المحاضرة الأولى

المتحول العشوائي

المتغير العشوائي: ان نموذج الانحدار الخطي البسيط c موجه اساسا لاختبار الفرضيات المتعلقة بالعلاقة بين المتغير التابع y و متغير مستقل او مفسر cx كما انه موجه للتنبؤ. ان دراسة الانحدار الخطي البسيط يبدا عادة بوضع على لوح او بيان مجموعة النقاط (x,y) ثم نتفحص هل توجد علاقة خطية تقريبية

$$Y_i=b_0+b_iX_i$$

مادام من المشكوك فيه ان تقع جميع النقاط الى الخط المستقيم المناسب، فان العلاقة الخطية الدقيقة السابقة لابد من تغييرها بادخال حد الاضطراب او حد الخطا او حد التشويش عليها ممثل في u_i

$$Y_i= b_0+b_iX_i + u_i$$

و هذا ما يدعونا الى دراسة اسباب ادخال حد الاضطراب او المتغير العشوائي.

اسباب ادخال المتغير العشوائي: هنالك في الواقع ثلاث اسباب ممكنة تدفعنا الى ادخال الحد u. و من اجل تثبيت الافكار بصورة احسن نفترض ان هناك معطيات بفعل استمارة حول الاسر حيث x يمثل الدخل المتاح و y تمثل الانفاق على الاستهلاك.

اولا: يمكننا القول ان الانفاق الاستهلاكي لكل اسرة من الاسر يمكن تفسيره كليا لو عرفنا جميع العوامل المؤثرة و المتفاعلة و كانت لدينا جميع البيانات الضرورية لكن من بين الاسر ذات الحجم نفسه و الركيب نفسه سوف تكون هناك تباينات في الاعمار و العادات و غير ذلك و كذلك اذا كان دخل الاسر في حالة تزايد او تناقص فبلاشك فان الاسر سوف تنفق اكثر او اقل.

و في تفسير السلوك الانساني في الانفاق من الممكن ان تكون العوامل المؤثرة و المتفاعلة مشكلة لقائمة تصل الى مالا نهاية و كثير من هذه العوامل غير قادرة على اعطاء مضمونا كميا و حتى و لو كان قياسها ممكن فليس من الممكن عمليا الحصول على بيانات عنها جميعا و حتى و لو كان ذلك مستطاع فان عدد العوامل سوف يتجاوز عدد المشاهدات الممكن الحصول عليها بحيث انه ليس هناك من ادوات احصائية يمكن تقدير اثارها و هذا الذي نقوله اذن بان التابع y يمكن كتابته على الشكل:-

$$Y=f(X_1,X_2,.....,X_n)$$

حيث n كبير جدا و غير عملي الامر الذي يدفعنا الى جعل y تابع ظاهريا لمجرد عدد صغير من المتغيرات التي نعتقد انها ذات تاثير اكبر من جميع المتغيرات x و سوف نعبر عن باقي المتغيرات (ذات الاثر الضئيل) ب u و يكون لدينا:

$$Y=f(x,u)$$

ملاحظة: و بما ان عوامل عديدة تتفاعل في اسرة ما فان بعض هذه العوامل يمكن ان تكون متجهة في اتجاه معاكس لبعضها البعض (بعضها يزيد الاستهلاك و بعضها ينقصه) و ما دام لا يوجد هناك سبب بصفة عامة لنتوقع تحيز في اتجاه معين عنه في اتجاه اخر و لذا علينا ان نتوقع قيما صغيرة لـ u متكررة اكثر من تكرر قيم كبيرة له و هذا يقودنا للتفكير بان u هو متغير ذو توزيع احتمالي يتمركز حول الصفر و ذو تباين محدود و ثابت اي:

$$E(u)=0$$

و بالاستناد الى نظرية النهاية المركزية فان مقدار الخطا العشوائي u و بالنظر الى العوامل المتفاعلة الكثيرة $u\sim N(0,\sigma^2)$

$$\text{Lim } z=\frac{\lim u-E(u)}{\sqrt{var(u)}} \sim N(0,1) \Rightarrow u\sim N(0,\sigma^2)$$

$$n \rightarrow \infty$$

$$var(z) = 1 \Rightarrow E[z-E(z)]^2=1$$

$$E(z)^2=1 \Rightarrow E\left(\frac{u-0}{\sqrt{var(u)}}\right)^2 = 1$$

$$E\left(\frac{u}{\sqrt{var\,(u)}}\right)^2 = 1 \Rightarrow E\left(\frac{u^2}{var(u)}\right) =1 \Rightarrow var\,(u) = E\,(u^2)= \sigma^2$$

ثانيا: السبب الثاني لظهور الحد العشوائي للعلاقة بين المتغيرات الاقتصادية هو احتواء السلوك الانساني على عنصر الصدفة بشكل اساسي و صعوبة التنبؤ به.

ثالثا: هناك سبب ثالث في الواقع يعود الى اخطاء المشاهدات او قياسها فمثلا من الممكن ان يكون المتغير z مرتبط تماما في شكل خطي مع المتغير x وفق العلاقة:

$$Z=\alpha+\beta x$$

لكن اخطاء القياس تحجب القيمة الحقيقية لـ z فبدلا من z فننا نشاهد z+u حيث z تعبر عن خطا القياس لدينا اذن:

$$Y=z+u$$

$$Y=\alpha+\beta x+u$$

مثال:

و من اخطاء القياس مثلا عدم ادلاء الاسرة بانفاقها الحقيقي .

ملاحظة: من الممكن من جهة اخرى ان يكون لدينا خطا القياس مفروض على خطا المعادلة.

خطا المعادلة ——— $z=\alpha+\beta x+u$

خطا القياس ——— $y=z+v$

$$Y=\alpha+\beta x+u+v$$

$$W=u+v$$

$$Y=\alpha+\beta x+w$$

الفرضيات الاساسية التي يقوم عليها التوزيع **الاحتمالي للمتغير العشوائي u:** ان هذه الفرضيات تتعلق اساسا بالوسط الحسابي و التباين و التباين المشترك.

الفرضية الولى: كما سبق ان ذكرنا في ظل افتراض ان u يمكنه ان ياخذ قيم سالبة و قيم موجبة و دون اي تحيز في اي اتجاه دون الاخر فان التوقع الرياضي لـ u

$$E(u)=0$$

الفرضية الثانية: افتراض ان U يتبع توزيع احتمالي معين فان نظرية النهاية المركزية فاننا نعتبر U يتبع التوزيع الطبيعي و منه:-

$$U\sim N(0,\sigma u^2) \qquad E(u)^2=\sigma^2_u$$

الفرضية الثالثة:نفترض ان القيم المختلفة لـ u تمثل متغيرات عشوائية مستقلة عن بعضها البعض. ففي اطار مثالنا المتعلق بالاسر فان هذه الفرضية تعني اذا كان الحد العشوائي المشاهد موجب بالنسبة لاسرة ما فان اختما لان نشاهد حد عشوائي موجب بالنسبة لاسرة اخرى فانه لا يتاثر بالحد العشوائي الاول الموجب بالنسبة للاسرة الاول.

$$E(u_i , u_j) = 0 \qquad i\neq j$$

و بالتالي فان التعبير الرياضي للنموذج الخط البسيط.

$$Y_i=\alpha+\beta x_i+u_i \qquad i=1,2,.....n$$

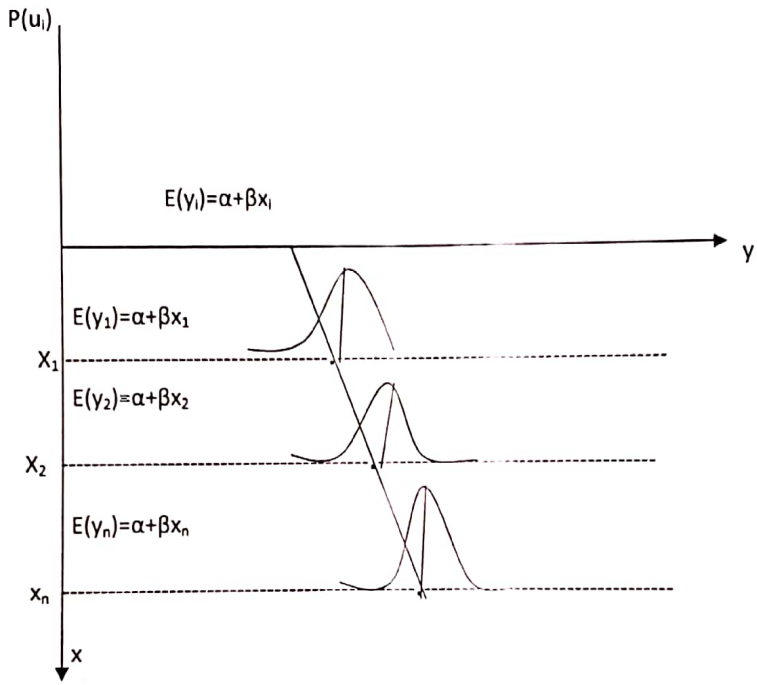
$$E(u_i)=0 \quad \forall i: 1,2,\ldots\ldots\ldots\ldots n$$

$$E(u_i , u_j) \begin{cases} 0 & \forall\ i \neq j \\ \sigma_2 & i=j \end{cases}$$

لان التباين ثابت مهما كانت x_i و بالتالي $E(x_i,u_i)=0$ لا توجد علاقة بين x_i و u_i :-

$$u \sim N(0,\sigma^2_u)$$

و يمكن توضيح النموذج بيانيا



و مما سبق نستنتج :

$$E(y_i) = E(\alpha+\beta x_i+u_i)$$

$$= \alpha+\beta x_i$$

لكن α و βx_i عبارة عن ثوابت .

$$Var(y_i) = E(\alpha+\beta x_i+u_i-\alpha-\beta x_i)^2$$