

الجمهورية العربية السورية
جامعة دمشق- كلية الزراعة
قسم الاقتصاد الزراعي-
السنة الخامسة



الاقتصاد القياسي

(Econometric)

السنة الخامسة – الفصل الثاني

د.خالد السلطان + د. رواد شتي

2022/2021

الفصل الأول

مدخل إلى الاقتصاد القياسي
Introduction To Econometric

مدخل إلى الاقتصاد القياسي

مفهوم الاقتصاد القياسي

إن مصطلح الاقتصاد القياسي (Econometrics) مكون من مقطعين: المقطع الأول (Econo) وهو مشتق من كلمة الاقتصاد (Economy)، والمقطع الثاني (Metrics) وتعني القياس، ولذلك فإن كلمة الاقتصاد القياسي (Econometrics) تعني حرفيا القياس في الاقتصاد، مثل قياس الناتج المحلي، البطالة، الدخل، الصادرات، الواردات، ..الخ..

والاقتصاد القياسي يعتبر فرع من فروع علم الاقتصاد يختص بالقياس (التقدير) الكمي للعلاقة بين المتغيرات مستخدما النظرية الاقتصادية والرياضيات والأساليب الاحصائية، بهدف إختبار النظريات الاقتصادية المختلفة من أجل المساعدة في رسم السياسات الاقتصادية من ناحية ومساعدة رجال الأعمال والحكومات في إتخاذ القرارات المناسبة من ناحية أخري.

تعريف الاقتصاد القياسي (Econometric):

هناك الكثير من التعريفات لهذا المصطلح وعلى الرغم من اختلافها بالصيغة إلا أنها تشترك بأنها تعتبر الاقتصاد القياسي نوعاً خاصاً من التحليل الاقتصادي يستخدم كلاً من النظرية الاقتصادية والرياضيات الاقتصادية والإحصاء للوصول إلى نتائجه.

تعريف (1): الاقتصاد القياسي هو فرع من فروع علم الاقتصاد يختص بقياس وتحليل العلاقات الاقتصادية مستعيناً بالنظرية الاقتصادية والأدوات الرياضية والأساليب الاحصائية بهدف تحليل واختبار النظريات الاقتصادية المختلفة، والتنبؤ بقيم المتغيرات الاقتصادية في المستقبل. وذلك للمساعدة في رسم السياسات الاقتصادية و اتخاذ القرارات.

تعريف (2): يعرف الاقتصاد القياسي بأنه أحد أحدث فروع علم الاقتصاد، يهتم بتقدير (قياس) العلاقة بين المتغيرات الاقتصادية باستخدام الطرق الرياضية و الإحصائية بغرض اختبار مدى صحة هذه العلاقات كما تقدمها النظرية الاقتصادية. وذلك بهدف رسم السياسات الاقتصادية والتنبؤ بسلوك بعض المتغيرات الاقتصادية.

1-1- تاريخ الاقتصاد القياسي

يعدّ علم الاقتصاد القياسي علماً حديثاً نسبياً إذا ما قورن بالعلوم الاقتصادية الأخرى.وهناك آراء مختلفة حول نشأة الاقتصاد القياسي منها ما يقول إنها تعود إلى الاقتصادي البريطاني William petty أواخر القرن السابع عشر.

ومنها ما يقول إنها تعود إلى الإحصائي الألماني Ernest engel عام (1821-1896) الذي توصل إلى مجموعة من قوانينه من خلال أبحاثه.

ونظراً لأهمية الاقتصاد القياسي في التحليل الاقتصادي، أسس بعض واضعي الفكر الاقتصادي الأوائل من أمثال مور H.More، وشولتز H.Schultz، وفريش وستون R. Stone الجمعية الدولية للاقتصاد القياسي International Econometrics Association في عام 1930 في أمريكا.

1-2- تطور وانتشار الاقتصاد القياسي

حقق الاقتصاد القياسي تقدماً سريعاً وتطوراً كبيراً في دراسة الظواهر الاقتصادية المختلفة وتقدير دوال (توابع) الإنتاج والتكاليف والنماذج القياسية التي تصف العلاقات الاقتصادية على مستوى الاقتصاد الكلي والجزئي. أيضاً أصبح علم الاقتصاد القياسي يستخدم أداة فعالة في حل المعضلات الاقتصادية وفي عمليات التخطيط الاقتصادي.

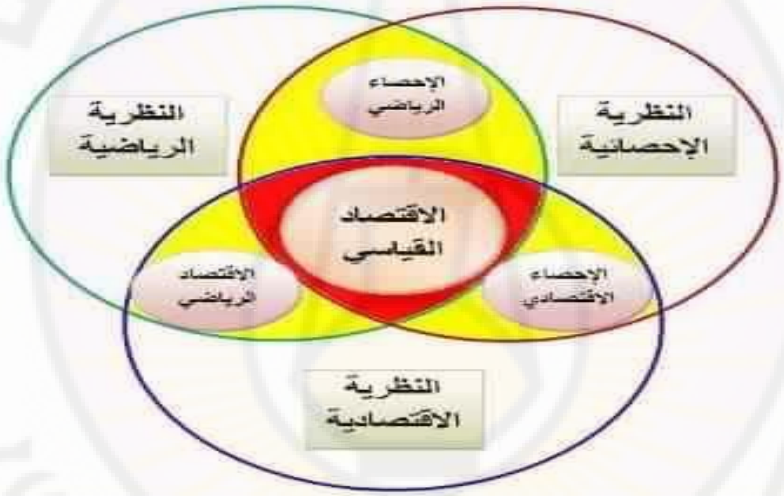
بدأ تطبيق مبادئ هذا العلم بالانتشار حديثاً في بلدان العالم الثالث. وساعد على انتشار طرائق الاقتصاد القياسي عاملان اثنان هما:

1. توافر الإحصـــاءات الاقتصـــادية بكميات أكبر وبدقة أفضـــل. وهي تؤلف المادة الأولية للبحث العلمي في الاقتصاد القياسي.

2. التطور الكبير والســـريع في مجال الحاســـبات الإلكترونية الذي مكن من التوســـع في النماذج الاقتصادية لتشمل عدداً كبيراً من المتغيرات بعد أن كان ذلك مقتصراً على التحليل النظري.

1-3- علاقة الاقتصاد القياسي بالعلوم الأخرى

علم الاقتصاد القياسي يعتمد على ثلاثة علوم هي: الاقتصاد والرياضيات والإحصاء كما واضح من الشكل التالي:



وفيما يلي نبين علاقة الاقتصاد القياسي بكل فرع من هذه العلوم:

1. **الاقتصـاد:** وظيفته هي دراسـة العلاقة بين المتغيرات وبين الظاهرة المدروسـة والعوامل المؤثرة عليها، أي إن النظرية الاقتصادية تزودنا بطبيعة واتجاه العلاقة بين المتغيرات، كالعلاقة بين الكمية المطلوبة من سلعة معينة وسعرها وأسعار السلع البديلة.

2. **الرياضـيات:** يقتصـر دوره على صـياغة العلاقات بين المتغيرات الاقتصـادية المدروسـة التي تم تحديدها اعتماداً على النظرية الاقتصـادية على شـكل رموز ومعادلات رياضـية، ومسـألة قياس متغيرات هذه المعادلات وإثبات ملائمتها للظاهرة المدروسة من مهمات الاقتصاد القياسي.

3. **الإحصـــاء:** ودوره في تجميع البيانات الإحصـــائية الخاصـــة بالمتغيرات المدروســـة، وكذلك تطبيق الاختبارات الإحصـــائية المختلفة على معالم النماذج لبيان معنوية تأثير كل عامل من العوامل على الظاهرة المدروسة، ومعنوية العلاقة ومعالجة أخطاء التقدير تمهيداً لتبني هذه العلاقات.

ويمكن تلخيص ما سبق فيما يلي:

الاحصاء الاقتصادي يختصر دوره على تجميع البيانال الاحصانية الخاصة بالمتغيرات الاقتصادية. والاحصاء الرياضي يمد الاقتصاد القياسي بطرق التقدير بعد تعديلها لتتناسب مع طبيعة العلاقات الاقتصادية. بينما ينصب دور الاقتصاد القياسي على استخدام هذه البيانات والتي يقدمها له الاحصاء الاقتصادي في تقدير قيم المعالم واختبار نوع العلاقة بين المتغيرات، وذلك باستخدام طرق القياس الاحصائي، والذي يقدمها له الاحصاء الرياضي.

1-4- النظرية الاقتصادية والاقتصاد القياسي

النظرية الاقتصـــادية توضـــح العلاقة بين المتغيرات الاقتصـــادية المختلفة وتفســـر ســـلوك بعض الظواهر الاقتصادية.

مثال: كلما ارتفع سعر سلعة معينة كلما قل الطلب عليها مع ثبات العوامل الأخرى.

هنا هذا الفرض يحدد العلاقة بين الكمية المطلوبة وسعر السلعة.

وبتوضيح أكثر فإن نظرية الطلب تشير إلى أن طلب المستهلك على سلعة معينة يتحدد بعدد من العوامل أهمها سعر السلعة، وأسعار السلع الأخرى، ودخل المستهلك، وذوق المستهلك.

ويأتي دور الاقتصاد الرياضي لإعادة صياغة العلاقات الاقتصادية كما تحددها النظرية من أسلوب لفظي إلى أســـلوب رياضـــي. وهذا يعني انه لا يوجد اختلاف بين النظرية الاقتصـــادية والاقتصـــاد الرياضي إلا في وسيلة التعبير عن العلاقات الاقتصادية، فالاقتصاد الرياضي يعبر عن العلاقات في

$$D1 = a_0 + a_1 p_1 + a_2 p_2 + a_3 Y + a_4 S$$

نظرية الطلب بالصيغة التاليةوالتي تسمى دالة الطلب وهي توضح أن التغير في أي من المتغيرات الأربعة التي تظهر بالجانب الأيمن يصاحبه تغير في طلب السلعة.

D_1	الكمية المطلوبة من السلعة
P_1	سعر السلعة 1
P_2	سعر سلعة أخرى
Y	الدخل
S	الذوق

1-5- الفرق بين النظرية الاقتصادية والاقتصاد القياسي

- النظرية الاقتصـــادية ومن ثم الاقتصـــاد الرياضـــي تنظر إلى كل المتغيرات التي تؤثر في الظواهر الاقتصـــادية على أنها متغيرات منتظمة تتغير بصـــورة مســـتمرة، ويمكن التعبير عنها بقيم محدودة. كما يمكن التنبؤ بسلوكها بدقة، مثل السعر والدخل.

وهي بذلك تهمل ما يسمى بالمتغيرات العشوائية التي لا تأخذ قيم محددة وليس هناك تأكد من حدوثها ولا تتغير بطريقة منتظمة، ولكنها تؤثر في سلوك أي ظاهرة مثل (انتشار الأوبئة، الزلازل والحروب، التقلبات الجوية، الأحداث السياسية ،الإشاعات).

وهكذا النظرية الاقتصادية تهمل المتغيرات العشوائية. فنظرية الطلب تشير إلى أن التغير في الطلب يرجع لسعر السلعة أو أسعار السلع الأخرى أو الدخل أو الذوق بحيث أن 100% من التغير يرجع للتغير في هذه المتغيرات المنتظمة وأن 0% في التغير في الطلب يرجع للتغير في المتغيرات العشوائية.

- أما الاقتصـــاد القياســـي يهتم بقياس العلاقات كما هي بالواقع مســـتخدماً بيانات فعلية. فإنه يأخذ اثر المتغيرات العشوائية بجانب المتغيرات المنتظمة.

ولما كان وجود المتغيرات العشـــوائية يجعل العلاقات الاقتصـــادية بين أي متغيرين غير مؤكدة، فإن الاقتصاد القياسي ينظر إلى العلاقات الاقتصادية على أنها علاقات احتمالية (غير مؤكدة).فارتفاع سعر

سلعة معينة قد لا يصاحبه في كل مرة انخفاض في الكمية المطلوبة كما تقرر النظرية الاقتصادية رغم ثبات المتغيرات الأخرى، ويحدث هذا إذا صاحب إشاعة بان السلعة سوف تختفي وهذا يعني انه وجود إشاعة كمتغير عشوائي يجعل احتمال أن تكون العلاقة عكسية بين السعر والكمية المطلوبة أقل من 100% ولهذا السبب فان دالة الطلب في الاقتصاد القياسي تأخذ الصيغة التالية.

$$D1 = a_0 + a_1 p_1 + a_2 p_2 + a_3 Y + a_4 S + u$$

حيث u يعتبر المتغير العشوائي.

وهكذا فإن النظرية الاقتصادية ومن ثم الاقتصاد الرياضي لا تحدد درجة العلاقة بين المتغيرات الاقتصادية كما هي في الواقع، وإنما تحدد فقط طبيعة واتجاه هذه العلاقة.

فهي لا تحدد مثلا نسبة الزيادة في الطلب التي يمكن أن تنجم عن ارتفاع الدخل بنسبة 10% ولعل هذا يرجع إلى أن النظرية الاقتصادية ومن ثم الاقتصاد الرياضي لا يحددان قيماً رقمية لمعاملات العلاقات الاقتصادية .

ولكن الاقتصاد القياسي يحدد درجة العلاقة بين المتغيرات الاقتصادية كما هي في الواقع من خلال قياس معاملات المرونة. فباستخدام الأساليب الإحصائية يمكن للاقتصاد القياسي تحديد قيم المعاملات من خلال معاملات واقعية متاحة وبتحديد هذه القيم يمكن حساب معاملات المرونة التي تحدد درجة العلاقة الاقتصادية بين المتغيرات.

1-6- أنواع الاقتصاد القياسي

هناك نوعان أساسيان من الاقتصاد القياسي هما:

1. الاقتصاد القياسي النظري (Econometric Theory):

ويعني بدراسة وتطوير الطرق الخاصة بقياس العلاقات الاقتصادية التي يتضمنها النموذج الاقتصادي القياسي.

2. **الاقتصاد القياسي التطبيقي (Applied Econometric):**

والذي يستعمل الطرق القياسية النظرية لدراسة موضوعات معينة في حقل الاقتصاد مثل دالة الاستهلاك، ودالة الإنتاج، ودالة الاستثمار، والنماذج القياسية للاقتصاديات القومية وقطاعاتها المختلفة.

1-7- أهداف الاقتصاد القياسي

يمكننا تلخيص أهداف الاقتصاد القياسي بالنقاط التالية:

1. اختبار النظرية الاقتصادية (هدف تحليلي)
2. تقدير قيم المعالم وتفسير بعض الظواهر الاقتصادية (هدف حسابي).
3. التنبؤ بسلوك المتغيرات الاقتصادية(هدف استشرافي)
4. رسم أو تقييم السياسات الاقتصادية(هدف تخطيطي)

أولاً: اختبار النظرية الاقتصادية (هدف تحليلي):

تحليل واختبار النظريات الاقتصادية يعتبر هدفاً رئيسياً من أهداف الاقتصاد القياسي، حيث نجد أن الاقتصاديين في المرحلة الأولى لتطور النظرية الاقتصادية استخدموا الأسلوب الوصفي لصياغة القوانين الاقتصادية، ووضعوا مجموعة من الفروض دون أن تختبر من الناحية التطبيقية، وعندما جاء الاقتصاد القياسي فقد هدف أساساً إلى التحقق من النظريات الاقتصادية عن طريق اللجوء للواقع وقياس العلاقة موضع الدراسة بهدف اختبار مدى صحة النظرية الاقتصادية ويوجد احتمالان بهذا الصدد:

أ- أن تتفق النظرية مع الواقع، وفي هذه الحالة تقبل النظرية على أنها صحيحة في ظل اظروف الراهنة.

ب- أن تتعارض النظرية مع الواقع، وفي هذه الحالة أن ترفض النظرية في صورتها القديمة، أو نعدلها ثم يتم اختبارها من جديد.

وبالتالي أصبح في واقع العلاقات الاقتصادية اليوم أنه ليست هنالك نظرية يمكن قبولها إلا إذا دعمها الاختبار التطبيقي للتحقق من مدى صحتها.

ثانياُ: تقدير قيم المعالم وتفسير بعض الظواهر الاقتصادية (هدف حسابي):

يتم تطبيق طرق الاقتصاد القياسي بهدف الحصول على قيم عددية لمعاملات العلاقات الاقتصادية، كالمرونات و الميول والمعلمات الفنية للإنتاج، وكلها أدوات لها أهميتها في تفسير العلاقات بين الظواهر الاقتصادية يساعد في اتخاذ القرارات، وفي صياغة السياسات الاقتصادية.

ثالثاً: التنبؤ بسلوك المتغيرات الاقتصادية(هدف استشرافي)

تكمن أهمية الاقتصاد القياسي في عملية التنبؤ في تحقيق تقديرات للنماذج الاقتصادية التي تستخدم في الحصول على تنبنؤات كمية للمتغيرات الاقتصادية، وهي مهمة في مجال تخطيط المستقبل، وتمكن واضعي السياسة ومتخذك القرار من تنظيم الحياة الاقتصادية.

فمثلاً ا إذا رغبت السلطات الاقتصادية في وضع سياسة للعمالة، لابد لها من دراسة الموقف الحالي للعمالة وما يجب أن يكون عليه مستوى العمالة في السنوال القادمة، وبالتالي فإن استخدام أسلوب الاقتصاد القياسي يمكننا من التنبؤ بحجم العمالة في المستقبل.

كذلك إذا ارادت السلطات الاقتصادية أن تفرض ضريبة جديدة، فعليها أن تعرف مقدماً الآثار المتوقعة لهذه الضريبة على التضخم والبطالة والناتج القومي، وبالطبع لا يمكننا عمل ذلك دون تقدير نموذج الدخل الذي يربط بين كل هذه المتغيرات.

رابعاً: رسم أو تقييم السياسات الاقتصادية(هدف تخطيطي)

إن معرفة القيم للمتغيرات الاقتصادية تساعد في اتخاذ القرار المناسب ورسم السياسات الاقتصادية، سواء أكان على مستوى المؤسسة أم على مستوى الدول، كذلك إن معرفة القيم العددية للمعاملات الخاصة بالعلاقات الاقتصادية يساعد في عملية المقارنة، ومن ثم تقييم السياسات الاقتصادية.

فمثلاً أذا أرادت منشاة أن ترسم سياسة سعرية ملائمة لزيادة ايراداتها الكلية فلابد من معرفة القيم الرقمية لمرونة الطلب على السعرية لسلعتها.

1-8- تطبيقات الاقتصاد القياسي

يعتبر مجال تطبيق الاقتصاد القياسي واسعاً جداً حيث يشمل كافة الظواهر الاقتصادية:

1. يتم تطبيق الاقتصاد القياسي في تحليل الدورات الاقتصادية، بهدف التنبؤ بمواعيدها والتصدي للأزمات الاقتصادية ومعالجتها أو التخفيف من حدتها قبل حدوثها وتقليص الخسائر الناجمة عنها.
2. أبحاث السوق وتحديد مرونة الطلب والعرض، ورسم السياسات السعرية من أجل توجيه أنماط الإنتاج و الاستهلاك باتجاهات مرغوب فيها.
3. وعلى صعيد أجهزة الدولة المسئولة عن تخطيط عملية التنمية فإن تطبيق تقنيات الاقتصاد القياسي يعتبر ذو أهمية خاصة، بالتنبؤ بالصادرات والواردات والتضخم والدخل والاستثمارات.
4. دراسة مستويات الإنتاج وعلاقتها بالتكلفة، وهي دراسات ذات أهمية في مسائل تخطيط الإنتاج على صعيد الوحدات والقطاعات الإنتاجية. إذ تبين هذه الدراسات الأهمية النسبية لكل عامل من عوامل الإنتاج في العملية الإنتاجية على صعيد المؤسسة وأهميته في النمو الاقتصادي على مستوى القطاع والمجتمع.

1-9- تطبيقات الاقتصاد القياسي على مستوى الاقتصاد الجزئي والكلي

← على مستوى الاقتصاد الجزئي:

حيث يمكن استخدام تطبيقاته لتحديد دوال (توابع) الإنتاج والتكاليف على مستوى المنشأة وكافة اشتقاقاتها مثل دوال الناتج المتوسط والناتج الحدي والتكلفة المتوسطة والحدية. وكذلك يقيس تأثير العوامل المؤثرة على الإنتاج كمياً، ويحدد الحدود المثلى من كل عامل التي يجب إدخالها في العملية الإنتاجية، ويحدد التوليفة المثلى من العوامل مجتمعة التي تحقق أفضل عائدية.

← على مستوى الاقتصاد الكلي:

يمكن باستخدام النماذج القياسية تقدير دوال الاستهلاك والطلب للسلع المختلفة على المستوى الكلي. وكذلك دوال الإنتاج (بصيغها غير الخطية المختلفة). كما يمكن بناء نماذج قياسية(متعددة المعادلات) توصف الاقتصاد ككل وتتضمن دوال الدخل القومي والاستثمار والاستخدام والاستهلاك والتجارة الخارجية (الصادرات والواردات).

1-10- النموذج الاقتصادي أنواعه ومتطلباته

تعريف النموذج: هو تمثيل مبسط لظاهرة واقعية أو لشيء معقد. والتبسيط يعني تلخيص الحقائق في صورة مركزة وهذه يعني فقدان جزء من المعلومات التفصيلية ذات الأهمية الأقل والتركيز على المعلومات والعلاقات ذات الأهمية الأكبر.

وسبب استخدام النماذج في معظم العلوم هو ان مختلف الظواهر تكون عادة في واقعها الفعلي معقدة ومتشابكة الى درجة يتعذر معها دراستها الا من خلال تمثيلها بشكل مبسط باستخدام النماذج.

تعريف النموذج الاقتصادي (Economic Model):

النموذج الإقتصادي هو الصورة المبسطة التي تمثل النشاط الإقتصادي للدولة أو القطاع الخاص في شكل صيغ رياضية وقيم عددية. وهذا النموذج يوضح سلوكية وطبيعة العلاقات بين المتغيرات الإقتصادية ، والغرض منه تسهيل وصف طبيعة تلك العلاقات (تحديد العوامل التي تؤثر فيها) بصورة خالية من التفاصيل والتعقيدات، وممثلة للواقع.

يهدف النموذج الإقتصادي إلى تبسيط الواقع من خلال بناء نموذج إقتصادي لا يحتوي علي جميع تفاصيل الظاهرة الإقتصادية المراد دراستها بل يتضمن العلاقات الأساسية بها، والغرض منها تسهيل وصف طبيعة تلك العلاقات بصورة خالية من التفاصيل والتعقيدات وممثلة للواقع بمعني أنها لا تعكس الواقع الإقتصادي وإنما تعطى صورة مقربة منه.

1-10-1- أنواع النماذج الاقتصادية

يمكن صياغة النموذج الاقتصادي على عدّة صور حيث يمكن الاكتفاء بالتعبير الأدبي (السردي) ليكون النموذج وصفيا، أو يمكن استخدام القياس الكمي ليكون النموذج رياضياً، كما يمكن استخدام أساليب اختبار العلاقات المفترضة ليكون النموذج قياسيا كما يأتي.

1- النموذج الوصفي (Descriptive Model)

في هذا النوع من النماذج يستخدم الأسلوب اللفظي (السردي) القائم على المنطق لشرح ظاهرة اقتصادية معينة أو لتقديم تحليلاً وصفياً لمختلف العلاقات الموجودة بين المتغيرات الاقتصادية بطريقة.

وهذا النموذج نفسه الذي تقترحه النظرية الاقتصادية في تحليلها وتفسيرها للظواهر الاقتصادية ونوع العلاقات الموجودة بين المتغيرات الاقتصادية بطريقة سردية كأن يقال مثلاً: إن استهلاك العائلات يتناسب طرداً مع دخلها، أو يقال إن الطلب على سلعة معينة يتناسب عكساً مع سعرها.

2- النموذج الرياضي (Mathematical Model):

هو عبارة عن تمثيل رياضي مبسط للنموذج الاقتصادي يقتصر على الرموز والمعادلات الرياضية التي تمثل العلاقة بين المتغيرات الاقتصادية. مثلاً: في حالة العلاقة بين الدخل و الاستهلاك يمكن كتابة النموذج الرياضي المعبر عن العلاقة الطردية بين الدخل والاستهلاك كما يلي: $Y = ax + b$

حيث Y هو الاستهلاك و x يمثل الدخل بينما a و الـ b معاملات النموذج. حيث تتصف المعادلات الجبرية بكون العلاقات فيه محددة أو مؤكدة Deterministic وليست احتمالية.

3- النموذج القياسي (Econometric Model):

النموذج القياسي هو عبارة عن نموذج اقتصادي رياضي يتضمن إضافة متغيرات عشوائية إلى العلاقات بين المتغيرات الاقتصادية، وهو يتكون من معادلة واحدة مثل معادلة الطلب أو عدة معادلات. ويفيد هذا النموذج في معرفة أو رصد سلوك بعض المتغيرات في الماضي، ثم التنبؤ بسلوكها المستقبلي.

مثلاً: يمكن كتابة النموذج القياسي في حالة العلاقة بين الدخل و الاستهلاك المعبر عن العلاقة الطردية بين الدخل والاستهلاك كما يلي: $Y = ax + b + u_i$

حيث Y هو الاستهلاك و x يمثل الدخل.

بينما a و الـ b معاملات النموذج.

u_i: تمثل متغير عشوائي يعبر عن الأخطاء التي يمكن أن تكون في النموذج ويحقق مجموعة من القرضيات تسمح لاحقاً بتقدير معالم النوذج (a, b).

الفرق بين النموذج الرياضي والنموذج القياسي هو المتغير العشوائي الذي يجعل العلاقات بين المتغيرات الاقتصادية في النموذج القياسي علاقات احتمالية غير مؤكدة كما يتميز النموذج القياسي بكونه يسمح بتقدير العلاقة الموجودة بين المتغيرات خلافاً للنموذج الاقتصادي والرياضي اللذان يكتفيان بتحديد اتجاهها فقط. أي أن النموذج الرياضي نموذج محدد (Deterministic) بينما النموذج القياسي نموذج جبري احتمالي (Stochastic).

6-2 تصنيف النماذج بناءً على علاقتها بالزمن

يمكن تصنيف النماذج الاقتصادية بناءً على علاقتها بالزمن إلى نوعين هما النماذج الساكنة والحركية:

النماذج الساكنة (Static Models)

هذا النوع من النماذج لا يعتمد على الزمن ولا يظهر الزمن فيه كمتغير مستقل

النماذج الحركية (Dynamic Models)

في النماذج الحركية يلعب الزمن دوراً هاماً في التأثير على بعض متغيراته. ومن الأمثلة على ذلك.

نموذج الاستهلاك الساكن.....	$Y = ax + b + u$
نموذج الاستهلاك الحركي	$Y_t = a_1 x_t + a_2 x_{t-1} + b + u$

ومن الواضـــح أن الزمن لا يؤثر في النموذج الأول، حيث أن الدخل (X) عند نقطة معينة يؤثر في الاسـتهلاك (Y) عند نفس النقطة. أما النموذج الحركي فيوضـح أن دخل الفترة السـابقة (X_{t-1})يؤثر أيضا في استهلاك الفترة الحالية (Y_t) بجانب دخل الفترة الحالية (X_t).

المحاضرة الثانية

منهجية البحث في الاقتصاد القياسي

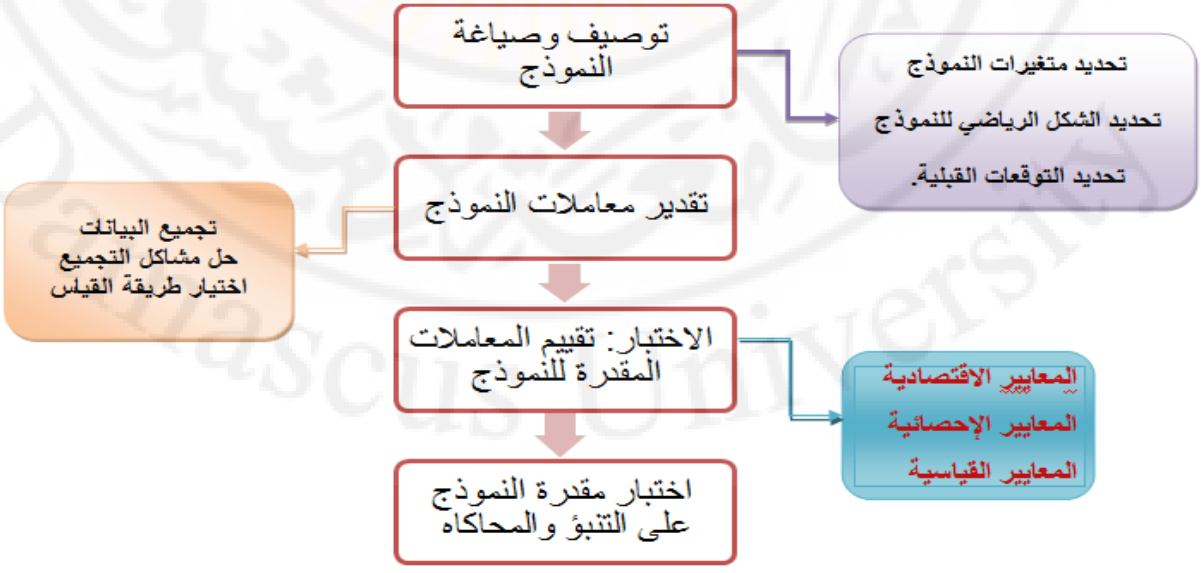
يعتبر الاقتصاد القياسي مثل أي علم له منهجية علمية خاصة لقياس العلاقات الاقتصادية

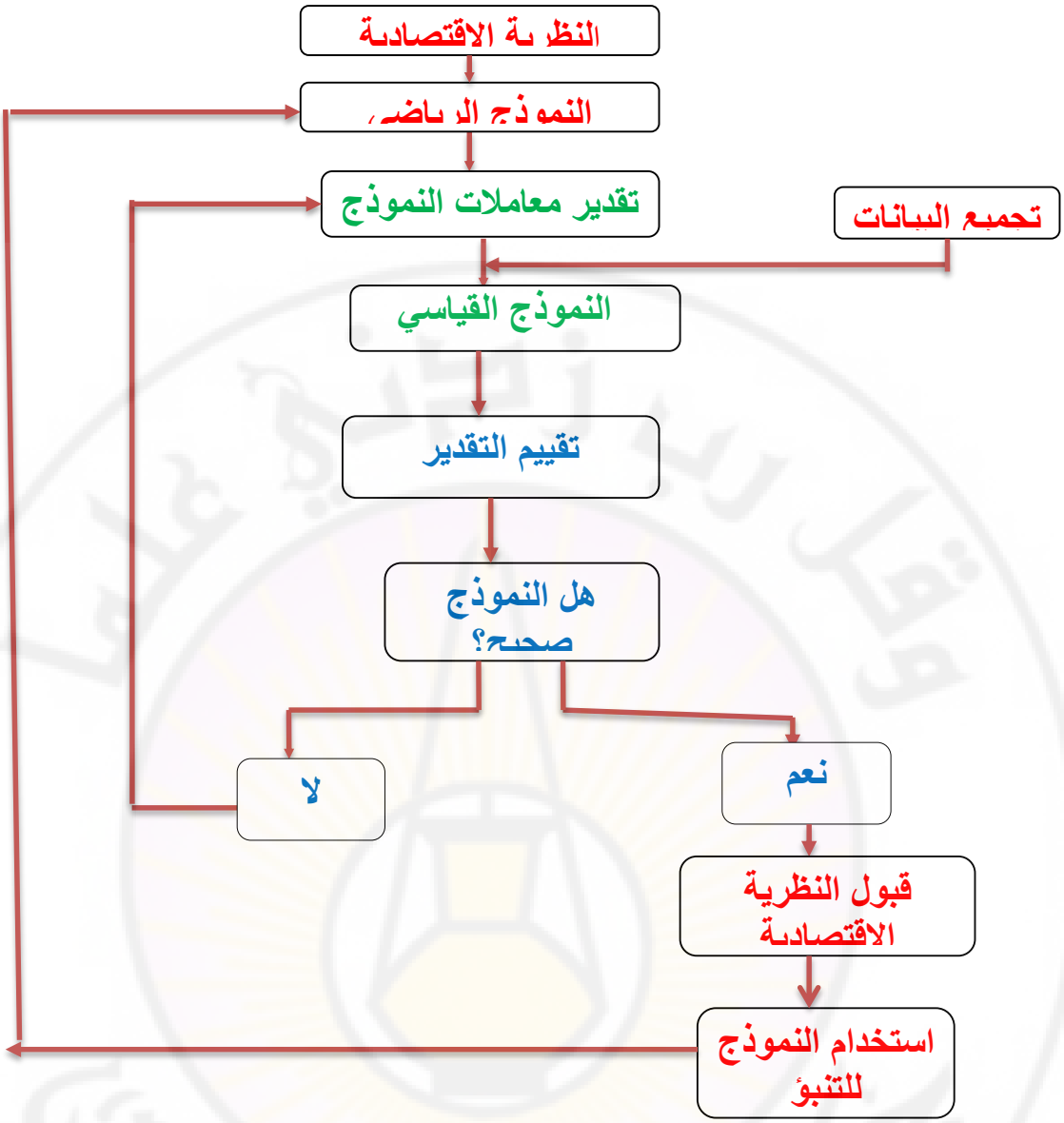
وبالتالى فإن أي بحث قياسي لابد وأن يمر بأربع مراحل هي:

2-1- مراحل تطبيق منهجية الاقتصاد القياسي

1. المرحلة الاولى: توصيف و صياغة النموذج (Specification of the Model)
2. المرحلة الثانية (القياس): تقدير معاملات النموذج (Estimation of the Model)
3. المرحلة الثالثة (الاختبار): تقييم المعاملات المقدرة للنموذج (Evaluation of the Model)
4. المرحلة الرابعة (التنبؤ): اختبار مقدرة النموذج على التنبؤ (Evaluation of the Forecasting). والشكل التالي يصف الخطوات السابقة:

و يمكن وصف المراحل السابقة بالمخطط التفصيلي التالي:





2-2- المرحلة الاولى: توصيف النموذج (Specification of the Model)

تسمى هذه مرحلة وضع الفروض، ويقصد بتوصيف النموذج صياغة العلاقات الاقتصادية محل البحث في صورة رياضية, أو صياغة العلاقات بين المتغيرات الاقتصادية المدروسة التي تم تحديدها اعتماداً على النظرية الاقتصادية على شكل رموز ومعادلات رياضية, حتى يمكن قياس معاملاتها باستخدام الطرق القياسية، وخطوات هذه المرحلة.

