_t + 0.0254\, X_{t-1} + 0.00185\, X_{t-2}$$

مؤشرات الابطاء (التأخر) الزمنية:

من نماذج الارتداد الزمنية يمكن الحصول على مؤشرات التأخر او الابطاء الزمنية والتي يمكن حسابها بعد تقدير النموذج وتتمثل بالآتي:

1- الوسيط المرتد زمنياً: The Median Lag

هو الوقت المطلوب للنصف الأول أو ما يمثل 50% من التغير الحاصل في المتغير المعتمد (y_t) الناتج من تغير (X_t) بمقدار وحدة واحدة ولنموذج كويك Koyck وتستخدم الصيغة التالية:

$$Median\ lag\ \ or\ Ml = \frac{\log(0.5)}{log\lambda}$$

$$Ml = -\frac{\log 2}{log\lambda}$$

ولو فرضنا $(\lambda = 0.2)$ فان:

$$Median\ lag\ \ = -\frac{0.301}{(-0.699)} = 0.4306$$

وإذا كانت $\lambda = 0.8$ فسيكون الوسيط هو:

$$Median\ lag\ \ = -\frac{0.301}{(-0.097)} = 3.1067$$

وهذا يعني انه كلما كبرت قيمة (λ) كلما ازدادت قيمة الوسيط. وان صغر قيمة (λ) معناها كبر سرعة التعديل أي قصر فترة الابطاء (الارتداد)، وكبر قيمة (λ) معناها صغر سرعة التعديل أي إطالة فترة الابطاء.

2- متوسط فترة الابطاء (الوسط الحسابي المرتد زمنياً): Average (mean) Lag

ويسمى ايضاً بالإبطاء الموزون (Lag – weighted Average) ويقيس ايضاً سرعة استجابة (y_t) للتغير في (X_t). ويحسب وفق الصيغة الاتية:

$$Al = \frac{\lambda}{1-\lambda}$$

فاذا كانت 0.5= λ فان المتوسط يساوي واحد.

ولبيانات المثال السابق وبما ان قيمة 0.0728= λ فان:

$$Ml = \frac{-log\ 2}{log\ \lambda} = \frac{\log(0.5)}{log\ \lambda} = -\frac{0.301}{(-1.1379)} = 0.265$$

وهذا هو الزمن المطلوب لنصــف أو 50% من التغير في مصــاريف الاســتهلاك (y_t) الناتج من تغير الدخل (X_t) بمقدار وحدة واحدة. أي ما يعادل ثلاثة أشهر تقريباً.
وان متوسط فترة الابطاء هو:

$$Al = \frac{\lambda}{1-\lambda} = \frac{0.0728}{0.9272} = 0.0785$$

ويقيس سرعة الاستجابة للتغير في الدخل والبالغة اقل من شهر واحد.

3- الأثر القريب المدى: Short Run Impact (SR)

ويحسب وفق الصيغة الاتية: $SR = B_0$

ويدل على ان كل زيادة في (X_t) بمقدار وحدة واحدة يؤدي الى زيادة (y_t) بمقدار (B_0). خلال السنة نفسها.

4- الأثر البعيد المدى: Long Run Impact (LR)

ويحسب وفق الصيغة الاتية: $LR = \frac{B_0}{1-\lambda}$

ويشير الى مجموع التأثيرات بدلالة مرونة التأخر الزمني (λ) بعد انتهاء فاعلية الأثر القريب المدى (SR).

5- تباين توزيع الابطاء (تباين فترة التأخير): Variance of Lag (VL)

ويحسب وفق الصيغة الاتية: $VL = \frac{\lambda}{(1-\lambda)^2}$

ويُعبّر تباين فترة التأخير عن الاختلاف بين الفترة الفعلية والمقدرة للتنفيذ.

6- معامل التقاطع المعدل.

ويحسب وفق الصيغة الاتية: $\alpha^* = \frac{\hat{\alpha}}{1-\lambda}$

7- اوزان المتغيرات المبطأة (المرتدة) زمنياً: يمكن حسابها من معامل (y_{t-1}) وبموجب الصيغة الاتية:

$$Wi = (1-\lambda)\lambda^i$$

حيث ان i=0,1,2,... n ، وان: $\Sigma Wi = 1$

مثال: باحث اقتصادي قدّر نموذج كويك باستخدام الصيغة اللوغاريتمية المزدوجة وبطريقة (OLS) وكانت النتائج كالاتي:

$$lny_t = 0.710 + 0.385\, lnX_t + 0.460 \ln y_{t-1}$$
$$(t) \quad (1.15) \quad (4.11) \quad (3.73)$$

$$R^2 = 0.71\ , \bar{R}^2 = 0.69\ , F = 30.4\ ,\ D-W = 2.14\ , h = -0.492$$
$$dl = 1.27\ , du = 1.56\ ,\ n = 29\ ,\ r_{xt}y_{t-1} = 0.66$$

حيث ان:

y_t : الإنتاج من الرز (بالألف طن) للسنة t.

X_t : المساحة المزروعة بالرز (ألف دونم) للسنة t.

والمطلوب:

1- تقييم نتائج تقدير النموذج؟

2- حساب جميع مؤشرات الابطاء (التأخر) الزمنية؟

ملاحظة: في حالة وجود المتغيرات المرتدة زمنياً كمتغيرات مستقلة فان اختبار (D-W) لا يصلح لاختبار وجود الارتباط الذاتي، ولذلك اقترح ديربن (J.Durbin) اختباراً مختلفاً لتشخيص وجود الارتباط الذاتي ويعرف باختبار (h) لاختبار فرضية العدم الاتية: $H_0:\ \hat{\mathcal{P}} = 0$

مقابل الفرضية البديلة: $H_1:\ \hat{\mathcal{P}} \neq 0$

ان إحصاءة اختبار (h) تأخذ الصيغة الاتية:

$$h = (1 - \frac{D-W}{2}) \sqrt{\frac{n}{1 - n\,.\,var(b_1)}}$$

اذ ان:

n: حجم العينة.

$var(b_1)$: تباين معلمة المتغير (y_{t-1}) في النموذج المقدّر.

وان (h) تقع قيمتها بين (∓ 1.96) عند مستوى معنوية 5%.

$\hat{\mathcal{P}} = 1 - \frac{D-W}{2}$: معامل الارتباط الذاتي

$$var(b_1) = 0.01520891$$

الحل:

1- جاءت نتائج تقدير النموذج متفقة مع منطق النظرية الاقتصادية، والاختبارات الإحصائية (R^2 , F , t) تؤكد معنوية المعالم المقدّرة ومعنوية النموذج عند مستوى معنوية (1%) وان المتغيرات المستقلة (المستعملة) تفسّر 71% من التغيرات الحاصلة في الإنتاج، فضلاً عن خلو الدالة من الارتباط الذاتي بحسب اختبار (h) واختبار (D-W)، كما يؤكد اختبار (كلاين Klien) خلو النموذج المقدر من مشكلة التعدد الخطي حيث كان:

$$R^2 > r^2_{xt}y_{t-1}$$

2- في ضوء نتائج تقدير النموذج يمكن حساب المؤشرات الاتية للإبطاء الزمنية:

أ- **الأثر القريب المدى:** $SR = b_0 = 0.385$

حيث يبين الأثر القريب المدى (SR) بان زيادة المساحة بنسبة (100%) سيؤدي الى زيادة الإنتاج بنسبة (38.5%) خلال السنة نفسها.

ب - الأثر البعيد المدى:

$$LR = \frac{b_0}{1-\lambda} = \frac{0.385}{1-0.460}$$

$$\therefore\ LR = \frac{0.385}{0.54} = 0.713$$

ويشير الى ان زيادة مجموع تأثيرات المساحة بنسبة (100%) سوف يؤدي الى زيادة الإنتاج بنسبة (71.3%) على المدى البعيد أي بعد انتهاء فاعلية الأثر القريب المدى (SR).

ت - وسيط فترة الابطاء (التأخير):

$$ML = \frac{-log\ 2}{log\ \lambda} = \frac{log\ (0.5)}{log\ \lambda} = \frac{log\ (0.5)}{log\ (0.460)} = \frac{-0.301}{-0.337} = 0.893$$

وهذا هو الزمن المطلوب لنصف أو 50% من التغير في الإنتاج الناتج من تغير المساحة بمقدار الوحدة الواحدة، وهو ما يعادل عشرة أشهر ونصف تقريباً.

ث - متوسط فترة التأخير (الابطاء):

$$AL = \frac{\lambda}{1-\ \lambda} = \frac{0.460}{1-0.460} = \frac{0.460}{0.54} = 0.851$$

وقيمته قريبة من قيمة الوسيط وهو يقيس سرعة الاستجابة للتغير في المساحة والبالغة عشرة أشهر.

ج - تباين فترة الابطاء:

$$VL = \frac{\lambda}{(1-\ \lambda)^2} = \frac{0.460}{(0.54)^2} = \frac{0.460}{0.2916} = 1.58$$

ويعبر عن الاختلاف بين الفترة الفعلية والفترة المقدرة للإنتاج النهائي (التنفيذ) والبالغة أكثر من سنة ونصف تقريباً.

ح - معامل (حد) التقاطع المعدل:

$$\alpha^* = \frac{\hat{\alpha}}{1-\lambda} = \frac{0.710}{0.540} = 1.315$$

خ - اوزان المتغيرات المرتدة (المبطأة) زمنياً:

$$W_i = (1-\lambda)\lambda^i \qquad , i = 0,1,2, \dots . n$$

وان $\Sigma wi = 1$

لذلك فان:

$$W_0 = (0.54)(0.460)^0 = 0.54$$

$$W_1 = (1 - 0.46)(0.460)^1 = (0.54)(0.460) = 0.2484$$

$$W_2 = (0.54)(0.460)^2 = 0.1143$$

$$W_3 = 0.0526 \quad , \quad W_4 = 0.0242 \quad , W_5 = 0.0111$$

ويلاحظ من حساب الاوزان المتأخرة زمنياً انها تتناقص تدريجياً سنة بعد أخرى على شكل متوالية هندسية، وان السنوات الثلاثة الأولى تشكل 90% من مجموع الاوزان المتباطئة. وهذا يُعد منطقياً وكدليل على محدودية المساحة، لان قرار المنتج تحدده عوامل أخرى غير المساحة مثل: توفر المياه والذي يعد محدداً رئيسياً لزراعة الرز فضلاً عن عامل الأسعار.

واستناداً الى النتائج المقدرة والاوزان المحسوبة يمكن إعادة كتابة نموذج (كويك Koyck) لدالة انتاج الرز على الشكل التالي:

$$\begin{aligned} lny_t = 1.315 + 0.385 \, ln\, X_t + 0.54 \, ln\, X_{t-1} + 0.2484 \, ln\, X_{t-2} \\ + 0.1143 \, ln\, X_{t-3} + 0.0526 \, ln\, X_{t-4} + 0.0242 \, ln\, X_{t-5} \\ + 0.0111 \, ln\, X_{t-6} \end{aligned}$$

2- نموذج التوقع المكيف لـ Cagan- Friedman

يقودنا هذا النموذج الى بقاء الارتداد الزمني بشكل فعلي متطابق مع نموذج كويك، غير ان المتغير (y_t) يعتمد على القيم المتوقعة او الدائمة للمتغير المستقل (X_t^*) وليست على القيم الحقيقية له (X_t).
لذلك يستخدم نموذج
(كاكان Cagan) للتوقع المكيف نموذج الانحدار الذاتي والذي يكون المتغير المستقل فيه له قيم متوقعة مثلى.

ويمكن كتابة هذا النموذج بالصيغة الاتية:

$$y_t = \alpha + BX_t^* + u_t \dots \dots (11)$$

اذ ان:

y_t : الكمية المطلوبة من سلعة معينة.

X_t^* : السعر المتوقع للسلعة.

t : الزمن ، u_t : الخطأ العشوائي.

ان التفسير الأكثر شيوعاً لــ (X_t^*) في النظرية الاقتصادية هو انه يمثل الدخل الدائم حسب نظرية (فريد مان) للاستهلاك ويعبر عن التوازن طويل الأمد للمتغير المستقل، وكذلك في دالة الطلب على النقود حيث يمثل توازن معدل الفائدة الطبيعية طويلة الأمد.

ويتم تحديد(X_t^*)حسب الصيغة التالية لعدم امكانية جمع بيانات مباشرة عنها:

$$X_t^* - X_{t-1}^* = \mu(X_t - X_{t-1}^*) \quad .... \quad (12)$$

اذ ان: $0 \leq \mu \leq 1$

وهذا معناه ازدياد مشتريات السلع عند توقع زيادة أسعارها في المستقبل، وان الدخل الحقيقي(X_t) في الفترة (t) يفوق الدخل المتوقع في الفترة السـابقة (X_{t-1}^*) وتعرف (μ) بمعامل التوقع المكيف (Adjustment coefficient of Expectation) وتعرف هذه الفرضـية ايضـاً بفرضـية التوقع المتطور (التلائمي) (Progressive Expectation) او فرضـية تعلم الخطأ حيث تشـير المعادلة (12) أعلاه الى ان الأدوات الاقتصــادية ســوف تُكّيف توقعاتها في ضــوء التجربة الســابقة أي في الواقع يتم التعلم من الأخطاء السابقة. ويمكن إعادة كتابة المعادلة (12) بالشكل التالي:

$$X_t^* = \mu\, X_t + (1-\mu)X_{t-1}^* \,.......(13)$$

والتي تُظهر القيمة المتوقعة لمعدل الفائدة او الدخل في الوقت (t) كمعدل موزون للقيمة الفعلية لمعدل الفائدة في الوقت (t) والقيمة المتوقعة للفترة الســــابقة بالأوزان $(\mu$ و $(1-\mu)$ على التوالي، وحيث ان(μ) هي معامل التوقع وان قيمة (μ) تقع بين الصفر والواحد الصحيح.

فعندما تكون $(\mu = 1)$ فان الدخل الحقيقي سـيتطابق مع الدخل المتوقع $(X_t^* = X_t)$ وهذا يعني ان التوقعات أنجزت بشكل فوري وتام ولنفس الفترة.

وعندما تكون $(\mu = 0)$ فان الدخل المتوقع يبقى بدون تغيير ولا يتأثر بالدخل الحقيقي مهما كان حجمه، أي ان التوقعات تبقى ســـاكنة. وبشـــكل عام كلما كانت قيمة (μ) اكبر كلما كان امتداد

التعديل اكبر. ويمكن ملاحظة ذلك وكالاتي: عند تعويض المعادلة (13) في المعادلة (11) أعلاه نحصل على:

$$y_t = \alpha + B\{\mu\, X_t + (1-\mu)X_{t-1}^*\} + u_t$$

$$y_t = \alpha + \ B\,\mu\, X_t + B(1-\mu)X_{t-1}^* + u_t \quad \ldots.(14)$$

وإذا عملنا ارتداد زمني للمعادلة (11) أعلاه لفترة واحدة نحصل على الصيغة التالية:

$$y_{t-1} = \alpha + \ B\, X_{t-1}^* + u_{t-1} \qquad \ldots.(15)$$

وبضرب طرفي المعادلة (15) في $(1-\mu)$ ينتج ان:

$$(1-\mu)y_{t-1} = \alpha(1-\mu) + \ \beta(1-\mu)X_{t-1}^* + (1-\mu)u_{t-1} \qquad \ldots.(16)$$

وبطرح المعادلة (16) من المعادلة (14) ينتج ان:

$$y_t = \mu\,\alpha + \ \mu\,\beta X_t + (1-\mu)y_{t-1} + V_t \qquad \ldots.(17)$$

حيث ان: $V_t = u_t(1-\mu)u_{t-1}$

مثال: على افتراض ان نتائج تقدير العلاقة (17) كانت كالاتي:

$$\hat{y}_t = 64 + 0.3\, X_t + 0.6\, y_{t-1}$$

وفي ضوء العلاقة (17) فان:

$$1 - \hat{\mu} = 0.6 \quad \Rightarrow \quad \therefore\ \hat{\mu} = 0.4$$

وان: $\hat{\mu}\,\hat{\beta} = 0.3 \quad \Rightarrow \quad \therefore\ \hat{\beta} = \frac{0.3}{\hat{\mu}} = \frac{0.3}{0.4} = 0.75$

وان: $\hat{\mu}\,\hat{\alpha} = 64 \quad \Rightarrow \quad \therefore\ \hat{\alpha} = \frac{64}{\hat{\mu}} = \frac{64}{0.4} = 160$

وفي ضوء ذلك فان تقدير المعادلة السلوكية في (11) سيأخذ الشكل الاتي:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}\, X_t^* + e_i$$

هو : $\hat{y}_t = 160 + 0.75\, X_t^*$

أي ان معامل التوقع (التعديل) المكيف في معادلة التوقعات المعدلة هو (0.4) وهذه القيمة يمكن استخدامها لتمثل (40%) لأي اختلاف بين الدخل الحقيقي والدخل المتوقع.

مثال(1): اذا توفرت البيانات الاتية عن الكمية المطلوبة من سلعة ما (y) والسعر الحقيقي لتلك السلعة (X) كما في الجدول الاتي:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_t	30.6	31.6	31.3	33.3	33.5	33.2	36.7	38.6	39.0	40.8	42.7	41.9	40.2	40.7	40.4
X_t	125	140	130	155	145	163	170	182	173	192	203	178	163	182	175

المطلوب: تقدير نموذج التوقع المكيف في العلاقة (11). وكيف يمكن إيجاد السعر المتوقع للسلعة X^*؟

الحل: أ- لا يمكن تقدير المعادلة (11) من بيانات الجدول أعلاه لان قيمة (X_t^*) غير معلومة. لذلك لا بد من ايجادها اولاً ومن ثم يمكن تقدير النموذج.

ب – إيجاد السعر المتوقع (X_t^*) من خلال الخطوات الاتية:

1- باستخدام طريقة (OLS) وتطبيقها على المعادلة (17) نحصل على الشكل المقدّر الاتي:

$$\hat{y}_t = 1.95825 + 0.0805855\, X_t + 0.598679\, y_{t-1}$$

ومنه نحصل على:

$$1 - \hat{\mu} = 0.598679 \quad \Rightarrow \quad \therefore\ \hat{\mu} = 1 - 0.598679 = 0.401321$$

وان:

$$\hat{\mu}\,\hat{\alpha} = 1.95825 \quad \Rightarrow \quad \therefore\ \hat{\alpha} = \frac{1.95825}{0.401321} = 4.87951$$

وان:

$$\hat{\mu}\,\hat{\beta} = 0.0805855 \quad \Rightarrow \quad \therefore\ \hat{\beta} = \frac{0.0805855}{0.401321} = 0.20080$$

وبالتعويض نحصل على تقدير المعادلة (11) وكالاتي:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}\, X_t^* + e_i$$

$$\therefore\ \hat{y}_t = 4.87951 + 0.2008\, X_t^*$$

2- باستخدام تقدير المعادلة (11) أعلاه يمكن إيجاد قيمة السعر المتوقع (X_t^*) وكما يلي:

بإعادة ترتيب المعادلة (11) نحصل على ان: $\hat{\beta}\, X_t^* = \hat{y}_t - \hat{\alpha}$

وبقسمة طرفي المعادلة على $\hat{B}$ نحصل على ان: $X_t^* = \frac{\hat{y}_t - \hat{\alpha}}{\hat{\beta}}$

وبالتعويض عن قيم (y_t) الواردة في الجدول نحصل على قيم السعر المتوقع (X_t^*) المقابل وكالاتي:

$$X_1^* = \frac{30.6 - 4.87951}{0.20080} = 128.1$$

$$X_2^* = \frac{31.6 - 4.87951}{0.20080} = 133.1$$

$$X_3^* = \frac{31.3 - 4.87951}{0.20080} = 131.6$$

.

.

.

وهكذا نستمر بالحصول على قيم (X_t^*) الى ان نحصل على (من الجدول السابق):

$$X_{15}^* = \frac{40.4 - 4.87951}{0.20080} = 176.9$$

مثال (2): البيانات التالية تمثل الكمية المطلوبة (y_t) من احدى السلع والسعر الحقيقي لتلك السلعة (X_t). **المطلوب:** إيجاد السعر المتوقع للسلعة باستخدام (نموذج Cagan) ؟

y_t	27.5	28.5	28.2	30.2	30.4	30.1	33.6	35.5	35.9	37.7	39.6	38.8	37.1	37.6	37.3
X_t	123	137	127	152	142	160	167	179	170	189	200	175	160	179	172

الحل: يتم إيجاد المعادلة التقديرية الاتية باستخدام طريقة المربعات الصغرى الاعتيادية (OLS):

$$y_t = \mu\,\alpha + \mu\,\mathrm{B}\,X_t + (1-\mu)y_{t-1} + V_t$$

من البيانات أعلاه نحصل على ان:

$$(\acute{X}X) = \begin{bmatrix} n & \Sigma X_t & \Sigma y_{t-1} \\ \Sigma X_t & \Sigma X_t^2 & \Sigma X_t y_{t-1} \\ \Sigma y_{t-1} & \Sigma X_t y_{t-1} & \Sigma {y^2}_{t-1} \end{bmatrix}, \begin{bmatrix} 15 & 2432 & 470.7 \\ & 401296 & 78596.1 \\ & & 16069.03 \end{bmatrix}$$

