المحاضرة الثانية

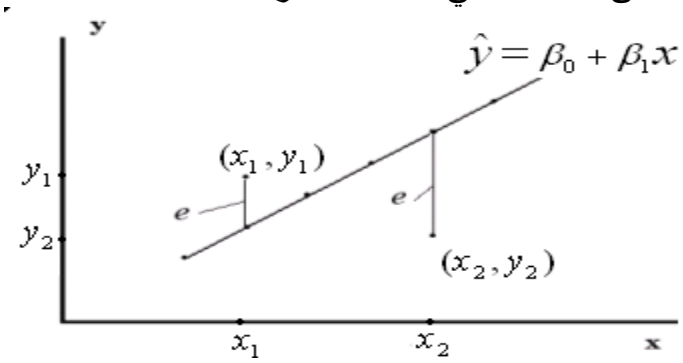
رابعاً: تحويل الصيغة الرياضية الى صيغة قياسية

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 وبذا يتحول النموذج من نموذج رياضي يعكس علاقة تامة الى نموذج قياسي يعكس علاقة احتمالية. وبالتالي يمكن صياغة معادلة الانحدار الحقيقي للنموذج الخطي البسيط قياسيا على النحو التالي:

$$Y=B_0+B_1X+E \qquad \text{---------- (2)}$$

حيث أن:

E هو الخطأ العشوائي، والذي يعبر عن الفرق بين القيمة الفعلية y ، والقيمة المقدرة $\hat{y}=\beta_0+\beta_1 x$ ، أي أن : $e=y-(\beta_0+\beta_1 x)$ ، ويمكن توضيح هذا الخطأ على الشكل التالي لنقط الانتشار.



شكل (3)

الخطأ العشوائي في نموذج الانحدار البسيط

ويلاحظ من الشكل (3) 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 . وهذا الوضع يدفعنا للتساؤل حول الأسباب التي تؤدي الى انحراف القيم المشاهدة للمتغير التابع عن خط الانحدار والذي يؤدي لظهور الأخطاء العشوائية. والاجابة على هذا السؤال تكمن في الأسباب التالية:

١- أخطاء توصيف الصيغة الرياضية للنموذج

قد يقوم الباحث بتقدير صيغة للعلاقة غير الصيغة الحقيقية، كأن يقدر خطأ علاقة خطية وهي ليست خطية. ويترتب على ذلك اختلاف المعادلة المستخدمة عن المعادلة الحقيقية والتي تعبر عن الظاهرة محل الدراسة.

٢- اسقاط بعض المتغيرات الهامة في النموذج

هناك بعض العوامل او المتغيرات التي يتعذر على الباحث قياسها كميا كالأذواق، او المتغيرات التي لا يمكن التنبؤ بحدوثها كالكوارث الطبيعية او المتغيرات المعروفة للباحث والقابلة للقياس الكمي، ولكن البيانات المتوفرة عنها غير كافية او غير دقيقة وبالتالي يقوم الباحث بإسقاطها ويترتب على ذلك اختلاف المعادلة المستخدمة عن المعادلة الحقيقية والتي تعبر عن الظاهرة محل الدراسة.

٣- الأخطاء الراجعة للتجميع

نجد على سبيل المثال ان البيانات التجميعية للاستهلاك الكلي والدخل الكلي تعبر عن قيم متعلقة باستهلاك ودخول الافراد دون ان تعكس هذه النوعية من البيانات الخاصة بهيكل التوزيع للاستهلاك والدخل في المجتمع المدروس. فنجد ان تساوي مجموع الدخل القومي لبلدين لا يعني تساوي مستوى الرفاهية الاقتصادية فيهما، نظرا لإمكانية اختلاف توزيع الدخول بين الافراد في كلا البلدين. وبالتالي فان هذه الاختلافات تؤثر على الظاهرة محل البحث واسقاطها يترتب عليه خطا في المعادلة.

٤- وجود خطأ في القياس

ويقصد بذلك حدوث أخطاء عند قياس المتغيرات المكونة للنموذج محل الدراسة عند اخذ المشاهدات. وتنتج هذه الأخطاء عند قيام الباحث بالمعالجة الإحصائية للبيانات، كما قد تنتج الأخطاء بسبب قصور في أساليب جمع البيانات. ومن ثم نجد ان الأخطاء السابقة تؤدي الى انحراف القيم المشاهدة عن الخط المستقيم والذي يمثله الخطأ العشوائي، مما يؤدي الى تحولها من علاقة تامة الى علاقة احتمالية بين متغيرات النموذج.