1- تحديد متغيرات النموذج

2- تحديد الشكل الرياضي للنموذج

3- تحديد التوقعات القبلية.

أولاً: تحديد متغيرات النموذج:

تتحد المتغيرات من النظرية الاقتصـــادية أو من دراســـات قياســـية ســـابقة، أو المعلومات المتاحة عن الظاهرة.

مثال: نريد وضع نموذج للطلب على سيارات بالاعتماد على النظرية الاقتصادية

يمكن وصف دالة الطلب على الشكل: $D1 = a_0 + a_1 p_1 + a_2 p_2 + a_3 Y + a_4 S$

D1	الكمية المطلوبة من السلعة
P1	سعر السلعة 1
P2	سعر سلعة أخرى
Y	الدخل
S	الذوق

وممكن تكون هناك متغيرات أخرى بالاعتماد على دراسات قياسية أخرى أجريت بنفس المجال مثل:

- أثبتت دراسات قياسية سابقة أن مستوى الدخل في فترات سابقة Y_{t-1} يؤثر على الطلب في الفترة الحالية.
- الإنفاق الحكومي G وتوزيع الدخل E
- ومن خلال معلومات خاصـــة متاحة يمكن تقدم متغيرات تفســـيرية تؤثر في الطلب مثل عدد الكيلومترات من الطرق المخصصة للفرد (K)
- والتعريفة الجمركية على واردات السيارات M

وبجمع كل هذه المتغيرات من مصادر مختلفة يمكن كتابة دالة الطلب كالتالي.

$$D_1 = f(P_1, P_2, Y, Y_{t-1}, S, G, E, K, M)$$

ثانياً: تحديد الشكل الرياضي للنموذج:

وهي وضـع المعادلات الرياضـية التي تعبر عن الموضـوع (معادلة واحدة أو عدة معادلات وهل هي خطية أم غير خطية) حيث إن النظرية الاقتصادية لا تقدم ما يوضح النموذج الرياضي الملائم.

ثالثاً: تحديد التوقعات القبلية:

لنأخذ دالة الطلب التالية $$D1 = a_0 + a_1 p_1 + a_2 p_2 + a_3 Y + a_4 S$$

وفقا لما هو متاح يمكن أن نتوقع أن تكون إشارة a_1 سالبة لان العلاقة عكسية بين الطلب وسعر السلعة بينما إشارة a_3 موجبة لأنها تمثل الدخل.

2-3- المرحلة الثانية: تقدير معاملات النموذج (Estimation of the Model)

هذه المرحلة تسمى مرحلة اختبار الفروض. إن عملية تقدير معاملات النموذج يتم بثلاث مراحل هي:

1- تجميع البيانات

2- حل مشاكل التجميع

3- اختيار طريقة القياس الملائمة.

1- **تجميع البيانات:** هناك أنواع مختلفة من البيانات كما يوضحها الجدول التالي:

1-	بيانات سلسلة زمنية	Time Series Data
2-	بيانات قطاعية	Cross- Section Data
3-	بيانات سلسلة قطاعية	Cross-Series Data
4-	بيانات تجريبية	Experimental Data

- **بيانات سلسلة زمنية (Time Series Data):**

عدد من القياسات لمتغير ما عند نقاط زمنية مختلفة وهي تصف سلوك المتغير الاقتصادي عبر الزمن(انتاج سوريا من القمح من 1990-2000).

- **بيانات قطاعية(Cross- Section Data):**

وهي القياسات التي يأخذها متغير ما عند نقطة زمنية معينة (انتاج المحافظات السولرية من القمح عام 2015)

- **بيانات سلسلة قطاعية(Cross- Series Data):**

وهي مزيج من بيانات السلسلة الزمنية والبيانات القطاعية. فهي تعطي بيانات عن مجموعة من المفردات عبر سلسلة زمنية(انتاج المحافظات السورية من القمح من 2000 حتى عام 2020)

- **بيانات تجريبية (Experimental Data):**

هذه تحصل في محلات السوبر ماركت عند تغير سعر سلعة معينة ورصد الكميات المباعة خلال أسبوع مع ثبات العوامل الأخرى.

2- حل مشاكل التجميع

مشكلة التجميع تنشأ عندما يحتاج الباحث لاستخدام متغيرات تجميعية في الدالة محل القياس مثل الناتج القومي والاستهلاك القومي. وعملية التجميع تتم على أكثر من مستوى فهناك التجميع على مستوى الأفراد مثل الدخل القومي الذي هو مجموع دخول الأفراد ومن المشاكل التي تواجه الباحث في هذه الحالة اختلاف محتوى الدخل فهناك الدخل النقدي والدخل العيني. وعند التجميع عادة الباحث يركز على الدخل النقدي ويهمل العيني.

كما أن التجميع قد يتم على مستوى الفترات الزمنية. مثلا ممكن مصادر إحصائية تنشر بيانات ربع سنوية في حين أنها مطلوب أن تكون سنوية في هذه الحالة لابد من تجميعها.

قد يتم التجميع على أساس جغرافي مثل تجميع بيانات الناتج أو السكان لمدن أو دول مختلفة.

على الباحث أن يتأكد من حل مشاكل التجميع قبل أن يبدأ في عملية تقدير المعاملات.

3- اختيار طريقة القياس الملائمة.

هناك طرق قياسية يمكن استخدامها في قياس العلاقات الاقتصادية.

❖ **طرق المعادلة الواحدة:**

وهي تطبق على كل معادلة من معادلات النموذج على حده، مثل طريقة المربعات الصغرى العادية.

❖ **طريقة المعادلات الأنية:**

ومن أمثلتها طريقة المربعات الصغرى ذات المرحلتين وطريقة المربعات الصغرى ذات الثلاث مراحل وغيرها.

وتختلف هذه الطرق في ملائمتها لعملية القياس من حالة لأخرى تبعا لعدة عوامل منها طبيعة العلاقة محل البحث: هل هي معقدة أم بسيطة، وخصائص المقدرات التي تعطيها كل طريقة: هل هي غير متحيزة ومتسقة وكافية أم غير ذلك، والهدف من البحث القياسي: هل هو اختبار نظرية ما، أم وضع سياسة، أم التنبؤ، أم تفسير ظاهرة.

كما تختلف هذه الظواهر من حيث كمية البيانات التي تتطلبها وتكاليف البحث. ويتعين على الباحث الاسترشاد بهذه العوامل عند اختياره لطريقة القياس.

2-4- المرحلة الثالثة: تقييم المعاملات المقدرة للنموذج (Evaluation of the Model)

بعد أن ينتهي الباحث من تقدير القيم الرقمية لمعاملات النموذج من خلال بيانات واقعية بعدها يشرع في تقييم المعاملات المقدرة. أي تحديد ما إذا كانت قيم هذه المعاملات لها مدلول أو معنى من الناحية

الاقتصـادية والاحصـائية والقياسـية. ويوجد هناك عدد من المعايير التي تسـتخدم في تقييم نتائج التقدير القياسي وهي:

- المعيار الاقتصادي
- المعيار الاحصائي
- المعيار القياسي

أولاً: تقييم التقديرات وفقاً للمعيار الاقتصادي (Economic Criteria)

تعتبر المعايير المستمدة من النظرية الاقتصادية أولى المعايير التي يجب أن تستخدم لتقييم النتائج، حيث تشير النظرية الاقتصادية إلى افتراضات محددة عن إشارات المعلمات المراد تقديرها، وهذه الافتراضات تستخدم للحكم على مدى سلامة التقديرات من الناحية الاقتصصادية.

هذه المعايير تأتي من خلال النظرية الاقتصادية وتتعلق هذه المعايير بحجم وإشارة المعاملات المقدرة.

مثال: نظرية الاستهلاك والدخل $Y = ax + b$ حيث Y هو الاستهلاك و x يمثل الدخل بينما a و الـ b معاملات النموذج.

من خلال هذه النظرية نحدد قيمة المعاملات كما يلي:

$(b\rangle 0)$ يعني أن المجتمع لابد أن يسـتهلك حتى إذا انخفض الدخل على الصـفر ويتم بالاعتماد على الاقتراض أو السحب من المدخرات السابقة.

$(0<a<1)$ أي أن الميل للاسـتهلاك يجب أن يكون موجبا وتتراوح قيمته بين الصـفر والواحد وهكذا فإن نظرية الاستهلاك قد وضعت معايير اقتصادية خاصة بإشارة وحجم المعاملتين (a, b) ويتعين على أي محاولة لقياس دالة الاستهلاك أن تعطي نتائج تتفق مع هذه المعايير حتى يمكن قبولها اقتصادياً.

ملاحظة: إذا جاءت المعلمات المقدرة على عكس ما تقرره النظرية الاقتصادية مسبقاً، فإن هذا يمكن أن يكون مبرراً لرفض هذه المعلمات، مالم يوجد هناك من المبررات المنطقية القوية للتسليم بصحة التقديرات، وفي مثل هذه الحالة يجب عرض هذه المبررات بوضوح.

ثانياً: تقييم التقديرات وفقاً للمعيار الاحصائي (Statistical Criteria)

تهدف هذه المعايير إلى اختبار مدى الثقة الإحصائية في التقديرات الخاصة بمعاملات النموذج مثل اختبارات المعنوية. تعتبر هذه المعايير من المعايير المهمة في دراسة قياس العلاقات الاقتصادية، وذلك للتعرف على معنوية التقديرات ومدى مطابقتها مع منطوق النظرية الاقتصادية وتمثيلها للمجتمع الذك تنتمي إليه، وتنقسم المعايير الاحصائية إلى نوعين من الاختبارات:

اختبار القوة التفسيرية للنموذج (اختبارات جودة التوفيق)

1- اختبار القوة التفسيرية للنموذج (اختبارات جودة التوفيق): (خاص بالنموذج)

2- اختبارات المعنوية الاحصائية: (خاص بالمعلمات المقدرة).

1- اختبار جودة التوفيق (The test of the Goodness of fit):

اختبار جودة التوفيق هو مقياس للمقدرة التفسيرية للنموذج، حيث يعكس هذا الاختبار درجة الانحرافات بين القيم المقدرة والقيم المشاهدة، ويوضح أنه كلما زادت انحرافات القيم المقدرة عن القيم المشاهدة للمتغير التابع، كلما قلت جودة التوفيق، وبالتالي انخفاض المقدرة التفسيرية للنموذج، والعكس صحيح.

من هنا يتضح لنا أن هناك ارتباط تام بين جودة التوفيق والمقدرة التفسيرية. ولذلك من المهم جداً عند إجراء تحليل الانحدار التعرف على نسبة مساهمة المتغيرات المستقلة في إحداث التغيرات في المتغير التابع. ويتم ذلك باستخدام معامل التحديد (R^2)، وكلما ارتفعت قيمة معامل التحديد كان ذلك دليلاً على قوة العلاقة، والعكس هو الصحيح.

فمثلاً إذا كانت قيمة معامل التحديد تساوي ($R^2=0.90$) هذا يعني أن معادلة الانحدار المقدرة ذات جودة توفيق عالية إذ إن المتغيرات المستقلة في المعادلة تفسر (90%) من التغيرات الكلية في المتغير التابع،

بينما يوجد (10%) يتم ارجاعها إلى حد الخطأ العشـوائي (القياسـي)، أي أنه يوجد متغيرات أخرى (غير مضمنة في المعادلة) تؤثر على المتغير التابع بنسبة (10%).

ولكن من عيوب معامل التحديد أنه يبالغ في حقيقة تأثير المتغيرات المســـتقلة على المتغير التابع، ومن أجل علاج هذه المشلكلة تم اقتراح معامل التحديد المعدل $(Adjusted - R^2)$

2- اختبارات المعنوية.

بعد تقدير قيم المعالم من بيانات العينة لابد من اختبار إلى أي مدى يمكن الأعتماد عليها كأســـاس جيد للوصـــول لمعلمات المجتمع، وســـوف يتم ذلك من خلال اختبار مدى ملائمتها الاحصـــائية باســـتخدام اختبارات المعنوية الاحصائية، ويوجد ثلاثة اختبارات يمكن استخدامها لهذا الغرض وهي:

اختبار (t) T-test

اختبار (z) Z = test

اختبار (F) F –test

وقد يتســـأل الدارس أيهاما أكثر أهمية في المعايير الاحصـــائية ؟ معامل التحديد أم اختبارات المعنوية، فعلى ســـبيل المثال أيهما أفضـــل أن يكون معامل التحديد مرتفعاً أم تكون الاخطاء المعيارية للمعلمات المقدرة منخفضة، بالطبع لن تكون عملية الحكم على النموذج المقدر صعبة أذا اتضح أن معامل التحديد مرتفعاً والأخطاء المعيارية منخفضـــة أو العكس، ففي مثل هذه الحالات يوجد هناك اتفاق بين المعيارين في الحكم على النموذج، ولكن تنشـــأ الصـــعوبة عندما يكون معامل التحديد مرتفعاً وفي الوقت نفســـه الأخطاء المعيارية مرتفعة أو العكس، ففي هذه الحالات لا يوجد اتفاق بين المعيارين في الحكم على النموذج.

وهنا يظهر تســـاؤل، هل نقبل المعلمات المقدرة أم نرفضـــها ؟ يرى البعض أن قبول أو رفض المعلمات المقدرة بناءً على معيار ما يعتمد أساساً على الهدف من تقدير النموذج، **إذا كان الهدف هو التنبؤ، فإن**

معامل التحديد يكون هو المعيار الأكثر أهمية، أما إذا كان الهدف من القياس هو تفسير بعض الظواهر الاقتصادية، فإن اختبار المعنوية يعتبر هو الأكثر أهمية.

وعموماً فإن الأولوية تعطى للمعايير الاقتصادية، ثم تأتي بعدها المعايير الاحصائية والقياسية، فإذا لم يجتاز النموذج المقدر اختبار المعايير الاقتصادية بنجاح، فلن يكون هناك أهمية كبرى للاختبارات الأخرى من وجهة النظر الاقتصادية.

ثالثاً: تقييم التقديرات وفقاً للمعيار القياسي (Econometric Criteria)

هذه المعايير تحددها نظرية الاقتصاد القياسي، وتهتم بإرشاد الباحث إلى ما تتصف به التقديرات من خصائص، كعدم التحيز والاتساق والكفاءة، وغيرها من الخصائص المرغوبة في المقدرات، كما تهدف أيضاً إلى البحث عن مدى مطابقة فروض الأساليب القياسية المستخدمة، والتي تختلف باختلاف الطرق القياسية، ولذلك يترتب على الباحث قبل اعتماد نتائج التقدير أن يتأكد من عدم وجود مشاكل قياسية في النموذج موضع الدراسة، وذلك نتيجة لعدم توافر واحد أو أكثر من الافتراضات الكلاسيكية، والتي تشمل خطية النوذج، واستقلالية المتغيرات المستقلة عن الأخطاء العشوائية، وعدم ارتباطها خطياً، وثبات تجانس الأخطاء العشوائية، وعدم ارتباطها ذاتياً.

وأخيراً يتضح من خلال استعراض معايير تقييم النتائج المتحصل عليها أن عملية التقييم في النماذج القياسية ليست بالأمر السهل، بل يتحتم على الباحث استخدام جميع المعايير الاقتصادية والاحصائية والقياسية قبل قبول أو رفض أي من التقديرات، ,غذا لم يتحقق فرض قياسي، فغالباً ما يعاد توصيف النموذج بإضافة أو حذف أو تعديل بعض المتغيرات أو بعض المعادلات، ويعاد اختبارها مرة أخرى.

2-5- المرحلة الرابعة: اختبار مقدرة النموذج على التنبؤ (Evaluation of the Forecasting). اختبار القدرة التنبؤية للنموذج:

يعتبر التنبؤ بقيم المتغيرات الاقتصادية في المستقبل أحد أهم أهداف الاقتصاد القياسي إذ يتم بموجبه التعرف على مسار الظاهرة في المستقبل ليساعد في عملية التخطيط واتخاذ القرارات.,ولذا يتعين اختبار مدى مقدرة النموذج القياسي على التنبؤ. ولاختبار مقدرة النموذج على التنبؤ لابد من اختبار مدى استقرار المعاملات المقدرة عبر الزمن, واختبار مدى حساسية هذه التقديرات للتغير في حجم العينة.

وقبل استخدام النموذج المقدر في عملية التنبؤ ينبغي اختبار مقدرته التبؤية، ففي كثير من الأحيان قد يكون النموذج ذا معنى اقتصادي واحصائي ومقبولاً من الناحية القياسية بالنسبة للفترة التي أخذت فيها العينة،إلا أنه قد لا يكون ملائماً للتنبؤ بسبب التغيرات السريعة في المعالم الهيكلية للعلاقات في الواقع، ولهذا فإنه من الأهمية أن يقوم الباحث باختبار مقدرة النماذج القياسية على التنبؤ، وعلى المستوى التطبيقي توجد عدة اختبارات تستخدم لتحقيق هذا الهدف من أهمها:

1- تطبيق اختبار كاي مربع (X2) لمقارنة التوزيع المتوقع بالتوزيع المشاهد.

2- تطبيق اختبار (t) لاختبار معنوية الفرق بين قيم التنبؤ والقيم الفعلية.

3- تطبيق معامل عدم التساوي ل ثايل.

الفصل الثاني

السلاسل الزمنية (Time series)

الفصل الثاني

السلاسل الزمنية (Time series)

مقدمة:

تعتمد دراسة السلاسل الزمنية بشكل أساسي على فهم السلوك التاريخي للظاهرة المدروسة، وتقدير الوضع الحالي والتنبؤ بمستقبلها.

مما جعل تحليل السلاسل الزمنية يشكل ركناً أساسياً وهاماً في التخطيط الاقتصادي والاداري، وبشكل خاص الجانب المتعلق بالتبؤ، فعلى سبيل المثال إذا كان الأمر مرتبط بإجراء التوسع العمراني أو شبكات النقل أو المؤسسات التعليمية فلا بد من دراسة الزيادة السكانية (معدل النمو) في البلد بالاضافة لدراسة أثار مكونات السلسلة الزمنية.

تعتبر السلاسل الزمنية أحد أساليب الاقتصاد القياسي المستخدمة في التبوء بالمستقبل من خلال وقائع الأمس واليوم، وعملية دراسة تحليل التوقع بحجم الإنتاج تتطلب الحصول على بيانات واحصائيات عن الكمية المنتجة خلال فترات زمنية سابقة تكون محل الدراسة.

ونموذج السلسلة الزمنية يعتمد بالدرجة الأولى على عنصر الزمن كعامل مستقل و قيم الظاهرة المدروسة تعتبر قيم المتغير التابع لفترات سابقة أي سلوك هذا المتغير في الماضي.

تعريف السلسلة الزمنية (Time Series):

لقد تعددت تعاريف السلسلة الزمنية وجميعها تلتقي حول تغير الظاهرة عبر الزمن، وأهم هذه التعاريف.

تعريف(1): السلسلة الزمنية (Time Series): هي مجموعة من المشاهدات أو القياسات المأخوذة عن متغير وأحد أو أكثر مرتبة وفقا لزمن حدوثها في فترات زمنية غالباً تكون متساوية.

تعريف(2): السلسلة الزمنية (Time Series) هي عبارة عن مجموعة من المشاهدات أو القياسات لظاهرة معينة (اقتصادية، اجتماعية، طبية، طبيعية،.) أخذت على فترات زمنية محددة ومتعاقبة وفي أغلب الأحيان تكون هذه الفترات الزمنية منتظمة ومتساوية (سنة، فصل، شهر، أسبوع، يوم، ... إلخ).

تعريف(3): السلسلة الزمنية (Time Series) السلسلة الزمنية هي مجموعة من المشاهدات مرتبة وفق حدوثها في الزمن كالسنين أو الفصول أو الأشهر أو الأيام أو أية وحدة زمنية . فهي بذلك عبارة عن سجل تاريخي يتم اعتماده لبناء التوقعات المستقبلية.

وتتكون السلسلة الزمنية من متغيرين أحداهما متغير مستقل وهو الزمن والاخر متغير تابع يمثل قيم الظاهرة محل الدراسة، مثل انتاج سوريا من القمح من 2000 إلى 2020.

رياضياً: نقول أن متغير الزمن المستقل (t) والقيم المناظرة له المتغير التابع (y) وإن كل قيمة في الزمن (t) يقابلها قيم للمتغير التابع y فإن y دالة في الزمن t أي $y = F(t)$.

وتستخدم السلاسل الزمنية في مجالات عدة منها: الدراسات السكانية والاقتصادية والطبية، حيث تستخدم نماذج السلاسل الزمنية عادة في التنبؤ بقيم متغير ما مثل: التنبؤ بحالة الجو، ودرجات الحرارة، وعدد السكان ومتوسط الدخل والصادرات. من أهم السلاسل الزمنية تلك الخاصة بالمؤشرات الاقتصادية والمبيعات السنوية للشركات بكافة أوجه نشاطاتها والتعليم وحجم السكان وما شابه ذلك.

2-1- وصف السلسلة الزمنية وتمثيلها بيانياً

يتم وصف السلسلة الزمنية عادة عن طريق تمثيلها بيانياً بغرض الحصول على صورة عامة وواضحة لسلوك السلسلة في الفترات الزمنية المختلفة، ويكون المحور الأفقي ممثلاً للزمن والمحور الرأسي للمتغير التابع ونتعرف من الرسم البياني على الأمور التالية:

1- يمكن تحديد اتجاه السلسلة هل هي متزايدة، متناقصة أم ثابتة.

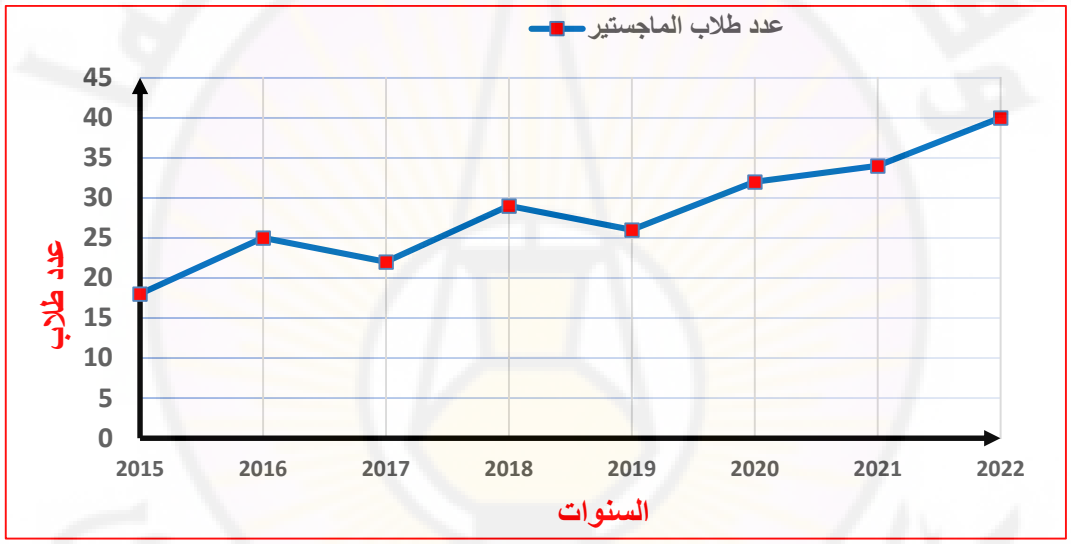
2- التعرف على القيم المتطرفة في السلسلة الزمنية، والتغيرات المتكررة ووصف هذه التغيرات.

3- معرفة الفترات التي يكون فيها لهذه السلسلة قمة (قيمة عظمى) وقاع (قيمة صغرى).

ويمكن تمثيل السلسة بيانياً باتخاذ المحور الأفقي للزمن والرأسي لقيم المتغير كما هو مبين بالشكل الآتي

لجدول البيانات الآتي والدال على عدد طلاب الماجستير في قسم معين لعدة سنوات.

السنة (الزمن)	2015	2016	2017	2018	2019	2020	2021	2022
عدد طلاب الماجستير	18	25	22	29	26	32	34	40



يتضح من الرسم البياني أن هناك تغيرات في عدد الطلاب من سنة لأخرى فقيم هذا المتغير ترتفع سنة وتنزل أخرى إلا أن الطابع العام يدل على زيادة في عدد الطلاب وبالتالي نتوقع زيادة في السنوات القادمة وبناء عليه يستلزم وضع الاستعدادات الخاصة بالمرحلة القادمة بزيادة الطلاب.

2-2- انواع السلاسل الزمنية:

يمكن تصنيف السلاسل الزمنية وفقاً لعدة معايير مثل قيم السلسلة أو طبيعة الزمن أو الاتجاه العام للسلسلة كما ياتي:

أولاً: حسب قيم السلسلة:

يمكن تصنيف السلاسل الزمنية وفقاً لقيم السلسلة إلى الفئتين التاليتين:

- السلاسل الزمنية المتصلة: وهي السلاسل الزمنية التي تكون مشاهدات الظاهرة قيم متصلة متغيره خلال فتره من الزمن الساعة, اليوم, الاسبوع, الشهر, ربع سنة ...الخ , مثل, الدخل العام , وكمية الامطار السنوية، درجات الحرارة وغيرها.
- السلاسل الزمنية غير المتصلة (المتقطعة): وهي السلاسل الزمنية التي تكون فيها قيم الظاهرة قيم منفصلة متغيرة عند لحظة من الزمن , ومن امثلة هذه السلاسل عدد السكان في مدينة ما في اليوم الاول من كل سنة، عدد السيارات في مدينة ما بداية كل عام.

ثانياً: حسب طبيعة الزمن:

يمكن تصنيف السلاسل الزمنية وفقاً لطبيعة الزمن الذي تحدث فيه قيم السلسلة الزمنية إلى نوعين:

- السلاسل الزمنية النقطية: وهي التي تقاس في أزمنة محدده مسبقه ومن امثلة هذه السلاسل سلسلة ارباح شركة الاسمنت في منتصف العام, وسلسلة معدل الدخل السنوي لأفراد والتي تقاس في نهاية كل عام.
- السلاسل الزمنية غير النقطية: وهي السلاسل التي تقاس قيمتها في ازمنة غير متوقعه مثل سلاسل الكوارث, سقوط الطائرات , حوادث القطارات, حوادث السيارات, سلسلة الهزات الارضية.

ثالثاً: حسب الاتجاه العام للسلسلة:

يمكن تصنيف السلاسل الزمنية وفقاً لتغيرات الاتجاه العام لنمو السلسلة إلى أربعة أنواع:

- ✓ السلاسل ذات الاتجاه المتزايد: وهي السلاسل التي يمثلها خط مستقيم متزايد (ميلة موجب) مثل السلاسل التي تمثل اعداد السكان والدخل القومي وسلاسل حوادث السيارات.
- ✓ السلاسل ذات الاتجاه التناقص: وهي السلاسل التي يمثلها خط مستقيم متناقص (ميله سالب) مثل مساحة الاراضي الزراعية في منطقه والتي تتناقص بسبب انتشار الابنية عليها.

- ✓ السلاسل ذات الاتجاه الثابت: وهي السلاسل التي يمكن ان يتوسط نقطها خط مستقيم ثابت (ميله صفر) ومن امثلة ذلك سلسلة الطاقة الكهربائية المستهلكة في اضاءة الاشارات الضوئية , والشوارع الرئيسية في احدى المدن.
- ✓ السلاسل ذات التغيرات المتكررة على فترات متباعدة: وهي السلاسل التي يمكن ان يتوسط نقطها خط يشبه منحني وذلك لان قيم السلسلة قد تتأثر بأمور فصلية او سنوية ومن امثلة ذلك سلسلة مبيعات الملابس الصوفية التي تتم في كافة ايام السنة ولكنها تزداد في فصل الشتاء وتنقص في الصيف.

2-3- مكونات السلسلة الزمنية:

الهدف من دراسة وتحليل السلاسل الزمنية هو التعرف على خصائص تلك السلاسل و التغيرات التي حدثت فيها خلال الفترة الماضية والتي تساعد على التنبؤ في المستقبل ويطلق على تلك الخصائص والمتغيرات (عناصر او مكونات السلاسل الزمنية) وهي:

- الاتجاه العام (Secular Trend)
- التغيرات الموسمية (Seasonal Variations)
- التغيرات الدورية (Cyclical Variations)
- التغيرات العشوائية أو العرضية (Irregular Variations)

إن هذه المكونات (العناصر) الأربع الخاصة بالسلسلة الزمنية والتي تتأثر بالعوامل الاقتصادية والبيئية والاجتماعية والسياسية وسنتعرض لكل من هذه العناصر بصورة تفي بالغرض المطلوب.

2-3-1- الاتجاه العام (Secular Trend):

يعد الاتجاه العام من أهم عناصر السلسلة الزمنية وغالبا ما يكون هو العنصر الوحيد الذي يتم الاعتماد عليه بالتنبؤ بالقيم المستقبلية للظاهرة أي عند استخدام السلاسل الزمنية لأغراض التنبؤ.

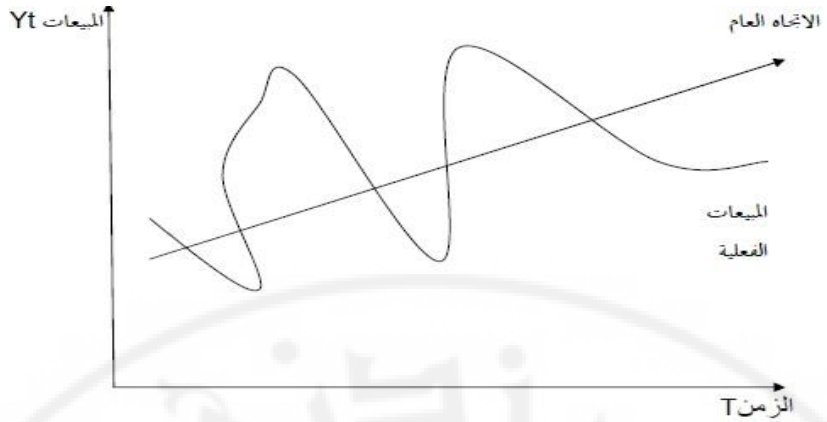
و الاتجاه العام للسلسلة يمثل تطور الظاهرة على المدى البعيد أو التوجه الذي تكون عليه السلسلة أي ميل الظاهرة نحو الزيادة أو النقصان أو الثبات خلال فترة طويلة من الزمن، فإذا مالت قيم الظاهرة إلى التزايد باستمرار يعني أن الاتجاه العام موجب، أما إذا مالت إلى التناقص باستمرار فإن الاتجاه العام سالب ويكون على شكل خط مستقيم، وفي حال مالت الظاهرة إلى التزايد ثم تحولت إلى النقصان يتمثل الاتجاه العام بخط منحني (قطع مكافئ).

- ورياضياً الاتجاه العام قد يكون خطياً ويتمثل رياضياً بخط مستقيم وفي هذه الحالة فإن الزيادة من فترة إلى أخرى قد تكون ثابتة ، كما يمكن أن يكون الاتجاه العام غير خطي مثل المنحنى الأسي أو منحنى يأخذ شكل قطع مكافئ وهو معادلة رياضية من الدرجة الثانية $y = a\,t^2 + b\,t + c$ وفي هذه الحالة فإن الزيادة من فترة إلى أخرى تكون بنسب مئوية ثابتة.

يختلف شكل الاتجاه العام في السلاسل الزمنية حسب طبيعة البيانات، وأحد أنواع الاتجاه العام هو الاتجاه الخطي.

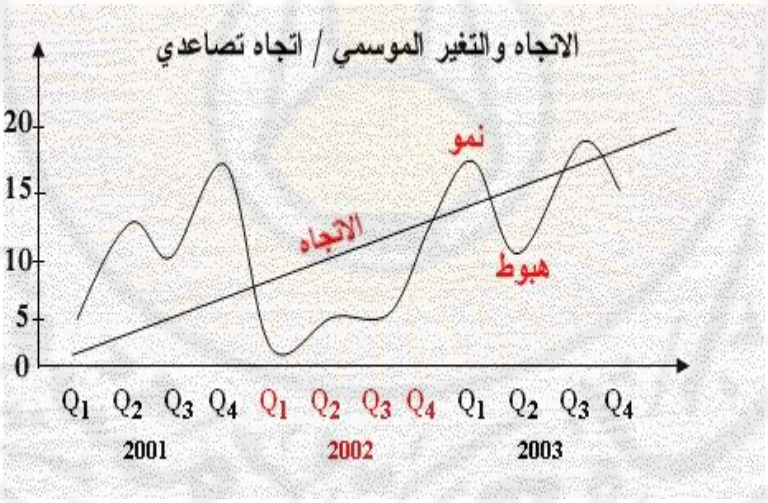
وقد نجد في الواقع أشكالا مختلفة أخرى للاتجاه العام ومعالجتها تتم بواسطة التحويل اللوغارتمي إلى اتجاها خطيا ، أو بواسطة تقنيات الانحدار غير الخطي.

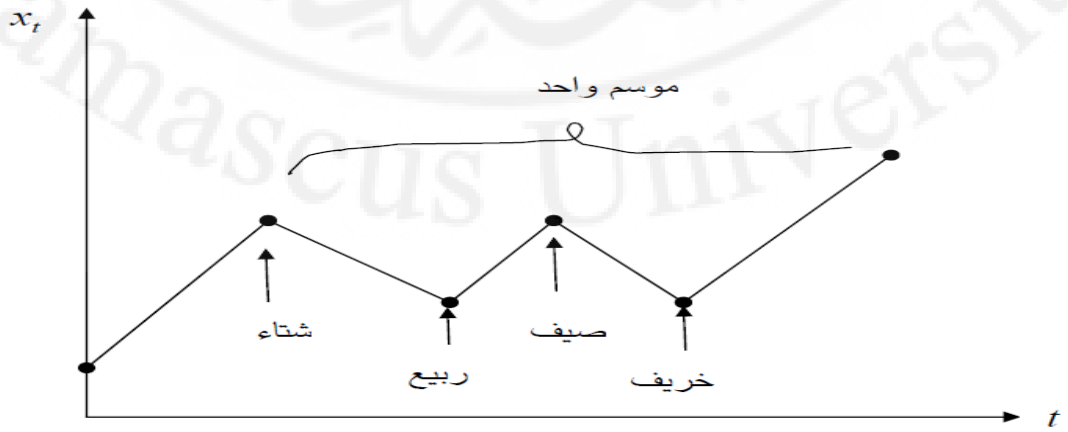
ويمكن رسم الشكل البياني الذي يجسد ظاهرتين أحدهما الزمن ولتكن الأخرى المبيعات، حيث يمثل المحور الأفقي الزمن بينما يمثل المحور الرأسي المبيعات على مدى السلسلة الزمنية ، ثم تمثل النقاط الموافقة ، ثم يمد خط يتوسط هذه النقاط أو يمر بأكبر عدد ممكن منها ، ممثلا بذلك خط الاتجاه العام ، سواء كان صاعدا أو نازلا ،كما يوضح الشكل التالي:



2-3-2- التغيرات الموسمية (Seasonal Variations)

التغيرات الموسمية هي التغيرات التي تحدث بشكل منتظم وتكرر نفسها في وحدات أو فترات زمنية شبه ثابتة، وهذه التغيرات تتكرر كل سنة وتأتي في مواعيد ثابتة (فصل أو شهر أو يوم بالسنة مثل الأعياد الدينية والحج (زيادة بيع الملابس والحلويات)، عيد رأس السنة، عيد الحب) ويرمز لها بالرمز اللاتيني (S). والتغيرات الموسمية لا تزيد طول فترته عن السنة والشكل التالي يبين نموذج لهذا المتغير (الموسمي).



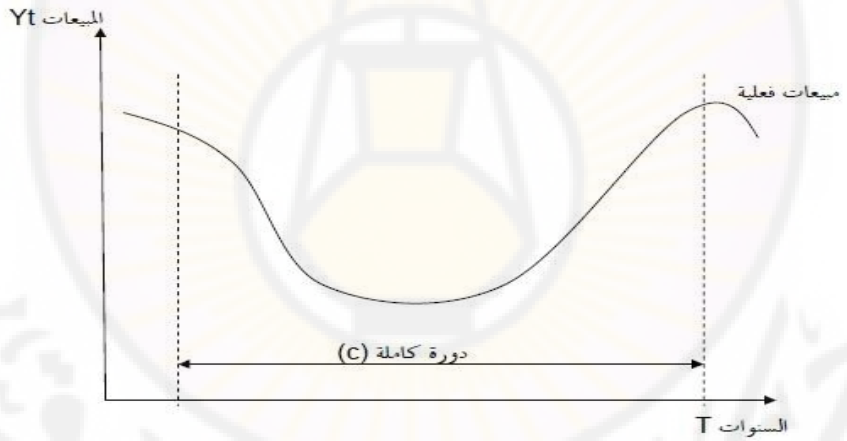


الشكل (1) :التغيرات الموسمية للسلسلة الزمنية

2-3-3- التغيرات الدورية (Cyclical Variations)

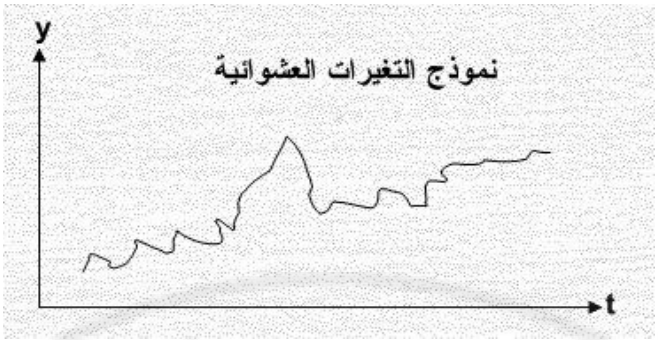
وهيَ التغيرات التي تحصل على قيمة الظاهرة بصورة دورية وتعيد نفسها خلال فترات زمنية تزيد عن السنة الواحدة (وقد تستغرق من سنة الى 10 أو أكثر) وتسمى هذه التغيرات بالتذبذبات الدورية وتكون هذه التذبذبات اقل انتظاماً من التغيرات الموسمية حيث ان الذبذبة الواحدة لا تتكرر بنفس الطول او القوة كفترات الركود والازمات الاقتصادية التي تحدث بصورة دورية ويرمز لهذه التغيرات بالرمز (C) . والدورة هيَ المسافة بين التقعرين او التحدبين في منحنى السلسلة الزمنية.

وبصورة عامة تتضمن الفترة عدة مراحل هي:(المرحلة الاولى وهي الارتفاع– مرحلة التراجع– مرحلة الانتعاش المحدود (الركود)– مرحلة الانتعاش أو الانفراج– مرحلة الارتفاع النهائي). تأخذ الفترة بين الارتفاع الأولي والارتفاع النهائي دورة كاملة كما يوضحه الشكل التالي:

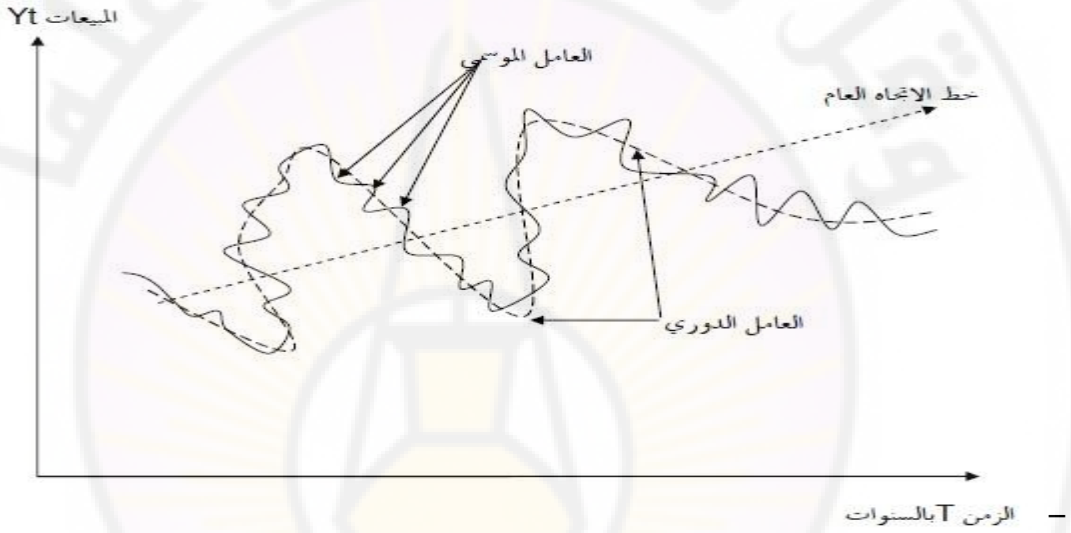


2-3-4- التغيرات العشوائية أو العرضية (Irregular Variations)

التغيرات العشوائية تكون نتيجة عوامل عشوائية خارج نطاق السيطرة ولا يمكن التحكم أو التبؤ بها كالأوبئة والزلازل والحروب والبراكين وهذه التغيرات مهمة نسبيًا بالرغم من صعوبة تفسيرها إلا انه يمكن إظهار تأثيرها وهي تظهر كتذبذبات صغيرة في بيانات السلسلة الزمنية، ويرمز لهذه التغيرات بالرمز (I) ، والشكل التالي يبين نموذج للتغير العشوائي.



ويمكن التعبير عن العناصر المكونة للسلسلة الزمنية بالشكل البياني التالي:



2-4- تحليل السلسلة الزمنية (Time series Analysis):

يعد أسلوب تحليل السلاسل الزمنية واحداً من أهم الاساليب الإحصائية التي تعتمد على تتبع الظاهرة المدروسة أو (المتغير) حدثت بتواريخ زمنيه متلاحقة في الماضي ودراسة التغيرات التي جرت عليها لغرض تحليل أسبابها ونتائجها بغية فهم الحاضر والتنبؤ بالمستقبل، بناء على القيم المختلفة التي ظهرت في السلسلة الزمنية وعلى نمط النمو في القيم.

والغرض الأساسي من تحليل السلاسل الزمنية هو تفكيك السلسلة إلى مكوناتها الأساسية (الاتجاه العام، المركبة الموسمية، المركبة الدورية، والمركبة العرضية) وهذا ما يدعى بتحليل السلسلة الزمنية أو تبسيط القيمة الإجمالية إلى العناصر الجزئية المكونة لها.

أهداف تحليل السلاسل الزمنية:

تتلخص أهداف تحليل السلاسل الزمنية بالنقاط الاتية:

1) وصف وتحديد طبيعة الظاهرة من خلال المشاهدات المتتابعة.
2) بناء نموذج لتفسير وشرح سلوك السلسلة الزمنية من خلال ربط القيم المشاهدة ببعض قواعد السلسلة.
3) استخدام النتائج من تحليل السلسلة للتنبؤ بسلوك السلسلة في المستقبل.

2-4-1- نماذج السلاسل الزمنية الرياضية:

كما نعلم أن مكونات السلسلة الزمنية هي مركبة الاتجاه العام (T)، والتغيرات الموسمية (S)، و التغيرات الدورية (C)، والتغيرات العشوائية (I).

بعد تحليل السلسة الزمنية إلى مكوناتها (الاتجاه العام – التغيرات الموسمية – التغيرات الدورية – التغيرات العشوائية) **يتم تحديد العلاقة** بين هذه المكونات من خلال نماذج الجمع والضرب الخاصة بالسلسة الزمنية، أي أن تحليل السلسلة الزمنية يتطلب صياغة نموذج رياضي يمثل السلسلة المعطاة، وقد طور الأخصائيون عدة نماذج رياضية تربط بين قيم المشاهدات، وقيم المركبات المختلفة للسلسلة الزمنية. ومن أبرز النماذج الرياضية التي تصف السلسلة الزمنية هي النموذج الضربي والنموذج الجمعي حيث أن النموذج الضربي يعد أكثر النماذج شيوعاً.

- **النموذج الضربي (Multiplicative Model)**

هذا النموذج يفترض أن قيمة الظاهرة (المشاهدة) عند أي نقطة زمنية يساوي حاصل ضرب أثار المركبات الأربعة، مركبة الاتجاه العام والمركبة الفصلية والمركبة الدورية و العشوائية، أي $Y=T*S*C*I$. ويستعمل

هذا النموذج غالباً في الحالات التي تكون فيها المركبات (S, C, I) معطاة أو متمثلة على صورة نسب مئوية أو أرقام قياسية (بدلاً من القيم الموجبة والسالبة) بينما قيمة الاتجاه العام (T) تتمثل بقيمها الأصلية.

النموذج الجمعي (Additive model):

هذا النموذج يفترض أن قيمة الظاهرة (المشاهدة) عند أي نقطة زمنية يساوي حاصل جمع أثار المركبات الأربعة أي $Y=T+S+C+I$. حيث يعبر عن المكونات بقيم عددية وليس نسب مئوية.

في هذا النموذج تكون المتغيرات مستقلة عن بعضها البعض، بمعنى أن حدوث أحدها لا يؤثر في حدوث المركبات الأخرى، ويستعمل هذا النموذج غالباً في الحالات التي تكون فيها وحدة قياس جميع المركبات متشابهة وتشابه وحدة قياس المشاهدات. وسوف نستخدم الرموز الآتية:

T قيمة الاتجاه العام (Secular Trend)

S قيمة التغيرات الموسمية (Seasonal Variations)

C قيمة التغيرات الدورية (Cyclical Variations)

I قيمة التغيرات العشوائية أو العرضية (Irregular Variations)

Y قيمة الظاهرة عند زمن معين (مشاهدات السلسلة الزمنية)

2-5- طرائق إيجاد الاتجاه العام:

إن عملية تقدير (ايجاد) الاتجاه العام تعتبر الخطوة الاولى في تحليل السلسلة الزمنية.

والاتجاه العام يحدد صفة تزايد السلسلة الزمنية أو تناقصها أو ثباتها، ولهذا فإن الاتجاه العام للسلسلة الزمنية يوضح نمو السلسلة أو انكماشها.

يوجد العديد من الطرق لتقدير الاتجاه العام للظواهر المختلفة، تختلف كل منها عن الأخرى من حيث طبيعتها ومدى دقتها ومدى مرونة استخدامها في التنبؤ من أهم بين هذه الطرق:

- طريقة الرسم البياني باليد (Scattered Method)

- طريقة المتوسطات النصفية (Semi Average method)
- طريقة المتوسطات المتحركة (Moving Averages Method)
- طريقة المربعات الصغرى (Least Square method)

2-5-1- طريقة الرسم البياني باليد (Scattered Method):

في هذه الطريقة يمكن تقدير خط الاتجاه العام بيانياً برسم خط أو منحنى باليد على ورق الرسم البياني يمثل حركة السلسلة الزمنية خلال فترة زمنية طويلة نسبياً، وقد يكون خط الاتجاه العام خط مستقيم أومنحنى يتوسط النقاط الحقيقة، وقد يكون الخط ذو ميل موجب أو ميل سالب. من مميزات هذه الطريقة أنها سريعة وسهلة التنفيذ، أما من عيوبها فإنها غير دقيقة تخضع للحكم الشخصي، أي أن التقدير قد يختلف من شخص لآخر.

2-5-2- طريقة المتوسطات النصفية أو أشباه المتوسطات (Semi Average method):

تتلخص هذه الطريقة بتقسيم السلسلة الزمنية لنصفين متساويين (زمنياً) واستبعاد السنة الوسطية (الواقعة في منتصف السلسلة الزمنية) في حال عدد السنوات فردي، ويحسب متوسط كل نصف ويوضع امام السنة الوسطية لنصف السلسلة فنحصل على نقطتين (الوسط الحسابي، السنة الوسطى) ثم يتم التوصيل بين النقطتين بخط مستقيم ليمثل خط الاتجاه العام.

وفي حالة السنوات الزوجية تكون النقطة (الوسط الحسابي، منتصف السنتين في الوسط) كما في المثالين الآتيين:

مثال(1): حدد خط الاتجاه العام بطريقة المتوسطات النصفية للجدول الآتي الذي يمثل قيمة استهلاك الماء بالمتر المكعب لأسرة للفترة 2000-2006م.

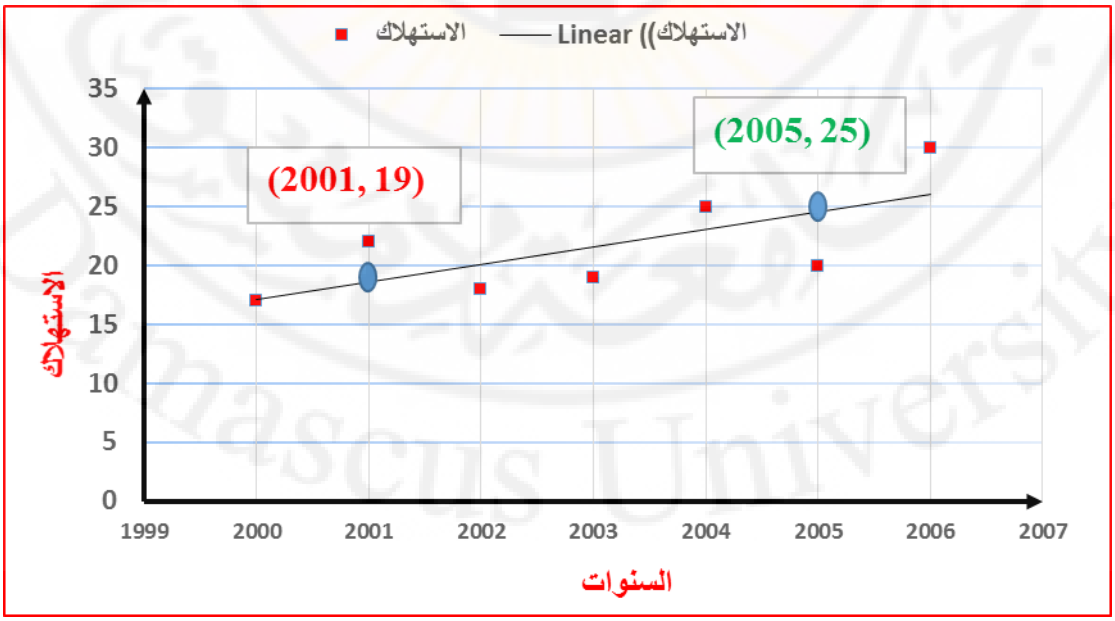
السنة	2000	2001	2002	2003	2004	2005	2006
الاستهلاك	17	22	18	19	25	20	30

الحل: نقسم السلسلة الزمنية إلى نصفين (تم حذف السنة الوسطى 2003) كما يلي: (يمكن حذف سنة 2000 أو 2006)

السنة	2000	2001	2002	2004	2005	2006
الاستهلاك	17	22	18	25	20	30

نحسب المتوسط لكل قسم: من 2000 إلى 2002 : المتوسط = (17 + 22 + 18) ÷ 3 = 57 ÷ 3 = 19 عند منتصف الفترة (النصف من السلسلة) عند 2001.

من 2004 إلى 2006 : المتوسط = (25 + 20 + 30) ÷ 3 = 75 ÷ 3 = 25 عند منتصف الفترة عند 2005. فنحصل على النقطتين (2001, 19) و (2005, 25) نقوم برسم خط الاتجاه العام كما بالشكل :



لمعرفة قيمة الاتجاه العام لسنة بأن نحسب الزيادة السنوية للاتجاه:

قيمة الاتجاه العام لسنة = متوسط النصف الواقعة فيه السنة + [(مجموع الزيادة بين المتوسطين) ÷ عدد السنين بين المتوسطين] × موقع السنة المطلوبة

قيمة الاتجاه العام لسنة 2002 = 19 + [(25 – 19) ÷ 3] × 3 =19 + 6=25

مثال(2): حدد خط الاتجاه العام بطريقة المتوسطات النصفية للجدول الآتي الذي يمثل أجور العاملين بملاين الليرات في أحدى المؤسسات للفترة 1992– 1999م.

السنة	1992	1993	1994	1995	1996	1997	1998	1999
الأجور	22	18	25	20	30	28	24	20

الحل: نقسم السلسلة الزمنية إلى نصفين كما يلي:

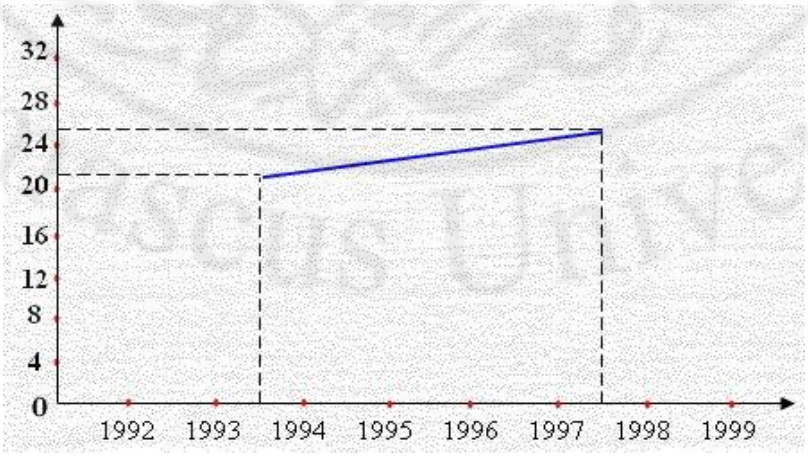
السنة	1992	1993	1994	1995	1996	1997	1998	1999
الأجور	22	18	25	20	30	28	24	20

نحسب المتوسط لكل قسم:

1992- 1995: المتوسط = (22 + 18 + 25 + 20) ÷ 4 = 85 ÷ 4 = 21.25 مقابل منتصف الفترة (النصف من السلسلة) بين 1993 ، 1994.

1996- 1999: المتوسط = (30 + 28 + 24 + 20) ÷ 4 = 102 ÷ 4 = 25.5 مقابل منتصف الفترة (النصف من السلسلة) بين 1997 ، 1998.

نحدد النقطتين ونصل بينهم بخط مستقيم يمثل خط الاتجاه العام كما هو واضح بالشكل الآتي:



الطريقة الرياضية لتحديد معادلة الاتجاه العام:

أما الطريقة الرياضية لتحديد معادلة خط الاتجاه العام وميل الخط هي كما يلي:

1- نقسم الجدول الذي يمثل السلسلة الزمنية إلى قسمين متساويين ثم نحسب الاوساط الحسابية كما يلي:

$$\bar{x}_1 = \frac{\sum x_1}{n_1} \qquad \bar{x}_2 = \frac{\sum x_2}{n_2} \qquad \bar{y}_1 = \frac{\sum y_1}{n_1} \qquad \bar{y}_2 = \frac{\sum y_2}{n_2}$$

حيث أن:

$\bar{x}_1$: متوسط العينة الأولى لمتغير الزمن.

$\bar{x}_2$: متوسط العينة الثانية لمتغير الزمن.

$\bar{y}_2$: متوسط العينة الاولى من المشاهدات.

$\bar{y}_2$: متوسط العينة الثانية من المشاهدات.

n_1 : عدد مشاهدات العينة الاولى.

n_2 : عدد مشاهدات العينة الثانية.

2- نحسب معاملات خط الاتجاه العام (a, b) كما يلي:

$$a = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} \qquad b = \bar{y} - a\bar{x}$$

3- معادلة خط الاتجاه العام المقدرة هي $\hat{y} = ax + b$

مثال(3): البيانات التالية تمثل سلسلة زمنية لتكلفة انتاج سلعة معينة، أوجد معادلة الاتجاه العام- قدر قيمة التكلفة المتوقعة للسلعة بسنة 2011، علماً أن سنة الاساس هي (2002).

السنة	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
التكلفة	15	16	18	18	20	21	22	24	28	27

الحل:

	السنة	التكلفة	X
النصف الاول	2000	15	-2
	2001	16	-1
	2002	18	0
	2003	18	1

2	20	2004	
3	21	2005	النصف الثاني
4	22	2006	
5	24	2007	
6	28	2008	
7	27	2009	

$$\bar{x}_1 = \frac{\sum x_1}{n_1} = \frac{0}{5} = 0 \qquad \bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{25}{5} = 5$$

$$\bar{y}_2 = \frac{\sum y_2}{n_2} = \frac{122}{5} = 24.4 \qquad \bar{y}_1 = \frac{\sum y_1}{n_1} = \frac{87}{5} = 17.4$$

نحسب معاملات خط الاتجاه العام (a, b) كما يلي:

$$a = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} = \frac{24.4 - 17.4}{5 - 0} = 1.4 \quad b = \bar{y}_1 - a\bar{x}_1 = 17.4 - 1.4 * 0 = 17.4$$

معادلة خط الاتجاه العام المقدرة هي $\hat{y} = 1.4x + 17.4$

ثم نشكل الجدول الاتي:

السنة	التكلفة	x	$\hat{y}$
2000	**15**	**-2**	**14.6**
2001	**16**	**-1**	**16**
2002	**18**	**0**	**17.4**
2003	**18**	**1**	**18.8**
2004	**20**	**2**	**20.2**
2005	**21**	**3**	**21.6**
2006	**22**	**4**	**23**
2007	**24**	**5**	**24.4**
2008	**28**	**6**	**25.8**
2009	**27**	**7**	**27.2**

المطلب الثاني: تقدير قيمة التكلفة لسنة 2011

$$\hat{y} = 1.4x + 17.4 = 1.4 * 9 + 17.4 = 30$$

2-5-3- طريقة المتوسطات المتحركة (Moving Averages Method)

تعتبر طريقة المتوسطات المتحركة من الأساليب البسيطة لتقدير الاتجاه العام عن طريق تخفيف التذبذبات للسلسلة وجعل بيانات السلسلة المدروسة أكثر انسياباً بتمهيد أو تنعيم السلاسل الزمنية. و يتم تمهيد البيانات

المشاهدة بتخليصها من التغيرات قصيرة الأجل التي تحدث بسبب التغيرات الفجائية غير المنتظمة (والتغيرات الموسمية إن وجدت). ولحساب المتوسطات المتحركة لابد أولاً من تحديد عدد الوحدات الزمنية التي يطلق عليه **طول الدورة** يكون أساساً لحساب هذه المتوسطات من أجل تمهيد البيانات بالتخلص من الذبذبات والتعرجات التي تعتريه.

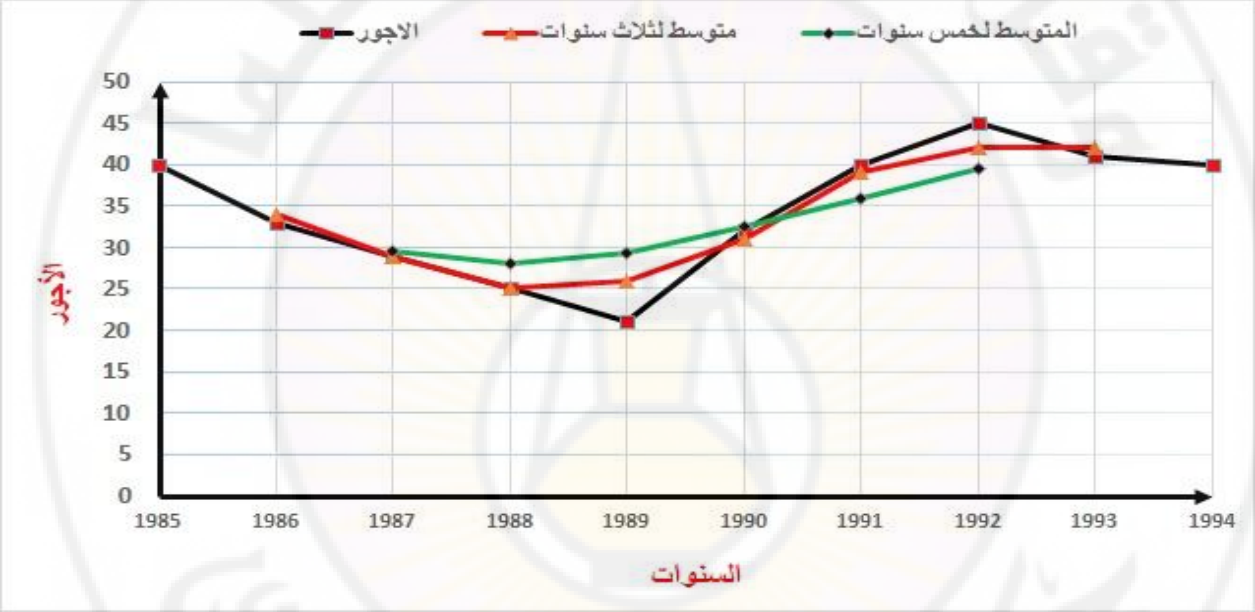
مثال(4): الجدول الآتي يبين طريقة حساب المتوسطات المتحركة على أساس ثلاث سنوات وخمس سنوات لسلسلة زمنية من عشرة سنوات.

السنة	الأجور بالآف	المجموع المتحرك لثلاث سنوات	المتوسط للسنوات الثلاث	المجموع المتحرك لخمس سنوات	المتوسط للسنوات الخمس
1985	40				
1986	33	102	34		
1987	29	87	29	148	29.6
1988	25	75	25	140	28
1989	21	78	26	147	29.4
1990	32	93	31	163	32.6
1991	40	117	39	179	35.8
1992	45	126	42	198	39.6
1993	41	126	42		
1994	40				

إن طريقة المتوسطات المتحركة تعتمد على وضع أساس معين يتم الحساب عليه مثل أساس ثلاثي أو رباعي أو خماسي أو سباعي، فالثلاثي يعني بأن نقوم بجمع القيم المقابلة للثلاث أزمنة الأولى في السلسلة الزمنية (40 + 33 + 29) = 102 ووضعه أمام الزمن الثاني (1986) وأخذ متوسطه 102 / 3 = 34) ووضعه أمام الزمن الثاني (1986) تحت عمود المتوسط وتجاهل القيمة المقابلة للزمن الأول في عمودي المجموع والمتوسط، نكرر ما سبق بجمع القيم المقابلة للأزمنة الثلاثة التالية بدءاً من الثاني (33 + 29

+ 25) =87 وتوضع مقابل سنة 1987 ونأخذ المتوسط 87 /3 = 29 ووضع المتوسط أمام سنة 1987 في عمود المتوسط وتجاهل القيمة المقابلة للزمن الثاني (1986) ، ... هكذا حتى مجموع الثلاثة الأخيرة.

إذا كان الأساس خماسي (جمع خمسة قيم) نقوم بنفس السابق مع جمع خمسة قيم الأولى ووضع المجموع والمتوسط أما السنة الواقعة في المنتصف 1987 ثم نجمع الخمسة التالية بدءاً من السنة الثانية ووضع المجموع والمتوسط أمام سنة المنتصف 1988وهكذا كما مبين بالجدول الآتي يبين حساب المتوسطات المتحركة على أساس ثلاث سنوات وخمس سنوات والتمثيل البياني له:



مثال(5): هذا المثال يبين خط الاتجاه على أساس أربع سنوات لحساب المتوسطات المتحركة وتمثيل بياني لها من خلال الجدول الآتي:

السنة والفصل		الأجور بالآف	المجموع المتحرك لأربع فصول	المتوسط للفصول الأربع	المتوسط المركزي	مجموع اخطاء التنبؤ مطلقة
1997	Q_1	40				
	Q_2	33				
	Q_3	29	127	31.75	29.375	0.375
	Q_4	25	108	27	26.875	1.875
1998	Q_1	21	107	26.75	28.125	
	Q_2	32	118	29.5	32	
	Q_3	40	138	34.5	37	

	40.5	39.5	158	45	Q_4
	71.375	41.5	166	41	1999 Q_1
	41.125	41.25	165	40	Q_2
		41	164	39	Q_3
				44	Q_4

نجمع الفصول الأربع لسنة 1997 (Q1 + Q2 + Q3 + Q4) والناتج (127) يوضع بين Q2 , Q3 من سنة 1997 كما مبين بالجدول باللون الأصفر.

نجمع الفصول الأربع 3 من سنة 1997 والرابع من 1998(Q2 + Q3 + Q4 + Q1) والناتج (108) يوضع بين Q3 و Q4 من 1997.

نكرر ذلك حتى مجموع الفصول الأربع الأخيرة ومجموعها 164 نأخذ المتوسط بقسمة المجموع المتحرك على 4 كما مبين في العمود الرابع.

نوجد المتوسط المركزي بأن نأخذ متوسط كل متوسطين متتاليين (27 + 31.75) ÷ 2 = 58.75 ÷ 2 = 29.375



مثال(6): اوجد المتوسطات المتحركة لبيانات السلسلة الزمنية التالية بافتراض أن طول الدورة يساوي (خمس وحدات زمنية) ثم أرسم معادلة الاتجاه العام المقدرة في مقابل القيم الحقيقية.

السنة	1	2	3	4	5	6	7	8	9	10	11	12
التكلفة	20	25	18	22	25	30	33	27	35	28	25	30

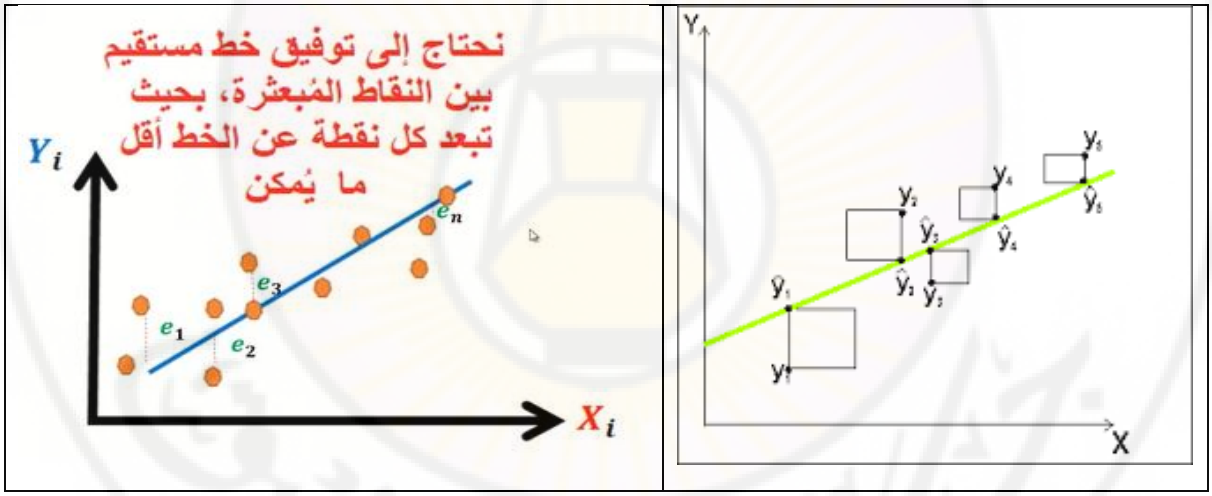
السنة	الأجور بالآف	المتوسط للسنوات الخمس
1	20	
2	25	
3	18	22
4	22	24
5	25	25.6
6	30	27.4
7	33	30
8	27	30.6
9	35	29.6
10	28	29
11	25	
12	30	

يتبين من الشكل أن منحني المتوسطات المتحركة (أو تقديرات الاتجاه العام) أكثر تمهيداً ونعومة من منحني السلسلة الزمنية الأصلية والسبب في ذلك أن المتوسطات المتحركة تضعف التغيرات غير النمطية لأنها تغيرات لا تستمر طويلاً وأحياناً تؤثر على الظاهرة بالزيادة وأحياناً بالنقصان، فضلاً عن ذلك يمكن ملاحظة أن منحنى الاتجاه العام المقدر قريب من الدرجة الثانية. كذلك منحني الاتجاه العام المقدر يكون أكثر تمهيدا ونعومة بزيادة طول فترة الدورة الزمنية، وذلك أذا كانت التأرجحات حول منحني الاتجاه العام عشوائية (غير منتظمة) أي أذا كانت السلسلة خالية من التغيرات الموسمية.



2-5-4- التقدير بطريقة المربعات الصغرى الاعتيادية (Ordinary Least Squares- OLS):

طريقة المربعات الصغرى العادية هي الأكثر استخداماً وهي تعتمد على أسلوب اختيار أفضل خط مستقيم يمر بأغلب النقاط المبعثرة لعينة المشاهدات X ، Y وبحيث تبعد كل نقطة أقل ما يمكن عن الخط المقدر. بشكل أكثر دقة فإن طريقة المربعات الصغرى تعتمد على مبدأ تصغير مجموع مربعات الفروق بين القيم الفعلية للمتغير التابع (Y) والقيم المقدرة الواقعة على الخط المستقيم (أي تصغير مجموع مربعات البواقي إلى آدنى قيمه لها). مع العلم أن البواقي هي الأخطاء العشوائية (ε) وتمثل الانحرافات لقيم السلسلة عن خط الاتجاه العام للبيانات باستثناء المتغيرات الموسمية.



ويرمز لهذا الخط بـ $\hat{Y}_i = \hat{a}\, X_i + \hat{b}$

$\hat{Y}_i$: القيمة المقدرة (التي تقع على خط الانحدار) المناظرة للمشاهدة الفعلية Y_i

$\hat{b}$: القيمة المقدرة للمعلمة الحقيقية المجهولة b

$\hat{a}$: القيمة المقدرة للمعلمة الحقيقية المجهولة a

ويمكن الحصول على قيم $\hat{b}$ و $\hat{a}$ بطريقة المربعات الصغيرة العادية.

من الشكل يمكننا تعريف البواقي (e_i) على أنه يساوي:

e_i = القيم الحقيقية للمشاهدة Y_i – القيم المقدرة للمشاهدة $\hat{Y}_i$ أي $(e_i = Y_i - \hat{Y}_i)$

و يمكن للبواقي أن تكون سالبة أو موجبة حسب موضع نقطة مشاهدة من الخط المقدر.

وتعطي طريقة المربعات الصغيرة العادية أفضل خط مستقيم يوفق مشاهدات العينة X ,Y بمعنى أنها تعطي أقل مجموع مربعات رأسية لانحرافات كل مشاهدة عن الخط المستقيم ويتضمن ذلك تصغير مجموع المربعات لانحرافات النقاط الرأسية على الخط لأدنى حد ممكن أي:

$$min \sum_{i=1}^{n} e_i^2$$

وبما أن:

$$min \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$
$$= \sum_{i=1}^{n} (Y_i - \hat{b} - \hat{a}X_i)^2$$

تبعاً لمبدأ المربعات الصغيرة يجري أختيار قيم $\hat{b}$ و $\hat{a}$ وتحديد الخط المقدر $(\hat{Y} = \hat{b} + \hat{a}Xi)$ بحيث يتم تصغير مجموع مربع حد الخطأ $(\sum_{i=1}^{n} e_i^2)$ إلى أدنى حد ممكن وتعتبر هذه مسألة في تحديد النهاية الصغرى لدالة $(\sum_{i=1}^{n} e_i^2)$ حيث يتم مساواة التفاضل الجزئي الأول للدالة بالنسبة لـ $\hat{b}$ و $\hat{a}$ بالصفر، ثم نحصل على معادلتين آنيتين يتم حلها آنيا للحصول على القيم المجهولة ($\hat{b}$ و $\hat{a}$).

وبعد إجراء العديد من عمليات الاشتقاق والتفاضل يمكن تقدير ميل خط الاتجاه العام (a) بإحدى المعادلتين التاليتين:

$$\hat{a} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i{}^2 - (\sum x_i)^2} \qquad \hat{a} = \frac{\sum yx - n\bar{y}\,\bar{x}}{\sum x_i^2 - \frac{(\sum x)^2}{n}}$$

أو يمكن تقدير ميل خط الاتجاه العام بالمعادلة التالي:

$$\hat{a} = \frac{\sum (y - \bar{y})(x - \bar{x})}{\sum (x - \bar{x})^2}$$

وكذلك يمكن تقدير ثابت الانحدار (b) من المعادلة الأتية: $\hat{b} = \bar{y} - a\bar{x}$

الخطأ المعياري لتقدير خط انحدار (Standard Error):

يمكن تعريف الخطأ المعياري لتقدير خط الاتجاه العام بأنه قيمة الجذر التربيعي لتباين الأخطاء.

ويرمز لخطأ التقدير معادلة الاتجاه العام بـــــ ($S_{y|x}$) وهو يساوي إلى الجذر التربيعي لتباين متوسط مربعات الخطأ ويمكن حسابه كما يلي: $S_{y|x}=\sqrt{\frac{\sum(y-\hat{y})^2}{n-2}}$

ويقيس مدى انحرافات القيم الفعلية عن القيم المقدرة $(y_i-\hat{y}_i)$. فإذا كان هذا التقدير كبير دل على أن انحرافات القيم الفعلية للمتغير التابع عن القيم المقدرة لها كبير وبالتالي يكون النموذج غير كفؤ.

مثال(7): الجدول التالي يمثل سلسلة زمنية لتكلفة إنتاج إحدى السلع (بالألف ليرة).

السنة	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
التكلفة	15	16	18	18	20	21	22	24	28	27

والمطلوب الآتي:

- تقدير معادلة الاتجاه العام للسلسلة الزمنية.
- تقدير التكلفة لعامي 2010 ، 2020.

الحل:

باعتبار إن السلسلة الزمنية تحتوى على متغيرين أحدهما مستقل X ويمثل الأعوام (2000-2009)، والأخر تابع Y ويمثل تكلفة الإنتاج، يتم تقدير الاتجاه العام الخطى للسلسة الزمنية بالمعادلة التالية:

$$\hat{y}=ax+b$$

x	y	x y	X^2
0	15	0	0
1	16	16	1
2	18	34	4
3	18	54	9
4	20	80	16
5	21	105	25
6	22	132	36
7	24	168	49
8	28	224	64
9	27	234	81
45	209	1058	285

يمكن تقدير ميل الانحدار (a) من المعادلة الأتية:

$$\hat{a}=\frac{n\sum x_i y_i-\sum x_i\sum y_i}{n\sum x_i{}^2-(\sum x_i)^2}=\frac{10*1058-45*209}{10*285-(45)^2}=1.42$$

وكذلك يمكن تقدير ثابت الانحدار (b) من المعادلة الأتية:

$$\hat{b}=\bar{y}-a\bar{x}=20.9-1.42*4.5=14.5$$

بالتعويض نحصل علي معادلة الاتجاه العام المقدرة

$$\hat{y}=ax+b=1.42x+14.5$$

معادلة خط الاتجاه العام يمكن استخدامها للتنبؤ بالتكلفة للسنوات القادمة.

ويمكن حساب معادلة الاتجاه العام بطريقة الانحرافات كما يلي:

X	Y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(x-\bar{x}) \times (y-\bar{y})$	$\hat{y}_i$	$\varepsilon_i = (y_i - \hat{y}_i)$	ε_i^2
0	**15**	-4.5	-5.9	20.25	26.55	14.50	-0.50	0.25
1	**16**	-3.5	-4.9	12.25	17.15	15.92	-0.08	0.01
2	**18**	-2.5	-2.9	6.25	7.25	17.35	-0.66	0.43
3	**18**	-1.5	-2.9	2.25	4.35	18.77	0.77	0.59
4	**20**	-0.5	-0.9	0.25	0.45	20.19	0.19	0.04
5	**21**	0.5	0.1	0.25	0.05	21.61	0.61	0.37
6	**22**	1.5	1.1	2.25	1.65	23.03	1.03	1.07
7	**24**	2.5	3.1	6.25	7.75	24.46	0.45	0.21
8	**28**	3.5	7.1	12.25	24.85	25.88	-2.12	4.51
9	**27**	4.5	6.1	20.25	27.45	27.30	0.30	0.09
45	209			82.5	117.5		0	7.55

تقدير ميل الانحدار $(\hat{a})$:

$$(\hat{a}) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{117.5}{82.5} = 1.422$$

ويمكن تقدير ثابت الانحدار $(\hat{b})$ من المعادلة الأتية:

$$\hat{b} = \overline{Y} - \hat{a}\overline{X} = 20.9 - 1.42 \times 4.5 = 14.5$$

وتكون معادلة الانحدار $Y = 1.42X + 14.5$

- ولتقدير التكلفة لعامي 2010 ، 2020، يتم حساب قيمة x للعامين بطرح كل عام من سنة الاساس، وعلي فرض أن سنة الأساس 2000، تصبح قيمة x للعامين 10 و 20 علي التوالي، وبالتعويض في معادلة الاتجاه، نحصل علي تقدير التكلفة بالألف ليرة كالآتي:

$$\hat{y} = ax + b = 1.42 * (10) + 14.5 = 28.7$$
$$\hat{y} = ax + b = 1.42 * (20) + 14.5 = 42.9$$

الاتجاه العام غير الخطي

قد نواجه حالات مغايرة للاتجاه الخطي عند وصف التغيرات للسلسلة بحيث لا يمكن معها استخدام معادلة الاتجاه الخطي خاصة مع الظواهر الاقتصادية التي تتصف بالتغير على الأمد الطويل، حينئذ نستخدم

معادلة غير خطية مناسبة لقياس منحنى الاتجاه . ومن أهم هذه الطرق:

معادلة الاتجاه العام الآسي:

إذا كانت الظاهرة تزيد (أو تنقص) بمعدل ثابت كل فترة زمنية فان معادلة الاتجاه العام تأخذ الشكل

$$\hat{y} = \hat{a}\,\hat{b}^{x}$$

نأخذ لوغاريتم الطرفين فتصبح المعادلة $\log \hat{y} = \log \hat{a} + x \log \hat{b}$:

وهي معادلة خطية وبنفس الاسلوب السابق نقدر قيمة $\log \hat{b}$ وقيمة $\log \hat{a}$ ومن ثم نوجد قيمة $\hat{a}$ و $\hat{b}$ ونحصل على معادلة الاتجاه الأسية التي يمكن استخدامها في التنبؤ.

معادلة الاتجاه التربيعي

إذا دل التمثيل البياني للسلسلة البيانية على وجود علاقة منحنية من الدرجة الثانية مثلًا (قطعاً مكافئاً) ، فإن معادلة الاتجاه العام تكون على صورة معادلة من الدرجة الثانية في متغير واحد وتأخذ الصورة التالية

$$\hat{y} = \hat{a}x + \hat{c}\,x^{2} + \hat{b}$$

وباستخدام طريقة المربعات الصغرى يتم تقدير كل من (a, b, c) وبالتعويض في المعادلة السابقة نحصل على معادلة الاتجاه العام التي يمكن استخدامها في التنبؤ بالقيم المستقبلية للظاهرة.

---------- **انتهى الفصل** ----------

الفصل الثالث

الانحدار (Regression)

الفصل الثالث

الانحدار (Regression)

مقدمة

من المعروف أن الارتباط الخطى بين متغيرين يقيس العلاقة الغير سببيه بينهما ليتعرف على درجتها واتجاهها، إن كلمه العلاقة الغير سببيه تعنى بها أن الارتباط لا يهتم بتصنيف المتغيرات إلى متغيرات مستقلة ومتغيرات تابعه ولا تهتم بمعرفة أي من المتغيرات يؤثر ومن منها يتأثر مثل:

1- سعر السلعة، والكمية المطلوبة منها

2- عدد مرات ممارسة نوع معين من الرياضة البدنية، ومستوى الكولسترول في الدم.

3- وزن الجسم، وضغط الدم.

أما إذا كان الاهتمام بدراسة أثر أحد المتغيرين على الآخر نستخدم أسلوب تحليل الانحدار، حيث أن تحليل الانحدار يحدد شكل تغير العلاقة الارتباطية بين متغيرين بواسطة معادلة تسمى معادلة الانحدار ومن الأمثلة على ذلك:

1- الاستهلاك والدخل.

2- كميات السماد المستخدمة، وكمية الإنتاج من القمح المسمد من هذا النوع من السماد.

الارتباط و الانحدار

- الارتباط : مقياس يحدد فيما إذا كانت هناك علاقة بين المتغيرين Xو Y
- الانحدار: إذا كانت هناك علاقة ارتباط بين المتغيرين Xو Y كيف يمكن أن نعبر عنها بمعادلة. أي أن الانحدار هو التمثيل الرياضي للعلاقة الارتباطية بين المتغيرين (X و Y) بمعادلة تسمى معادلة الانحدار.

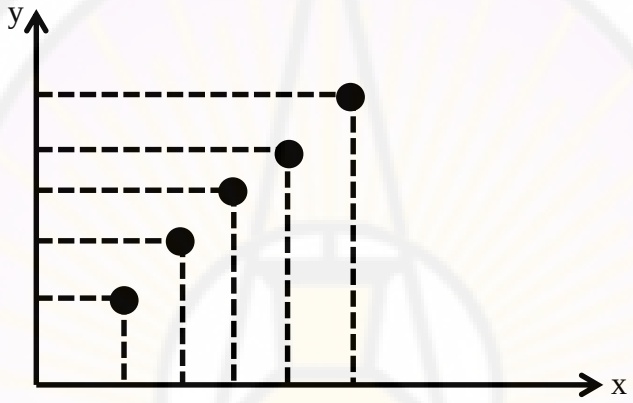
← تحليل الارتباط البسيط: هو تحديد نوع وقوة العلاقة الارتباطية بين متغيرين X و Y.

← تحليل الانحدار الخطي البسيط: هو دراسة وتحليل أثر المتغير المستقل (x) على المتغير التابع (y).

في الارتباط تتم دراسة العلاقة بين المتغيرين بحساب معامل الارتباط و تحديد حالته (موجب أو سالب) ودرجة قوة الارتباط (قوية جداً ،قوية، متوسطة، ضعيفة، ضعيفة جداً)

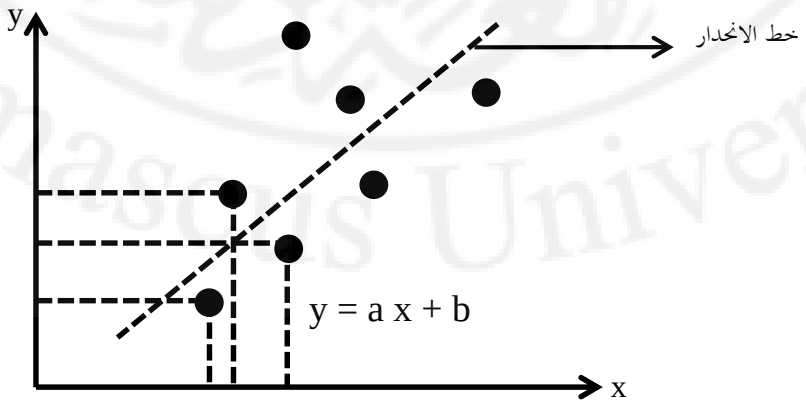
أما الانحدار هو تمثيل نقاط الانتشار التي تمثل الارتباط بين المتغيرين (X و Y) على شكل معادلة رياضية (معادلة الانحدار)، الهدف الأساسي من هذه المعادلة هو التنبؤ بقيمة المتغير التابع (Y) بمعلومية المتغير المستقل (X).

يتمثل الارتباط بيانياً بانتشار نقاط تعبر عن العلاقة الارتباطية بين المتغيرين (x, y) كما يلي:



أما التمثيل البياني للانحدار فهو ايجاد خط الانحدار (مستقيم) الذي يمر بأغلب نقاط الانتشار من العلاقة الارتباطية بين المتغيرين (X) و (Y).

أي انه تمثيل العلاقة الارتباطية بين المتغيرين (X) و (Y) بمعادلة مستقيم تسمى معادلة الانحدار وتأخذ الشكل y=ax+b



وتستخدم معادلة الانحدار للتنبؤ بقيم المتغير التابع (Y) بدلالة المتغير المستقل (X).

الفرق بين الارتباط و الانحدار

أوجه الاختلاف

- الارتباط يحدد درجة ارتباط المتغيرين دون أن يوضح وجود أي علاقة سببية بينهما أي لا يوضح أي المتغيرات تابع وأيهما مستقل ونظراً لان معامل الارتباط يتحدد في قيمة واحدة فهو لا يساعدنا على التنبؤ بقيمة أي متغير بدلالة الأخر.
- أما الانحدار يفترض وجود علاقة سببية بين المتغيرين محل البحث ويوضح أيهما المتغير التابع وأيهما المستقل ومن ثم يمكن التنبؤ بقيم المتغير التابع بدلالة المتغير المستقل باستخدام العلاقة المقدرة (كمية القمح مع كمية السماد).

أوجه الاتفاق

معامل الانحدار الخطي ومعامل الارتباط الخطي يتماثلان بالإشارة:

1. فإذا كان معامل الارتباط (r) موجباً بين متغيرين فلابد أن يكون معامل الانحدار (a) موجباً
2. وإذا كان معامل الارتباط سالباً فلابد أن يكون معامل الانحدار سالباً
3. إذا كان معامل الارتباط =0 فلابد أن يكون معامل الانحدار مساوياً للصفر.

الهدف من تحليل وقياس الانحدار

- قياس الانحدار يهدف إلى ايجاد الشكل الرياضي للعلاقة التي تربط المتغير المستقل (X) بالمتغير التابع (Y) وهذه العلاقة تسمى معادلة الانحدار
- يهدف إلى ايجاد أثر المتغير المستقل (X) على المتغير التابع (Y).
- الهدف الأساسي من تحليل الانحدار هو التنبؤ بقيمة المتغير التابع(Y) بمعلومية المتغير المستقل (X) بواسطة معادلة الانحدار.

مثال: من المتوقع أن يصل عدد سكان الصين في عام 2025 حوالي 3 مليار ومن المتوقع ان يبلغ إنتاج القمح إلى 4 مليون طن في سوريا عام 2020.

أنواع نماذج الانحدار (Regression Models)

يوجد أربعة أنواع رئيسيه من نماذج الانحدار Regression Models كل منها يمثل طريقه معينه لمعالجة مشكلة الانحدار.

1) الانحدار الخطي البسيط Simple Linear Regression تشير تسمية هذا المعامل "بسيط" الى أنه يتضمن متغير تابع y يعتمد على متغير واحد مستقل x وكلمة خطي تشير الى أن العلاقة ين المتغيرين y و x هي علاقة خطية. $Y = \alpha + \beta X$

2) الإنحدار الخطي المتعدد Multiple Linear Regression هذا النوع من الانحدار يتضمن اعتماد المتغير التابع y على أكثر من متغير مستقل مثل x1 و x2 ... الخ.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

3) الانحدار غير الخطي البسيط Simple non-Linear Regression إذا كانت العلاقة بين المتغير y والمتغير المستقل غير خطية مثل علاقة أسية أو لوغاريتمية أو تربيعية ... الخ. $Y = \alpha + \beta_1 X^2$

4) الانحدار غير الخطي المتعدد (Non-Linear Multiple Regression)

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2^2 + \beta_3 X_3^3$$

وهنالك أنواع أخرى مثل الانحدار الهرمي Hierarchical Regression والإنحدار التدريجي Stepwise Regression وغيرها.

نموذج الانحدار الخطي البسيط (Simple Linear Regression)

الانحدار الخطي البسيط والذي يسمى التنبؤ Prediction وهو تقدير القيمة المستقبلية لمتغير واحد (المتغير التابع) بناءً على معرفة قيم متغير أو أكثر (متغيرات مستقلة).

تهدف دراسة الانحدار الخطي البسيط إلى التنبؤ بقيمة المتغير التابع **(dependent) ويرمز له بالرمز**(Y) بدلالة وبمعلومية المتغير المستقل **(Independent) ويرمز له بالرمز**(x).

- إن المتغير التابع (Y) هو متغير غير المسيطر عليه والمعروف أيضاً باسم متغير الاستجابة
- إن المتغير المستقل (X) هو متغير مسيطر عليه.

النموذج الرياضي للانحدار الخطي البسيط:

إن العلاقة الخطية بين المتغير التابع Y والمتغير المستقل X هي علاقة خطية تأخذ الشكل التالي:

$$Y = ax + b$$

حيث أن:

Y: هو المتغير التابع (الذي يتأثر).

X: هو المتغير المستقل (الذي يؤثر).

معاملات نموذج الانحدار الخطي البسيط

إن لنموذج الانحدار الخطي البسيط معاملان هما:

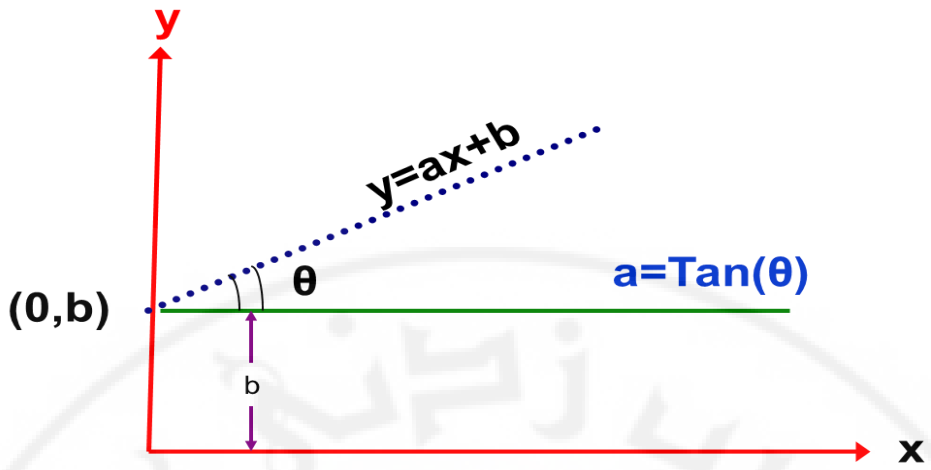
- **المعامل (a):** وهو يمثل ميل الانحدار
- **المعامل (b):** وهو يمثل ثابت الانحدار

a: ميل الانحدار: وهو يمثل ميل خط الانحدار، ويعكس مقدار التغير في y إذا تغيرت x بمقدار وحدة واحدة. والميل هو tan أي الزاوية التي يحصرها المستقيم مع الاتجاه الموجب للمحور x.

b: ثابت الانحدار : وهي تمثل قيمة الجزء المقطوع من المحور الرأسي (y) بالمستقيم(y = ax+b) أو تمثل قيمة التابع في حالة انعدام قيمة المتغير المستقل x أي في حالة x=0.

عندما $x = 0 \Leftarrow y = b$

هو الجزء المقطوع من المحور الرأسي y، وهو يعكس قيمة المتغير التابع في حالة انعدام قيمة المتغير المستقل x، أي في حالة $x = 0$ أي عندما $x = 0 \Leftarrow y = b$

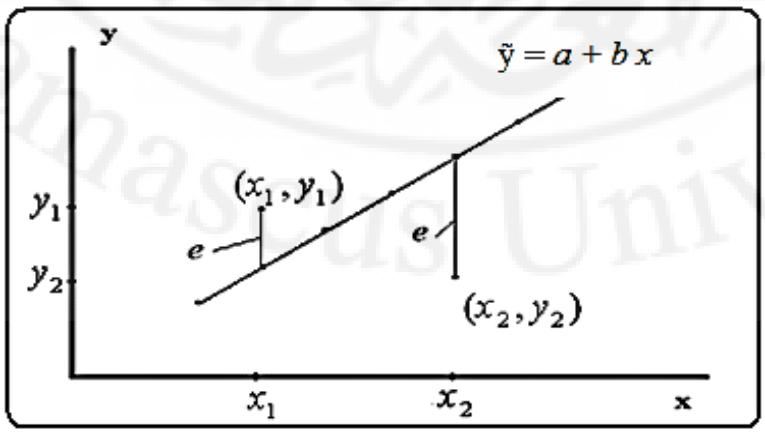


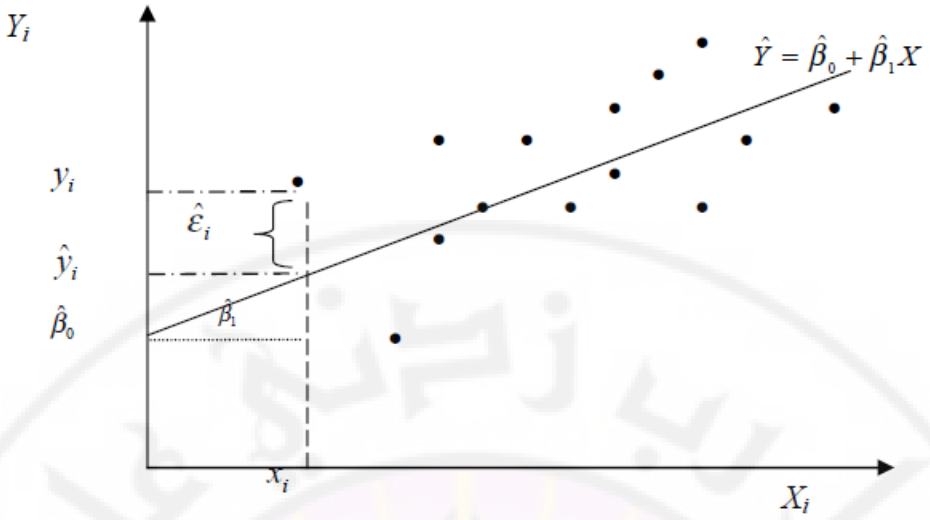
النموذج القياسي للانحدار الخطي البسيط:

نلاحظ ان الصيغة الرياضية للنموذج الخطي البسيط تحتوي على متغيرات منتظمة فقط، وهذا يعني اننا ننظر للعلاقة بين متغيرات النموذج باعتبارها علاقة تامة (غير احتمالية) حيث ان المتغير المستقل هو المتغير الوحيد المؤثر في المتغير التابع. ولكن في واقع الامر فان العلاقة ليست بهذه الصورة اذ يتعين ان يؤخذ في الحسبان أثر المتغيرات العشوائية بجانب المتغيرات المنتظمة. وذلك بإضافة حد عشوائي للصيغة الرياضية او ما يسمى بالخطأ العشوائي ويرمز له بالرمز e وعندها يتحول النموذج من نموذج رياضي الى نموذج قياسي يعكس علاقة احتمالية. وبالتالي يمكن صياغة معادلة الانحدار الحقيقي للنموذج الخطي البسيط قياسيا على النحو التالي: $Y=aX+b+e$

حيث أن:

e : الخطأ العشوائي، والذي يعبر عن الفرق بين القيمة الفعلية y ، والقيمة المقدرة $\hat{y} = \hat{a}x + \hat{b}$ ، أي أن : $e = y - \hat{y} = y - (\hat{a}x + \hat{b})$ ، ويمكن توضيح هذا الخطأ على الشكل التالي لنقط الانتشار.





ويلاحظ من الشكل السابق اذا كان الخطأ العشوائي مساوي الى الصفر فان القيم المشاهدة سوف تنطبق على الخط المستقيم, بمعنى اننا نستطيع تفسير المتغير التابع (Y) تفسيرا كاملا بالأثر الخطي للمتغير المستقل (X).

أسباب اضافة الخطأ العشوائي لنموذج الانحدار الخطي البسيط

حد الخطأ (e) أو العنصر العشوائي: يقيس انحراف القيمة المشاهدة عن خط الانحدار الحقيقي غير المشاهد. وتنشأ هذه الانحرافات التي تشكل الخطأ العشوائي للأسباب التالية:

1- وجود عدة متغيرات ذات تأثير ضئيل أو غير منتظم على (Y) وقد استبعدت من العلاقة الخطية التامة.

2- الأخطاء الراجعة لتجميع البيانات.

3- حدوث أخطاء ممكنة عند قياس المتغيرات المكونة للنموذج محل الدراسة عند اخذ المشاهدات.

4- السلوك الإنساني العشوائي.

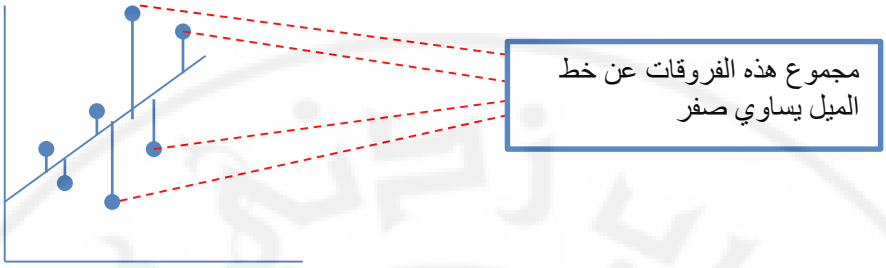
أسلوب تحليل الانحدار

الأساس العلمي الذي بني عليه أسلوب تحليل الانحدار يعتمد على طريقة المربعات الصغرى **(OLS)** Ordinary Least Squares وتتلخص هذه الطريقة في أن تجعل مجموع مربعات الأخطاء أقل ما يمكن.

تفترض أن معادلة الانحدار الخطي البسيط بين متغيرين أحدهما تابع والآخر مستقل على الصورة التالية:

$$y = ax + b + e$$

يبقى أن نعرف كيف نقدر (a) و (b) لذلك يجب أن يختفي الخطأ العشوائي(e) من المعادلة ، وذلك من خلال طريقة المربعات الصغرى التي تجعل مجموع مربعات الأخطاء أقل ما يمكن (أي صفر) كما في الشكل التالي:



لتصبح المعادلة هكذا: $\hat{y} = \hat{a}x + \hat{b}$

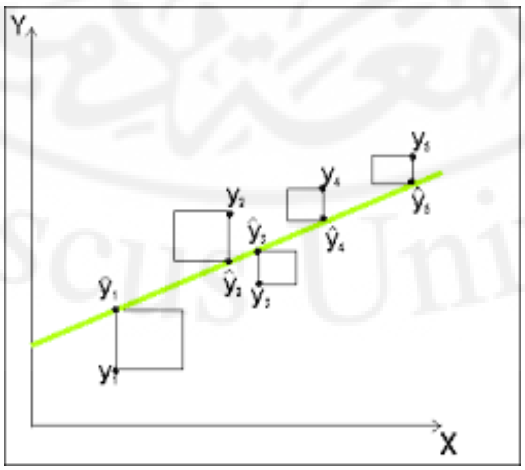
تقدير قيم معاملات الانحدار (a, b) باستخدام طريقة المربعات الصغرى (OLS)

هنالك عدة طرق لتقدير أو حساب نموذج الانحدار الخطي البسيط وكل الطرق تعتمد على حساب قيم معاملات الانحدار (a و b)، وتعد طريقة المربعات الصغرى (Ordinary Least Squares) من أفضل الطرق لأنها تجعل مجموع مربعات الأخطاء العشوائية أقل ما يمكن.

طريقة المربعات الصغرى:

استند العالم الألماني جاوس على ما يسمى بطريقة المربعات الصغيرة العادية لتقدير معالم النموذج المجهولة. وهي أسلوب لتوفيق (أفضل) خط مستقيم لعينة مشاهدات Y و X.

ففي الشكل التالي يكون خط الانحدار الحقيقي والخاص بالمجتمع مجهول المكان والشكل ولذلك نحاول تقدير خط آخر من العينة.



ويرمز لهذا الخط بـ $\hat{Y}_i = \hat{a}\,X_i + \hat{b}$

$\widehat{Y}_i$: القيمة المقدرة المناظرة للمشاهدة الفعلية Y_i

$\hat{b}$: القيمة المقدرة للمعلمة الحقيقية المجهولة b

$\hat{a}$: القيمة المقدرة للمعلمة الحقيقية المجهولة a

ويمكن الحصول على قيم $\hat{b}$ و $\hat{a}$ بطريقة المربعات الصغيرة العادية.

من الشكل يمكننا تعريف البواقي (e_i) على أنه يساوي:

e_i = القيم الحقيقية للمشاهدة Y_i – القيم المقدرة للمشاهدة $\widehat{Y}_i$ أي $\left(e_i = Y_i - \widehat{Y}_i\right)$

و يمكن للبواقي أن تكون سالبة أو موجبة حسب موضع نقطة مشاهدة من الخط المقدر.

وتعطي طريقة المربعات الصغيرة العادية أفضل خط مستقيم يوفق مشاهدات العينة X ,Y بمعنى أنها تعطي أقل مجموع مربعات رأسية لانحرافات كل مشاهدة عن الخط المستقيم في الرسم ويتضمن ذلك تصغير مجموع المربعات لانحرافات النقاط الرأسية على الخط لأدنى حد ممكن أي:

وبما أن: $min \sum_{i=1}^{n} e_i^2$

$$min \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$= \sum_{i=1}^{n} (Y_i - \hat{b} - \hat{a}X_i)^2$$

تبعاً لمبدء المربعات الصغيرة يجري أختيار قيم $\hat{b}$ و $\hat{a}$ وتحديد الخط المقدر $(\hat{Y} = \hat{b} + \hat{a}Xi)$ بحيث يتم تصغير مجموع مربع حد الخطأ $(\sum_{i=1}^{n} e_i^2)$ إلى أدنى حد ممكن وتعتبر هذه مسألة في تحديد النهاية الصغرى لدالة $(\sum_{i=1}^{n} e_i^2)$ حيث يتم ايجلد الشرط الضروري وهو مساواة التفاضل الجزئي الأول للدالة بالنسبة لـ $\hat{b}$ و $\hat{a}$ بالصفر، ثم نحصل على معادلتين آنيتين يتم حلها آنيا للحصول على القيم المجهولة ($\hat{a}$ و $\hat{b}$).

وعادة ما ناخذ الانحرافات الرأسية لأننا نحاول أن نفسر ونتنبأ بالتغيرات في Y التي تقيس على المحور الرأسي، ولا يمكن أخذ مجموع الانحرافات لكل مشاهدة عن خط المربعات الصغرى OLS لأن

الإنحرافات المتساوية في الحجم والمختلفة في الاشارة يلغي بعضها البعض فيكون مجموع الانحرافات مساوياً للصفر $\sum_{i=1}^{n} e_i = 0$

تبعاً لمبدء المربعات الصغيرة يجري أختيار قيم $\hat{b}$ و $\hat{a}$ بحيث تصغير المجموع إلى $\sum_{i=1}^{n} e_i^2$ إلى أدنى حد ممكن وهذه مسألة في تحديد النهايات الصغرى للدالة: $\sum_{i=1}^{n} e_i^2$

$$min \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$min \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{b} - \hat{a}X_i)^2$$

وبعد أجراء العديد من عمليات الاشتقاق والتفاضل يمكن تقدير ميل الانحدار (a) بإحدى المعادلتين التاليتين أيضاً:

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

$$\hat{a} = \frac{\sum yx - n\bar{y}\,\bar{x}}{\sum x_i^2 - \frac{(\sum x)^2}{n}}$$

أو يمكن حساب ميل الانحدار بالمعادلة التالي:

$$\hat{a} = \frac{\sum (y-\bar{y})(x-\bar{x})}{\sum (x-\bar{x})^2}$$

وكذلك يمكن تقدير ثابت الانحدار (b) من المعادلة الأتية:

$$\hat{b} = \bar{y} - a\bar{x}$$

الخطأ المعياري لتقدير خط انحدار (Standard Error):

يمكن تعريف الخطأ المعياري لتقدير خط انحدار المتغير التابع (Y) على المتغير المستقل (X) بأنه قيمة الجذر التربيعي لتباين الأخطاء.

ويمكن توضيح مفهوم الخطأ المعياري بشكل أكثر تفصيلاً، إنه في العلاقة الارتباطية بين المتغيرين (X) و(Y) يمكن أن يمر عدد من خطوط الانحدار، لذلك عند اختيار خط انحدار معين، وبســـبب أن خط

الانحدار لا يمر بكل النقاط حيث يوجد هناك نقاط مشـتتة حوله فإن اختيار خط انحدار معين يتم بخطأ تقدير يسمى خطأ التقدير (Standard Error).

ويرمز لخطأ التقدير في حالة معادلة انحدار المتغير التابع (Y) على المتغير المستقل (X) بـــ ($S_{y|x}$) وهو يساوي إلى الجذر التربيعي لتباين متوسط مربعات الخطأ ويمكن حسابه كما يلي:

$$S_{y|x} = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}}$$

ويقيس مدى انحرافات القيم الفعلية عن القيم المقدرة $(y_i - \hat{y}_i)$. فإذا كان هذا التقدير كبير دل ذلك إلى أن انحرافات القيم الفعلية للمتغير التابع عن القيم المقدرة لها كبير وبالتالي يكون النموذج غير كفؤ.

مؤشرات جودة نموذج الانحدار الخطي البسيط

يمكن استخلاص مؤشرين لجودة النموذج وهما:

1- معامل التحديد (Determination Coefficient): هو مؤشـر يسـاوي مربع معامل الارتباط الخطي البسـيط ويرمز له بالرمز (R^2) ويعبر عن نسـبة مجموع مربعات الانحدار (SSR) إلى مجموع المربعات الكلي (SST) أي أن $R^2 = \frac{SSR}{SSTO}$. وهذا المعامل يدل على مدى قوة العلاقة بين القيم المقدرة و القيم الفعلية. فإذا كان هذا المعامل كبير دل ذلك على جودة النموذج، لأن معنى ذلك أن المتغير المسـتقل (المفسـر) يؤثر بنسـبة كبيرة من التغيرات الكلية التي تحدث في المتغير التابع.

أي يستخدم معامل التحديد (R^2) كمقياس يحدد القوة التفسيرية لنموذج الانحدار الخطي حيث يشير إلى النسبة المئوية للتغير الكلي في المتغير التابع (Y) والتي يمكن تفسيرها بواسطة المتغير المستقل (X_1). اذاً معامل التحديد(R^2) يقيس نسـبة تأثير المتغير المسـتقل(X_1) على المتغير التابع (Y)، بمعنى أخر يقيس نسـبة مسـاهمة المتغير المسـتقل في تغير المتغير التابع، فهو مقياس للقدرة التفسـيرية للنموذج أي يختبر جودة التوفيق والارتباط.

عند تقدير معادلة نموذج الانحدار يكون من الضـروري اسـتخدام معامل التحديد (R^2) لمعرفة مدى جودة توفيق خط الانحدار المقدر لكل من المتغير التابع والمتغير المسـتقل، لذلك فإن اسـتخدام معامل التحديد

(R^2) الذي له علاقة بانحراف المشـــاهدات عن أفضـــل تقدير وتتراوح قيمته بين الصـــفر والواحد $(0 \ll R^2 \leq 1)$

أي عندما يأخذ معامل التحديد (R^2) أكبر قيمة له وهي (1) عندها تقع كل نقاط المشـــاهدات (x_i, y_i) على خط الانحدار المقدر $\hat{y} = \hat{a}x + \hat{b}$ فتكون القدرة التفسـيرية للنموذج عالية جداً، أي هناك جودة في التوفيق والارتباط بين المتغير التابع والمستقل.

أما أذا كان (R^2) يأخذ أصـــغر قيمة له وهي (0) فليس هناك جودة في التوفيق والارتباط بين المتغيرين التابع والمستقل أي ليس للنموذج قدرة تفسيرية على الاطلاق، ويعود ذلك إلى سببين إما العلاقة الموجودة بين المتغيرين غير خطية أو غياب السببية بينهما.

2- المؤشـــر الثاني وهو تباين الخطأ العشـــوائي وتقديره هو $\hat{\sigma}^2 = MSE = \frac{SSE}{n-2}$ ويقيس مدى انحرافات القيم الفعلية عن القيم المقدرة $(y_i - \hat{y}_i)$. فإذا كان هذا التقدير كبير دل ذلك إلى أن انحرافات القيم الفعلية للمتغير التابع عن القيم المقدرة لها كبير وبالتالي يكون النموذج غير كفؤ.

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

مثال (1):

المتغير (X) يمثل كميات الامطار والمتغير (Y) يمثل كمية التفاح مدرجة بالجدول التالي:

كمية الأمطار (X)	2	3	1	5	9	المجموع
كمية التفاح (Y)	4	7	3	9	17	40

المطلوب:

- احسب معامل الارتباط الخطي بين كمية الامطار وكمية التفاح.
- قدر معادلة انحدار كمية التفاح (Y) على كمية الامطار(X).
- فسر معادلة الانحدار.
- احسب الخطأ المعياري للتقدير.
- اوجد معامل التحديد.

الحل:

نشكل الجدول التالي:

X	Y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x}) \times (y-\bar{y})$	X^2	$\hat{y}_i$	$\varepsilon_i = (y_i - \hat{y}_i)$	$(\hat{y}_i - \bar{y})^2$	ε_i^2
2	**4**	-2	-4	4	16	8	4	4.5	0.5	12.25	0.25
3	**7**	-1	-1	1	1	1	9	6.25	-0.75	3.06	0.56
1	**3**	-3	-5	9	25	15	1	2.75	-0.25	27.56	0.06
5	**9**	1	1	1	1	1	25	9.75	0.75	3.06	0.56
9	**17**	5	9	25	81	45	81	16.75	-0.25	76.56	0.06
20	40	0	0	40	124	70	120	40	0	122.5	1.5

$\bar{x} = 4$ $\bar{y} = 8$

حساب معامل الارتباط:

$$r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}} = \frac{70}{\sqrt{40*124}} = \frac{70}{70.124} = 0.994$$

ارتباط موجب قوي جداً بين كمية الامطار وكمية التفاح.

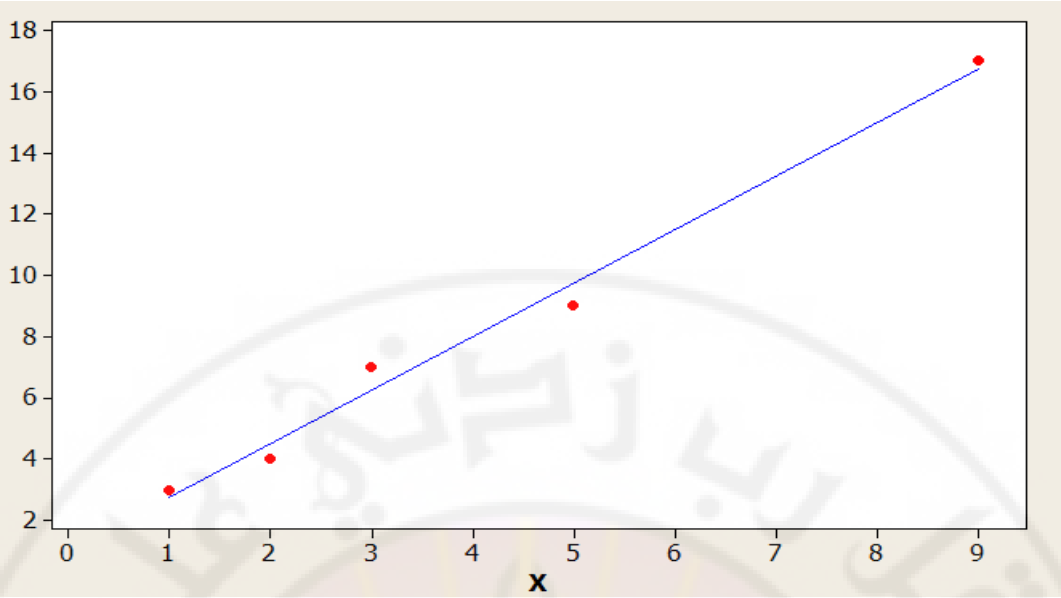
تقدير معادلة انحدار

$$\hat{a} = \frac{\sum (y-\bar{y})(x-\bar{x})}{\sum (x-\bar{x})^2} = \frac{70}{40} = 1.75$$

$$\hat{b} = \bar{y} - a\bar{x} = 8 - 1.75*4 = 1$$

وتكون معادلة الانحدار $Y = 1.75X + 1$

ويمكن رسم بيانات التمرين وخط معادلة الانحدار Y على X كما يلي:



تفسير معادلة انحدار

يتضـح من الشـكل السـابق أن خط الانحدار لا يمر بكل النقاط حيث يوجد هناك نقاط مشـتتة حول خط الانحدار. وبالرغم من ذلك يعد هذا الخط من أفضل الخطوط التي حصلنا عليها للتعبير عن العلاقة بين المتغيرين ولكن بخطأ معين يسمى خطأ التقدير (Standard Error).

الخطأ المعياري للتقدير:

ويرمز لخطأ التقدير في حالة معادلة الانحدار Y على X بـــــ($S_{y|x}$) وهو يشـير إلى الجذر التربيعي لمتوسط مربعات الخطأ ويمكن حسابه من الجدول السابق كما يلي:

$$S_{y|x}=\sqrt{\frac{\sum(y_i-\hat{y}_i)^2}{n-2}}=\sqrt{\frac{\sum\varepsilon_i^{\,2}}{n-2}}=\sqrt{\frac{1.5}{3}}=0.707$$

رابعاً: حساب معامل التحديد:

بما أن معامل الارتباط يساوي $r=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}}=\frac{70}{70.427}=0.993$

و بما أن معامل التحديد يساوي مربع معامل الارتباط فإن معامل التحديد

$$r^2=(0.993)^2=0.986$$

مثال (2): فيما يلي بيانات عن عدد ساعات الدراسة والعلامة التي يحصل عليها 5 طلاب في مقرر الاحصاء.

اسم الطالب	ريم	حسين	تامر	خالد	نور
ساعات الدراسة (X)	3	4	5	8	10
الدرجة في الاحصاء (Y)	8	12	15	18	22

المطلوب:

1- احسب قيمة معامل الارتباط بين عدد ساعات الدراسة والدرجات في الاحصاء.

2- ارسم نقاط الانتشار (الشكل الانتشاري) للعلاقة بين (X) و (Y).

3- احسب قيمة معامل التحديد واشرح معناه.

4- قدّر معادلة انحدار (y)على (x) [معادلة التنبؤ] $\hat{y} = \hat{a}(x) + \hat{b}$ مع التفسير.

5- ارسم لشكل الانتشاري للعلاقة بين المتغيرين، مع رسم خط ميل الانحدار.

6- استخدم معادلة الانحدار في التنبؤ بقيمة العلامة عند عدد ساعات 15.

7- احسب الخطأ المعياري للتقدير.

الحل:

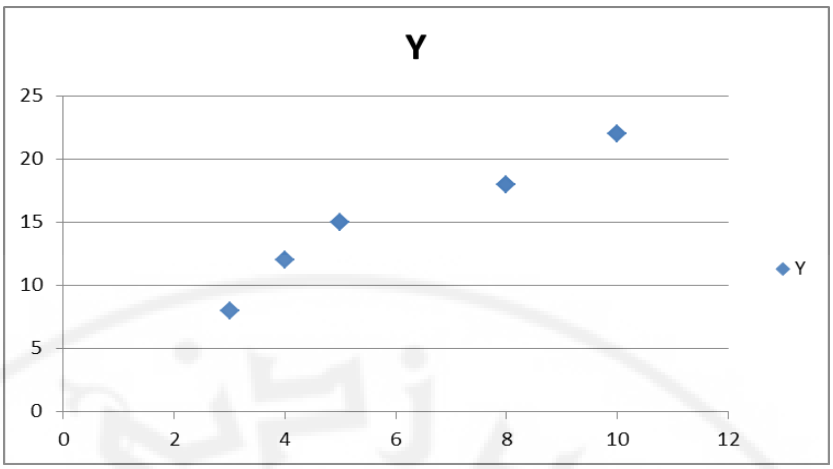
X	**Y**	**X^2**	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x}) \times (y-\bar{y})$	$\hat{y}_i$	$\varepsilon_i = (y_i - \hat{y}_i)$	ε_i^2	$(\hat{y} - \bar{y})^2$
3	8	9	-3	-7	9	49	21	9.61	1.6	2.56	28.98
4	12	16	-2	-3	4	9	6	11.41	-0.61	0.372	12.89
5	15	25	-1	0	1	0	0	13.20	-1.82	3.312	3.22
8	18	64	2	3	4	9	6	18.59	0.55	0.303	12.89
10	22	100	4	7	16	49	28	22.18	0.13	0.017	51.45
30	75	214	0	0	34	116	61	74.99	-0.15	6.564	109.44

1- معامل الارتباط

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{61}{\sqrt{34*116}} = \frac{61}{62.8} = 0.971$$

ارتباط موجب قوي جداً بين عدد ساعات الدراسة وعلامات الامتحان.

2- نقاط الانتشار للعلاقة بين (X) و (Y) ممثلة بالشكل التالي:



3- معامل التحديد (R^2) هو مربع معامل الارتباط ويساوي (0.943) وتفسيره ان المتغير (X) الذي يمثل عدد ساعات الدراسة تؤثر بنسبة (0.94) على علامة الطالب وهناك (0.06) لعوامل اخرى.

4- قدّر معادلة انحدار (y)على (x) [معادلة التنبؤ] $\hat{y} = \hat{a}(x) + \hat{b}$ مع التفسير

$$\hat{a} = \frac{\sum(y-\bar{y})(x-\bar{x})}{\sum(x-\bar{x})^2} = \frac{61}{34} = 1.79$$

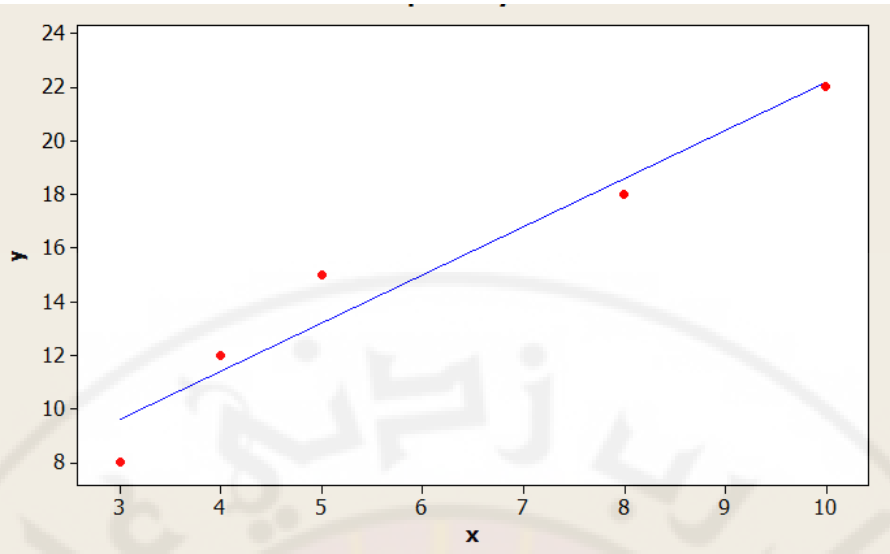
$$\hat{b} = \bar{y} - a\bar{x} = 15 - 1.79 * 6 = 4.23$$

معادلة الانحدار Y=1.79X+4.23

تفسير معادلة الانحدار: كلما زاد (x) ساعات الدراسة بمقدار ساعة واحدة تزداد علامات الطالب بمقدار (1.79) علامة. وعندما يكون (x=0) أي أنه عندما لا يكون ولا ساعة دراسة تكون علامة الطالب = 4.23.

5- الرسم للشكل الانتشاري للعلاقة بين المتغيرين، مع رسم خط ميل الانحدار.

يمكن رسم بيانات التمرين وخط معادلة الانحدار Y على X كما يلي:



6- استخدم معادلة الانحدار في التنبؤ بقيمة العلامة عند عدد ساعات (x=15)

من معادلة الانحدار Y=1.79X+4.23

Y=1.79X+4.23=1.79*15+4.23=31.08

7- احسب الخطأ المعياري للتقدير

ويرمز لخطأ التقدير في حالة معادلة الانحدار Y على X ب ($S_{y|x}$) وهو يشـــير إلى الجذر التربيعي لمتوسط مربعات الخطأ ويمكن حسابه من الجدول السابق كما يلي:

$$S_{y|x}=\sqrt{\frac{\sum(y_i-\hat{y}_i)^2}{n-2}}=\sqrt{\frac{\sum\varepsilon_i^{\,2}}{3}}=\sqrt{\frac{6.564}{3}}=1.479$$

ثانياً: الانحدار غير الخطي البسيط (Simple non-Linear Regression)

ليست كل الحالات تعكس حالة معامل الانحدار البسيط أن يتوافق مع الدالة الخطية (المستقيمة), حيث تكون العلاقة مستقيمة طردية أو عكسية بين العاملين التابع والمتغير، حيث تبدأ العلاقة طردية ثم تأخذ العلاقة الشكل العكسي لبدايتها . وعندها تصبح الدالة غير خطية وتأخذ شكل القطع المكافئ. ويسمى معامل الانحدار في هذه الحالة بغير المستقيم ومعادلة حسابه هي :

$$\mathbf{y = c + b\,x + a\,x^2}$$

والمثال التالي يبين كيفية حساب معامل الانحدار البسيط غير المستقيم .

مثال 4:

سجلت بيانات من تسعة حقول عن تأثير الري على الإنتاج في الجدول التالي :

رقم الحقل	1	2	3	4	5	6	7	8	9	مج
X	1	0	2	3	4	5	6	7	8	36
Y	4	3	3	6	9	8	11	14	5	63

لحساب معامل الانحدار البسيط غير المستقيم ننشئ الجدول التالي لأن عملية الحساب ستتم وفق طريقة المربعات الصغرى :

الحقل	y	x	y^2	x^2	x^3	x^4	Y x	yx^2
1	4	1	16	1	1	1	4	4
2	3	0	9	0	0	0	0	0
3	3	2	9	4	8	16	6	12
4	6	3	36	9	27	81	18	54
5	9	4	81	16	64	256	36	144
6	8	5	64	25	125	625	40	200
7	11	6	121	36	216	1296	66	396
8	14	7	196	49	343	2401	98	686
9	5	8	25	64	512	4096	40	320
مج	**63**	**36**	**557**	**204**	**1296**	**8772**	**308**	**1816**

وهنا نحتاج إلى ثلاثة معادلات هي :

1 - $\sum Y = a \sum X^2 + b \sum X + n\,c$

2 - $\sum Y\,X = a \sum X^3 + b \sum X^2 + c \sum X$

3 - $\sum Y\,X^2 = a \sum X^4 + b \sum X^3 + c \sum X^2$

من الجدول السابق نستبدل الرموز بقيمها في كل معادلة وفق التالي :

نأخذ المعادلتين (1 و 2) ونستبدل الرموز بقيمها .

1 - 63 = 204 a + 36 b + 9 c

2 - 308 = 1296 a + 204 b + 36 c

نضرب طرفي المعادلة 1 ب 4 ونطرحها من المعادلة 2 فنحصل على المعادلة رقم 4 .

4- 56 = 480 a + 60 b + 0

ثم نأخذ المعادلتين (1 و 3) ونستبدل الرموز بقيمها

1 - 63 = 204 a + 36 b + 9 c

3 - 1816 = 8772 a +1296 b +204 c

ثم نضرب طرفي المعادلة (1) ب (68) والمعادلة (3) ب (3)، ثم نطرح فنحصل على المعادلة (5) .

$$5448 = 26316\,a + 3888\,b + 612\,c$$

$$4284 = 13872\,a + 2448\,b + 612\,c$$

5- $1164 = 12444\,a + 1440\,b$

نأخذ المعادلتين (4 و 5)

$$56 = 480\,a + 60\,b$$

$$1164 = 12444\,a + 1440\,b$$

ونضرب طرفي المعادلة (4) ب (24) ثم نطرح من (5) .

$$1344 = 11520\,a + 1440\,b$$

$$1164 = 12444\,a + 1440\,b$$

ومنه : $-180 = 924\,a$ أي أن : $\mathbf{a = -0.195}$

نستبدل (a) بقيمتها بالمعادلة (4) وفق ما يلي :

$$156 = 60\,b + 480 \times (-0.195)$$

ومنه : $(b) = 2.493$

ثم نستبدل كلا من (a و b) بقيمهما بالمعادلة (1) وفق ما يلي :

$$63 = 204 (- 0.195) + 36 (2.483) + 9 c$$

ومنه : $(c) = 1.448$

وعليه تصبح معادلة الانحدار غير المستقيم لهذه الظاهرة :

$$\mathbf{Y = - 0.195 \ X^2 + 2.493 \ X + 1.488}$$

الفصل الرابع

الانحدار الخطي المتعدد

(Multiple Linear Regression)

الانحدار الخطي المتعدد

4-1- تعريف و أهمية الانحدار الخطي المتعدد

4-2- الفروض والشروط الأساسية لنموذج الانحدار الخطى المتعدد

4-3- التوصيف الرياضي لنموذج الانحدار الخطى المتعدد

4-4- تقدير معاملات نموذج الانحدار الخطي المتعدد

4-4-1- طريقة المربعات الصغرى

4-5- الاختبارات المعنوية لمعاملات الانحدار

4-6- اختبار جودة نموذج الانحدار الخطي المتعدد

4-7- معامل التحديد المتعدد ومعامل التحديد المعدل

الانحدار الخطى المتعدد(Multiple Linear Regression (MLR

يعد الانحدار الخطي المتعدد او النموذج الخطي العام من الاساليب الاحصائية المتقدمة والتي تضمن دقة الاستدلال عن طريق الاستخدام الامثل للبيانات في ايجاد علاقات سببية بين الظواهر موضوع البحث، والانحدار الخطي المتعدد عبارة عن انحدار للمتغير التابع (Y) على العديد من المتغيرات المستقلة $(X_1, X_2 ..... X_k)$ لذا فهو يستخدم في التنبؤ بتغيرات المتغير التابع الذي يؤثر فيه عدة متغيرات مستقلة، والانحدار الخطي المتعدد ليس مجرد اسلوب واحد وانما مجموعة من الاساليب التي يمكن استخدامها لمعرفة العلاقة بين متغير تابع مستمر وعدد من المتغيرات المستقلة التي عادة ما تكون مستمرة.

4-1- تعريف نموذج الانحدار الخطي المتعدد :

يعتبر نموذج الانحدار الخطي المتعدد ويسمى أحيانا النموذج الخطي العام هو امتداد لنموذج الانحدار الخطي البسيط، حيث أن الانحدار الخطي البسيط يتضمن متغيرين متغير تابع والآخر متغير مستقل، أما في حالة الانحدار الخطي المتعدد فإنه يتضمن عدة متغيرات مستقلة في النموذج (أكثر من متغير مستقل) والتي يعتقد أنها تؤثر في المتغير التابع.

ويمكن تعريف نموذج الانحدار الخطي المتعدد بأنه عبارة عن معادلة رياضية تعبر عن العلاقة بين متغيرين أو أكثر من المتغيرات المستقلة مع متغير تابع. أوبمعنى أخر هو عبارة عن انحدار للمتغير التابع (Y) على العديد من المتغيرات المستقلة $(X_1, X_2 ..... X_k)$ ويستخدم في التنبؤ بالمتغير التابع الذي تؤثر فيه عدة متغيرات مستقاة.

أهمية الانحدار الخطي المتعدد

تكمن أهمية الانحدار الخطي المتعدد في معرفة أكثر المتغيرات المستقلة تأثيراً على المتغير التابع، حيث إذا كنا بصدد دراسة مجموعة متغيرات مستقلة ومدى تاثيرها على المتغير التابع، ونود أن نعرف ترتيب

المتغيرات المستقلة في تأثيرها على المتغير التابع بدءاً بالمتغير الأكثر تأثير فالأقل، فإننا نستعين بالانحدار الخطي المتعدد الذي من خلال حساب قيمة المعامل (β_i) ومعرفة اشارتها واختبار دلالتها الاحصائية، فإننا يمكن معرفة مدى تأثير كل متغير مستقل على المتغير التابع واتجاه هذا التأثير هل هو سلبي أم تأثير ايجابي وذلك حسب أشارة المعاملة المقدرة.

4-2- الفروض والشروط الأساسية لنموذج الانحدار الخطي المتعدد:

توجد مجموعة من الفروض والشروط الأساسية التي يجب تحققها في نموذج الانحدار الخطي المتعدد وهي:

1- أن تكون العلاقة خطية بين المتغيرات المستقلة والمتغير التابع.

2- يجب أن يتم اختيار العينة عشوائياً، ويجب أن يكون كل فرد مستقل عن باقي أفراد العينة.

3- يجب أن يكون حجم العينة مساوياً على الأقل لأربعة أضعاف عدد المتغيرات المستقلة، فمثلاً إذا كان لدينا (3) متغيرات مستقلة، فإن حجم العينة يجب أن يساوي على الأقل ($n=3\times4=12$).

4- يجب أن يكون المتغير التابع متصلاً أي مصنف ضمن المقاييس الكمية (النسبي أو الفتري) بينما يمكن أن تكون المتغيرات المستقلة متصلة أو منفصلة أي مصنفة ضمن المقياس الأسمي أو الرتبي (ولكن بعد ترميز البيانات الأسمية وتحويلها إلى أرقام).

5- يجب أن تكون المتغيرات المستقلة غير عشوائية وثابتة (محددة)، وهذا يعني أن قيم المتغيرات المستقلة يمكن مراقبتها ومن أجل ضمان أن قيم المتغيرات المستقلة لا تتغير من حين لأخر أي

$$Cov\left(x_i, \quad \varepsilon_j\right) = 0 \ \ \forall\, i = 1..n$$

و يجب أن يكون تباين أي متغير مستقل أكبر من الصفر $(S^2>0)$، والغرض من هذا الشرط هو أن يسهم كل متغير مستقل في تفسير التباين للمتغير التابع. أي لها تأثير على المتغير التابع، وهذا يستلزم عدم ادخال المتغيرات المستقلة التي ليس لها أي تأثير على المتغير التابع في نموذج الانحدار الخطي المتعدد.

6- يجب أن تكون البواقي (الاخطأ العشوائية) لها توزيع اعتدالي (طبيعي) بمتوسط صفر وتباين ثابت ومشترك (σ^2)، أي $E(e) = 0$، أي أن مجموع الاخطاء العشوائية مساويا للصفر.

7- يجب أن تكون بيانات المتغير التابع المتنبأ بها $(\hat{Y})$ موزعة توزيعاً طبيعياً.

8- استقلالية حد الخطأ عن قيم المتغيرات المستقلة، هذا يعني تباين الأخطاء العشوائية مقدار ثابت لكل المشاهدات ومساويا σ^2، أي ثبات (تجانس) تباين الاخطاءالعشوائية، أي عدم وجود مشكلة عدم التجانس (Homoscedasticity).

9- استقلالية المتغيرات المستقلة فيما بينها، أي عدم وجود علاقة خطية بين المتغيرات المستقلة أو ما يسمى بعدم وجود علاقات متداخلة (مشكلة التعدد الخطي Multicollinearity) بين المتغيرات المستقلة.

10- عدم وجود مشكلة الارتباط الذاتي بين البواقي (الاخطأ العشوائية) (Autocorrelation)، أي أن تكون البواقي مستقلة عن بعضها البعض.

3-4- التوصيف الرياضي لنموذج الانحدار الخطى المتعدد (Multiple Linear Regression)

يمكن كتابة المعادلة الرياضية التي تصف الانحدار الخطي المتعدد على الشكل التالي:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_k X_k + u_i$$

حيث:

- $(X_1, X_2 \ldots\ldots X_k)$ هي المتغيرات المستقلة وعددها (k).
- β_0: هذه المعلمة تسمى المعلمة التقاطعية وهي تشير إلى قيمة المتغير التابع عندما نعزل أثر المتغيرات المستقلة للنموذج (أي تمثل قيمة المغير التابع عندما تكون جميع المتغيرات المستقلة في النموذج تساوي الصفر).

← ($\beta_1, \beta_2 \ldots . \beta_k$) هذه المعاملات تسمى معاملات الانحدار الجزئية، فمثلا المعاملة (β_1) تمثل انحدار المتغير التابع (Y) على المتغير المستقل (X_1). وهذه المعاملات تقيس التغير في المتغير التابع (Y) نتيجة للتغير في احد المتغيرات المستقلة بوحدة واحدة مع ثبات المتغيرات الأخرى فمثلاً المعلمة (β_1) تشير إلى التغيرات في قيمة المتغير التابع (Y) نتيجة للتغير في قيمة المتغير المستقل (X_1) بوحدة واحدة مع ثبات المتغيرات المستقلة الأخرى.

ويكثر استخدام النماذج المتعددة في الاقتصاد لأنه من الصعب أن نجد متغير تابع يتأثر بمتغير مستقل واحد فقط، وعلى الأغلب أنه يوجد كثير من التأثيرات، فمثلاً دالة الاستهلاك تتأثر بمتغير الدخل والثروة والأسعار ويمكن صياغتها على الشكل التالي $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u_i$.

حيث (Y) يمثل الاستهلاك و(X_1) يمثل الدخل و(X_2) يمثل السعر(X_3) يمثل الثروة.

و بما أن ($\beta_0, \beta_1, \beta_2 \ldots . \beta_k$) نعبر عن معاملات الانحدار، و (u_i) يعبر عن الخطأ العشوائي للمشاهدة رقم (i) حيث (i=1,2…..n) وحيث أن عدد المشاهدات هي (n) يكون لدينا (n) من المعادلات يمكن صياغتها على الشكل المصفوفي التالي:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

$$Y=XB+U$$

حيث أن:

Y: يعبر من متجه المشاهدات التابعة وهو من الدرجة (n-1) والعنصر رقم (i) في هذا المتجه هو (y_i).

X: تمثل مصفوفة المشاهدات المستقلة وهي من الدرجة ((k+1)× n) والصف رقم (i) في هذه المصفوفة هو $(x_{11}, x_{12}, \ldots \ldots x_{1k})$.

B: يعبر من متجه معملات الانحدار وهو من الدرجة (1× (k+1)).

U: يعبر من متجه الأخطاء العشوائية وهو من الدرجة (n×1) والعنصر رقم (i) في هذا المتجه هو الخطأ العشوائي (u_i).

4-4- تقدير معلمات نموذج الانحدار الخطي المتعدد

عند توفر بيانات المتغير التابع والمتغيرات المستقله فان تقدير معاملات معادلة الانحدار الخطي المتعدد يمكن ان يتم بعدة طرق أهمها طريقة المصفوفات و طريقة المربعات الصغرى (Least Squares Method).

4-4-1-طريقة المربعات الصغرى (Least Squares Method).

تعتبر طريقة المربعات الصغرى الأسهل والأكثر استخداماً، وهي تعرف على أنها اسلوب قياسي لتوفيق أفضل خط مستقيم لعينة من المشاهدات، وهي تعطى أحسن معادلة خط انحدار بين المتغير التابع والمتغيرات المستقله، وتعتبر أحسن معادلة لآنها تعطى أقل خطأ ممكن.