وان:

$$\acute{X}Y \begin{bmatrix} \Sigma y_t \\ \Sigma X_t y_t \\ \Sigma y_t y_{t-1} \end{bmatrix} = \begin{bmatrix} 508 \\ 83577.6 \\ 16398.07 \end{bmatrix}$$

$$\therefore\ b = (\acute{X}X)^{-1}\,\acute{X}Y$$

ومنها نحصل على المعادلة التقديرية الاتية:

$$\hat{y}_t = 9.03 + 0.13\,X_t + 0.12\,y_{t-1}$$

ومنها نحصل على ان: $(1-\hat{\mu}) = 0.12 \ \Rightarrow\ \therefore\ \hat{\mu} = 1 - 0.12 = 0.88$

$$\hat{\mu}\,\hat{\alpha} = 9.03 \ \Rightarrow\ \therefore\ \hat{\alpha} = \frac{9.03}{\hat{\mu}} = \frac{9.03}{0.88} = 10.26$$

وان: $\hat{\mu}\,\hat{\beta} = 0.13 \ \Rightarrow\ \therefore\ \hat{\beta} = \frac{0.13}{0.88} = 0.15$

وبالتعويض عن معالم العلاقة الاتية: $y_t = \alpha + B\,X_t^* + e_i$

نحصل على المعادلة التقديرية: $\hat{y}_t = 10.26 + 0.15\,X_t^*$

وبإعادة ترتيب العلاقة أعلاه نحصل على السعر المتوقع وفق الصيغة الاتية:

$$\hat{\beta}\,X_t^* = \hat{y}_t - \hat{\alpha}$$

$$\therefore\ X_t^* = \frac{\hat{y}_t - \hat{\alpha}}{\hat{B}}$$

وبالتعويض عن قيم $(B^{\hat{}}\ ,\ \hat{\alpha}\ ,\hat{y}_t)$ نحصل على قيم (X_t^*) الخمسة عشر وكما يلي:

$$X_1^* = \frac{27.5 - 10.26}{0.15} = 114.93$$

$$X_2^* = \frac{28.5 - 10.26}{0.15} = 121.6 \quad , \ X_3^* = 119.6\ , \ X_4^* = 132.93$$

$$\vdots$$

$$X_{15}^* = \frac{37.3 - 10.26}{0.15} = 180.27$$

3- نموذج التعديل الجزئي لـ بيرلوف: Marc Nerlove

ويسمى في بعض الأحيان بنموذج التكيف الجزئي أو نموذج تعديل الرصيد (The Stock Adjustment). ولتوضيح ذلك نعود الى النظرية الاقتصادية التي تفترض في بعض نماذجها ان هناك توازناً طويل الأمد في كميات خزينها من رأس المال المطلوب لإدامة الإنتاج تحت حالة التقدم العملي والتكنولوجي ولسهولة العمل نفترض ان المستوى المطلوب (المرغوب) من رأس المال (y_t^*) هو دالة خطية للمخرجات (الإنتاج) (X_t) خلال الفترة (t) وكما يلي:

$$y_t^* = \alpha + \beta X_t + u_t \quad \ldots\ldots (18)$$

وبما ان مستوى رأس المال لا يمكن استخدامه مباشرة فان العالم (نيرلوف Nerlove) اقترح هذا النموذج وفق الفرضيات التالية:

ان التغير في خزين رأس المال الى المستوى الأمثل (y_t^*) يمثل في حقيقته التغير المتحقق في خزين رأس المال في كل فترة، لذلك من الطبيعي ان $(y_t$ و $y_t^*)$ ليست نفسها، بمعنى اخر فان تعديل (y_t) الحقيقية الى التغير في (X_t) لا يحصل بشكل فوري بل ربما يكون حصيلة التقدم التكنولوجي والقرارات الإدارية والمالية التي تقدم عليها المنشأة او الدولة في مجال الاستثمار. وعليه فان معادلة التعديل ستأخذ الصيغة الرياضية الاتية:

$$y_t - y_{t-1} = \lambda(y_t^* - y_{t-1}) + V_t \quad \ldots\ldots (19)$$

اذ ان: $0 < \lambda < 1$

λ : معامل التعديل (التكيف) (Coefficient of Adjustment) وتقيس مدى اقتراب او ابتعاد (y_t) الحالية عن السابقة (y_{t-1}) ، وتتراوح قيمتها بين الصفر والواحد الصحيح.

وان:

$y_t - y_{t-1}$: التغير الحقيقي (الفعلي).

$y_t^* - y_{t-1}$: التغير المطلوب (المرغوب فيه).

V_t : حد الخطأ.

بتعويض المعادلة (18) في المعادلة (19) نحصل على:

$$y_t - y_{t-1} = \lambda[(\alpha + \beta X_t + u_t) - y_{t-1}] + V_t \quad \dots\dots (20)$$

$$y_t - y_{t-1} = \lambda\,\alpha + \lambda\,\beta X_t - \lambda\, y_{t-1} + \lambda\, u_t + V_t$$

$$y_t = \lambda\,\alpha + \lambda\,\beta X_t + (1-\lambda)\, y_{t-1} + w_t \quad \dots\dots (21)$$

حيث ان: $w_t = (\lambda\, u_t + V_t)$

وهذه المعادلة توضح ان (y_t) الرصيد من رأس المال في الفترة (t) يعتمد في جزء منه على مستوى الإنتاج في تلك الفترة (X_t) وفي الجزء الاخر على التوسع في الرصيد في الفترة (t). كما ان درجة التعديل والسرعة التي يمكن تعديل (y) الى (X) عبر الفترة الزمنية تعتمد على حجم معامل التعديل (λ). فعندما $(\lambda = 1)$ معنى ذلك ان الرصيد الحقيقي لرأس المال سوف يكون دائماً مساوياً للرصيد المخطط له، واذا كانت $(\lambda = 0)$ فان $(y_t = y_{t-1})$ مما يعني عدم وجود أي تغيير في الرصيد (لا يكون هناك أي تعديل على الاطلاق).

وبمقارنة نموذج (نيرلوف Nerlove) (المعادلة 21) مع نموذج (كويك Koyck) (المعادلة 7) يتضح بان كلاهما يحتويان على المتغيرات نفسها $(y_{t-1}\,,\ X_t\,,\ y_t)$. غير ان حد الخطأ في نموذج (نيرلوف Nerlove) لا يحتوي على أي شكل من اشكال التعدد الخطي كما هو الحال في نموذج (كويك Koyck)، أي ان حد الخطأ $(\lambda\, u_t + V_t)$ في نموذج (نيرلوف Nerlove) لا يرتبط خطياً.

مثال3: البيانات التالية تمثل الانفاق الاستثماري المخطط (y_t) والمبيعات (X_t) باستخدام نموذج (نيرلوف Nerlove) قدّر معالم نموذج التعديل الجزئي قصيرة وطويلة الأمد لدالة الطلب على الانفاق الاستثماري؟

y_t	19	17	20	21	24	38	45	36	46	49
X_t	38	36	40	43	44	51	57	61	65	76

الحل: نتبع الخطوات التالية:

أ- صياغة نموذج التعديل الجزئي لنيرلوف كما في المعادلة (21) وهي:

$$y_t = \lambda\alpha + \lambda\beta X_t + (1-\lambda)\, y_{t-1} + w_t$$

ب – تطبيق طريقة (OLS) على المعادلة أعلاه مستخدماً البيانات في الجدول أعلاه ومنها نحصل على ان:

$$(\acute{X}X) = \begin{bmatrix} 10 & 511 & 435 \\ & 27697 & 23991 \\ & & 21921 \end{bmatrix}, \qquad \acute{X}Y \begin{bmatrix} 315 \\ 17506 \\ 23991 \end{bmatrix}$$

وان: $$\mathrm{b} = (\acute{X}X)^{-1}\, \acute{X}Y$$

ومنها نحصل على ان: $$\hat{y}_t = -8.7 + 0.88\, X_t + 0.03 y_{t-1}$$

لذلك فان: $$(1-\hat{\lambda}) = 0.03 \;\Rightarrow\; \therefore\; \hat{\lambda} = 0.97$$

وان: $$\hat{\lambda}\,\hat{\alpha} = -8.7 \;\Rightarrow\; \therefore\; \hat{\alpha} = \frac{-8.7}{0.97} = -1.03$$

$$\hat{\lambda}\,\hat{\beta} = 0.88 \;\Rightarrow\; \therefore\; \hat{\beta} = \frac{0.88}{0.97} = 0.91$$

لذلك فان: $$y_t^* = -1.03 + 0.91\, X_t$$

ملاحظة: توجد نماذج أخرى للارتداد الزمني.

تمارين:

1- البيانات التالية مصاريف الاستهلاك (y_t) والدخل (X_t) وان مصاريف الاستهلاك مرتدة زمنياً لفترة واحدة (y_{t-1}) .

المطلوب إيجاد:

أ- الميل الحدي للاستهلاك وقدّر نموذج كويك Koyck:

$$y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda y_{t-1} + V_t$$

ب- الوسط الحسابي والوسيط المرتد زمنياً. (البيانات بالألف دينار).

y_t	11.4	12.0	12.1	12.2	14.3	12.44	12.7	13.1
X_t	14.2	14.7	15.1	15.2	15.3	15.4	15.7	16.7

2- البيانات التالية تمثل الكمية المطلوبة (y_t) من احدى السلع والسعر الحقيقي لتلك السلعة (X_t) .

المطلوب إيجاد:

أ- السعر المتوقع للسلعة باستخدام نموذج (كاكان Cagan) .

$$y_t = \mu\,\alpha + \mu\,\beta X_t + (1-\mu)\,y_{t-1} + V_t$$

ب- الوسط الحسابي والوسيط المرتد زمنياً؟ ومؤشرات التأخر الزمني الأخرى؟

y_t	25.3	26.4	26.3	31.2	31.4	31.1	32.7	34.4	35.8	36.7	37.7	37.8	38.1	37.9	37.6
X_t	112	126	117	142	133	169	168	169	171	179	201	178	167	189	174

3- البيانات التالية تمثل الانفاق الاستثماري المخطط (y_t) والمبيعات (X_t) .

y_t	17	15	21	20	23	37	44	35	47	48
X_t	37	35	41	42	43	50	56	60	64	75

والطلوب:

أ- باستخدام نموذج (نيرلوف Nerlove) قدّر معالم نموذج التعديل الجزئي قصيرة وطويلة الأمد لدالة الطلب على الانفاق الاستثماري:

$$y_t = \lambda \alpha + \lambda \beta X_t + (1-\lambda)\, y_{t-1} + w_t$$

ب- استخرج الوسط الحسابي والوسيط المرتد زمنياً؟ ومؤشرات الارتداد الزمنية الأخرى؟

الفصــــل العاشر

الســلاســل الــزمنيــة في الاقتصاد القياسي

أولا: مفهوم السببية : Causality

يشير مفهوم السببية الى معاني ودلالات فلسفية عميقة اختلف الكتاب في توضيحها، الا انه يمكن توضيحها من خلال العبارة (كل شيء يسبب كل شيء)، هذه العبارة تشير الى ان احداث الماضي يمكن ان تكون سببا رئيسيا في احداث اليوم، كما يمكن ان تكون عنصرا مهما في احداث المستقبل، هذه المقدمة توضح بشكل مبسط مفهوم السببية العالم الاقتصادي أنجل كرانجر، وأول من قدم انموذج العلاقة السببية عالم الاقتصاد كرانجر عام 1969 ثم تم تطويره على يد العالم سيمز (Sims) عام 1972.

