

الانحدار Regression

يعد الانحدار أحد أهم الأساليب الإحصائية. يختص بقياس العلاقة بين متغير يسمى المتغير التابع، ومتغير آخر أو مجموعة متغيرات تسمى المتغيرات المستقلة.

وتقسم نماذج الانحدار إلى عدة أنواع:

1. الانحدار الخطي البسيط

$$Y = \alpha + \beta X$$

2. الانحدار الخطي المتعدد

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

3. الانحدار غير الخطي البسيط

$$Y = \alpha + \beta X^2$$

4. الانحدار غير الخطي المتعدد

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

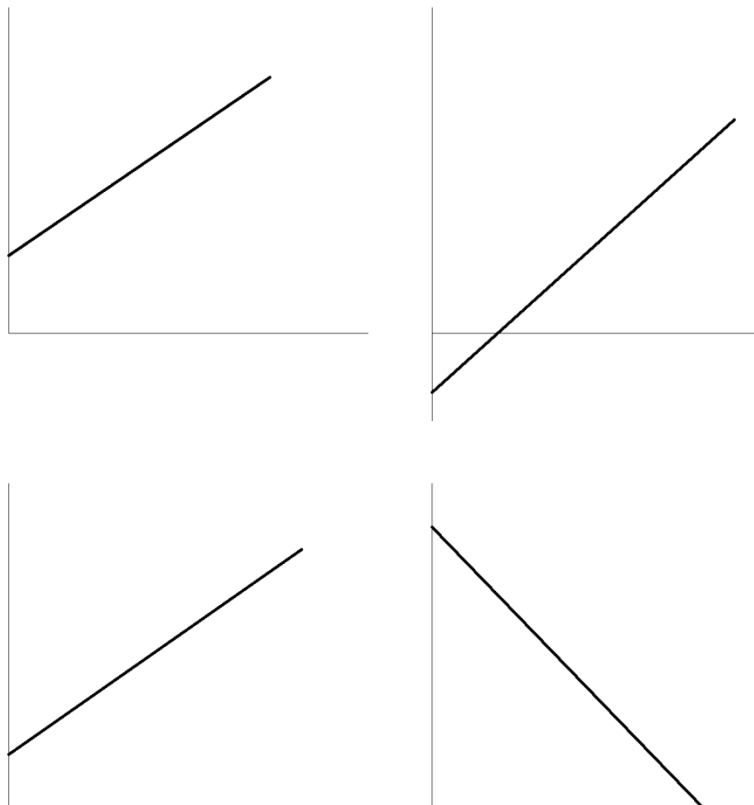
وسنركز في هذا المقرر على الانحدار الخطي لأهميته وسهولة تقديره.

SIMPLE LINEAR REGRESSION **الإنحدار الخطي البسيط**

يعتبر أبسط أنواع نماذج الانحدار. وتأخذ المعادلة فيه الشكل:

$$Y = \alpha + \beta X$$

- **المعلمة α** : تشير إلى الجزء المقطوع الذي يختلف حسب نوع الدالة. ففي دالة الاستهلاك يمثل الحد الأدنى للإنفاق الاستهلاكي، وفي دالة الادخار يمثل الادخار السالب اللازم لتغطية حد الكفاف، وفي العرض تشير للحد الأدنى من الكمية المعروضة لو كان السعر صفراً، وفي الطلب يمثل الكمية القصوى المطلوبة في حال كان السعر صفراً، وفي دالة التكاليف قصيرة الأجل تمثل الحد الأدنى من التكاليف التي تتحملها المنشأة إذا توقفت عن الإنتاج.



- **المعلمة β** تسمى المعلمة الانحدارية. وتمثل ميل الخط المستقيم الممثل للعلاقة. وتشير لمقدار التغير في المتغير التابع Y نتيجة التغير بوحدة واحدة في المتغير المستقل X . وهذه المعلمة تمثل المعامل الحدي إي أنها تمثل الميل الحدي للاستهلاك في دالة الاستهلاك، والميل الحدي للاادخار في دالة الادخار.

القسم الأول: تعيين نموذج الانحدار

سندرس نموذج الاستهلاك كمثال لشرح هذه المرحلة.

A. تحديد المتغيرات:

بالعودة للنظرية الاقتصادية وحسب النظرية الكينزية: (يوجد علاقة طردية بين مستوى الاستهلاك وحجم الدخل). مما يعني أن الدخل أحد المحددات الرئيسية للاستهلاك. وتشير النظرية الكلاسيكية إلى أن سعر الفائدة هو عائد الادخار، أي أن ارتفاعه يؤثر سلباً على الاستهلاك بفرض ثبات الدخل. كما تشير المشاهدات الواقعية لوجود علاقة طردية بين توقعات الأسعار ومستوى الاستهلاك.