- حيث يتضمن هذا الاسلوب تصغير مجموع المربعات لانحرافات النقاط الفعلية عن خط التوفيق الى ادنى حد ممكن اي

$$\boldsymbol{min} \sum_{i=1}^{n} \mathbf{e_i^2} = \sum_{i=1}^{n} (\mathbf{Y_i} - \hat{\mathbf{Y}}_\mathbf{i})^2$$

حيث $\boldsymbol{e_i} = (\mathbf{Y_i} - \hat{\mathbf{Y}}_\mathbf{i})$ يمثل الخطأ العشوائي أو بواقي المربعات الصغرى (Residuals)

وتجدر الاشارة هنا الى انه لابد من اخذ مجموع مربعات الانحرافات عن خط التوفيق لان بعض القيم الفعلية لـ (Yi) تقع أعلى خط الانحدار فيتولد عنها انحرافات موجبة، والبعض الثاني لقيم لـ (Yi) تقع أسفل خط الانحدار فيتولد عنها انحرافات سالبة، في حين ان البعض الآخر من قيم (Yi) تقع على خط الانحدار فلا يتولد عنها انحرافات، وبكل الاحوال فإن مجموع الانحرافات يساوي صفر أي:

$$\sum_{i=1}^{n} \boldsymbol{e_i} = \sum_{i=1}^{n} (\mathbf{Y_i} - \hat{\mathbf{Y}}_\mathbf{i}) = \mathbf{0}$$

وبذلك فان هذا المعيار سيكون مرفوضاً لعدم قدرته على التمييز بين التوفيق الجيد من غيره وبذات الوقت فهو لا يعني ان الانحرافات سوف تختفي بتوفيق خط المربعات الصغرى لذلك لابد من ايجاد مربعات الانحرافات ومحاولة جعل هذه المربعات اصغر مايمكن أي $\boldsymbol{min} \sum_{i=1}^{n} \mathbf{e_i^2}$ ومن هنا جاءت تسمية طريقة المربعات الصغرى.

لتقدير معادلة الانحدار الخطي المتعدد للمجتمع وهي: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_k X_k + u_i$

وباستخدام الفروض السابقه تصبح أفضل معادلة لتقدير قيمة المتغير التابع بدلالة المتغيرات المستقله هى: $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_k X_k$

المعيار الذي تعتمد علية المربعات الصـغرى في الحصـول على المقدرات حيث يتطلب المعيار تصـغير مجموع مربعات البواقي ألي أدنى قيمه لها.

$$\sum u_i^2 = \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_1 \ldots . \hat{\beta}_k X_k)^2$$

تفاضل البواقي بالنسبه لـ $\hat{\beta}_0$ ويساوى بالصفر ويعاد كذلك لقيم $\hat{\beta}_1, \hat{\beta}_2 \ldots \hat{\beta}_k$ وهكذا

$$\frac{\partial \sum u_i^2}{\partial \hat{\beta}_0} = +2\sum (Y_I - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2 \ldots \hat{\beta}_K X_K) = 0$$

$$\frac{\partial \sum u_i^2}{\partial \hat{\beta}_1} = -2\sum X_{I1} (Y_I - \hat{\beta}_0 - \hat{\beta}_2 X_2 - \hat{\beta}_3 X_3 \ldots \hat{\beta}_K X_K) = 0$$

$$\frac{\partial \sum u_i^2}{\partial \hat{\beta}_K} = -2\sum X_{IK} (Y_I - \hat{\beta}_0 - \hat{\beta}_2 X_2 - \hat{\beta}_3 X_3 \ldots \hat{\beta}_K X_K) = 0 .$$

بفك الأقواس والقسمة على 2 وإعادة كتابة المعادلات الطبيعية مقابلة للنموذج الخطي العام

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u_i$$

$$\sum u_i^2 = \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2)^2$$

نحصل على مقدرات النموذج العام وعلى سبيل المثال سنكتفي بنموذج بثلاث متغيرات (متغير تابع ومتغيرين مستقلين).

$$\sum Y_I = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{1i} + \hat{\beta}_2 \sum X_{2i}$$

$$\sum X_{1I} Y = \hat{\beta}_0 \sum X_{1I} + \hat{\beta}_1 \sum X_{1I}^2 + \hat{\beta}_2 \sum X_{1I} X_{2I}$$

$$\sum X_{2I} Y = \hat{\beta}_0 \sum X_{21I} + \hat{\beta}_1 \sum X_{1I} X_{2I} + \hat{\beta}_2 \sum X_{2I}^2$$

وباستخدام الانحرافات نتحصل على

$$\hat{\beta}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$\hat{\beta}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2$$

مثال(1): الجدول التالي يوضح بيانات خاصة بحجم الاستهلاك (Y)، وكل من الدخل (X_1)، وحجم الثروة (X_2).

الاستهلاك (Y)	40	44	46	48	52	58	60	68	74	80
الدخل (X_1)	6	10	12	14	16	18	22	24	26	32
الثروة (X_2)	4	4	5	7	9	12	14	20	21	24

المطلوب: تقدير دالة الاستهلاك (تقدير قيم معالم نموذج الانحدار المتعدد)

الحل: إن صيغة معادلة الانحدار على الشكل التالي:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

وعلينا تقدير قيم المعاملات.

Y	X_1	X_2	y	x_1	x_2	$x_1 y$	$x_2 y$	$x_1 x_2$	x_1^2	x_2^2
40	6	4	-17	-12	-8	204	136	96	144	64
44	10	4	-13	- 8	-8	104	104	64	64	64
46	12	5	-11	- 6	-7	66	77	42	36	49
48	14	7	- 9	- 4	-5	36	45	20	16	25
52	16	9	- 5	- 2	-3	10	15	6	4	9
58	18	12	+ 1	0	0	0	0	0	0	0
60	22	14	+ 3	+4	+2	12	6	8	16	4
68	24	20	+11	+6	+8	66	88	48	36	64
74	26	21	+17	+8	+9	136	153	72	64	81
80	32	24	+23	+14	+12	322	276	168	196	144
570	**180**	**120**	**0**		**0**	**956**	**900**	**524**	**576**	**504**

$$\hat{\beta}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(956)(504) - (900)(524)}{(576)(504) - (524)^2} = 0.65$$

$$\hat{\beta}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(900)(576) - (956)(524)}{(576)(504) - (524)^2} = 1.11$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2 = 57 - (0.65)(18) - (1.11)(12) = 31.98$$

$$\hat{Y} = 31.98 + 0.65X_1 + 1.10X_2 + u$$

5-4- اختبار جودة نموذج الانحدار الخطي المتعدد (اختبار القوة التفسيرية للنموذج (جودة التوفيق))

تبعا لمنهجية القياس الاقتصادي، بعد الحصول على مقدرات معلمات النموذج تأتي مرحلة اختبار جودة النموذج أي قياس مدى قدرة المتغيرات المستقلة التي يتضمنها النموذج في تفسير تغيرات المتغير التابع أو الظاهرة محل الدراسة.

تطبيق اختبار تحليل التباين لااختبار جودة (صلاحية) نموذج الانحدار الخطي المتعدد

يسمى هذا الاختبار اختبار جودة نموذج الانحدار الخطي المتعدد بواسطة أسلوب "تحليل التباين" أو بتطبيق اختبار (F). الهدف من هذا التطبيق هو اختبار صلاحية النموذج المقترح للتنبؤ، و التوصل إلى قرار حول صلاحية النموذج في تمثيل العلاقة بين المتغير التابع والمتغيرات المستقلة تمثيلاً جيداً، وفيما يلي خطوات اجراء هذا الاختبار.

1- تنظيم جدول تحليل التباين كما يلي

مصادر التباين Source of Variance (S.O.V)	مجموع مربعات الانحرافات Sum of squares (SS)	درجات الحرية df	متوسط المربعات (التباين) Mean squares (MS)	F (المحسوبة) Calculated	F (الجدولية) Tabulated
التباين بسبب الانحدار Regression	SSR	k-1	$MSR = \frac{SSR}{k-1}$	$F_C = \frac{MSR}{MSE}$	$F_{\alpha\,(k-1\,,\,n-k)}$
التباين بسبب الاخطاء (البواقي) Errors –	SSE	n-k	$MSE = \frac{SSE}{n-k}$		
المجموع Total	SSTO = SSR + SSE	n-1			

حيث يتم حساب مجموع مربعات الانحرافات كما يلي:

$$SSTO = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

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

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

ثانياً: حساب قيمة احصاءة الاختبار: تحسب احصاءة الاختبار من العلاقة $F_C = \frac{MSR}{MSE}$

حيث إن MSR تمثل متوسط مربعات الانحدار و MSE تمثل متوسط مربعات الأخطاء وهو في نفس الوقت يمثل تقدير تباين الخطأ العشوائي الذي ذكر سابقا.

ثالثاً: المقارنة و اتخاذ القرار: تتم مقارنة قيمة (F المحسوبة) مع (F الجدولية)، فاذا كانت قيمة (F المحسوبة) F_c أصغر من قيمة (F الجدولية) F_t هنا يتم قبول الفرض العدم القائل بأن النموذج غير مناسب ويكون تقدير الباحث للنموذج غير صحيح. أما اذا كانت (F المحسوبة) F_c أكبر من قيمة(F الجدولية) F_t هنا يتم رفض الفرض العدم وقبول الفرض البديل القائل بأن النموذج مناسب لتمثيل العلاقة بين المتغيرين ومناسب للتنبؤ.

4-6- معامل التحديد المتعدد (Multiple Coefficient Of Determination)

يشير معامل التحديد المتعدد إلى النسبة التي يمكن تفسيرها من التغير الكلي في المتغير التابع بدلالة المتغيرات المستقلة المدرجة في دالة الانحدار الخطي المتعدد. وهو يساوي نسبة التغيرات المفسرة إلى التغيرات الكلية، ويمكن حسابه باحدى العلاقتين التاليتين:

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

$$R^2 = \frac{\beta_1 \sum y_i x_1 + \beta_2 \sum y_i x_2 + \cdots \ldots + \beta_k \sum y_i x_k}{\sum y_i^2}$$

و العلاقة الثانية تصبح في حالة متغيرين مستقلين فقط على الشكل التالي:

$$R^2 = \frac{\beta_1 \sum x_1 y + \beta_2 \sum x_2 y}{\sum y^2}$$

وفي حالة الانحدار الخطي البسيط يتم حساب معامل التحديد كما يلي:

$$R^2 = \hat{\beta}\frac{\sum y_i x_i}{\sum y_i^2}$$

يلاحظ أنه مع كل إضافة لمتغير تفسيري جديد، نضيف حداً في البسط يمثل أثر هذا المتغير على العلاقة الكلية .وبشكل عام غالباً ما تزداد قيمة معامل التحديد المتعدد كلما أضفنا متغيراً مستقلاً (تفسيرياً) جديداً، ويندر أن تنقص .مما يعني أن هذا المقياس يتأثر بعدد المتغيرات المستقلة في الدالة.

لتلافي هذا القصور يتعين أن نصحح قيمة معامل التحديد بحيث لا يتأثر بعدد المتغيرات التفسيرية.

ويمكن ذلك عبر أخذ درجات الحرية في الحسبان (n–k) حيث تقل درجات الحرية مع زيادة عدد المتغيرات التفسيرية .ويقدر معامل التحديد المعدل (Adjusted R^2) أعتماداً على القانون:

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

حيث (n) تمثل عدد المشاهدات، و (k) عدد المعالم.

تتراوح قيمة معامل التحديد المعدل بين (0 و 1) وكلما اقترب الرقم من الواحد ازدادت المقدرة التفسيرية للنموذج.

مثال(1): الجدول التالي يوضح العلاقة بين الكمية المطلوبة من سلعة معينة (Y)، وكل من السعر (X_1)، والدخل (X_2).

الكمية المعروضة (Y)	7	6	3	2	7
السعر (X_1)	4	5	6	7	3
الدخل (X_2)	18	15	12	10	15

المطلوب:

1- تقدير قيم معالم نموذج الانحدار المتعدد
2- تقدير تباين المتغير العشوائي
3- تقدير معامل التحديد المتعدد والمعدل
4- اختبار المعنوية الكلية للنموذج عند مستوى معنوية (0.05).

الحل: إن صيغة معادلة الانحدار على الشكل التالي:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

وعلينا تقدير قيم المعاملات.

Y	X_1	X_2	y	x_1	x_2	$x_1 y$	$x_2 y$	$x_1 x_2$	x_1^2	x_2^2
7	4	18	2	-1	4	**-2**	8	-4	1	16
6	5	15	1	0	1	0	1	0	0	1
3	6	12	-2	1	-2	-2	4	-2	1	4
2	7	10	-3	2	-4	-6	12	-8	4	16
7	3	15	2	-2	1	-4	2	-2	4	1
25	25	70	0	0	0	-14	27	-16	10	38

$$\hat{\beta}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(-14)(38) - (27)(-16)}{(10)(38) - (-16)^2} = -0.81$$

$$\hat{\beta}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(27)(10) - (-14)(-16)}{(10)(38) - (16)^2} = 0.37$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}_1 - \hat{\beta}_2 \overline{X}_2 = 5 - (-0.81)(5) - (0.37)(14) = 3.87$$

$$\hat{Y}_1 = 3.87 - 0.81 X_1 + 0.37 X_2 + u$$

يتم تقدير تباين المتغير العشوائي وفقاً للصيغة التالية:

$$\boldsymbol{\sigma^2 = \frac{\sum u_i^2}{n-k} = \frac{\sum Y_i^2 - \beta_1 \sum x_i y - \beta_2 \sum x_2 y}{n-k}}$$

$$\boldsymbol{\sigma^2 = \frac{22 - (-0.81)(-14) - (-0.37)(27)}{5-3} = 0.335}$$

2- تقدير معامل التحديد المتعدد والمعدل

معامل التحديد المتعدد يحسب بالعلاقة التالية:

$$\boldsymbol{R^2 = \frac{\beta_1 \sum x_i y + \beta_2 \sum x_2 y}{\sum y^2} = \frac{-0.81(-14) + 0.37(27)}{22} = \frac{21.33}{22} = 0.97}$$

$$\boldsymbol{\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} = 1 - (1 - 0.97)\frac{5-1}{5-3} = 0.94}$$

حيث (n) تمثل عدد المشاهدات، و (k) عدد المعالم.

Y	X_1	X_2	$\hat{Y}$	u	u^2 $(Y - \hat{Y})^2$	y^2 $(Y - \overline{Y})^2$

7	4	18	7.29	-0.29	0.0841	4
6	5	15	5.37	0.63	0.3969	1
3	6	12	3.45	-0.45	0.2025	4
2	7	10	1.9	0.1	0.01	9
7	3	15	6.99	0.01	0.0001	4
25	25	70	34.52	0	SSE=0.693	SSTO= 22

SSR=SSTO-SSE=22-0.693=21.307

$$R^2 = \frac{SSR}{SSTO} = \frac{\sum \hat{y}_i}{\sum y_i} = \frac{21.307}{22} = 0.968$$

تطبيق اختبار تحليل التباين لاختبار جودة (صلاحية) نموذج الانحدار الخطي المتعدد

مصادر التباين Source of Variance (S.O.V)	مجموع مربعات الانحرافات Sum of squares (SS)	درجات الحرية df	متوسط المربعات (التباين) Mean squares (MS)	F (المحسوبة) Calculated	F (الجدولية) Tabulated
التباين بسبب الانحدار Regression	SSR =21.307	3-1=2	$MSR = \frac{21.307}{2} = 10.65$	$F_C = \frac{10.65}{0.346} = 30.78$	$F_{0.05\,(2\,,\,2)}=19$
التباين بسبب الاخطاء (البواقي) Errors –	SSE =0.693	n-k=5-3=2	$MSE = \frac{0.693}{2} = 0.346$		
المجموع Total	SSTO = 22	n-1=4			

حيث يتم حساب مجموع مربعات الانحرافات كما يلي:

$$SSTO = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

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

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

ثانياً: حساب قيمة احصاءة الاختبار: تحسب احصاءة الاختبار من العلاقة

$$F_C = \frac{10.65}{0.346} = 30.78$$

ثالثاً: المقارنة و اتخاذ القرار: بما أن قيمة F المحسوبة $F_C = 30.78$ أكبر من F الجدولية $F_t = 19$، هنا يتم رفض الفرض العدم وقبول الفرض البديل القائل بأن النموذج مناسب لتمثيل العلاقة بين المتغيرين ومناسب للتنبؤ.

القسم الثاني

د. رواد شتي

المتغيرات الصورية (الصماء)

Dummy Variables

أولاً- مقدمة Introduction:

تم التأكيد في المحاضرات السابقة على أن الدور الأساسي للاقتصاد القياسي ينحصر في تقدير معالم النموذج الاقتصادي من خلال مشكلة عدم التأكد Uncertainty ، وتجدر الاشارة إلى أن النموذج الاقتصادي قد يتضمن العديد من المعادلات ، كما وقد تتضمن كل معادلة على الكثير من المتغيرات. ويتوقف عدد المعادلات وعدد المتغيرات التي يتضمنها أي نموذج اقتصادي على طبيعة الظاهرة الاقتصادية التي يرمي الباحث إلى تفسير اختلافاتها. والجدير بالذكر، أن معالجة مثل هذه النماذج الاقتصادية المعقدة يصبح أمراً بسيطاً فيما لو أتقن الباحث معالجة النماذج الاقتصادية البسيطة التي تتناول تفسير العلاقة بين متغيرين أحدهما تابعDependent variable والآخر مستقل Independent variable ، ومثال ذلك تفسير العلاقة بين الاستهلاك الشخصي (المتغير التابع C_t) والدخل الشخصي المتاح (المتغير المستقل Y_t) والتي يمكن صياغتها في نموذج بسيط كالآتي:

$$C_t = a + bY_t \pm U_t$$

فإذا كانت المتغيرات التي تؤثر في ظاهرة اقتصادية معينة تمثل متغيرات وصفية أو نوعية كالجنس واللون والديانة والمهنة والمستوى التعليمي وغيرها أطلقنا عليها المتغيرات الصورية أو الصماء ، وقد يشار اليها في بعض المصادر بالمتغيرات الوهمية ، أو المتغيرات الثنائية Binary Variables أو المتغيرات النوعية Qualitative Variables ، وفي بعض الكتابات يشار اليها بالمتغيرات الفئوية Categorical Variables . وتأخذ هذه المتغيرات قيمتين إجباريتين فقط هما: الصفر والواحد. فهي تأخذ القيمة واحد عند وجود خاصية معينة ، وتأخذ القيمة صفر عند غياب هذه الخاصية. فإذا رمزنا الى المتغير الصوري بالرمز (D) فإن **1**=D إذا كان الشخص أبيض مثلاً , D= صفر إذا كان الشخص أسوداً ، باعتبار أن المتغير الصوري (D) هنا يرمز إلى اللون. أو أن 1=D إذا كان الفرد ذكراً ،

D = صفر إذا كان الفرد أنثى وذلك على اعتبار أن (D) هنا ترمز إلى الجنس ، وهكذا بالنسبة لبقية الأمثلة.

تستخدم المتغيرات الصورية في نماذج الانحدار إما كمتغيرات تفسيرية (مستقلة) أو كمتغيرات تابعة ، ولكن التركيز الأكبر عليها كمتغيرات تفسيرية كما سنرى لاحقاً.

ثانياً- كيفية استخدام المتغيرات الصورية:**How to use dummy variables**

يوجد هناك أكثر من طريقة لاستخدام المتغيرات الصورية في نماذج الانحدار (Regression) كمتغيرات تفسيرية (مستقلة). نشـير في هذه المحاضرة لبعض منها كما يلي:

1-2 متغير تفسـيري نوعي واحد :

من الممكن أن يحتوي نموذج الانحدار على متغير تفسـيري نوعي واحد دون وجود أي متغيرات كمية تفسـيرية. فإذا أردنا مثلاً اختبار أثر الجنـس (ذكر أو أنثى) على مسـتوى الأجور في مجتمع العاملين بمجال التدريـس ، فمن الممكن القيام بذلك من خلال قياس علاقة الانحدار بين الأجر Y_iكمتغير تابع و D_i كمتغير تفسيري صوري يمثل الجنس. وتأخذ هذه العلاقة الصيغة التالية:

$$Y_i = a + b\, D_i + e_i$$

حيث Y_i= الراتب السنوي للمدرس خريج الجامعة

D_i = الجنس بحيث: D_i= 1 إذا المدرس ذكراً

D_i = صفر إذا كان المدرس أنثى

مثال: العلاقة بين الأجر والجنس:

إذا كان لدينا عينة من المدرسـين خريجي الجامعات عددهم 10 وكانت رواتبهم على النحو الموضح في الجدول (1-1) ، فمن الممكن تقدير معادلة الانحدار الخاصة بالعلاقة بين الاجر والجنس باسـتخدام بيانات العمودين الثالث والرابع بنفس الطريقة التي تم شـرحها في حالة الانحدار الخطي البسـيط ، وذلك كما هو موضح بالجدول (2-1).

جدول (1-1): الراتب السنوي لعينة من المدرسين خريجي الجامعات

المشـاهدة	الجنس	الراتب السنوي (Y)	المتغير الصوري D
1	ذكر	1000	1
2	ذكر	1100	1
3	أنثى	800	0
4	أنثى	900	0
5	ذكر	1050	1

6	أنثى	950	0
7	أنثى	850	0
8	ذكر	1300	1
9	ذكر	1500	1
10	أنثى	950	0

Damascus University

جدول (2-1) حســابات أثر الجنــس على الأجر

d^2	y * d	(d) = D-$\bar{D}$	(y) = Y-$\bar{Y}$	D	Y
0.25	-20	+0.5	-40	1	1000
0.25	+30	+0.5	+60	1	1100
0.25	+120	-0.5	-240	0	800
0.25	+70	-0.5	-140	0	900
0.25	+5	+0.5	+10	1	1050
0.25	+45	-0.5	-90	0	950
0.25	+95	-0.5	-190	0	850
0.25	+130	+0.5	+260	1	1300
0.25	+230	+0.5	+460	1	1500
0.25	+45	-0.5	-90	0	950
$\sum_{d} 2 = 2.5$	$\sum y*d = 750$			$\sum D = 5$	$\sum Y = 10400$

$$\bar{Y} = \frac{\sum Y}{n} = \frac{10400}{10}$$

1040 =المتوسط العام للرواتب

$$\bar{D} = \frac{\sum D}{n} = \frac{5}{10} = \mathbf{0.5}$$

$$b = \frac{\sum y*d}{\sum_d 2} = \frac{750}{2.5} = 300$$

$$a = \bar{Y} - b\bar{D} = 1040 - (300)*(0.5) = 890$$

يمكن أن نســتنتج مما ســبق المعادلة الخطية بمعالمها a و b كما يلي:

$$Y_i = 890 + 300\ D_i + e_i$$

وبحساب متوسط رواتب الإناث والذكور كل على حدة من الجدول (2-1) نجد أن:

$\bar{Y}_f = \frac{4450}{5} = 890$ متوسط رواتب الإناث

$\bar{Y}_m = \frac{5950}{5} = 1190$ متوسط رواتب الذكور

ويلاحظ من المعادلة الأساسية: $Y_i = a + b\,D_i + e_i$

ما يلي: - القيمة المتوقعة لراتب المُدرسة $= E(Y \backslash D_i = 0) = a$

- القيمة المتوقعة لراتب المدرس $= E(Y \backslash D_i = 1) = a+b$

وذلك بفرض أن القيمة المتوقعة للحد العشوائي e_i = صفر ، ويعني الخط المائل (\) بشرط.

ومما سبق نجد أن المعلمة التقاطعية a تشير إلى متوسط راتب المُدرسة ، ويتضح ذلك من المعادلة:

$$Y_i = 890 + 300\,D_i + e_i$$

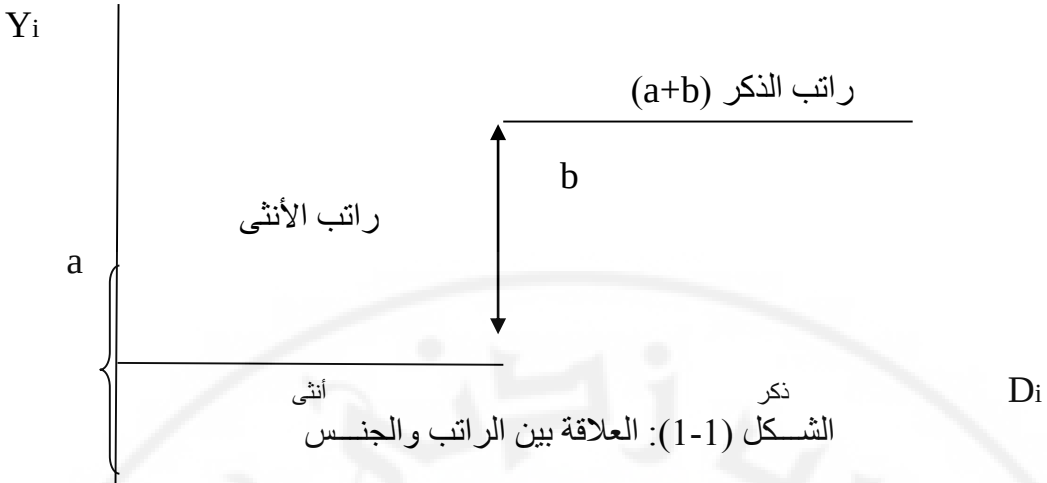
حيث نجد أن $a = 890$ وهي نفسها $\bar{Y}_f$ ، أما المعلمة الانحدارية b فهي تشير إلى الفرق بين راتب المدرس الذكر والمُدرسة ، حيث يلاحظ أن متوسط راتب المدرس $= a + b$. وفي المعادلة:

$$Y_i = 890 + 300\,D_i + e_i$$

نجد أن $a+b = 890 + 300 = 1190$ وهو نفسه $\bar{Y}_m$.

ومن ثم فإن الفرق بين المتوسطين $= \bar{Y}_m - \bar{Y}_f = (a + b) - (a) = b = 300$.

وإذا أردنا تمثيل ذلك بيانياً فإن الشكل (1-1) أدناه يعبر عن العلاقة بين الراتب والجنس:



الشـــكل (1-1): العلاقة بين الراتب والجنـــس

ومن الممكن التعميم في هذه الحالة بالقول أن المعلمة التقاطعية (a) في نموذج يأخذ الصيغة:

$$Y_i = a + b D_i + e_i$$

تشـــير إلى متوســـط المتغير التابع في حالة الصـفة التي يكون عندها (صفر.(=Di

أما المعلمة الانحدارية (b) فهي تشـير إلى الفرق بين متوسط المتغير التابع في حالة الصفة التي يكون عندها (=Di1) ، والمتوسط في حالة الصفة التي يكون عندها (صفر.(=Di

وفي هذه الحالة فإنه من الممكن عندها اختبار الفرض القائل " أنه يوجد تمييز في الأجر وفقاً للجنس " من خلال اختبار: - فرض العدم : $b = 0$

في مواجهة: - الفرض البديل: $b \neq 0$

وذلك باســتخدام اختبار (t) ، وإذا ثبت من الاختبار أن (b) لها معنوية احصائية فإن هذا يعني أن الاختلاف بين رواتب المدرســين الذكور والمدرسـات جوهري ، ومن ثم نقبل الفرض القائل " أنه يوجد هناك تمييز في الراتب وفقاً للجنس" ، والعكـس صحيح .

2-2 أكثر من متغير تفسـيري (مسـتقل) نوعي واحد:

لاحظنا في المحاضرة السـابقة أن معادلة الانحدار يمكن أن تُبنى على أسـاس متغير تفسـيري نوعي واحد سـواء كان هذا المتغير ممثلاً في متغير صوري واحد أو أكثر من متغير صوري. ولكن بالإضافة إلى ذلك ، فإنه من الممكن أن يُبنى نموذج الانحدار على أسـاس أكثر من متغير تفسـيري نوعي.

إذا افترضنا مثلاً أن راتب الخريج يعتمد على الأسـاسـيين التاليين:

1- المسـتوى التعليمي للخريج: حيث يمكن أن يكون معهد متوسط ، إجازة جامعية أو ماجستير = D_i

2- نوع المؤسـسة التي يعمل بها: يمكن أن يكون قطاع عام أو قطاع خاص = H_i

ففي هذه الحالة يكون لدينا عدد من المتغيرات الصورية يمثل الآتي:

- عدد المتغيرات الصورية الممثلة للمسـتوى التعليمي = 3-1=2
- عدد المتغيرات الصـورية الممثلة لنوع المؤسـسة = 2-1=1
- عدد المتغيرات الصـورية الكليـة بنموذج الانحـدار= 2+1= 3

ويمكن توضيح الصفات المتعلقة بقيم المتغيرين النوعيين السـابقين في الجدول التالي:

جدول (6-1) صفات المتغيرين النوعيين

نوع المؤسسة / المستوى التعليمي	قطاع عام H_1	قطاع خاص H_2
معهد متوسط D_1	1	2
إجازة جامعية D_2	3	4
ماجسـتير D_3	5	6

وبالتالي فإن معادلة الانحدار التي تمثل العلاقة بين راتب الخريج والمتغيرات التفسـيرية الصورية المؤثرة فيه يمكن صياغتها على النحو التالي:

$$Y_i = b_1 + b_2D_2 + b_3D_3 + a_2H_2 + U_i$$

حيث: Yi = الراتب السـنوي للخريج

D_2=1 إذا كان الخريج من حاملي الإجازة الجامعية

D_2= 0 إذا كان الخريج يحمل درجة أخرى

D_3= 1 إذا كان الخريج من حاملي درجة الماجسـتير

D_3= 0 إذا كان الخريج من غير حاملي درجة الماجستير

H_2= 1 إذا كان الخريج يعمل في القطاع الخاص

H_2= 0 إذا كان الخريج يعمل في القطاع العام

يتضح لنا مما سبق أن فئة الأساس التي تنعكس في المعلمة التقاطعية هي الخريجين من حملة المعهد المتوسط D_1 العاملين في القطاع العام H_1 .

متوسط راتب العاملين في القطاع العام من حملة المعهد المتوسط:

$$E (Yi\backslash D_2= D_3 = H_2 =0) = b_1$$

متوسط راتب العاملين في القطاع العام من حملة الإجازة الجامعية:

$$E (Yi\backslash D_2=1, D_3 = H_2=0) = b_1+b_2$$

متوسط راتب العاملين في القطاع العام من حملة الماجستير:

$$E (Yi\backslash D_3=1, D_2 = H_2=0) = b_1+b_3$$

متوسط راتب العاملين في القطاع الخاص من حملة المعهد المتوسط:

$$E (Yi\backslash H_2=1, D_2 = D_3=0) = b_1+a_2$$

متوسط راتب العاملين في القطاع الخاص من حملة الإجازة الجامعية:

$$E (Yi\backslash H_2= D_2 = 1, D_3=0) = b_1+b_2 + a_2$$

متوسط راتب العاملين في القطاع الخاص من حملة الماجستير:

$$E (Yi\backslash H_2= D_3 = 1, D_2=0) = b_1+b_3 + a_2$$

ويمكن إجراء عدد من الاختبارات هنا أهمها:

1- هل هناك فروق جوهرية بين رواتب الخريجين العاملين في القطاع العام من مختلف المستويات التعليمية؟ أو بين رواتب العاملين في القطاع الخاص من مختلف المستويات التعليمية؟ ويمكن عمل ذلك من خلال اختبار الفرض: $b_2= b_3 = 0$

2- هل هناك فروق جوهرية بين رواتب الخريجين العاملين في القطاع الخاص ورواتب العاملين في القطاع العام من مختلف المستويات التعليمية؟ ويمكن عمل ذلك باختبار الفرض: $a_2= 0$

بالإضافة إلى تلك الاختبارات فإنه يمكن أيضاً استخدام اختبار T لإتمام الاختبارات السابقة ، حيث نختبر فروض العدم : $b_2= 0$ ، $b_3= 0$ ، $a_2 = 0$ ، في مواجهة الفروض البديلة: $b_2 \neq 0$ ،

$a_2 \neq 0$ ، $b3 \neq 0$. ومن نتيجة هذه الاختبارات يمكن تحديد ما إذا كان المستوى التعليمي أو نوع المؤسسة التي يعمل بها الخريج له تأثير جوهري على مستوى الراتب أم لا.

3-2 متغيرات تفسيرية (مستقلة) نوعية وكمية:

يمكن أن يبنى نموذج الانحدار على أساس وجود عدد من المتغيرات التفسيرية النوعية إلى جانب عدد آخر من المتغيرات التفسيرية الكمية. وسوف نتعرض في هذا الصدد لأكثر من نموذج على النحو التالي:

2-3-1 متغير كمي ومتغير نوعي واحد بصفتين:

إذا افترضنا أن رواتب العاملين من حاملي درجة الإجازة الجامعية في مجال معين يتحدد بعنصرين:

أ- عدد سنوات الخبرة كمتغير كمي X_i

ب- الجنس (ذكر أو أنثى) كمتغير نوعي D_i

فمن الممكن في هذه الحالة صياغة العلاقة بين المتغير التابع Y_i ، والمتغيرين المستقلين X_i و D_i في صورة دالة الانحدار التالية:

$$Y_i = a_1 + a_2 D_2 + bX_i + U_i$$

حيث: $D_2 = 1$ إذا كان العامل ذكراً ، و $D_2 =$ صفر إذا كان العامل أنثى.

ويلاحظ في هذه الحالة أن نموذج الانحدار يحتوي على متغيرين تفسيريين ، أحدهما كمي وهو عدد سنوات الخبرة ، والآخر نوعي ذو صفتين وهو الجنس (ذكر أو أنثى) . ومن ثم فإن:

القيمة المتوقعة لراتب الخريج الأنثى $= E(Y_i \backslash X_i, D_2 = 0) = a_1 + bX_i$

القيمة المتوقعة لراتب الخريج الذكر $= E(Y_i \backslash X_i, D_2 = 1) = (a_1 + a_2) + bX_i$

وبمعاينة القيمتين المتوقعتين السابقتين نلاحظ ما يلي:

أ- أن متوسط الراتب لكل فئة ممثلاً في القيمة المتوقعة ليس ثابتاً وإنما متغير. فهو يتغير في الحالتين تبعاً لتغير عدد سنوات الخبرة.