ويمكن توضيح مفهوم العلاقة السببية من خلال المثال الاتي:

لنفترض وجود متغيرين هما (X,Y) حيث ان (X) متغير مستقل والمتغير التابع (Y) فان المتغيرات السابقة في المتغير (X) تساعد في تفسير التغيرات الحالية في قيمة المتغير (Y) هذه العلاقة يمكن التعبير عنها بالعلاقة السببية تتجه من المتغير (X) الى المتغير (Y). ويحدث العكس عندما يكون المتغير (Y) مستقل والمتغير (X) تابع فان العلاقة السببية في هذه الحالة تتجه من المتغير (Y) الى المتغير (X).

وتساعد العلاقات السببية في النماذج الاقتصادية على التمييز بين المتغيرات الداخلية والمتغيرات الخارجية وهناك مسلمتان لابد من توفرهما في السببية هما:

أ- السببية لا تنطبق الا على المتغيرات العشوائية

ب- الماضي والحاضر يمكن ان يسبب المستقبل والعكس غير ممكن.

ويمكن تحديد العلاقة السببية تبعاً لصيغة انموذج الانحدار الذاتي بالصيغ الاتية:

$$Y_t = \sum_{i=1}^{n} a_i X_{t-i} + \sum_{J=1}^{n} \beta_J Y_{t-J} + U_1t \quad \ldots\ldots\ldots 1$$

$$X_t = \sum_{i=1}^{n} c_i X_{t-i} + \sum_{J=1}^{n} d_J Y_{t-J} + U_2t \quad \ldots\ldots\ldots 2$$

اذ ان:

U_1t , U_2t: حدي الخطأ العشوائي

$d_J - c_i - \beta_J - a_i$: معلمات مطلوب تقديرها.

وتندرج ضمن سببية كرانجر ثلاثة احتمالات للسببية واتجاهها تقسم وفق الاتي:

1- السببية أحادية الاتجاه:

تتجـه من المتغير $(Y \leftarrow X)$عنـدمـا تكون المعلمـات المقـدرة تختلف احصــــائيـا عن الصـــفر$(\sum_{i=1}^{n} a_i \neq 0)$ في المعادلة رقم (1) أعلاه، وان المعلمات المقدرة في المعادلة (2) لا تختلف احصائيا عن الصفر $(\sum_{J=1}^{n} d_J = 0)$ وهذه الحالة تدل على ان ماضي (X) يحسن من القدرة التنبؤية للمتغير (Y) في اللحظة t وتكون الحالة معاكســـة عندما تتجه العلاقة من $(X \leftarrow Y)$

2- السببية الثنائية او التبادلية $(Y \leftrightarrow X)$

في هذه الحالة يكون التأثير متبادل لكلا المتغيرين ، احدهما يسبب الاخر، وتكون المعلمات لكلا المتغيرين تختلف احصائيا عن الصفر في المعادلتين (1 و 2) أي ان $(\sum_{i=1}^{n} a_i \neq 0)$ و $(\sum_{J=1}^{n} d_J \neq 0)$.

هذه الحالة تبين ان ماضي المتغير (X) يحسن من القدرة التنبؤية للمتغير (Y) وان ماضي المتغير (Y) يحسن القدرة التنبؤية للمتغير (X) في اللحظة t

3- عدم وجود علاقة سببية بين المتغيرين (x,y)

في هذه الحالة تكون المعلمات لكلا المتغيرين لا تختلف احصائيا عن الصفر في المعادلتين (1و 2) أعلاه أي ان $(\sum_{i=1}^{n} a_i = 0)$ وكذلك $(\sum_{J=1}^{n} d_J = 0)$ وهذا يدل على عدم وجود علاقة سببية بين المتغيرين في اللحظة t .

ويتم التوصل الى نتائج اختبار السببية لكرانجر من خلال اختبار احصاءة (F) اذ يتم المقارنة بين قيمة (F) المحتسبة وقيمة (F) الجدولية فاذا كانت قيمة (F) المحتسبة أكبر من قيمة (F) الجدولية عند مستوى معنوية معين (5% او 1%) ترفض فرضية العدم (H_o) وتقبل الفرضية البديلة (H_1) التي تدل على وجود علاقة سببية بين المتغيرين (X,Y)

ان دراسة السببية تتطلب تحديد مدة الابطاء المثلى، كون اختبار فترة الابطاء غير صحيحة تؤدي الى تحيز في النتائج في حالة اختبار مدة ابطاء اقل من مدة الابطاء المثلى، في حين تكون المعلمات المقدرة غير كفوه في حالة كون مدة الابطاء المختارة أكبر من المدة المثلى.

ثانيا: سكون السلاسل الزمنية: Stationary of time Series

تعاني اغلب السلاسل الزمنية للمتغيرات الاقتصادية من عدم الاستقرار أي تحتوي على جذر الوحدة الذي يتجلى بان متوسط وتباين المتغير غير مستقلين عن الزمن، مما يؤدي الى وجود ارتباط زائف ومشكلات في التحليل والاستدلال القياسي بسبب صعوبة نمذجة تلك السلاسل الزمنية، لذا يتم اللجوء الى اختبار سكون السلاسل الزمنية. ويقصد بسكون السلسلة الزمنية عدم وجود اتجاه عام للظاهرة صعودا او نزولا في المسار الزمني للسلسلة فضلا عن عدم وجود تقلبات موسمية أي لا تتغير خصائصها عبر الزمن.

ويعد سكون السلاسل الزمنية من عدمها اهم المشاكل التي تواجه الباحثين في الشأن القياسي، ذلك لان غياب صفة السكون عن السلاسل الزمنية للمتغيرات الاقتصادية المستخدمة في النماذج القياسية تؤدي الى مشكلة الانحدار الزائف (Spurious Regression) التي تؤثر على نتائج الدراسة وتكون النتائج مضللة وغير واقعية. وهنالك عدة مؤشرات تدل على ان تقدير النموذج القياسي يكون زائفا بسبب عدم سكون السلاسل الزمنية منها:

أ- قيمة معامل التحديد (R^2) اكبر من احصاءة ديربن واتسن (D-W) أي ان $(R^2 > D-W)$

ب- وجود ارتباط تسلسلي ذاتي يبرز في قيمة ديربن واتسن (D-W)

ت- زيادة المعنوية الإحصائية للمعلمات المقدرة بدرجة كبيرة.

خصائص السكون ومميزاته:

تطلق صفة السكون على السلاسل الزمنية إذا توافرت فيها الخصائص الإحصائية الاتية:

- ثبات الوسط الحسابي لقيم السلسلة الزمنية عبر الزمن ويعبر عن ذلك بالمعادلة الاتية:
$$E(Y_t) = \mu$$
- ثبات التباين لقيم السلسلة عبر الزمن ويمكن التعبير عن ذلك بالمعادلة التالية:
$$Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$$
- امتلاك السلسلتين الزمنيتين تباين مشترك معتمدا على الازاحة (K) فقط (ثبات التغاير) أي ان:

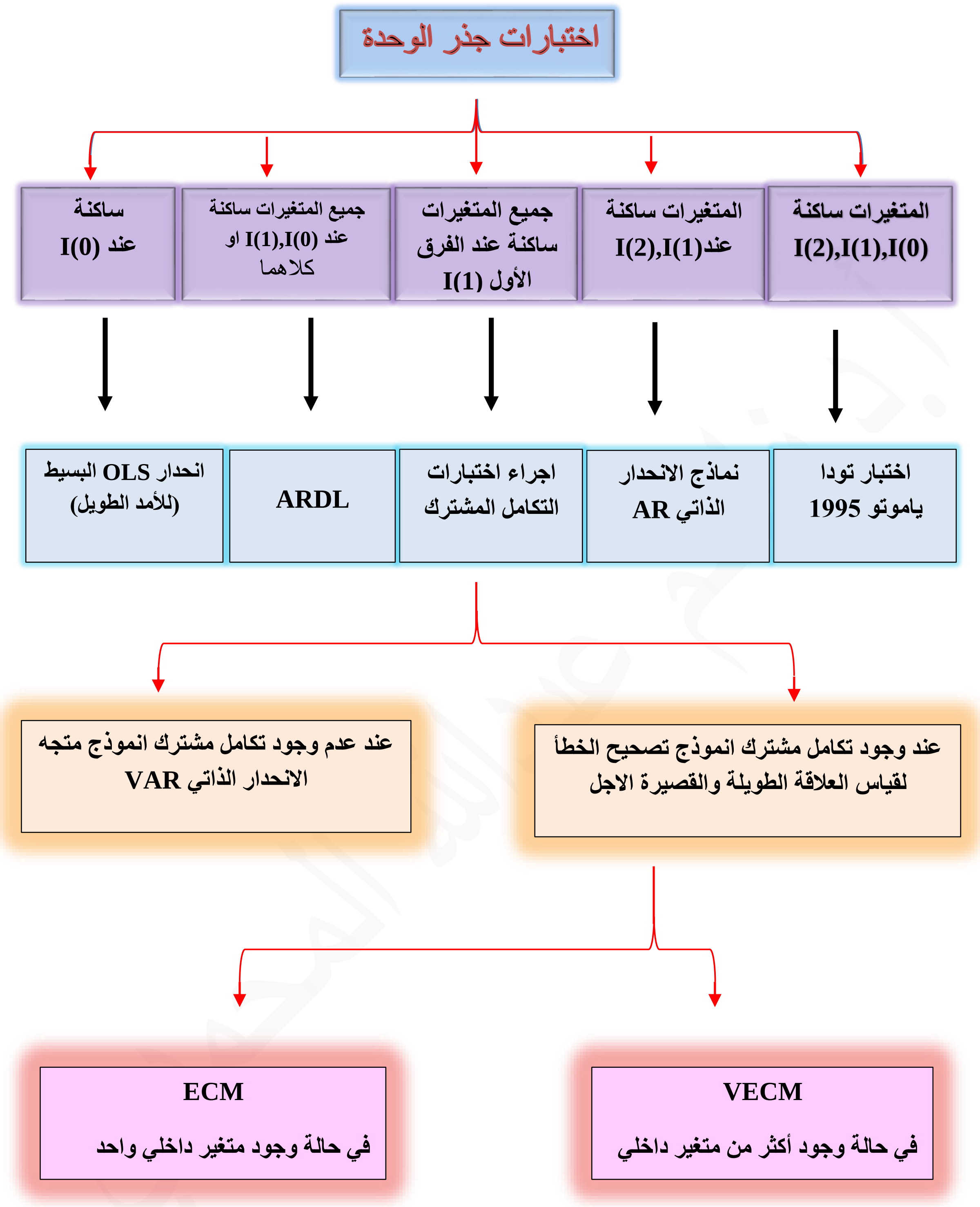
$$Cov(Y_t , Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k$$

حيث ان:

μ: الوسط الحسابي ، σ^2: التباين ، γ_k: معامل التباين المشترك(التغاير).

ومن الجدير بالذكر فان سكون السلسلة الزمنية عند المستوى الأصلي I(0) او عند الفرق الأول I(1) او عند كليهما او الفرق الثاني I(2) هو الذي يحدد نوع الانموذج او الاختبار الاحصائي (اختبار التكامل المشترك) الذي يستخدم في القياس والتقدير،

وكما موضح في الشكل الاتي:



ألشكل(1): مخطط هيكلية ألإختبارات ألقياسية.

2- اختبار سكون السلاسل الزمنية:

هناك عدة طرق يمكن بواسطتها اختبار سكون السلاسل الزمنية منها:

أ- الرسم البياني:

يعطي الشكل البياني الزمني تصور اولي عن مدى سكون السلاسل الزمنية لأي متغير، فضلا عن قدرته على بيان وجود اتجاه عام تصاعدي او تنازلي في السلسلة الزمنية ناجم عن اختلاف متوسطات العينات الزمنية للسلسلة ككل وهذا يعني عدم سكون السلسلة الزمنية. وفي بعض الأحيان يصعب تحديد طبيعة السلسلة الزمنية ما إذا كانت ساكنه او غير ساكنه بالملاحظة او حتى الرسم البياني.

ب- دالة الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي(DACF):

تفسر دالة الارتباط الذاتي والجزئي لسلسلة زمنية معينة الارتباط الموجود بين المشاهدات لمختلف الفترات الزمنية، وهي ذات أهمية كبيرة في ابراز بعض الخصائص العامة للسلسلة الزمنية، وتعد دالة الارتباط الذاتي والجزئي من الاختبارات المهمة والبسيطة المستخدمة في التحليل والتحقق من مدى استقرارية السلسلة الزمنية لمتغير ما.

ت - اختبارات جذر الوحدة:

يهدف اختبار جذر الوحدة الى فحص خواص السلسلة الزمنية لكل متغير من متغيرات الدراسة خلال مدة زمنية معينة، والتأكد من مدى استقراريتها وتحديد رتبة تكامل كل متغير على حده، فاذا كانت السلسلة الزمنية مستقرة في قيمتها الاصلية (المستوى) عندئذ يقال انها متكاملة من الرتبة صفر ويعبر عنها بـ I(0) ، اما اذا استقرت السلسلة بعد اخذ الفرق الأول فان السلسلة الاصلية تكون متكاملة من الرتبة الأولى أي I(1) ، اما اذا استقرت السلسلة بعد اخذ الفرق الثاني فان السلسلة الاصلية تكون متكاملة من الرتبة الثانية أي I(2) .

ويوجد نوعين من السلاسل الزمنية غير الساكنة هما:

1- سلاسل زمنية غير ساكنة من نوع Trend stationary (TS)

هذه السلاسل تبرز عدم سكونيه تحديدية Deterministic ويستخدم غالبا طريقة المربعات الصغرى (OLS) من اجل اعادتها مستقرة، وتأخذ الصيغة الاتية:

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

2- سلاسل زمنية غير ساكنة من النوع Difference stationary (DS)

هذه السلاسل تبرز عدم سكونيه عشوائية (Random) في مركبة الاتجاه العام، وغالبا ما يستخدم معادلة الفروق الأولى من اجل اعادتها مستقرة، وتأخذ الصيغة الاتية:

$$Y_t = Y_{t-1} + \beta + \varepsilon_t$$

ان نقطة البدء للكشف عن وجود جذر الوحدة في بيانات السلسلة الزمنية تنطلق من خلال المعادلة الاتية:

$$Y_t = \mathcal{P}Y_{t-1} + u_t \qquad\qquad -1 \leq \mathcal{P} \leq 1$$

وتعني هذه المعادلة اجراء الانحدار لمتغير السلسلة الزمنية (Y_t) بوصفه المتغير التابع مع المتغير نفسه ولكن بتباطؤ مدة واحدة (المتغير المتخلف زمنيا) فاذا كانت $(\mathcal{P} = 1)$ فهذا يعني ان السلسلة الزمنية (Y_t) تحتوي على جذر الوحدة Unit Root ومن ثم افتقار السلسلة الزمنية لصفة السكون (عدم استقرار السلسلة الزمنية).

ولو كانت $(\mathcal{P} > 1)$ فان السلسلة تصبح سلسلة متفجرة أي متجه نحو الأعلى Explodes . ولو كانت $(\mathcal{P} = 0)$ فان السلسلة تعتبر عشوائية بحتة White noise .

وكل السلاسل السابقة تعتبر سلاسل غسر ساكنه.

اما اذا كانت المعلمة $(\mathcal{P} < 1)$ فذلك يعني عدم وجود تغيرات دائمة في السلسلة وتكون ساكنة (مستقرة) وبالتالي يمكن استخدامها في التقدير ، وهذه هي فكرة اختبار جذر الوحدة بصورة عامة ومبسطة.

ومن الجدير بالذكر ان الباحثون في القياس الاقتصادي يهتمون بدراسة جذر الوحدة لأسباب التالية:

أ- لتجنب حدوث مشكلة الارتباط الذاتي، لأنه أحيانا مشكلة الارتباط الذاتي تنتج من كون واحد او أكثر من متغيرات النموذج هو غير مستقر أي يحتوي على جذر الوحدة (non-stationary).

ب- أحيانا قد نحصل على قيمة R^2 عالية جدا خلال تقدير معادلة تحتوي على متغيرين مثلا بالرغم من عدم وجود علاقة منطقية بين هذين المتغيرين مما يقود الى ظهور مشكلة (التقدير) الانحدار الزائف Spurious regression ، لذلك فان اختبار جذر الوحدة يعتبر أساسي لتجنب هذه الحالة.

ت- في حالة اجراء اختبار العلاقة السببية Granger causality test او VAR test فأننا نفترض بان المتغيرات هي مستقرة Stationary.

ان من اهم اختبارات جذر الوحدة المستخدمة لاختبار سكون السلاسل الزمنية وتحديد رتبة تكاملها هي:

1- اختبار ديكي-فوللر الموسع: Augmented Dickey-Fuller test:(ADF)

هو احد الاختبارات المهمة المستخدمة في تشخيص وجود جذر الوحدة في بيانات السلاسل الزمنية للمتغيرات تحت الدراسة ، حيث طور كل من (Dickey-Fuller) صيغتهم البسيطة لاختبار طبيعة السلاسل الزمنية وذلك من اجل تفادي السلبيات التي تحتويها الصيغة البسيطة والمتمثلة بعدم اهتمامها بمشكلة الارتباط الذاتي في حد الخطأ العشوائي، اذ ان البواقي في نموذج الانحدار الخطي البسيط غالبا ما تكون مرتبطة ذاتيا، ولتفادي ذلك يتم اجراء اختبار (ADF) عن طريق تضمين دالة الاختبار عددا معينا من فروقات المتغير التابع وتطبيق طريقة المربعات الصغرى (OLS) لتقدير واحد او اكثر من النماذج الاتية:

- بدون الحد الثابت والاتجاه الزمني ويعبر عنه بالمعادلة الاتية:

$$\Delta Y_t = \beta Y_{t-1} + \sum_{i=1}^{P} \lambda i \, \Delta Y_{t-i} + U_t$$

- وجود حد ثابت بدون الاتجاه الزمني ويعبر عنه بالمعادلة الاتية:

$$1\Delta Y_t = \alpha_o + \beta Y_{t-1} + \sum_{i=1}^{P} \lambda i \, \Delta Y_{t-i} + U_t$$

- وجود حد ثابت واتجاه زمني معبرا عنه بالمعادلة التالية:

$$\Delta Y_t = \alpha_o + \alpha_1 t + \beta Y_{t-1} + \sum_{i=1}^{P} \lambda i \, \Delta Y_{t-i} + U_t$$

اذ ان:

Δ: الفرق الأول

$\alpha_o, \alpha_1, \beta$: المعلمات المطلوب تقديرها

Y: المتغير المراد اختباره

P: عدد فترات الابطاء

U_t: حد الخطأ العشوائي

t: الزمن على شكل اتجاه عام Time trend.

ويستند اختبار ديكي-فوللر تبعا للصيغ أعلاه الى فرضيتين أساسيتين هما:

فرضـية العدم $(H_o: \beta = 0)$ التي عندها تكون بيانات السـلسـلة Y_t غير سـاكنة وتحتوي على جذر الوحدة . والفرضية البديلة $(H_1: B > 0)$ التي تنص على سكون السلسلة الزمنية.

ويتم اختبار فرضية العدم (H_o) من خلال مقارنه إحصائية (t) المقدرة للمعلمة (β) مع القيم الجدولية لـ(ديكي-فولير) والمطورة بواسطة (Mackinnon1991) عند حجم عينة (n) ومستوى معنوية (1% - 5% - 10%) فاذا كانت القيمة المطلقة لإحصائية (t) المقدرة تتجاوز (اكبر من) القيمة المطلقة لـ(DF) او قيمة (Mackinnon) فأنها تكون معنوية احصائيا وعليه ترفض فرضية العدم (H_o) بوجود جذر الوحدة وتقبل الفرضية البديلة (H_1) أي ان السلسلة الزمنية للمتغير المدروس مستقرة (ساكنه) . اما إذا كانت القيمة المطلقة المحتسبة لإحصائية (t) أصغر من القيمة الجدولية فانه لا يمكن رفض فرضية العدم (فرض وجود جذر الوحدة) مما يعني ان السلسلة الزمنية غير ساكنة (غير مستقرة) وبالتالي نقوم باختبار سكون الفرق الأول (First difference) للسلسلة الزمنية وإذا كان غير ذلك نكرر الاختبار للفرق الثاني وهكذا. ....

ومن الجدير بالذكر فانه وفي ظل التطور الحاصل في المنهجية الحديثة للتحليل القياسي من خلال استعمال برنامج افيوز (Eviews) يمكن التعرف على المعنوية الإحصائية للمعلمة (β) لتحديد القبول او الرفض للفرضيتين (العدم والبديلة) من خلال قيمة الاحتمال (Probability) فاذا كانت اقل من 5% فهذا يدل على ان القيمة المحسوبة لاحصاءة (t) اكبر من القيمة الجدولية ، لذلك ترفض فرضية العدم وتقبل الفرضية البديلة (سكون السلاسل الزمنية لمتغيرات النموذج) .

واذا كانت السلسلة الزمنية ساكنة عند المستوى عندئذ يقال عنها متكاملة من الدرجة صفر أي I(0) وعندئذ يمكن تطبيق طريقة المربعات الصغرى الاعتيادية (OLS) مباشرة وبدون الخشية من الحصول على الانحدار الزائف، اما اذا كانت السلسلة الزمنية مستقرة عند الفرق الأول فان السلسلة الزمنية تكون متكاملة من الدرجة الأولى I(1) واذا تحققت صفة السكون عند الفرق الثاني تكون السلسلة الزمنية متكاملة من الدرجة الثانية I(2) .

2- اختبار فيلبس بيرون (PP): Phillips-Perron test

قدم هذا الاختبار من قبل العالم فيلبس بيرون عام 1988 ويستند هذا الاختبار على تقدير المعادلات في اختبار ديكي-فولير (ADF) نفسها، الا انه يتميز عن اختبار (ADF) في طريقة معاجلة وجود مشكلة الارتباط التسلسلي، حيث يقوم بعملية تصحيح غير معلميه (nonparametric) لاحصاءة (t) للمعلمة (β) بينما اختبار (ADF) يواجه مشكلة الارتباط التسلسلي بعملية تصحيح معلميه من خلال إضافة حدود الفروق المبطأة للمتغير على يمين

المعادلة. ومن المعلوم ان اختبار (ADF) قائم على فرضية ان السلسلة الزمنية متولدة بواسطة عملية الانحدار الذاتي (Autoregressive Process-AR-) بينما اختبار (PP) قائم على افتراض اكثر عمومية وهي ان السلسلة الزمنية متولدة بواسطة عملية (ARIMA) (Autoregressive Integrated Moving Average) ويستند اختبار فيلبس بيرون (PP) الى نفس الفروض الخاصة بفرضية (العدم والبديلة) لاختبار (ADF) وفي حالة تضارب النتائج بين الاختبارين (ADF) و(PP) يفضل الاعتماد على نتائج اختبار (PP) لأنها ذات قدرات اختبارية افضل (عالية) ودقيقة لاسيما عندما يكون حجم العينة صغير.

3- اختبار كبس Kpss test (Kwiatkowski Phillip's Schmidt and shin):

اقترح هذا الاختبار عام 1992 من قبل Kwiatkowski Phillip's Schmidt and shin. ويعتمد على مضاعف لاجرانج (LM) ويرتكز على فرضية العدم التي تنص على سكون (استقرارية) السلسلة الزمنية خلافا للاختبارين (ADF) و (PP) السابقين، فهو يعالج بعض أوجه الضعف في فاعلية الاختبارين (ADF) و (PP) في حالة وجود ارتباط ذاتي بين الأخطاء العشوائية، ويمكن القول ان نتائج هذه الاختبارات تكمل بعضها البعض، فضلا عن انه في حالة اتفاقها على نتيجة واحدة تصبح النتيجة اكثر دقة.

ويستند اختبار (Kpss) على الخطوات الاتية:

أ- احتساب المجموع الجزئي للبواقي بعد تقدير النماذج في الاختبارين (ADF) و (PP) وفق الصيغة الاتية:

$$S_t^2 = \sum_{t=1}^{n} e_t$$

ب- تقدير التباين طويل الاجل (s_1^2) بنفس طريقة اختبار فيلبس بيرون (PP) وفق الصيغة الاتية:

$$S_1^2 = \frac{1}{n}\sum_{t=1}^{n} e_t^2 + 2\sum_{i=1}^{n}\left(1-\frac{i}{t+1}\right)\frac{1}{n}\sum_{t=i+1}^{n} e_t + e_{t-i}$$

ت- تحسب إحصائية اختبار (Kpss) من العلاقة الاتية:

$$LM = \frac{1}{S_1^2} * \frac{\sum_{t=1}^{n} S_t^2}{n^2}$$

وتعتمد القيم الحرجة لهذا الاختبار على قيم (*LM statistic*) مقارنة مع قيم (*Kpss*) المحتسبة، حيث ترفض فرضية العدم (السكون) إذا كانت احصاءة (*LM*) المحتسبة أكبر من القيمة الحرجة المستخرجة من الجدول المعد من قبل (*Kpss*) عام 1992 وتقبل الفرضية العدم (السكون) إذا كانت احصاءة (*LM*) المحتسبة أصغر من القيمة الحرجة لها. حيث ان فرضية

العدم (H_o) تنص على ان السلسلة الزمنية للمتغير ساكنة على العكس من اختبار (ADF) و (PP) الذي تنص فيها فرضية العدم على عدم سكون السلسلة الزمنية للمتغير (عدم استقرارية السلسلة الزمنية).

ومن الجدير بالذكر فان من بين الأساليب المستخدمة في تثبيت سكون السلسلة الزمنية هي:

1. **في حالة عدم ثبات التباين:** من اهم التحويلات المستخدمة في تثبيت تباين السلسلة الزمنية هو الحصول على اللوغاريتم الطبيعي لبيانات السلسلة او الحصول على الجذر التربيعي لها او استخدام مقلوب بيانات السلسلة الزمنية.
2. **في حالة الاتجاه العام:** من الطرق المستخدمة للتخلص من الاتجاه العام هي:

- طريقة الانحدار الخطي في تقدير الاتجاه العام ثم عزله والتعامل مع البواقي كسلسلة زمنية ساكنة وتسمى هذه العملية $(Detrending)$.
- طريقة الفروق من خلال طرح قيم المشاهدات من بعضها البعض لفترات ابطاء معينة، كالفروق من الدرجة الأولى او الثانية. وقد يلجأ الى تطبيق (d) من الفروق للتخلص من الاتجاه العام للحصول على سلسلة زمنية ساكنة.

3. إزالة التقلبات الموسمية باستخدام طريقة الفرق الموسمي وذلك بطرح القيم من بعضها البعض حسب فترات الابطاء المتسقة مع نوع البيانات للحصول على الفروق ربع السنوية او فروق شهرية.

ثالثا: اختبار التكامل المشترك: *Cointegration test*

يعد اختبار التكامل المشترك من الاختبارات الأساسية التي يجب القيام بها قبل البدء بعملية تقدير النماذج القياسية بهدف تجنب حالات التقدير او الانحدار الزائف *Spurious Regression* ويرجع أصل الفكرة الى *Granger* عام 1981 وتم تطوير هذه الفكرة وتوضيحها واجراء الاختبارات والامثلة التجريبية من قبل انجل وجرانجر عام 1987. ويمكن تعريف التكامل المشترك بانه يمثل حالة تصاحب بين سلسلتين زمنيتين مثل$(X_t\ , Y_t)$ او اكثر اذ تؤدي التقلبات في احدهما الى الغاء التقلبات في الأخرى بطريقة تجعل النسبة بين قيمتها ثابتة عبر الزمن، أي ان بيانات السلاسل الزمنية قد تكون غير ساكنة اذا ما اخذت كل على حده ولكنها تكون ساكنة كمجموعة، أي ان السلاسل الزمنية سوف يكون بينها اتجاه عشوائي مشترك في الجل الطويل. ولكي يكون التفسير الاقتصادي للفرضية التي تنص على وجود العلاقة السببية بين متغيرين مقبولا، لابد من ان تكون البيانات لهذه المتغيرات متكاملة من الرتبة الأولى $I(1)$ وهذا يعني ان العلاقة طويلة الاجل بين المتغيرين $(X_t\ , Y_t)$ تكون معنوية في الحالة التي يكون فيها حد الخطأ المقدر متكاملة من الرتبة صفر $I(0)$.

في لغة الاقتصاد القياسي ان أي متغيران يكونان متكاملين تكاملا مشتركا فقط إذا كان بينهما علاقة طويلة الأمد او توازن في الأمد الطويل.

ويمكن الربط بين مفهوم التكامل المشترك ومفهوم النظرية الاقتصادية وخاصة فيما يتعلق بفكرة العلاقة التوازنية في الاجل الطويل اذ ينص انموذج التكامل المشترك على ان المتغيرات الاقتصادية التي تفترض النظرية الاقتصادية وجود علاقة توازنيه بينهما في الاجل الطويل لا تتباعد عن بعضها البعض بشكل كبير في الاجل الطويل، مع إمكانية ان تتباعد هذه المتغيرات عن التوازن في الاجل القصير ويصحح هذا التباعد عن التوازن بفعل قوى اقتصادية تعمل على إعادة هذه المتغيرات الاقتصادية للتحرك نحو التوازن طويل الاجل. وهكذا فان فكرة التكامل المشترك تحاكي وجود توازن في الاجل طويل يؤول اليه النظام الاقتصادي.

وهناك شرط لابد من توفره في منهجية التكامل المشترك وهو ان السلاسل الزمنية لمتغيرات الانموذج (Y_t , X_t) يجب ان تكون متكاملة من نفس الدرجة، فضلا عن ان تقدير البواقي الناتجة عن تقدير العلاقة بينهما تكون متكاملة من الدرجة *صفر* I*(0)*.

وان من اهم الاختبارات المستخدمة لتحليل التكامل المشترك للسلاسل الزمنية هما اختبار انجل وجرانجر *Engle and Granger* عام 1987 واختبار جوهانسن جيسليوس *Johansen and Juselius* عام 1990 في تحديد التكامل المشترك، الا ان هناك شرط مسبق لتطبيق هذين الاختبارين هو ان تكون السلاسل الزمنية لمتغيرات الانموذج متكاملة من الدرجة *I(1)* نفسها في تحليل العلاقات في الاجل الطويل.

فضلا عن هذين الاختبارين صدر حديثا اختبار ثالث على يد العام بيسيران *Pesaran* عام 2001 لتحديد التكامل المشترك يعرف باختبار الحدود *The bounds testing procedure* ويتميز عن الاختبارين السابقين بانه لا يشترط ان تكون المتغيرات محل الدراسة متكاملة من الدرجة نفسها.

1- **اختبار انجل وجرانجر:** ***Engle and Granger Test***

قدم كل من انجل وجرانجر عام 1987 دراسة مشتركة ومنهجية جديدة لاختبار التكامل المشترك بين المتغيرات الاقتصادية، ويقتصر هذا الاختبار على متغيرين فحسب (فقط) ويتم اجراء هذا الاختبار وفق الخطوات الاتية:

أ- اجراء اختبارات جذر الوحدة (*Kpss , PP , ADF*) لمعرفة استقرارية بيانات السلاسل الزمنية لمتغيرات الدراسة ، فاذا كانت مستقرة عند مستوى *I(0)* يمكن استخدام الطرق الإحصائية التقليدية للتقدير للحصول على دلالات ومؤشرات إحصائية عالية الدقة . اما اذا كانت السلاسل الزمنية غير مستقرة عند المستوى وتصبح مستقرة بعد اخذ الفرق الأول لها او الثاني وتكون متكاملة من الرتبة نفسها *I(1)* ، فيتم تقدير انحدار العلاقة طويلة الاجل بين المتغيرين بطريقة (*OLS*) الصيغة الاتية:

$$Y_t = a_o + a_1 X_t + U_t$$

ب- تحتسب البواقي (U_t) والتي تقيس انحراف العلاقة المقدرة في الاجل القصير عن اتجاهها التوازني في الاجل الطويل وفق المعادلة الاتية:

$$u_t = Y_t - a_o - a_1 X_t$$

ت- بعد تقدير بواقي الانحدار المقدرة $(\hat{u}_t)$ الناجمة عن انحدار العلاقة التوازنية طويلة الاجل ، يتم اجراء احد اختبارات السكون (الاستقرارية) على بواقي الانحدار وفق الصيغة الاتية:

$$\Delta \hat{u}_t = b_o + b_1 \hat{u}_{t-1} + \Delta \hat{u}_{t-1} + e_t$$

ويتم تحديد نتائج الاختبار من خلال مقارنة احصاءة (t) المحتسبة للمعلمة $(\hat{u}_{t-1})$ بالقيمة الحرجة لها من جداول اعدها انجل وجرانجر ، فاذا تم قبول فرضية العدم $(H_o: \beta_1 = 0)$ فهذا يعني وجود جذر الوحدة في سلسلة البواقي أي انها غير ساكنة (وذلك عندما تكون قيمة t المحتسبة اقل من الجدولية) ، ومنه نستنتج عدم وجود تكامل مشترك بين السلاسل الزمنية للمتغيرات المكونة للأنموذج. اما اذا كانت قيمة احصاءة (t) المحتسبة اكبر من الجدولية لذلك ترفض فرضية العدم ونقبل الفرض البديل $(H_1: \beta_1 \neq 0)$ فهذا يعني ان سلسلة البواقي ساكنة ومتكاملة من الرتبة I(0) ، أي ان السلاسل الزمنية لمتغيرات النموذج تتصف بخاصية التكامل المشترك أي وجود علاقة توازنيه طويلة الاجل بينهما ، مما يجعل انموذج تصحيح الخطأ (Error correction model –ECM-) اكثر ملائمة لتقدير العلاقة بينهما.

ومن الجدير بالذكر ان برنامج Eviews يحتوي على ايعاز مستقل ومباشر لغرض احتساب هذا الاختبار Engle-Granger test cointegration

ان اختبار انجل وجرانجر يعاني من عدة مشكلات أهمها:

أ- يفترض وجود علاقة تكامل واحدة بين متغيرات النموذج، وهذا غير دقيق في نظام مكون من عدد كبير من المعادلات.

ب- يحدد التكامل المشترك باتجاه واحد ولمتغيرين فقط، وهذا يمثل قيدا كبيرا خاصة مع وجود أكثر من متغيرين ومع العلاقات التبادلية.

2- اختبار جوهانسن وجيسيليوس للتكامل المشترك: Johansen and Juselius test

اقترح كل من Johansen and Juselius عام 1990 اختبارا جديدا يعتمد على مقدرات الإمكان الأعظم لتلافي عيوب اختبار انجل وجرانجر. ويمتاز هذا الاختبار بكونه يتناسب مع العينات صغيرة الحجم، وإمكانية استخدامه لأكثر من متغيرين في النموذج، وكذلك يعد الأفضل حتى في حالة استخدام متغيرين فقط. وتعتمد منهجية الاختبار على رتبة المصفوفات والتي نستطيع من خلالها تحديد إمكانية تحقيق التكامل المشترك من عدمه، إذا ان رتبة المصفوفة تمثل عدد متجهات التكامل المشترك والتي يمكن اختبارها كالاتي:

أ- اذا كانت رتبة المصفوفة π مساوية للصفر $\pi = 0$ فان المتغيرات محل الدراسة لها جذر الوحدة، مما يدل على عدم وجود تكامل مشترك ويستوجب استخدام الفرق الأول.

ب- اذا كانت رتبة المصفوفة π مساوية للواحد الصحيح $\pi = 1$ فانه يوجد متجه وحيد للتكامل المشترك.

ت- اذا كانت المصفوفة π تامة الرتبة $\pi = n$ وان (n) تساوي عدد متغيرات النموذج المقدر، فان جميع المتغيرات ليس لها جذر الوحدة أي انها متغيرات ساكنة(مستقرة).

ث- اما الحالات الأخرى لرتبة المصفوفة π فهي الحالات التي تكون فيها رتبة المصفوفة $(1 < Rank\ \pi < n)$ فهذا يدل على وجود عدة متجهات متكاملة تكاملا مشتركا.

اذا ان π: مصفوفة التباين والتباين المشترك للاخطاء العشوائية التي من خلالها يمكن حساب القيم الذاتية Eigen Values .

ولتحديد عدد متجهات التكامل المشترك يوجد اختبارين احصائيين هما:

1- اختبار الأثر: Trace Test

ويتم احتسابه وفق الصيغة الاتية:

$$\lambda_{trace} = -T \sum_{i=r+1}^{n} \log(\hat{\lambda}_i)$$

حيث ان r: عدد متجهات التكامل المشترك (رتبة المصفوفة)

و$\hat{\lambda}_i$: القيم الذاتية لمصفوفة التباين والتباين المشترك

وn: عدد متغيرات النموذج

حيث يتم اختبار فرضية العدم $(r=0)$ مقابل الفرضية البديلة $(r \geq 1)$ فاذا كانت قيمة اختبار الأثر المحتسبة اصغر من القيمة الحرجة تقبل فرضية العدم والتي تعني ان متجهات التكامل

المشترك تساوي صفرا (لا يوجد تكامل مشترك) اما اذا كانت قيمة اختبار الأثر المحتسبة اكبر من القيمة الحرجة تقبل الفرضية البديلة والتي تعني ان عدد المتجهات اكبر من واحد مما يعني وجود التكامل المشترك بين متغيرات النموذج تحت الدراسة (أي انه يوجد علاقة طويلة الأمد بين المتغيرات)

2- اختبار القيم المميزة العظمى: *Maximum Eigen Values Test*

يتم احتسابه وفق الصيغة الاتية: $\lambda_{max} = -T\log(1-\hat{\lambda}_i)$

ويتم اختبار فرضية العدم مقابل الفرضية البديلة بمقارنة قيمة *Maximum Eigen Values* المحتسبة مع القيمة الحرجة (الجدولية) الواردة في جدول جوهانسن جيسيليوس وبدلالة قيمة *p-value* عند مستوى معنوية معين، فاذا كانت القيم المحتسبة اكبر من القيمة الجدولية ترفض فرضية العدم وتقبل الفرضية البديلة مما يعني وجود تكامل مشترك بين المتغيرات ويمكن معرفة عدد متجهات هذا التكامل، والعكس صحيح في حالة قبول فرضية العدم ورفض الفرضية البديلة.

رابعا: إنموذج متجه الانحدار الذاتي: *Vector Autoregressive Model(VAR)*

الانحدار الذاتي *Autoregressive* في الاقتصاد القياسي تعني ان النموذج المراد تقديره يحتوي على قيم متخلفة (مرتدة) زمنيا للمتغير التابع كأحد المتغيرات المستقلة في النموذج كما في النموذج الاتي: $C_t = \beta_o + \beta_1 Y_t + \beta_2 C_{t-1} + U_t$

حيث ان:

c_t: الاستهلاك في السنة t

Y_t: الدخل في السنة t

خامسا: انموذج متجه تصحيح الخطأ: *Vector Error Correction Model (VECM)*

يطلق على هذا الانموذج تسمية حد تصحيح الخطأ لأنه انموذج انحدار ذاتي مضافا اليه مقدار الخطأ في التوازن الى معلمات النموذج. ويساعد هذا الانموذج في فهم الالية التي يتكيف فيها المتغير التابع ليستعيد العلاقة طويلة الأمد نتيجة لاستجابة حد تصحيح الخطأ، ويطبق هذا النموذج على السلاسل الزمنية المتكاملة من الرتبة نفسها (أي يوجد بين السلاسل الزمنية علاقة طويلة الامد) وغير المستقرة (غير ساكنة).

ويمكن من خلال هذا النموذج تحديد وتفسير العلاقات التوازنية طويلة الاجل التي تتجسد بخاصية التكامل المشترك والعلاقة قصيرة الاجل من خلال مدة الابطاء، فضلا عن تجنب الأخطاء عند توصيف النموذج.

سادسا: اختبار التكامل المشترك باستخدام انموذج الانحدار الذاتي للفجوات الزمنية الموزعة المتباطئة: Autoregressive Distributed Log Model (ARDL)

تتطلب اختبارات التكامل المشترك السابق ذكرها (انجل وجرانجر، جوهانسن وجيسيليوس) ان تكون السلاسل الزمنية للمتغيرات محل الدراسة متكاملة من الرتبة نفسها، كما ان هذه الاختبارات تعطي نتائج غير دقيقة في حالة كون حجم العينة صغيرا، وهذا يضع شرطا على استخدام هاتين الطريقتين في تحليل العلاقات طويلة الاجل بين المتغيرات. ونتيجة لهاتين المشكلتين أصبحت منهجية (*ARDL*) للتكامل المشترك شائعة الاستخدام في السنوات الأخيرة، والذي قدم من قبل بيسران وآخرون (*Pesaran et.al*) عام 2001، حيث تم في هذه المنهجية دمج نماذج الانحدار الذاتي (*AR*) مع نماذج فترات الابطاء الموزعة (*Distributed Log Model*) في انموذج واحد، وفي هذه المنهجية تكون السلسلة الزمنية دالة في ابطاء قيمتها وقيم المتغيرات التفسيرية الحالية وابطائها بمدة واحدة او أكثر.

وتعد هذه المنهجية من اهم الاختبارات المستعملة في تحديد التكامل المشترك والعلاقات الطويلة الاجل الخاصة بالنماذج الاقتصادية، وهناك شروط مسبقة لتطبيق هذا الاختبار أهمها:

- ان تكون البيانات خالية من مشكلة الارتباط الذاتي.
- ان تكون البيانات خالية من مشكلة عدم تجانس التباين.
- ان تكون البيانات موزعة توزيعا طبيعيا.
- لا يوجد أي متغير ساكن بالفرق الثاني $I(2)$

ويتميز انموذج (*ARDL*) عن بقية اختبارات التكامل المشترك بعدة مزايا أهمها:

1- يعطي انموذج (*ARDL*) مرونة عالية في تحديد التكامل المشترك عندما يكون هناك خليط من السلاسل الزمنية للمتغيرات متكاملة من الرتبة نفسها $I(0)$ ومن الرتبة الأولى $I(1)$ ، بشرط ان لا تكون متكاملة من الرتبة الثانية $I(2)$ فقط.
2- إمكانية تقدير معلمات الاجل الطويل والقصير في آن واحد، فضلا عن تقدير معامل تصحيح الخطأ
3- ان نتائج تطبيقه دقيقه في حالة كون حجم العينة صغيرا، فضلا عن بساطة الانموذج في تقدير التكامل المشترك باستخدام طريقة المربعات الصغرى الاعتيادية، كونه يحتوي على معادلة واحدة مما يجعل الانموذج سهل في التنفيذ والتفسير

4-يمكن تنفيذه بتخلفات زمنية مختلفة لكل من متغيراته الداخلية والخارجية قسم منها بتخلف واحد او متخلفين او بدون تخلف، مع إمكانية ان يتولى الانموذج تحديد عدد هذه المتخلفات ذاتيا للحصول على أفضل مجموعة من البيانات من انموذج الإطار العام.

5-إمكانية التمييز بين المتغيرات التابعة والمتغيرات التفسيرية في الانموذج والسماح باختبار العلاقة بين المتغيرات الاصلية (في المستوى -*Level*) بغض النظر فيما اذا كانت المتغيرات التفسيرية هي $I(0)$ و $I(1)$ او مزيج منها .

6-ان استخدام هذا الانموذج يساعد على التخلص من المشكلات المتعلقة بحذف المتغيرات ومشكلة الارتباط الذاتي، مما يجعل المقدرات الناتجة كفوءة وغير متحيزة.

7-ان انموذج (*ARDL*) يعطي أفضل النتائج للمعلمات في الاجل الطويل، وان اختبارات التشخيص يمكن الاعتماد عليها بشكل كبير.

8-من خلال انموذج (*ARDL*) يمكن تحديد العلاقة التكاملية بين المتغير التابع والمتغيرات المستقلة، فضلا عن تحديد حجم تأثير كل من المتغيرات المستقلة على المتغير التابع، وتعد معلماته المقدرة للأجل القصير والطويل أكثر اتساقا من تلك المقدرة بالطرق الأخرى لاختبار التكامل المشترك.

والصيغة العامة للأنموذج مكون متغير تابع (Y) و K من المتغيرات التفسيرية $(X_1,X_2,\ldots\ldots X_K)$ يكتب الانموذج $(P,q_1,q_2,\ldots\ldots q_k)$ (*ARDL*) بالشكل الاتي:

$$\Delta Y_t = c + \beta_1 Y_{t-1} + \beta_2 X_{1t-1} + \beta_3 X_{2t-1} + \cdots + \beta_{k+1} X_{kt-1} + \sum_{i=1}^{p-1} \lambda_{1i} \Delta Y_{t-1} + \sum_{i=0}^{q_1-1} \lambda_{2i} \Delta X_{1t-i} + \sum_{i=0}^{q_2-1} \lambda_{3i} \Delta X_{2t-i} + \cdots + \sum_{i=0}^{q_k-1} \lambda_{(k+1)i} \Delta X_{kt-1} + \mathrm{u}_t$$

اذ ان:

Δ: الفروق الأولى

C: الحد الثابت

u_t: حد الخطأ العشوائي

β: معلمات العلاقة طويلة الاجل

λ: معلمات العلاقة قصيرة الاجل

$(p, q_1, q_2, \ldots qk)$: فترات الابطاء للمتغيرات $(Y, X_1, X_2, \ldots X_k)$ على الترتيب.

ولتطبيق منهج تحليل التكامل المشترك في إطار انموذج (ARDL) يستلزم اجراء واتباع الخطوات الاتية:

1- اختبار مدد الابطاء المثلى للفروق:

يتم تحديد فترات التخلف (الابطاء) المثلى لقيم المتغيرات من خلال تحديدها اوتوماتيكيا من قبل البرنامج اذ يختار البرنامج الفترات المناسبة للتخلف لكل من المتغير التابع والمتغيرات المستقلة وقد يكون ذلك بفترات متباينة وهذه هي الخاصية المميزة لبرنامج Eviews. وهناك خمسة معايير مختلفة يوفرها البرنامج لتحديد فترات الابطاء وهي:

- معيار خطأ التنبؤ النهائي: Final Prediction Error(FPE)
- معيار معلومات اكيكي: Akaike Information Criterion (AIC)
- معيار معلومات شوارز: Schwarz Criterion (SC)
- معيار معلومات حنان وكوين: Hannan and Quinn Criterion (H-Q)
- معيار نسبة الإمكان الأعظم: Likelihood Ratio Test (LR)

وتتفق جميع تلك المعايير على ان فترة الابطاء المثلى (P) هي تلك الفترة التي تعطي أدنى(اقل) قيمة لمعظم هذه المعايير عند الاختبار.

2- منهج اختبار الحدود: Bounds Testing Approach

لاختبار مدى وجود علاقة توازنيه طويلة الاجل بين المتغير التابع والمتغيرات المستقلة الداخلة في النموذج يستخدم اختبار (WALD) او احصاءة اختبار (F) لاختبار الفرضية الاتية:

H_o: عدم وجود علاقة توازنية طويلة الاجل بين المتغيرات (انعدام التكامل المشترك بين المتغيرات) أي ان

$$H_o: \beta_1 = \beta_2 = \cdots = \beta_{k+1} = 0$$

H_1: وجود علاقة توازنية طويلة الاجل بين المتغيرات (وجود تكامل مشترك بين المتغيرات)

$$H_1: \beta_1 \neq \beta_2 \neq \cdots \neq \beta_{k+1} \neq 0$$

وتحسب قيمة احصاءة (F) وفق الصيغة الاتية:

$$F = \frac{(sse_R - sse_U)/\mu}{sse_U/(n-k)}$$

اذ ان:

sse_R: مجموع مربعات البواقي للنموذج المقيد (تحت فرضية العدم H_0)

sse_U: مجموع مربعات البواقي للنموذج غير المقيد(تحت الفرضية البديلة H_1)

μ: عدد معلمات النموذج المقيد ، n: حجم العينة (عدد المشاهدات)

K: عدد المتغيرات المستقلة

وبعد احتساب قيمة احصاءة (F) تتم مقارنتها بقيمة (F) الجدولية المحتسبة من قبل (Pesaran et.al عام 2001) ونظرا لان اختبار (F) له توزيع غير معياري فان هناك قيمتين حرجتين له:

- قيمة الحد الأدنى: وتفترض ان كل المتغيرات ساكنة في مستواها الأصلي، أي متكاملة من الرتبة صفر I(0)
- قيمة الحد الأعلى: وتفترض ان كل المتغيرات ساكنة في فرقها الأول، أي متكاملة من الرتبة الأولى I(1)

ويكون القرار على ثلاث حالات هي:

أ- اذا كانت قيمة احصاءة (F) المحتسبة اقل من الحد الأدنى لقيمة (F) الجدولية تقبل فرضية العدم (H_o) القائلة بعدم وجود تكامل مشترك بين المتغيرات أي (لا توجد علاقة توازنيه طويلة الاجل)

ب- اذا كانت قيمة احصاءة (F) المحتسبة اكبر من الحد الأعلى لقيمة (F) الجدولية تقبل الفرضية البديلة (H_1) القائلة بوجود تكامل مشترك بين المتغيرات أي (وجود علاقة توازنيه طويلة الاجل)

ت- إذا كانت قيمة احصاءة (F) المحتسبة تقع بين الحدين الأعلى والادنى لقيمة (F) الجدولية فان نتائج الاختبار تكون غير محسوسة.

3- تقدير معلمات انموذج(ARDL) ومعلمة تصحيح الخطأ(VECM)

بعد التأكد من وجود علاقة توازنيه طويلة الاجل بين المتغير التابع والمتغيرات المستقلة يتم تقدير معلمات النموذج (ARDL) للأجلين القصير والطويل ومعلمة متجه تصحيح الخطأ

(VECM) باستخدام طريقة المربعات الصغرى الاعتيادية (OLS) استنادا الى عدد فترات الابطاء المحددة.

وقبل اعتماد هذا الانموذج المقدر وتطبيقه ينبغي التأكد من جودة أداء هذا الانموذج، ويتم ذلك من خلال اجراء الاختبارات التشخيصية الاتية:

أ- اختبار مضروب لاجرانج للارتباط التسلسلي بين البواقي: (Brush-Godfrey) Lagrange Multiplier Test of Residual

ب- اختبار عدم ثبات التباين المشروط بالانحدار الذاتي: (ARCH) Autoregressive Conditional Heteroscedasticity

ت- اختبار التوزيع الطبيعي للأخطاء العشوائية: Jarque Bera (JB)

ث- اختبار مدى ملائمة تحديد الانموذج من حيث الشكل الدالي: Ramsey(AESET)

ج- اختبار الازدواج الخطي (التعدد الخطي بين المتغيرات المستقلة): Multicolinearty Test

4- اختبار استقرارية معلمات انموذج (ARDL)

للتأكد من خلو البيانات المستخدمة من وجود أي تغيرات هيكلية منها، ومعرفة مدى استقرار وانسجام معاملات الاجل الطويل مع تقديرات معلمات الاجل القصير يمكن استخدام أحد الاختبارين الآتيين:

أ- اختبار المجموع التراكمي للبواقي المعاودة: Cumulative sum of Recursive Residual (cusum)

ب- اختبار المجموع التراكمي لمربعات البواقي المعاودة:Cusum-SQ

Cumulative sum of squares Recursive Residual

ووفقا لهذان الاختباران يتحقق الاستقرار الهيكلي للمعلمات المقدرة بصيغة تصحيح الخطأ للأنموذج (ARDL) إذا وقع الخط البياني لاختبار كل من (cusum) و (cusum-sq) داخل الحدود الحرجة (الحد الأعلى والحد الادنى) عند مستوى معنوية 5% في حين تكون المعلمات لا تتسم بالاستقرارية الهيكلية إذا وقع الخط البياني للاختبارات خارج الحدود الحرجة.

5- اختبار الأداء التنبؤي لأنموذج تصحيح الخطأ غير المقيد المقدر:

تستخدم عدة معايير للتأكد من تمتع الانموذج بقدرة جيدة على التنبؤ خلال الفترة الزمنية للتقدير من أهمها:

أ- معامل ثايل (معامل عدم التساوي لثايل): Theil Inequality coefficient

يعد من المعايير الشائعة في قياس واختبار القدرة التنبؤية للأنموذج القياسي والتحقق من دقة التنبؤات. ويحسب وفق الصيغة الاتية:

$$T = \sqrt{\frac{\Sigma(Si - di)^2}{\Sigma di^2}}$$

اذ ان:

T: معامل ثايل.

Si: التغير المتوقع في القيمة المتنبأ بها للظاهرة (المتغير التابع).

di: التغير الفعلي في قيم المتغير التابع.

فاذا كانت (di=Si) فان المعامل (T=0) وهذا يدل على مقدرة عالية للنموذج على التنبؤ، اما اذا كانت (Si=0) فان المعامل (T=1) وهذا يعكس ضعف مقدرة النموذج على التنبؤ، أي لا يوجد تغير متوقع عبر الزمن ويكون ثابتا لهذا فان ($\hat{y}_F = a$) . اما اذا كانت قيمة المعامل (T) اكبر من الواحد الصحيح ، فهذا يعني انخفاض قدرة النموذج على التنبؤ وبذلك فان ($0 \leq T \leq \infty$).

ب- معيار نسبة عدم التساوي (مصادر الخطأ):

ويتكون من ثلاث نسب هي:

- نسبة التحيز: Bias proportion (BP).
- نسبة التباين: Variance proportion (VP).
- نسبة التغاير: Covariance proportion (CP).