مما سبق نحدد متغيرات النموذج كما يلي: المتغير التابع (Y) الإنفاق الاستهلاكي، المتغيرات المستقلة: الدخل (X)، سعر الفائدة (r)، مستوى الأسعار المتوقعة (P)

أي أن دالة الاستهلاك تأخذ الشكل: $Y = f(X, r, P)$

لكن المتغيرات المستقلة السابقة لا تمتلك نفس الأهمية. حسب الدراسات السابقة يكون التأثير الأقوى للدخل. لذا سنركز عليه كعامل مستقل وحيد. فتكون الدالة: $Y = f(X)$

B. تحديد الشكل الرياضي للنموذج:

يوجد أكثر من شكل رياضي يمكن استخدامه لقياس العلاقة الخطية بين الاستهلاك والدخل. سنفرق بين صيغتين:

⊕ دالة الاستهلاك الخطية غير النسبية: من الشكل

$$Y = \alpha + \beta X$$

تشير β هنا إلى الميل الحدي للاستهلاك وبما أن الدالة خطية فقيمتها ثابتة. لكن خطية الدالة تؤدي لأن الميل الحدي للاستهلاك لدى ذوي الدخل المرتفعة يساوي نظيره لدى أصحاب الدخل المنخفضة. ومن ناحية أخرى بما أنها دالة غير نسبية فإن الزيادة في الدخل بنسبة معينة تؤدي للزيادة في الاستهلاك بنسبة أقل حيث

$$\frac{Y}{X} = \frac{\alpha}{X} + \beta$$

وقيمة $\frac{Y}{X}$ تمثل الميل المتوسط للاستهلاك، وهذه القيمة تتخفض مع زيادة الدخل X . وهذه العلاقة قد تكون صحيحة بشكل أكبر لأصحاب الدخل المرتفعة.

⊕ دالة الاستهلاك النسبية: من الشكل

$$Y = \beta X$$

وتختلف عن الدالة السابقة في أن الحد الثابت يساوي الصفر. مما يعني أن انخفاض الدخل للصفر يؤدي لانخفاض الاستهلاك للصفر. وهذا قد يكون صحيحاً على المدى الطويل، فالمجتمع الذي لا ينتج خلال فترة طويلة ينهار ويموت. لكن ذلك غير صحيح بالنسبة للبيانات القطاعية أو بيانات السلسلة الزمنية القصيرة، مما يجعل الدالة الأولى أكثر ملاءمة في هذه الحالات. وبما أن هذه الدالة نسبية، فذلك يعني أن زيادة الدخل بنسبة معينة تؤدي لزيادة الاستهلاك بنفس النسبة.

وحتى نحدد الشكل الأكثر ملاءمة للدالة، لا بد أن ننظر لشكل الانتشار الذي يوضح العلاقة بين Y و X فإذا كان الشكل يقارب الخط المستقيم تكون الدالة خطية. وإذا كانت البيانات قطاعية أو سلسلة زمنية قصيرة تستخدم الدالة غير النسبية، في حين تستخدم النسبية لبيانات السلسلة الزمنية الطويلة.

C. تحديد التوقعات القبلية للمعاملات:

تمثل α حد الكفاف الذي لا بد أن يستهلكه المجتمع حتى إذا انخفض الدخل إلى الصفر. لذا فمن المتوقع أن تكون $\alpha > 0$. بينما تمثل β الميل الحدي للاستهلاك ومن المتوقع أن تكون العلاقة بين الاستهلاك والدخل طردية وأقل من الواحد لأن الزيادة تتوزع بين زيادة في الادخار وزيادة في الاستهلاك أي أن $0 < \beta < 1$.

D. تعيين الحد العشوائي:

الشكل المفترض للدالة لا يحتوي على حد عشوائي. أي أننا ننظر للعلاقة بين الدخل والاستهلاك على أنها علاقة مؤكدة! (كل التغيرات في Y تعود للتغيرات في X) ولو كان ذلك صحيحاً لانطبقت كل النقاط في شكل الانتشار على الخط المستقيم. لكن الواقع مغاير لذلك فدالة الاستهلاك دالة احتمالية تشير للقيمة الأكثر احتمالاً للاستهلاك عند مستويات الدخل المختلفة. والقيم المشاهدة تنحرف عن خط الانحدار المقدر. مما يعني أنه لا بد للدالة الاحتمالية أن تحتوي حداً عشوائياً (e) فتصبح الدالة من

$$Y = \alpha + \beta X + e \quad \text{الشكل:}$$

الحد العشوائي يسمى بالخطأ العشوائي ويمكن التفرقة بين مصدرين له:

⊗ خطأ المعادلة أو خطأ الحذف Equation error:

ينشأ عن بعض العوامل التي تؤدي لاختلاف شكل المعادلة المقدر عن الحقيقية، وأسبابه:

- I. حذف بعض المتغيرات المؤثرة
- II. عدم كمال تعيين الشكل الرياضي للنموذج، مثل افتراض خطية الدالة والواقع أنها غير خطية.
- III. أخطاء التجميع عند استخدام بيانات تجميعية كالدخل الكلي والاستهلاك الكلي والادخار الكلي، فهذه المجاميع لا تعكس الاختلافات بين الأفراد وتوزيعها.

⊗ خطأ المشاهدة أو خطأ القياس Measurement error:

وتحدث نتيجة عدم كمال أساليب جمع البيانات أو نتيجة خطأ في المعالجة الإحصائية.

تؤدي الأخطاء السابقة لانحراف القيم المشاهدة عن الخط المستقيم النظري. وتتحوّل العلاقة من علاقة مؤكدة إلى احتمالية، والأخيرة تمثل العلاقة الحقيقية.