ب- أن عدد سـنوات الخبرة يؤثر على متوسـط رواتب العاملين الذكور بنفـس المعدل الذي يؤثر به على متوسـط رواتب العاملين الإناث. ويتضح ذلك من تسـاوي المعلمة الانحدارية الخاصة بعدد سـنوات الخبرة في معادلتي الانحدار السـابقتين حيث:

$$\frac{\partial Y_m}{\partial X} = \frac{\partial Y_f}{\partial X} = b$$

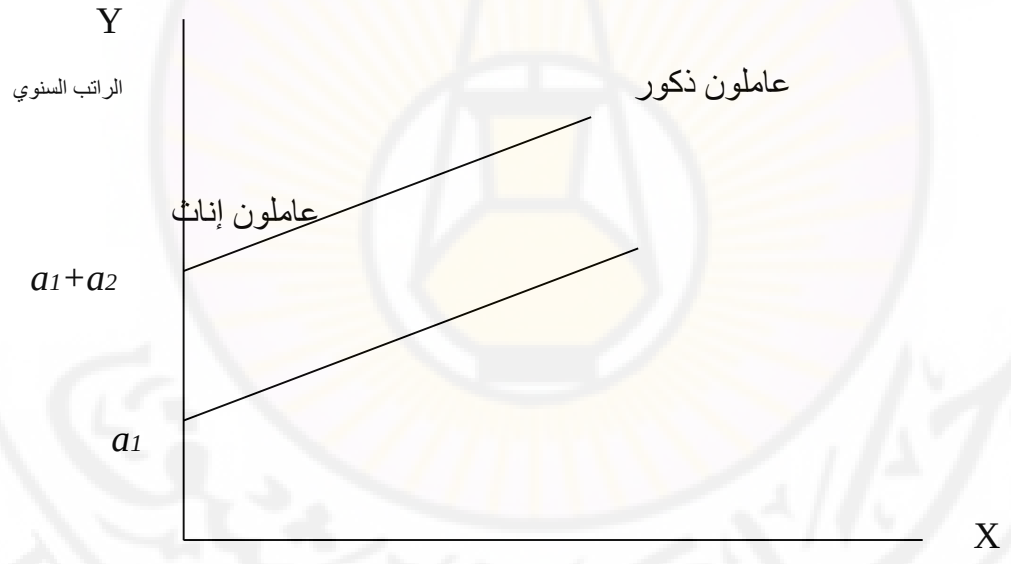
أي أن ميل معادلتي الانحدار متماثل.

جـ- بالإضافة إلى ما سـبق ، يلاحظ أن متوسـط رواتب العاملين الذكور ممثلاً في علاقة الانحدار:

$Y_m = (a_1 + a_2) + bX_i$

يفوق متوسـط رواتب العاملين الإناث ممثلاً في علاقة الانحدار: $Y_f = a_1 + bX_i$

بالمقدار a_2 . أي أن المعـلمة التقاطعية في علاقة الانحـدار الأولى أكبر منها في علاقة الانحـدار الثانية بالمقدار a_2 ، بافتراض أن: a_2 > الصفر. ويمكن توضيـح ذلك بالشـكل (1-2).



الشـكل (1- 2) : أثر سـنوات الخبرة والجنـس على الراتب السـنوي

د- من الممكن التوصـل لمعادلتي الانحدار الموضحتين بالشـكل (1-2) من خلال تقدير معادلة انحدار واحـدة هي المعادلة: $Y_i = a_1 + a_2D_2 + bX_i + U_i$

ويتضح من هذه المعادلة أن فئة الأساس هي العاملين الإناث ولذلك فإن المعلمة التقاطعية بهذه الدالة وهي a_1 تمثل المعلمة التقاطية لمعادلة انحدار العاملات. أما المعلمة الانحدارية للمتغير الصوري D_2 وهي a_2 فهي تمثل الفرق بين متوسط رواتب العاملين الذكور ومتوسط رواتب فئة الأساس وهي العاملات. ومن المتوقع أن تكون a_2 > الصفر في هذه الحالة.

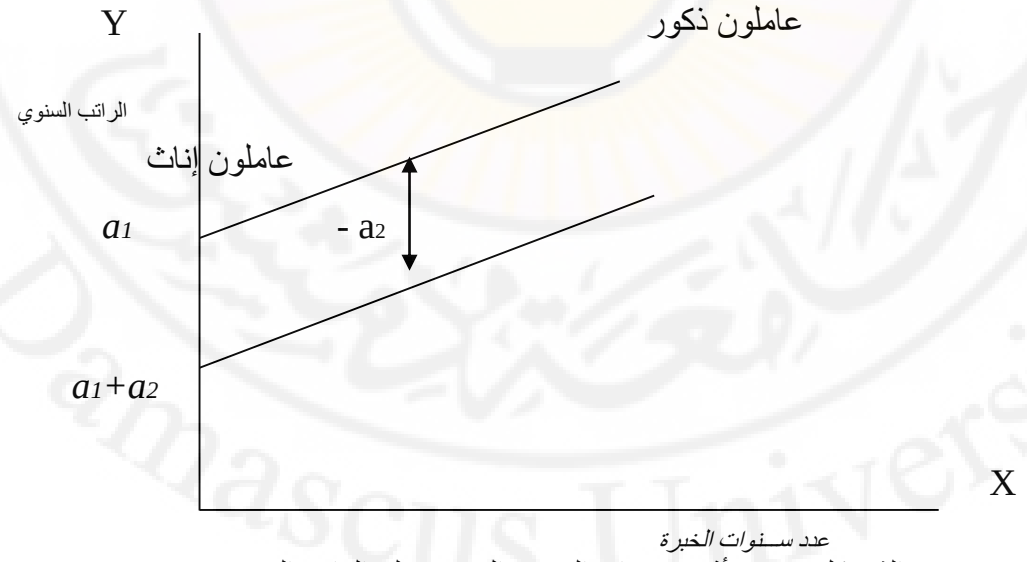
هـ- إذا أردنا اختبار مدى وجود تمييز في الراتب يرجع للجنس فإننا نحتبر الفرض: $a_2 = 0$ في مواجهة الفرض البديل: $a_2 <$ الصفر ، وذلك باستخدام اختبار T . فإذا كانت a_2 المقدرة من بيانات عينة لها معنوية إحصائية فإن هذا يرجح صحة الفرض القائل بوجود تمييز أجري يرجع للجنس (ذكر ، أنثى).

و- يعتبر اختيار الفئة التي تمثل فئة الأساس أمراً تحكمياً يرجع لتقدير الباحث. ففي المثال السابق جعلنا فئة العاملات هي فئة الأساس وفئة العاملين هي فئة المقارنة . ومن الممكن أن نعكس الوضع حيث نحدد من البداية أن: $D_2 = 1$ إذا كان العامل أنثى ، و $D_2 = 0$ إذا كان العامل ذكراً. ومن ثم فإن فئة الأساس تصبح هي العاملين من الذكور حيث يتم إعطاء المتغير الصوري D_2 قيمة صفر بالنسبة لها. ومما سبق نجد أن:

القيمة المتوقعة لراتب الأنثى $= E(Y_i \backslash X_i, D_2=1) = (a_1+a_2) + bX_i$

القيمة المتوقعة لراتب الذكر $= E(Y_i \backslash X_i, D_2=0) = a_1 + bX_i$

ويلاحظ أن الفرق بين القيمة المتوقعة لراتب العاملة وراتب العامل هو a_2 ، غير أن من المتوقع أن تكون a_2 أصغر من الصفر في هذه الحالة. وإذا أردنا اختبار مدى وجود تمييز في الراتب يرجع للجنس في هذه الحالة فإننا نختبر الفرض: $a_2 =$ صفر ، في مواجهة الفرض $a_2 \neq$ صفر . ويمكن توضيح ذلك من خلال الشكل (2-2).



الشكل (2-2): أثر سنوات الخبرة والجنس على الراتب السنوي

2-3-2 متغير كمي ومتغير نوعي بأكثر من صفتين:

لنفترض أن الإنفاق على المواصلات كمتغير تابع Y_i يتأثر بعاملين:

1- مستوى الدخل الفردي كمتغير تفسيري كمي X_i

2- موقع العمل كمتغير نوعي Di ممثلاً في : شمال المدينة D_1

وسط المدينة D_2

جنوب المدينة D_3

وفي هذه الحالة يمكن قياس العلاقة بين الإنفاق على المواصلات وبين مستوى الدخل وموقع العمل من خلال معادلة الانحدار التالية:

$$Y_i = a_1 + a_2D_2 + a_3D_3 + bX_i + U_i$$

حيث: Y_i = الانفاق السنوي على المواصلات للعامل

X_i = الدخل السنوي للعامل

$D_2 = 1$ إذا كان العامل يعمل في وسط المدينة

$D_2 = 0$ إذا كان العامل يعمل في مكان آخر

$D_3 = 1$ إذا كان العامل يعمل جنوب المدينة

$D_3 = 0$ إذا كان العامل يعمل في مكان آخر

ويلاحظ في هذه الحالة أن فئة الأساس هي مجموعة العمال الذين يعملون شمال المدينة ، وذلك فإن المعلمة التقاطعية بالمعادلة السابقة وهي a_1 تشير إلى معلمة التقاطعية بمعادلة انحدار هذه الفئة. وبافتراض أن القيمة المتوقعة للحد العشوائي = صفر ، يمكن أن نحدد التالي من المعادلة:

القيمة المتوقعة للإنفاق على المواصلات من جانب العاملين شمال المدينة =

$$E(Y_i\backslash X_i, D_2 = D_3 = 0) = a_1 + bX_i$$

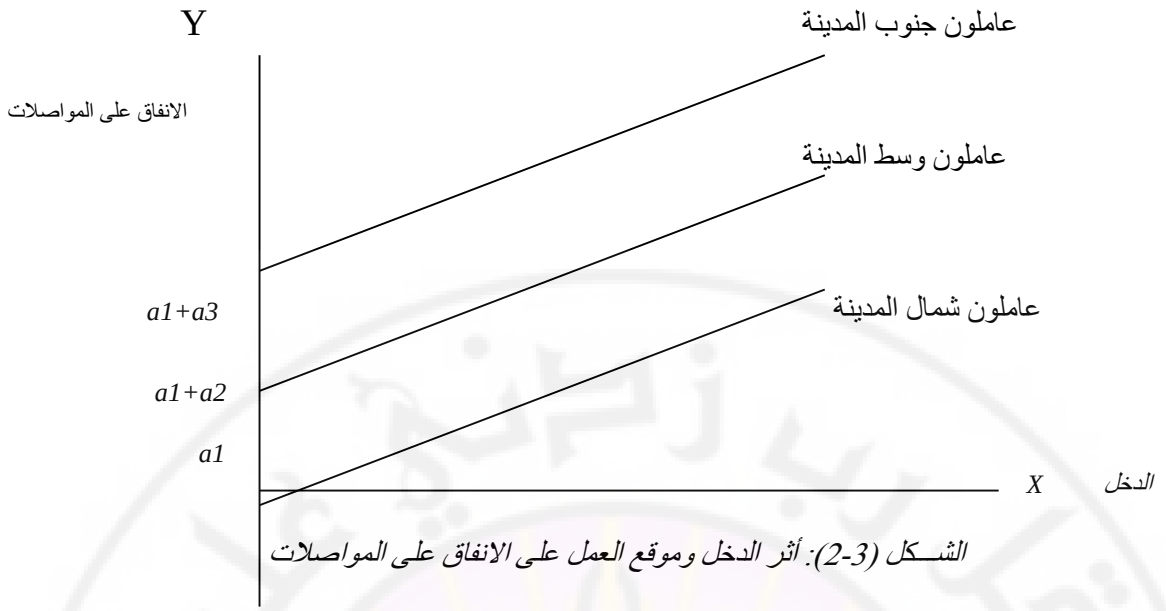
القيمة المتوقعة للإنفاق على المواصلات من جانب العاملين وسط المدينة =

$$E(Y_i\backslash X_i, D_2 = 1 = D_3 = 0) = (a_1 + a_2) + bX_i$$

القيمة المتوقعة للإنفاق على المواصلات من جانب العاملين جنوب المدينة =

$$E(Y_i\backslash X_i, D_3 = 1 = D_2 = 0) = (a_1 + a_3) + bX_i$$

ويمكن تمثيل هذه الحالات الثلاثة السابقة بالشكل (3-2):



الشـــكل (3-2): أثر الدخل وموقع العمل على الانفاق على المواصلات

ومن الممكن اختبار هل هناك اختـلاف جوهري في أعباء المواصـلات المالية بين هذه الفئات الثلاث باسـتخدام اختبار T من خلال اختبار الفروض التـالية:

فرضي العدم : a_2 = صفر ، a_3 = صفر

الفرضين البديلين: a_2 ≠ صفر ، a_3 ≠ صفر

2-3-3 متغير كمي ومتغيرين نوعيين:

إذا افترضنا أن راتب الخريج Y_i يتأثر بثـلاث متغيرات:

أ- عدد سـنوات الخبرة كمتغير كمي X_i

ب- نوع المؤسـسة التي يعمل بها (قطاع عام أو قطاع خاص) كمتغير نوعي H_i

جـ- المسـتوى التعليمي (مؤهل عالي أو مؤهل متوسط) كمتغي نوعي D_i

فمن الممكن صيـاغة دالة الانحـدار التي تمثل هذه العـلاقة فيما يلي:

$$Y_i = a_1 + a_2H_2 + a_3D_2 + bX_i + U_i$$

حيث: H_2= 1 إذا كان الخريج يعمل بالقطاع الخاص

H_2 = 0 إذا كان الخريج يعمل بالقطاع العام

D_2= 1 إذا كان الخريج يحمل مؤهل عالي

D_2 = 0 إذا كان الخريج يحمل مؤهل متوسط

ومن هذه المعلومات يتضح لنا أن فئة الأســاس هي فئة العاملين في القطاع العام من حملة المؤهلات المتوســطة. وبافتراض أن القيمة المتوقعة للحد العشوائي = صفر ، يمكن اشــتقاق معادلات الانحدار التالية من المعادلة الســابقة: $Y_i = a_1 + a_2H_2 + a_3D_2 + bX_i + U_i$

القيمة المتوقعة لراتب الخريج العامل في القطاع العام من حملة المؤهلات المتوسطة =

$$E\ (Y_i \backslash X_i,\ H_2=D_2=0) = a_1 + bX_i$$

القيمة المتوقعة لراتب الخريج العامل في القطاع العام من حملة المؤهلات العليا =

$$E\ (Y_i \backslash X_i,\ H_2=0,\ D2=1) = (a_1 + a_3) + bX_i$$

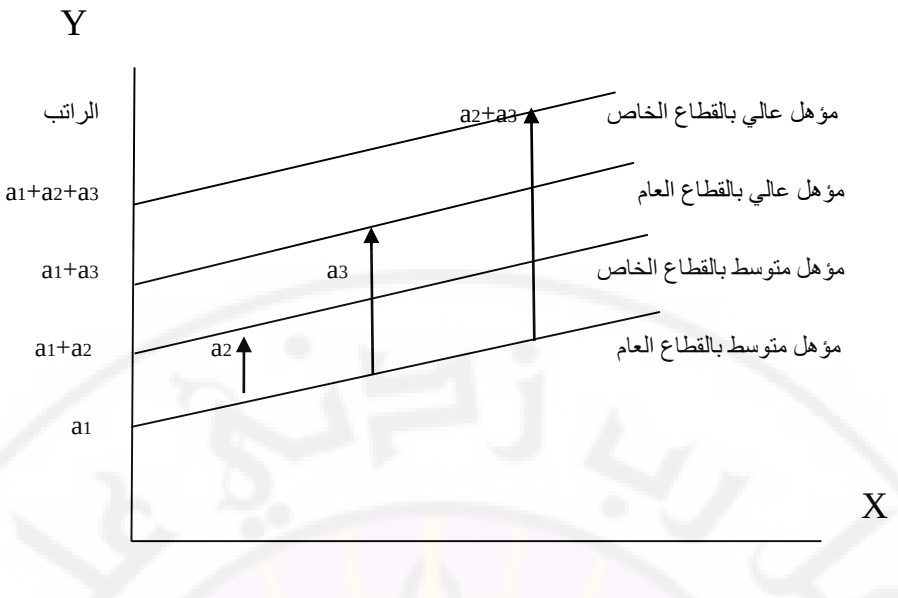
القيمة المتوقعة لراتب الخريج العامل في القطاع الخاص من حملة المؤهلات المتوسطة =

$$E\ (Y_i \backslash X_i,\ H_2=1,\ D2=0) = (a_1 + a_2) + bX_i$$

القيمة المتوقعة لراتب الخريج العامل في القطاع الخاص من حملة المؤهلات العليا =

$$E\ (Y_i \backslash X_i,\ H_2=D_2=1) = (a_1 + a_2 + a_3) + bX_i$$

ويمكن توضيح ذلك باســتخدام الشــكل (4-2).



الشـكل (4-2): أثر المستوى التعليمي ونوع المؤسسة على الراتب

ولو اختبرنا الفرضين: a_2 = صفر ، a_3 = صفر ، في مواجهة الفرضين: a_2 > صفر ، a_3 > صفر

واتضح أن $\hat{a}_2$ لها معنوية إحصائية فإن هذا يعني أن نوع المؤسـسة التي يعمل بها الفرد تؤثر تاثيراً جوهرياً على مستوى الراتب. كما أنه إذا اتضح أن $\hat{a}_3$ لها معنوية إحصائية فإن هذا يعني أن المستوى التعليمي يؤثر بصورة جوهرية على مسـتوى الراتب.

2-3-4 أكثر من متغير كمي وأكثر من متغير نوعي (تعميم):

لنفترض أننا نريد تحديد العوامل التي تؤثر على الدخل الكلي لمدرس ثانوي من خريجي الجامعة Y_i ، واتضح لنا ان العوامل التالية هي التي يعتقد أنها تؤثر في هذا الدخل:

1- الراتب الأسـاسي الذي يأخذه من المدرسـة كمتغير كمي = X_{1i}
2- عدد سـنوات الخبرة في التدريـس كمتغير كمي X_{2i}
3- نوع المدرسة التي يدَّرس فيها (خاصة أو حكومية) كمتغير نوعي H_i
4- مكان المدرسة التي يعمل بها (حضر أو ريف) كمتغير نوعي D_i
5- الجنس (ذكر أو أنثى) كمتغير نوعي F_i

ومما سـبق نسـتنتج أن هناك خمسـة متغيرات تفسيرية يوجد منها متغيرين كميين ، وثلاث متغيرات نوعية. وبناءاً على هذه المعلومات يمكن صياغة معادلة الانحـدار التي تمثل هذه العلاقة فيما يلي:

$$Y_i = a_1 + a_2H_2 + a_3D_2 + a_4F_2 + b_1X_{1i} + b_2X_{2i} + U_i$$

حيث: $H_2 = 1$ إذا كان المدرس يعمل بمدرسة خاصة

$H_2 = 0$ إذا كان المدرس يعمل بمدرسة حكومية

$D_2 = 1$ إذا كان المدرس يعمل بمدرسة في الحضر

$D_2 = 0$ إذا كان المدرس يعمل بمدرسة في الريف

$F_2 = 1$ إذا كان المدرس ذكراً

$F_2 = 0$ إذا كان المدرس أنثى

ومن المعلومات السـابقة يتضح أن فئة الأسـاس هي المدرسـات اللائي يعملن في مدارس حكومية في الريف. وإذا قمنا بتقدير العلاقة السابقة من بيانات واقعية واتضح لنا أن هذه العلاقة كانت على النحو التالي:

$$Y_i = 2 + 3H_2 + 5D_2 + 4F_2 + 0.75X_{1i} + 1.5X_{2i}$$

فهذا يعني أن:

القيمة المتوقعة لدخل الساعة لمُدرسة تعمل بمدارس حكومية في الريف =

$$E\,(Y_i \backslash X_{1i}, X_{2i}, H_2=D_2=F_2=0) = 2 + 0.75X_{1i} + 1.5X_{2i}$$

القيمة المتوقعة لدخل الساعة لمُدرس يعمل بمدارس خاصة بالحضر =

$$E\,(Y_i \backslash X_{1i}, X_{2i}, H_2=D_2=F_2=1) = 14 + 0.75X_{1i} + 1.5X_{2i}$$

ماذا يعني ذلك ؟

وقل رب زدني علما
جامعة دمشق
Damascus University

أهم اســتخدامات المتغيرات الصــورية

Main uses of Dummy Variables

اســتعرضنا في المحاضرتين الســابقتين طرق تقدير معالم نموذج الانحدار الخطي عندما يحتوي على متغيرات صورية *Dummy Variables* مســتقلة (تفســيرية). وقد درســنا ثلاث حالات أســاســية:
1- متغير صوري مستقل واحد 2- أكثر من متغير صوري مستقل 3- متغيرات كمية وصورية.

وعلى الرغم من التركيز على الجزء النظري وكيفية تقدير معالم معادلة الانحدار في الحالات الســابقة, إلا أننا لم نتطرق إلى التطبيقـات العملية التي يمكن أن تواجهنا ونســتفيد من اســتخدام المتغيرات الصورية في دراســتها. ســنتناول في هذه المحاضرة دراســة أهم اســتخدامات المتغيرات الصورية بشــيء من التفصيل.

1) قياس التغيـر في الميـول الحـدية:

لقد وضحنا سـابقاً كيف يؤثر المتغير الصـوري على المعلمة التقاطعية لعـلاقة الانحدار, ولكن لم نوضح كيف يؤثر على المعلمة الانحدارية فيجعلها تختلف من علاقة انحدار لأخرى. ولكي يتم توضيح كيف يؤثر المتغير الصوري على الميل الحدي للدالة أو المعلمة الانحدارية نورد المثال التالي:

لنفرض أننا نريد دراسـة مقارنة دالة الاسـتهلاك في الريف بنظيرتها في الحضـر. فمن الأسـاليب التي يمكن إتباعها لعمل ذلك هو أن نقوم بأخذ عينة من أسـر الريف حجمها (N_1) ثم نقدر دالة الاسـتهلاك منها، ونأخذ عينة من أسـر الحضر حجمها (N_2) ثم نقدر دالة الاسـتهلاك منها، فنحصل على دالتي الاسـتهلاك التاليتين:

الريف اسـتهلاك دالة $Y_r = a_1 + a_2X_r$

الحضر اسـتهلاك دالة $Y_u = b_1 + b_2X_u$

حيث: Y = الإنفاق الاستهلاكي

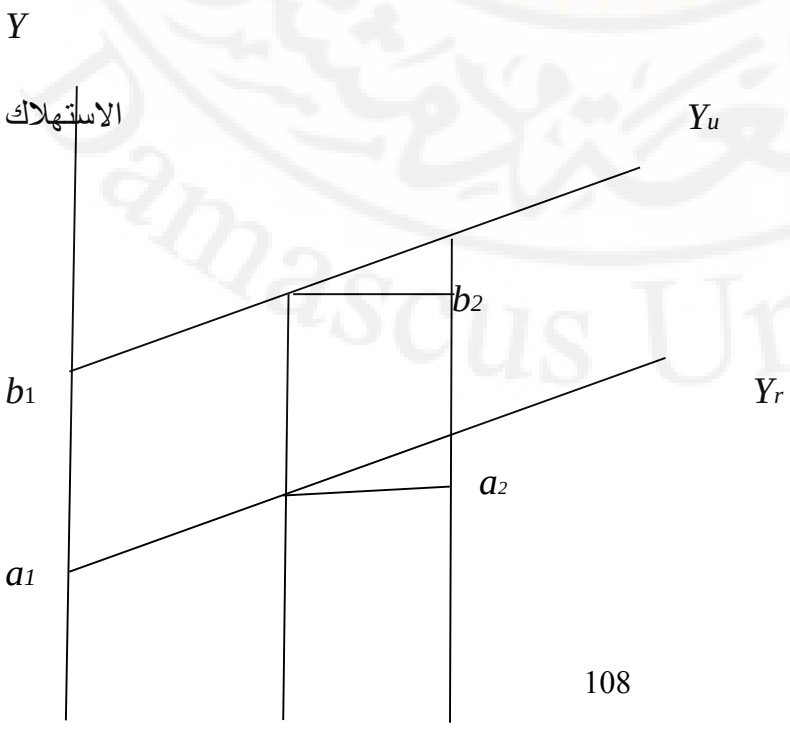
X = الدخل

r تشـير للريف u تشـير للحضر

وبمقارنة دالتي الاسـتهلاك السـابقتين (الريف والحضر) نجد أن هناك أربع احتمالات ممكنة:

1- $a_1 = b_1$ ، $a_2 = b_2$ وبالتالي تكون دالتي الاسـتهلاك متماثلتين تماماً.

2- $a_1 \neq b_1$ ، $a_2 = b_2$ وبالتالي تكون دالتي الاسـتهلاك مختلفتين فقط في المعلمة التقاطعية التي تمثل حد الكفاف ، ولكنهما متماثلتين في الميل الحدي للاسـتهلاك. الشـكل (7-1) يوضح ذلك.

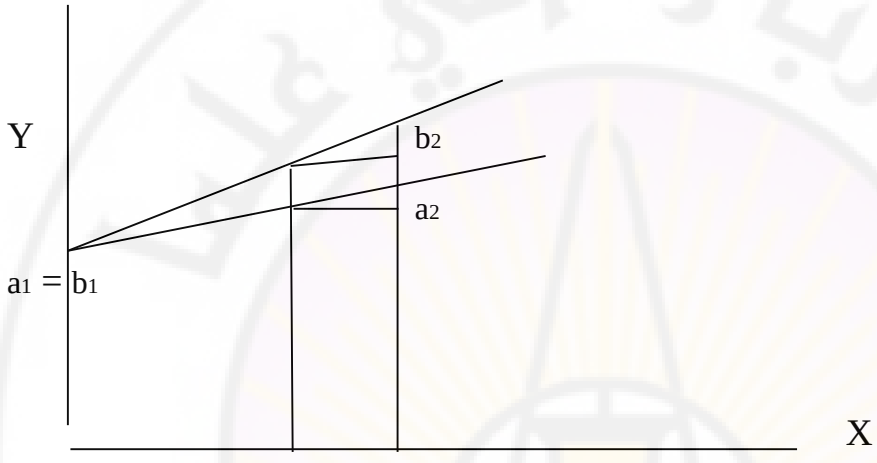


الدخل

X ———————————————————————

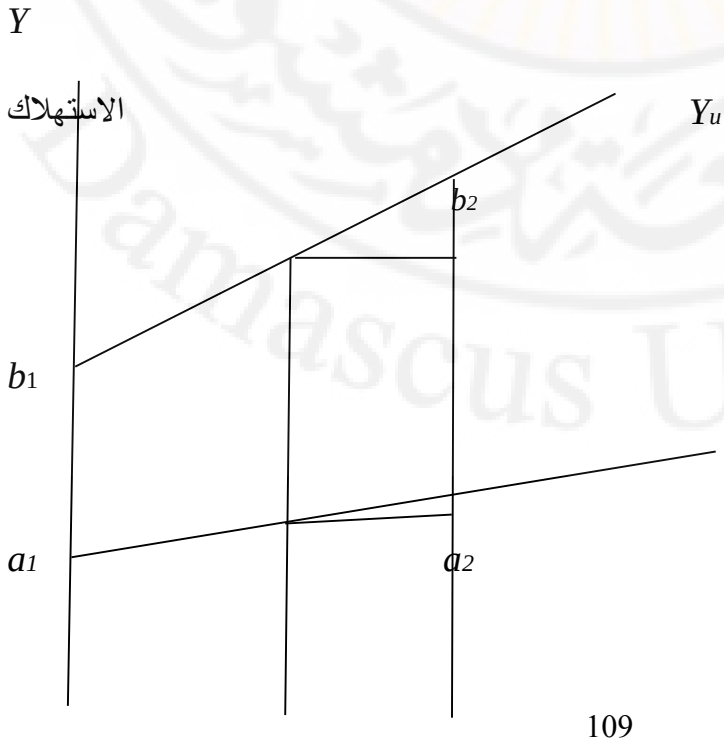
الـشـكل (7-1) اختلاف حد الكفاف بين الحضر والريف

3- $a_1 = b_1$ ، $a_2 \neq b_2$ ومن ثم تكون دالتي الاستهلاك مختلفتين في الميل الحدي للاســتهلاك ومتماثلتين في المعلمة التقاطعية. الشــكل (7-2).



الشكل (7-2): اختلاف الميل الحدي للاستهلاك بين الحضر والريف

4- $a_1 \neq b_1$ ، $a_2 \neq b_2$ أي أن دالتي الاســتهلاك مختلفتين في كل من المعلمة التقاطعية والمعلمة الانحدارية التي تمثل الميل ، ويتضح هذا من الشــكل (7-3).



الدخل

X ———————————————————————

الشـكل (7-3) اختلاف حد الكفاف والميل الحدي للاستهلاك بين الحضر والريف

ويســاعد اســتخدام المتغيرات الصورية على اختبار كل الاحتمالات الســابقة من خلال تقدير معادلة انحدار واحدة بدلاً من تقدير معادلتين ومقارنتهما. وفي هذه الحالة يتم تقدير معادلة انحدار واحدة من كل البيانات المتاحة عن الريف والحضر والتي تمثل عينة كبيرة حجمها $N= n_1+ n_2$. وتأخذ معادلة الانحدار في هذه الحالة الصيغة التالية:

$$Y_i= a_1+ a_2D_2 + b_1X_i + b_2D_2X_i + U_i$$

حيث: Y_i = الإنفاق الاســتهلاكي

X_i = الدخل

D_2 = موطن الأســرة.

$D_2 = 1$ إذا كانت الأســرة تقطن في الحضر.

$D_2 = 0$ إذا كانت الأســرة تقطن في الريف.

ومن هذه المعلومات يتضح أن فئة الأســاس هي الأســر التي تقطن في الريف.
وتحتوي المعادلة الســابقة على متغير مركب D_2X_i يترتب على وجوده اختلاف ميول الانحدار المختلفة التي يمكن اشــتقاقها من المعادلة: $Y_i= a_1+ a_2D_2 + b_1X_i + b_2D_2X_i + U_i$ حيث يتضح من هذه المعادلة أن:

القيمة المتوقعة للإنفاق الاســتهلاكي للأســرة التي تقطن في الريف =

$$E (Y_i\backslash X_i, D_2 = 0) = a_1+ b_1X_i$$

القيمة المتوقعة للإنفاق الاســتهلاكي للأســرة التي تقطن في الحضر =

$$E (Y_i\backslash X_i, D_2 =1) = (a_1+ a_2) + (b_1+ b_2) Xi$$

وبمقارنة دالتي الاســتهلاك الســابقتين (لإنفاق الأســر التي تقطن الريف مع الأســر التي تقطن الحضر) يتضح ما يلي:

(1) أن حد الكفاف للاســتهلاك في الريف الريف هو a_1 وفي الحضر هو $(a_1 + a_2)$. اي أن مقدار اختلاف حد الكفاف في الحضر عنه في الريف = a_2. وللتأكد فيما إذا كان هذا الاختلاف جوهرياً أم لا يتعين اختبـار: فرض العدم: a_2 = صفر ، في مواجهة الفرض البديل a_2> صفر.

(2) أن الميل الحدي للاســتهلاك بالريف =b_1 ، أما في الحضر فهو يســاوي $(b_1 + b_2)$ ، ومن ثم فإن مقدار الاختلاف في الميل الحدي للاســتهلاك = b_2 . ولاختبار ما إذا كان هذا الاختلاف جوهرياً أم لا ، نختبر فرض العدم: = b_2صفر ، في مواجهة الفرض البديل: b_2> صفر.

(3) أن صياغة معادلة الانحدار على النحو الســابق:

$$Y_i = a_1 + a_2D_2 + b_1X_i + b_2D_2X_i + U_i$$

يتضمن افتراض أن دالة الاســتهلاك في الحضر تختلف عنها في الريف في كل من المعلمتين التقاطعية والانحدارية. أما إذا افترضنا أن الدالتين مختلفتين في المعلمة الانحدارية فقط ومتماثلتين في المعلمة التقاطعية فإن معادلة الانحدار تأخذ الصيغة التالية:

$$Y_i = a_1 + b_1X_i + b_2D_2X_i + U_i$$

ومن هذه المعادلة يتضح أن:

القيمة المتوقعة للإنفاق الاســتهلاكي للأســر التي تقطن في الريف =

$$E(Y_i \backslash X_i, D_2 = 0) = a_1 + b_1 X_i$$

القيمة المتوقعة للإنفاق الاســتهلاكي للأســر التي تقطن في الحضر =

$$E(Y_i \backslash X_i, D_2 = 1) = a_1 + (b_1 + b_2) X_i$$

ومن ثم فإن الاختلاف في هذه الحالة يكون فقط في المعلمة الانحدارية ومقداره b_2. ويلاحظ عموماً أن من أهم مزايا اســتخدام المتغيرات الصــورية أنها تمكننا من تقدير أكثر من معادلة واحدة. ولاشــك أن نتائج المعادلة:

$$Y_i = a_1 + a_2D_2 + b_1X_i + b_2D_2X_i + U_i$$

ســوف تكون هي نفســها نتائج المعادلتين:

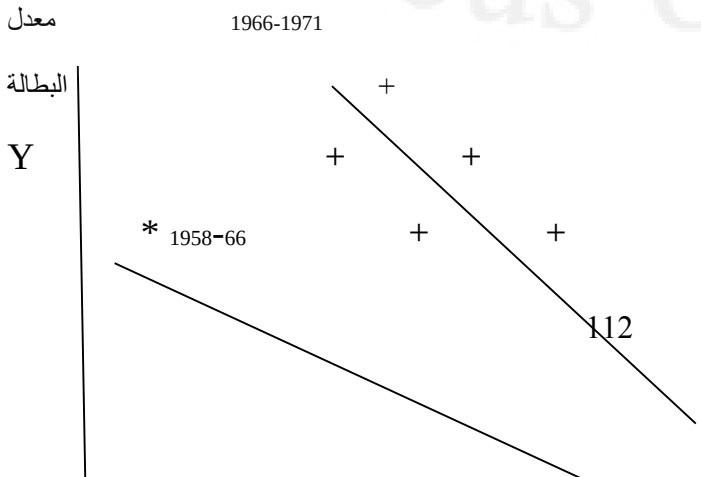
$Y_r = a_1 + a_2X_r$ الريف استهلاك دالة

$Y_u = b_1 + b_2X_u$ الحضر استهلاك دالة

2) قياس التغيرات الهيكلية:

عادة ما تواجهنا حالة تقدير دالة انحدار من بيانات سلسلة زمنية تخص فترة زمنية طويلة نسبياً تمتد على سبيل المثال من عام 1900 إلى العام 1995، ففي هذه الحالة فإن استخدام المتغيرات المستقلة (التفسيرية) الكمية فقط لتقدير هذه الدالة لا يعطي صورة حقيقية لسلوك الظاهرة موضوع البحث خلال هذه الفترة. ففي الفترات الزمنية الطويلة تحدث هناك تغيرات جوهرية نتيجة لبعض الأحداث الكبيرة مثل الحرب العالمية الأولى أو الثانية او الكساد العالمي الكبير عام 1929 أو قيام الثورات أواحلال السلام. ولاشك أن مثل هذه الأحداث الكبيرة تقسم التاريخ إلى مراحل وتجعل سلوك الظواهر مختلفاً في كل مرحلة من هذه المراحل. ومن ثم يصبح من الضروري تقدير معادلة انحدار مستقلة لكل مرحلة لاختلاف طبيعة الظاهرة بينها. ولكن يلاحظ في هذه الحالة أن استخدام المتغيرات الصورية يساعدنا على تقدير معادلة انحدار واحدة لكل المراحل ويمّكننا من اشتقاق معادلة مستقلة لكل مرحلة مع تحديد وجه الاختلاف في سلوك الظاهرة عبر الفترات.

ومن الأمثلة على ذلك ما قام به العالم الاقتصادي Damodar Gujarati في دراسته للعلاقة بين معدل البطالة ومعدل الوظائف الشاغرة في بريطانيا خلال فترة طويلة نسبياً (1958-1971). فمن خلال تمثيل شكل الانتشار الخاص بالعلاقة بين هذين المتغيرين بيانياً اتضح للمؤلف أن تغيراً هيكلياً أو جذرياً حدث في هذه العلاقة منذ العام 1966 كما يوضح الشكل (7-4). حيث أنه من الواضح بالشكل أن الخط الممثل للعلاقة بين المتغيرين قد انتقل لأعلى بعد عام 1966 وأصبح مختلفاً في شكله خلال الفترة 1966-1971 , وذلك من حيث كل من المعلمتين التقاطعية والانحدارية. ويلاحظ أن اختلاف شكل العلاقة على هذا النحو يعني أنه لكل معدل وظائف شاغرة أصبح هناك مستوى أعلى للبطالة. وبمعنى آخر لكل وظيفة شاغرة أصبح هناك عدد أكبر من المتعطلين.





الشــكل (7-4): التغير الهيكلي

وبالبحث في أسـباب هذه الظاهرة اتضح للمؤلف أن حكومة العمال في بريطانيا قد عَّدلت قانون التأمينات عام 1966 بما يتيح مزايا أكثر للعمال الذين يعانون من بطالة ، الأمر الذي شـجع العمال العاطلين على قضـاء وقت أطول في البحث عن وظيفة جديدة مع التحرر من ضغط الحاجة.

ولتقدير معادلة الانحـدار الممثلة لهذه العلاقة خلال كل الفترة 1958-1971 اسـتخدم المؤلف الصيغة التالية:

$$Y_i = a_1 + a_2D_2 + b_1X_i + b_2D_2X_i + U_i$$

حيث: Y_i = معدل البطالة % ، X_i = معدل الوظائف الشاغرة %

D_2 = 1 للفترة من تشـرين أول 1966-1971

D_2 = 0 للفترة من تشـرين أول 1958 – أيلول 1966

ويتضح من هذه المعلومات أن الباحث اعتبر الفترة 1958-1966 هي فترة الأســاس. كما أن الاختلاف بين الفترتين كان في كل من المعلمة الانحدارية والمعلمة التقاطعية. وبتقديره للعلاقة حصل على الصيغة التالية:

$$Y_i = 2.75 + 1.15D_2 – 1.53\ X_i – 0.85\ D_2X_i + U_i$$

فمن المعادلة الســابقة نجد أن:

العلاقة المقدرة للفترة 1958-1966 =

$$E\ (Y_i\backslash X_i,\ D_2= 0) = 2.75 – 1.53\ X_i$$

العلاقة المقدرة للفترة 1966-1971 =

$$E\ (Y_i\backslash X_i,\ D_2= 1) = (2.75+1.15) – (1.53+0.85)\ X_i$$

$$= 3.9 – 2.38X_i$$

كما يوجد اختلاف جوهري في المعلمة التقاطعية بالمقدار $a_2 = 1.15$ وآخر في المعلمة الانحدارية بالمقدار $b_2 = 0.85$ بين علاقتي الفترتين.

3) قياس أثر التقلبات الموسـمية:

تعتبر التقلبات الموســمية من بين العوامل التي تؤثر في الظواهر الاقتصادية إلى جانب العوامل الأخرى. فيلاحظ مثلاً أن الموجة الباردة في فصل الشـتاء تزيد من الطلب على الملابس الصـوفية والثقيلة بوجه عام ، كما أن الموجة الحارة في فصل الصـيف تقلل من الطلب على هذا النوع من الملابس. ويلاحظ أيضاً أن الموجة الحارة في فصـل الصيف تزيد من الطلب على المثلجات والآيس كريم والمشـروبات الباردة بمحتلف أنواعها ، في حين أن الموجة الباردة بفصل الشتاء تقلل من الطلب على هذه المنتجات. وفي المواسم والأعياد يزداد الطلب على الحلوى والهدايا بصورة ملحوظة بالمقارنة مع الأوقات الأخرى من السـنة. ومن الممكن قياس أثر التقلبات الموسمية على المتغيرات الاقتصادية باسـتخدام المتغيرات الصورية. ويلاحظ في هذه الصـدد أن التقلبات الموسـمية قد تؤثر على المعلمة التقاطعية لعلاقة الانحدار فقـط ، أو قد تؤثر على المعلمة الانحدارية فقط ، أو قد تؤثر على كليهما معاً.

3-1 تأثير التقلبات الموسـمية على المعلمة التقاطعية:

إذا فرضنا أن مسـتوى الأرباح للشـركات الصناعية Y_i يتأثر بعاملين هما: حجم المبيعات ربع السـنوية X_i ، والتقلبات الموسـمية D_i ، وأن هذه التقلبات الموسـمية تحدث عبر أربع فصـول بالعام هي: الربيـع D_1 ، الصيف D_2 ، الخريف D_3 ، والشـتاء D_4 ، فمن الممكن صياغة دالة الانحدار الممثلة لهذه العلاقة كما يلي:

$$Y_i = a_1 + a_2 D_2 + a_3 D_3 + a_4 D_4 + b_1 X_i + u_i$$

حيث: $D_2 = 1$ في فصل الصيف ، و $D_2 = 0$ في أي فصل آخر.

$D_3 =$ 1 في فصـل الخريف ، $D_3 = 0$ في أي فصل آخر.

$D_4 =$ 1 في فصل الشـتاء ، $D_4 = 0$ في أي فصل آخر.

ومن الواضح أن الفصـول الأربعة مانـعة بالتبادل. فوجودنا في فصل الصيف يمنع من وجودنا في أي فصـل آخر ، ولذا نعطي في هذه الحالة فصل الصـيف قيمة 1 ، والفصول الأخرى القيمة صـفر، وهكـذا. ويلاحـظ هنا أن فصل الأسـاس هو فصل الربيـع ولذا فإن المعلمة التقاطعية a_1 تخصه. فالمعلمة a_1 تشـير إلى مسـتوى الأرباح عندما $D_2 = D_3 = D_4$ = صـفر ، وذلك لا يكون إلا في فصل الربيع. ومن المعادلة السـابقة يمكن اشـتقاق معادلات الانحدار الخاصة بالأرباح في الفصول الاربعة كما يلي:

القيمة المتوقعة للأرباح في فصل الربيع =

$$E\,(Y_i \backslash X_i, D_2 = D_3 = D_4 = 0) = a_1 + b_1 X_i$$

القيمة المتوقعة للأرباح في فصل الصيف =

$$E\,(Y_i \backslash X_i, D_2 = 1, D_3 = D_4 = 0) = (a_1 + a_2) + b_1 X_i$$

القيمة المتوقعة للأرباح في فصل الخريف =

$$E\,(Y_i \backslash X_i, D_3 = 1, D_2 = D_4 = 0) = (a_1 + a_3) + b_1 X_i$$

القيمة المتوقعة للأرباح في فصل الشـتاء =

$$E\,(Y_i \backslash X_i, D_4 = 1, D_2 = D_3 = 0) = (a_1 + a_4) + b_1 X_i$$

وهكذا فإن:

a_2 = الفرق بين مستوى الأرباح في فصل الصيف وفصل الربيع.

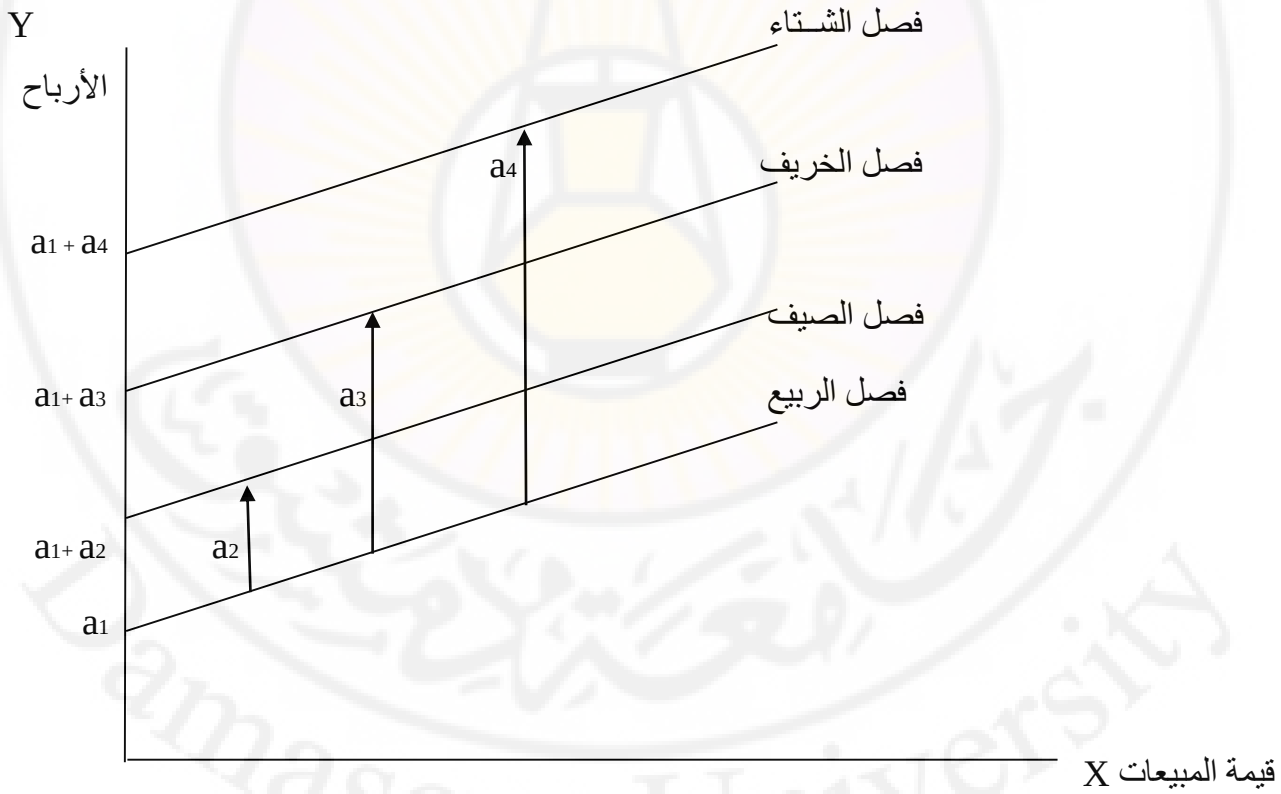
a_3 = الفرق بين مستوى الأرباح في فصل الخريف وفصل الربيع.

a_4 = الفرق بين مستوى الأرباح في فصل الشتاء وفصل الربيع.

ويمكن اختبار ما إذا كانت هذه الفروق جوهرية أم لا من خلال اختبار فروض العدم التالية:

a_2 = صفر ، a_3 = صفر ، a_4 = صفر ، في مواجهة الفروض البديلة:

a_2 > صفر ، a3 > صفر ، $a4$ > صفر ، ويوضح الشكل (7-5) معادلات الانحدار السابقة.



الشكل (7-5) أثر التقلبات الموسمية على الأرباح

2-3 تأثير التقلبات الموسمية على المعلمة الانحدارية:

إذا افترضنا أن التقلبات الموسـمية تؤثر على الميول الحدية ، أي على المعلمات الانحدارية فقط ، فإن معادلة الانحـدار التي تصف العلاقة بين الأرباح من ناحية والتقلبات الموسـمية وحجم المبيعات من ناحية أخرى تأخذ الصيغة التالية:

$$Y_i = a_1 + b_1 X_i + b_2 D_2 X_i + b_3 D_3 X_i + b_4 D_4 X_i + u_i$$

وبافتراض أن فصل الربيع هو فصل الأســـاس نجد أن:

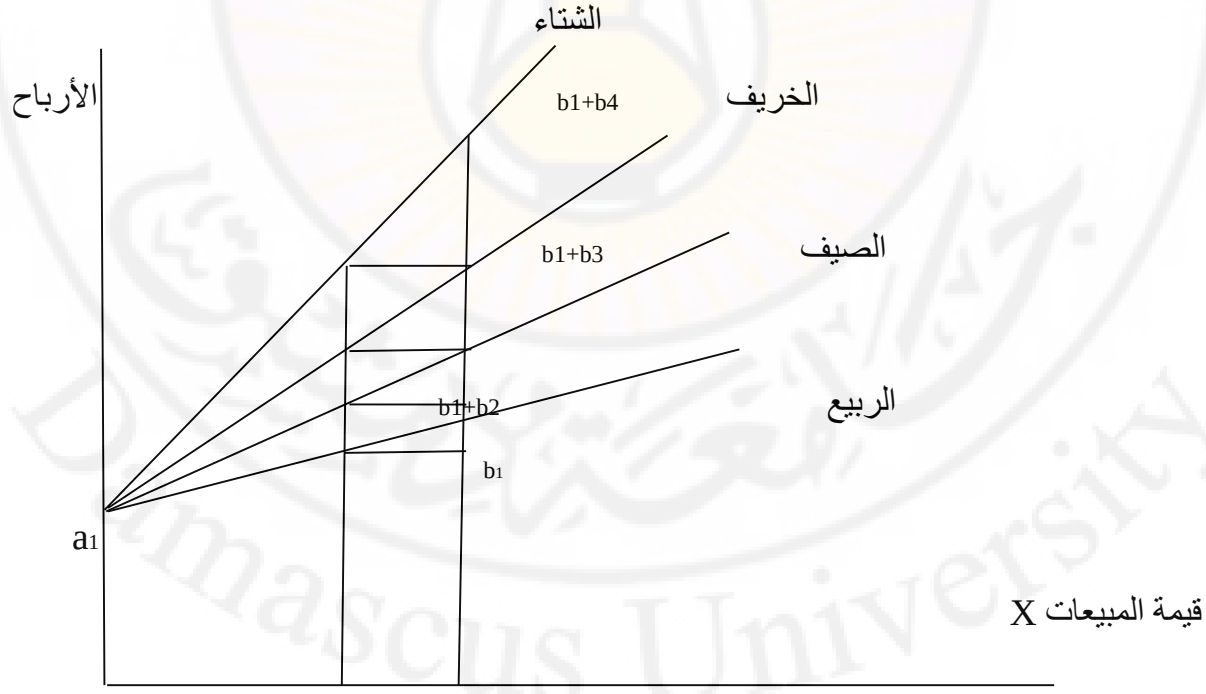
القيمة المتوقعة للأرباح في فصل الربيع $= Y_1 = a_1 + b_1 X_i$

القيمة المتوقعة للأرباح في فصل الصيف $= Y_2 = a_1 + (b_1 + b_2) X_i$

القيمة المتوقعة للأرباح في فصل الخريف $= Y_3 = a_1 + (b_1 + b_3) X_i$

القيمة المتوقعة للأرباح في فصل الشتاء $= Y_4 = a_1 + (b_1 + b_4) X_i$

ومن ثم فإن التقلبات الموسـمية تؤثر فقط على المعلمات الانحدارية الخاصة بالفصول المختلفة دون المعلمة التقاطعية. ويمكـن توضيح الاختلاف بين دوال الأرباح بالنسـبة للفصـــول المختلفة في هذه الحالة في الشـكل (6-7)



الشـــكل (7-6) تأثير التقلبات الموسـمية على المعلمة الانحدارية

3-3 تأثير التقلبات الموسـمية على المعلمتين التقاطعية والانحدارية معاً:

إذا فرضنا أن التقلبات الموســمية تؤثر ليــس فقط على مســتوى الربح فتجعله في فصل مختلفاً عنه في فصل آخر، ولكن أيضاً على معدل الزيادة فيه فتجعله ينمو في فصل بمعدل مختلف عنه في فصل آخر، فإن معادلة الانحدار التي تصف العلاقة بين الربح وحجم المبيعات في ظل هذه الظروف تأخذ الصيغة التالية:

$$Y_i = a_1 + a_2D_2 + a_3 D_3 + a_4D_4 + b_1Xi + b_2D_2Xi + b_3 D_3 X_i + b_4 D_4 X_i + u_i$$

ومن المعادلة نجد أن:

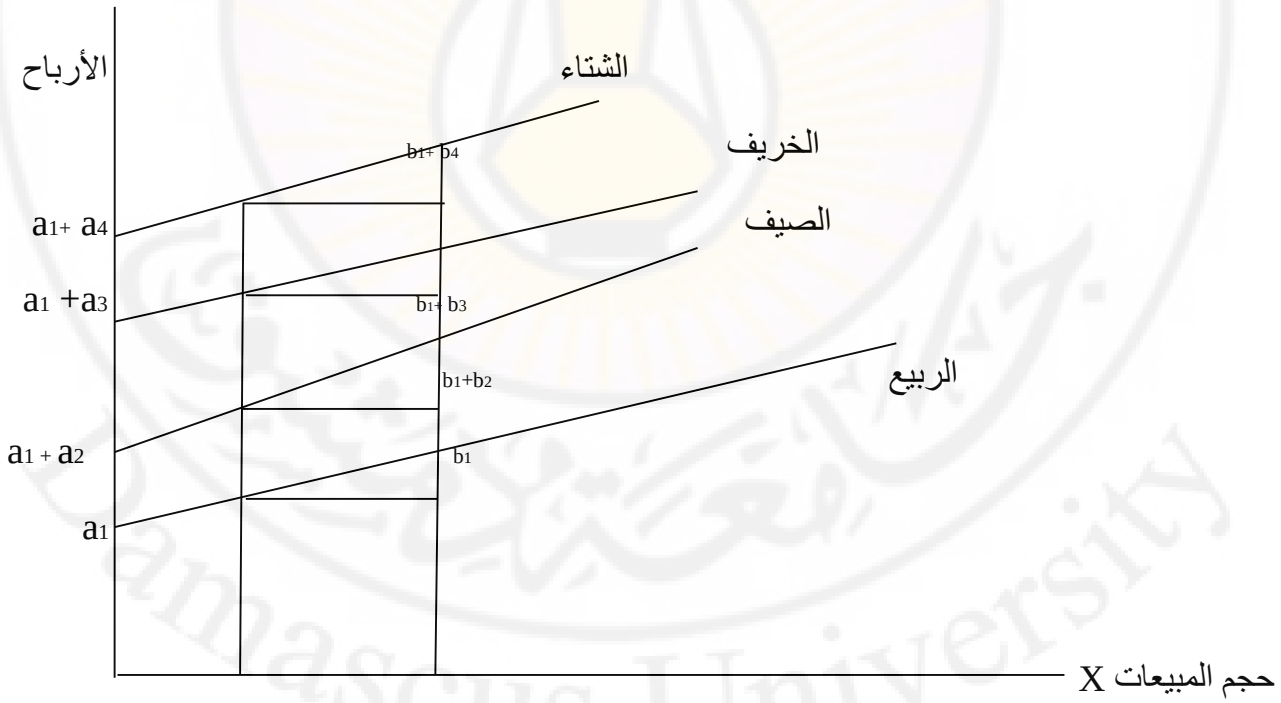
القيمة المتوقعة للأرباح في فصل الربيع $= Y_i = a_1 + b_1X_i$

القيمة المتوقعة للأرباح في فصل الصيف $= Y_i = (a_1+a_2) + (b_1+b_2)X_i$

القيمة المتوقعة للأرباح في فصل الخريف $= Y_i = (a_1+a_3) + (b_1+b_3)X_i$

القيمة المتوقعة للأرباح في فصل الشتاء $= Y_i = (a_1+ a_4) + (b_1+b_4)X_i$

ومن ثم تكون التقلبات الموســمية قد أدت إلى اختلاف كل من المعلمة التقاطعية والمعلمة الانحدارية لدالة الربح من فصل لآخر. ويمكن توضيح ذلك بالشــكل (7-7).



الشــكل (7-7) تأثير التقلبات الموسمية على كل من المعلمة التقاطعية والانحدارية

المتغيرات المتباطئة

Lagged Variables

1- مقدمة:

لقد كانت نماذج الانحدار التي استخدمناها في محاضرات سابقة تفترض أن التغير في المتغير المستقل يؤثر تأثيراً مباشراً وفورياً على المتغير التابع. وهي بذلك لم تعطِ أي اعتبار للفجوة الزمنية التي تمر قبل أن يبدأ المتغير التابع في الاستجابة للتغير في المتغير المستقل (التفسيري) ، أو للفترة الزمنية التي يحدث عبرها التغير في المتغير التابع كاستجابة لتغير ما في المتغير المستقل. ويلاحظ عموماً أن التغير في المتغيرات المستقلة كثيراً ما لا يحدث آثاره بصورة مباشرة وفورية على الظواهر الاقتصادية ، وإنما يحتاج الأمر لفترة زمنية قد تكون طويلة حتى يمكن لهذه المتغيرات أن تمارس آثارها كاملة على مثل هذه الظواهر. فتخفيض قيمة العملة مثلاً لا يمارس آثاره مباشرة على الصادرات والواردات وإنما يحتاج لفترة زمنية طويلة نسبياً حتى تتم آثاره كاملة ، وكذلك الأمر بالنسبة لتغير معدلات الضرائب وما تمارسه من آثار على الاستهلاك أو الانتاج أو الاستثمار.

ومن هنا ظهرت الحاجة لضرورة استخدام النماذج ذات الفجوات الزمنية أو نماذج المتغيرات المتباطئة ، وهي نماذج تستخدم عندما توجد هناك متغيرات مستقلة تمتد آثارها عبر عدد من الفترات الزمنية.

2- أنواع نماذج المتغيرات المتباطئة:

يمكن تقسيم نماذج المتغيرات المتباطئة وفقاً لمعيارين: نوع المتغير المستقل المتباطئ ، وطول الفجوة الزمنية.

أ- نوع المتغير المستقل المتباطئ:

تنقسم نماذج المتغيرات المتباطئة لنوعين وفقاً للمتغير المستقل المتباطئ:

(1) نماذج المتباطئات الموزعة *Distributed-lag models*

وهي نماذج تحتوي على قيم سابقة Past values لمتغيرات خارجية كمتغيرات مستقلة ، مثال ذلك دالة الاستثمار التالية:

$$I_t = a + b_1 Y_t + b_2 r_t + b_3 r_{t-1} + U_t$$

I_t= حجم الاستثمار بالفترة الحالية

Y_t= مســتوى الناتج الكلي بالفترة الحالية

r_t= ســعر الفائــدة بالفترة الحالية

r_{t-1} = ســعر الفائــدة بالفترة الســابقة

وباعتبار أن ســعر الفائدة متغير خارجي فإن الاســتثمار الحالي يكون دالة في قيمة ســعر الفائدة في الفترة الحالية وقيمته في الفترة الســابقة ، ومن ثم فإن هذا النموذج يكون ذو متغير مســتقل متباطئ موزع.

(2) نماذج الانحدار الذاتي *Autoregressive models*

وتسمى أيضاً نماذج المتغير التابع المتباطئ ، وهي نماذج تحتوي على قيم ســابقة لمتغيرات تابعة كمتغيرات تفــسيرية ، بمعنى أنه يتم اســتخدام المتغير التابع بصــورته المتباطئة ضمن المتغيرات المســتقلة على الجانب الأيمن من معادلة الانحــدار ، مثال ذلك دالة الطـلب التاليـة:

$$Q_t = a + b_1Y_t + b_2Q_{t-1} + b_3P_t + U_t$$

حيث : Q_t = الكمية المطلوبة من الســلعة في الفترة الحالية

Y_t = دخــل الفتــرة الحالية

Q_{t-1}= الكمية المطلوبة من الســلعة في الفترة الســابقة

P_t = ســعر الســلعة في الفترة الحالية

وتصــف هذه الدالة حالة الطـلب على الســلع المعمرة أو الســلع غير المعمرة التي يتكون لدى المســتهلك عادة عند اســتهلاكها (كالســجائر أو البن وغيرها). فمثل هذه الســلع تتأثر الكمية المطــلوبة منها في الفترة الحالية بالكمية المطــلوبة منها في الفترات الســابقة. ويلاحــظ هنا أن الكمية المطـلوبة في الفترة الســابقة تُســتخدم كمتغير تفســيري (مســتقل).

ب- طول الفجوة الزمنية:

تنقســم نماذج المتغيرات المتباطئة وفقاً لطول الفجوة الزمنية الى نوعين:

(أ) نماذج ذات عدد محدود من الفجوات:

وفي هذه الحالة يمتد أثر المتغير التفســيري عبر عدد محدد من الفترات أقل من ما لا نهاية. ومن الأمثلة على ذلك الصيغة التالية:

$$Y_t = a + b_1X_t + b_2X_{t-1} + \ldots\ldots + b_mX_{t-m} + U_t$$

حيث أن عدد الفترات التي يمتد عبرها تأثير المتغير التفسـيري X= m ويلاحظ هنا أن:

$b_1 = \frac{\partial yt}{\partial xt}$ = تأثير التغير في x بمقدار وحدة واحدة على y خلال الفترة الحالية

$b_2 = \frac{\partial yt}{\partial xt-1}$ = تأثير التغير في قيمة xبالفترة السابقة بمقدار وحدة واحدة على y_tخلال الفترة الحالية

$b_m = \frac{\partial yt}{\partial xt-m}$ = تأثير التغير في قيمة x بالفترة t-m بمقدار وحدة واحدة على قيمة y_tخلال الفترة الحالية

$\sum_{i=0}^{m} bi =$

مجموع تأثيرات التغير في قيمة x بمقدار وحدة واحدة على قيمة Y خلال فترة من الزمن طولها m

(ب) نماذج ذات عدد لا نهائي من الفجوات:

وفي هذه الحالة يمتد أثر المتغير المستقل المتباطئ (ذو الفجوة الزمنية) عبر عدد غير محدود من الفترات الزمنية ، وتأخذ معادلة الانحدار الصيغة التالية:

$$Y_t = a + b_1X_t + b_2X_{t-1} + b_3X_{t-2} + \ldots\ldots\ldots + U_t$$

وبالطبع حتى يمكن تقدير مثل هذه النماذج لابد من وضع قيود معينة على عدد الفجوات الزمنية.

3- أمثلة اقتصادية لنماذج المعادلات ذات المتغيرات المتباطئة:

1-3 دالة الاسـتهلاك:

يلاحظ عموماً أن الشـخص لا يغير من عاداته الاسـتهلاكية بصورة سـريعة أو فورية ، وإنما يقتضي الأمر أن يمر وقتاً طويلاً نسـبياً قبل أن تتغير هذه العـادات ، وغالباً ما يتم هذا بصـورة تدريجية. فإذا افترضنا أن شـخصاً ما زاد دخله السـنوي بمقدار 1000 دولار بصفة دائمة فإنه من المتوقع أن يزداد اسـتهلاكه بصورة تدريجية عبر فترة زمنية طويلة نسـبياً. فهو قد يزيد اسـتهلاكه في السـنة الأولى بمقدار 400 دولار وفي السـنة الثانية بمقدار 300 دولار أخرى وفي السـنة الثالثة بمقدار 200 دولار

أخرى. ومن ثم فإن الزيـادة الدائمة في الدخـل بمقدار 1000 دولار تـكون قد أدت لزيادة نهائية في الاسـتهلاك بمقدار = 400+300+200= 900 دولار على مدى 3 ســنوات. ويمكن في هذه الحالة التعبير عن العلاقة بين الاســتهلاك C والدخل Y باســتخدام الصيغة التالية:

$$C_{t+2} = a + b_1Y_t + b_2Y_{t+1} + b_3Y_{t+2} + U_{t+2}$$

ويمكن كتابة نفـس الصيغة بالنســبة للفترة t كما يلي:

$$C_t = a + b_3Y_{t-2} + b_2Y_{t-1} + b_1Y_t + U_t$$

وبتعديل حدود دالة الاســتهلاك تصبح هذه الدالة كما يلي:

$$C_t = a + b_1Y_t + b_2Y_{t-1} + b_3Y_{t-2} + U_t$$

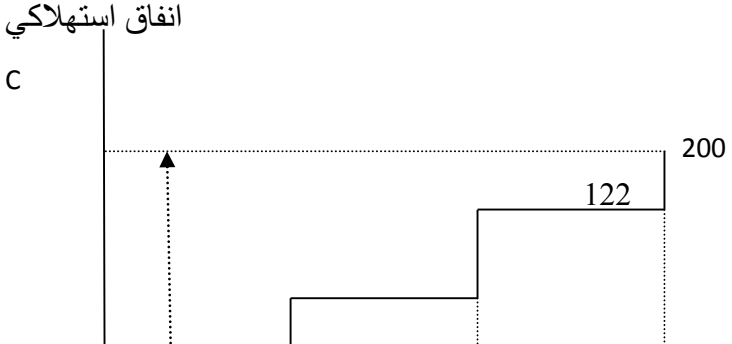
وباســتخدام بيانات المثال الســابق يمكن كتابة دالة الاســتهلاك على النحو التالي:

$$C_t = a + 0.4Y_t + 0.3Y_{t-1} + 0.2Y_{t-2} + U_t$$

ويتضح من المعادلة الســابقة أن أثر الدخل على الاســتهلاك موزع على 3 فترات زمنية. وباســتخدام المعـادلتيـن الســابقتين يمكن توضيح ما يلي:

- مضـاعف الفترة القصـيرة $= b_1 = 0.4$ وهو يشـير في هذه الحالة إلى الميل الحدي للاســتهلاك في الفترة القصيـرة.
- مضاعف الفترات الســابقة $= \sum_{i=2}^{m} bi = b_2 + b_3 = 0.2 + 0.3 = 0.5$
- مضاعف الفترة الطويلة $= \sum_{i=1}^{m} bi = b_1 + b_2 + b_3 = 0.4 + 0.2 + 0.3 = 0.9$

وهو يشــير في هذه الحالة للميل الحدي للاســتهلاك في الفترة الطويلة. أي أن زياة دائمة في الدخـل مقدارها 1 دولار تؤدي إلى زيادة في الاســتهلاك تســاوي 90 ســنتاً في الفترة الطويلة. ويمكن توضيح هذه الفكرة باســتخدام الشــكل (4-1).



900

300

400

الزمن

t1 t3t2

الشكل (4-1) المسار الزمني للاستهلاك

2-3 نموذج التضخم:

يلاحظ بشكل عام أن زيادة كمية النقود لا تؤدي لارتفاع المستوى العام للأسعار بصورة مباشرة وفورية وإنما تمارس تأثيرها خلال فترة زمنية طويلة نسبياً. ولذلك يتم التعبير عن العلاقة بينهما باستخدام الصيغة التالية:

$$P_t = a + b_1M_t + b_2M_{t-1} + b_3M_{t-2} + b_4W_t + U_t$$

حيث: P_t = الرقم القياسي للأسعار

M_t = كمية النقود

W_t = الرقم القياسي للأجور

- ومن الممكن بشكل عام أن نلخص أهم العوامل التي تؤدي إلى وجود فجوات زمنية (تباطؤ في الاستجابة) في مجال العلاقات الاقتصادية فيما يلي:

أ- **عوامل نفسية**: فالفرد كثيراً ما يتعود على نمط من السلوك دون أن يكون على استعداد للتخلي عن هذا السلوك بصورة فجائية لمجرد تغير الأسعار أو الدخول. فلابد أن تمر هناك فترة حتى يتأكد أن هذا التغير الذي حدث هو تغير دائم وليس تغير مؤقت سرعان ما يزول. فإذا تأكد له أن

التغير في الأســـعار أو الدخول أو غيرها هو تغير دائم فإنه يبدأ في تغيير ســـلوكه أو عاداته الاســـتهلاكية بصورة تدريجية عبر فترة زمنية طويلة نســـبياً.

ب- **عوامـــل فنية**: هذه العوامل كثيراُ ما تعوق استجابة المتغير التابع للمتغيرات المســـتقلة. فعلى سبيل المثال كثيراً ما تتباطأ استجابة رأس المال لتغييرات الإنتاج وذلك لأسباب فنية بحتة ، منها طلب رأس المال وتركيبه ثم الشروع في عملية الانتاج ، مما يستغرق بعضاً من الوقت.

جـ- **عوامل مؤسسية وقانونية**: فمثلاً كثيراً ما يدخل رجال الأعمال في تعاقدات طويلة الأجل نســـبياً مع موردين لبعض المواد أو مع مشـــترين لبعض المنتجات ، ومن ثم فإن حدوث تغييرات في الأســـعار قد لا تحفزهم على إحداث تغيـيرات فورية في طلبهم على المواد أو في عرضهم للمنتجات وذلك لارتباطهم بتعاقدات قانونية معينة.