خامساً: التقدير Estimation

تتمثل المرحلة الثانية من مراحل البحث في القياس الاقتصادي في تقدير النموذج أي قياس القيم الرقمية لمعالم النموذج للوصول الى خط الانحدار الممثل للعلاقة بين متغيرات النموذج الذي تم توصيفه في المرحلة السابقة. ويعتمد الباحث أساساً في تقديره على ما يلي:

- بيانات واقعية تخص متغيرات النموذج
- طرق قياسية تستخدم في عملية التقدير

وتنطوي هذه المرحلة على ما يلي:

١. تحديد النموذج

سبق وان بينا ان نموذج الانحدار الخطي البسيط يمكن كتابته كما يلي:

$$Y=B_0+B_1X+E$$

ان هذا النموذج يعكس العلاقة الحقيقية بين متغيرات النموذج، وهذه المتغيرات يمكن الحصول عليها عند جمع بيانات واقعية عن كل القيم المشاهدة ويطلق على هذه المعادلة معادلة الانحدار

الحقيقي. الا ان خط الانحدار الحقيقي لا يكون معروفا على وجه التحديد وذلك لان النقاط التي نحصل عليها عادة ما تكون عينة صغيرة من مجتمع كبير أي انها لا تمثل سوى جزء صغير جدا من كل النقاط التي يمكن الحصول عليها، لذلك نقوم بتقدير خط الانحدار ونطلق عليه خط الانحدار المقدر ومعادلته:

$$y=B_0+B_1x+e$$

٢. **فروض النموذج** Assumptions of the model

لكي يكون نموذج الانحدار الخطي البسيط قابلا للتقدير يتوجب توفر الفروض التالية:

- **(U_i)متغير عشوائي حقيقي.**

ان هذا الافتراض يعني ان (U_i) هو متغير حاله حال المتغيرات المستقلة الداخلة في الأنموذج ، الا انه استبعد من الانموذج المراد تقديره ، وهو بذلك يكون مع المتغيرات المستقلة تركيبة خطية تؤثر في المتغير التابع كما في الصيغة ( ) ، من جانب آخر ، فهو عشوائي وكل قيمة يأخذها في اية مدة زمنية ، تعتمد على المصادفة Depends on chance وذات تأثير متباين على (Y) او معدوم ولكل مشاهدة. وعندما نقول عشوائي يفترض انه استنزف من أي متغير احتواه ذي تأثير فعال على المتغير التابع (Y) ، وما يحتويه من متغيرات فإنها جاءت بطريقة غير عمدية ، اما اذا كان تحرك قيم احد او اغلب المتغيرات المحذوفة ، والمجتمعة بافتراض جدلي في (U_i) ، بنمط منتظم systematic pattern في الفترات المختلفة ، فان (U_i) لا يتمتع بافتراض العشوائية. ومع هذا نقول انه لا توجد صيغة اختبار عملية لذلك الافتراض لسبب بسيط هو ان (U_i) غير مشاهد، وحتى عندما يتم تقديره ويصبح رمزه (e_i) فان هذا التقدير اعتمد على احدى الطرائق المستخدمة في التقدير القائمة على افتراض العشوائية اصلاً، فكيف نستخدم مقدراً بني على افتراض العشوائية لاختبار ان كان هذا المتغير المقدر عشوائياً ام لا. من جانب آخر، فان (U_i) متغير حقيقي، لما يحتويه من متغيرات هي في الاصل حقيقية Real، وكان من المفترض ان تنظم للجزء التوضيحي في الانموذج، الا انه لأسباب كثيرة احتواها هذا المتغير تم استبعادها منه، ومن هنا نسلم بأهمية ان يكون (U) عشوائياً غير منتظم وخالٍ من المتغيرات ذات التأثير الفعال على (Y). وفي حالة كونه متغيراً مستقلاً غير عشوائي فان هذا يعني تحول العلاقة من علاقة قياسية الى علاقة رياضية.

- **متوسط القيمة المتوقعة للمتغير العشوائي يساوي صفراً.**

ان هذا الفرض يمكن صياغته رياضيا كما يلي: -

$$E(U_i)=0$$

والمقصود بهذه الفرضية ان قيم (U_i) المقابلة لقيم x_i والتي أخذت من المجتمع الإحصائي، احتوى هذا المجتمع على قيمها التي تكون أكبر، أصغر، مساوية للصفر، وان تلك القيم تتساوى في مجاميعها، فيكون متوسط (U_i) صفراً.

- **خاصية ثبات تجانس تباين الخطأ Homoscedsticity**

معنى هذه الفرضية ان تباين قيم (U_i) حول متوسط هذه القيم ثابتاً constant لجميع قيم (X) ، وهذا يعني ان قيم (U_i) سيكون لها التشتت Dispersion نفسه حول متوسطها لجميع قيم (X) . أي ان: