المحور الرابع: التعدد (الازدواج، الاشتراك، التداخل) الخطي

المحاضرة 6:

تظهر هذه المشكلة عندما يكون أكثر من متغير مستقل، أي تكون في نموذج الانحدار المتعدد، وتكون في قمتها عندما يكون معامل الارتباط بين المتغيرات المستقلة يساوي واحد، وتتمثل هذه المشكلة في نقض فرضية من فرضيات نموذج الانحدار المتعدد التي تنص على ان المتغيرات المستقلة منفصلة عن بعضها البعض.

1- أسباب التعدد الخطي واثاره:

من اهم أسباب التعدد الخطي يمكننا التمييز ما يلي:

- ارتباط المتغيرات الاقتصادية فيما بعضها حيث تتغير مع بعض على مرور الزمن على غرار الدخل، الاستهلاك، الاستثمار والادخار هذا ما يعد ويعتبر دليل هام وواضح لوجود التعدد الخطي بين المتغيرات الاقتصادية.
- استخدام الإبطاءات والفجوات الزمنية للمتغيرات المستقلة على غرار استعمال تأخر المتغير المستقل X بفترة زمنية واحدة مثل: استخدام دخل الفترة الحالية والمعتمد على دخل الفترة السابقة.

في حالة وجود التعدد الخطي فان ذلك سيترتب عنه:

- ارتفاع قيمة التباين والتباين المشترك للمقدرات بدرجات كبيرة مما يعني ابتعادها عن القيم الحقيقية والتأثير على التنبؤ.
- القيم المقدرة للمعاملات لن تكون محدد ولا دقيقة.
- عدم ثبات المعلمات عبر الزمن حيث تصبح كثيرة التغير مع مرور الوقت.
- تضخم قيمة الخطأ المعياري للمقدرات.
- انعدام المعنوية الاحصائية لإحدى المعالم المقدرة أو أكثر رغم أن قيمة معامل التحديد عالية.

2- أنواع الارتباط الخطي:

هناك نوعين من الارتباط الذاتي بين المتغيرات المستقلة الذي ينتج لنا التعدد الخطي:

الارتباط الخطي التام: مثلا $[x_{1t} = 2x_{2t}]$ أو $[x_{1t} = 3x_{3t}]$ أو $[x_{2t} = x_{3t}]$ ..... وفي حال حدوث

ارتباط خطي تام بين المتغيرات المستقلة في النموذج، فإنه يستحيل أصلا تقدير النموذج، لأن محدد المصفوفة (X'X) يكون مساويا للصفر [det (X'X) = 0] وبالتالي لا يمكن تقدير النموذج لأن معكوس المصفوفة (X'X) لا يمكن حسابه.

- **الارتباط الخطي غير التام:**

مثلا $[x_{1t} = 1 + x_{2t}]$ في حالة الارتباط الخطي غير التام يمكن تقدير معالم النموذج، لأنه يمكننا ايجاد معكوس المصفوفة، طالما أن محددها غير معدوم. ولكن لا يمكننا التأكد من وجود مشكلة التعدد الخطي إلا باختبارات الكشف عن هذه المشكلة.

3- طرق الكشف عن وجود التعدد الخطي:

هناك اختبارات عديدة للكشف عن مشكلة التعدد الخطي نذكر منها:

3-1- اختبار Klein :

لا يتعلق الأمر باختبار احصائي وانما مجرد خاصية يمكن الاعتماد عليها لكشف مشكلة التعدد الخطي، حيث يتم أولا تقدير نموذج الانحدار الخطي المتعدد وحساب معامل التحديد الخاص به R^2. بالموازاة مع ذلك يتم حساب معاملات الارتباط الخطي بين كل متغيرين مستقلين x_i ،x_j , ، مع: i≠j، أي: r x_i , x_j

إذا كان: $R^2 >$ r x_i , x_j نقول ان نموذج الانحدار الخطي يعاني من مشكلة التعدد الخطي.

3-2- اختبار Farrar-Glauber:

يتم الاختبار كالتالي:

1- حساب محدد مصفوفة معاملات الارتباط بين المتغيرات المستقلة:

$$D = \begin{vmatrix} 1 & r_{X_1X_2} & r_{X_1X_3} & \cdots & r_{X_1X_k} \\ r_{X_2X_1} & 1 & r_{X_2X_3} & \cdots & r_{X_2X_k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{X_kX_k} & r_{X_kX_2} & r_{X_kX_3} & \cdots & 1 \end{vmatrix}$$

عندما تكون قيمة المحدد تقترب من الصفر، فإن هناك دليل على وجود تعدد خطي.

2- نستخدم اختبار كاي مربع χ^2 لتحديد وجود او عدم وجود مشكلة التعدد الخطي في النموذج المقدر وذلك بوضع الفرضيات التالية:

$$\begin{cases} H_0: |D| = 1 \text{ (حالة عدم وجود تعدد خطي)} \\ H_1: |D| < 1 \text{ (حالة وجود تعدد خطي)} \end{cases}$$

نقوم بحساب احصائية Farrar-Glauber القيمة المحسوبة تعرف كما يلي:

$$\chi^{2*} = -\left[n - 1 - \frac{1}{6}(2K + 5)\right].ln|D|$$

حيث n : عدد المشاهدات، K: عدد المتغيرات المستقلة إضافة الى الحد الثابت (K= k+1)، ln: اللوغاريتم النيبيري

مقارنة إحصائية كاي تربيع المحسوبة والجدولية وفق قاعدة القرار التالية:

$\chi^{2*} > \chi^2$ نرفض الفرضية الصفرية ونقبل الفرضية البديلة أي وجود تعدد خطي.

$\chi^{2*} < \chi^2$ نرفض الفرضية البديلة ونقبل الفرضية الصفرية أي عدم وجود تعدد خطي.

حيث ان درجات الحرية للإحصائية كاي تربيع الجدولية هي $\frac{1}{2}K(K-1)$ ومستوى معنوية α.

3-3- شرط الاعداد condition Numbers

3-4- معامل تضخم التباين " Variance Inflation Factor "

ما يهمنا نحن باستخدام البرنامج الاحصائي EVIEWS، هو استخدام معامل تضخيم التباين (VIF) الذي يحدد درجة التعدد الخطي. حيث يوفر برنامج EVIEWS هذا الاختبار.

يعطى معامل تضخيم التباين بالعلاقة التالية:

$$V.I.F\ (\hat{\beta}_i) = \frac{1}{1 - R_i^2}$$

ملاحظات:

- القيمة $[1 - {R_i}^2]$ تسمى بمقياس التحمل TOLERANCE. يتم حسابه من خلال اجراء معادلة انحداريه لكل متغير مستقل بدلالة المتغيرات المستقلة الأخرى وحساب مربع معامل الارتباط المتعدد.

- إذا كان معامل تضخيم التباين يساوي 10 أو أكثر، فذلك يدل على وجود مشكلة التعدد الخطي.
- معامل تضخيم التباين يساوي 10 يعني أن التباين تضاعف 10 مرات عن القيمة التي يمكن أن يحصل عليها إذا كانت المتغيرات المستقلة غير مرتبطة أي مستقلة عن بعضها البعض.

4- معالجة مشكلة التعدد الخطي:

يوجد عدد من طرق علاج التعدد الخطي بين المتغيرات المستقلة، منها ما يلي:

- زيادة حجم العينة.
- حذف بعض المتغيرات المستقلة من النموذج، عندما يكون هناك تداخل خطي بين متغيرين مستقلين، يلجأ الباحث أحيانا إلى حذف أحد المتغيرين للتخلص من هذا التداخل. ويجب أن يكون حذف المتغيرات عملية تكرارية تبدأ بالمتغير الذي له أكبر قيمة VIF.
- إجراء بعض التحويلات على المتغيرات، مثل مزج المتغيرات في متغير واحد، أو تحويل متغيرات النموذج إلى الفروق الأولى: $\Delta x = x_t - x_{t-1}$

مثال تطبيقي: باستخدام البيانات التالية في برنامج eviews اختبر وجود مشكلة التعدد الخطي

y	x_1	x_2	x_3
49	1300	7,5	0,012
50,2	1300	9	0,012
50,5	1300	11	0,0115
48,5	1300	13,5	0,013
47,5	1300	17	0,0135
44,5	1300	23	0,012
28	1200	5,3	0,04
31,5	1200	7,5	0,038
34,5	1200	11	0,032
35	1200	13,5	0,026
38	1200	17	0,034
38,5	1200	23	0,041
15	1100	5,3	0,084
17	1100	7,5	0,098
20,5	1100	11	0,092
29,5	1100	17	0,086

الحل:

1- حساب معامل VIF:

كأول خطوة يجب استخراج جدول تقدير النموذج،حيث تحصلنا على النتائج التالية:

Dependent Variable: Y
Method: Least Squares
Date: 11/14/25 Time: 21:15
Sample (adjusted): 2000 2015
Included observations: 16 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-121.2696	55.43571	-2.187572	0.0492
X1	0.126854	0.042182	3.007300	0.0109
X2	0.348158	0.177017	1.966802	0.0728
X3	-19.02170	107.9282	-0.176244	0.8630

R-squared	0.919815	Mean dependent var	36.10625
Adjusted R-squared	0.899769	S.D. dependent var	11.89877
S.E. of regression	3.767073	Akaike info criterion	5.702792
Sum squared resid	170.2901	Schwarz criterion	5.895939
Log likelihood	-41.62233	Hannan-Quinn criter.	5.712682
F-statistic	45.88452	Durbin-Watson stat	0.775471
Prob(F-statistic)	0.000001		

المتغيرين X2 و X3 غير معنويين عند 5% ومعامل التحديد قيمته عالية وتقترب من 1، وهذا مؤشر مبدئي عن وجود مشكلة التعدد الخطي. وللتأكد نقوم بحساب معامل VIF، ونستخدم التعليمة التالية:

Table Estimation→view→Coefficient Diagnostics→variance Inflation Factors.

Variance Inflation Factors
Date: 11/14/25 Time: 21:29
Sample: 2000 2016
Included observations: 16

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	3073.118	3464.903	NA
X1	0.001779	2961.596	12.22505
X2	0.031335	6.532563	1.061838
X3	11648.50	33.66826	12.32496

وكما نرى فإن نتائج معامل تضخيم التباين أكثر من 10 مما يدل على وجود تعدد خطي.

كما قلنا، من بين الطرق لمعالجة التعدد الخطي نقوم بحذف المتغير الذي له أكبر قيمة في VIF في حالتنا هذه نقوم بحذف المتغير X3، وتصبح نتائج التقدير كالتالي:

Dependent Variable: Y
Method: Least Squares
Date: 11/14/25 Time: 21:45
Sample (adjusted): 2000 2015
Included observations: 16 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-130.6899	14.14571	-9.238835	0.0000
X1	0.133961	0.011908	11.25008	0.0000
X2	0.351061	0.169553	2.070504	0.0589

R-squared	0.919607	Mean dependent var	36.10625
Adjusted R-squared	0.907239	S.D. dependent var	11.89877
S.E. of regression	3.623968	Akaike info criterion	5.580377
Sum squared resid	170.7308	Schwarz criterion	5.725237
Log likelihood	-41.64301	Hannan-Quinn criter.	5.587795
F-statistic	74.35306	Durbin-Watson stat	0.792878
Prob(F-statistic)	0.000000		

أما نتائج اختبار الكشف عن التعدد الخطي باستخدام VIF فهي:

Variance Inflation Factors
Date: 11/14/25 Time: 21:46
Sample: 2000 2016
Included observations: 16

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	200.1010	243.7815	NA
X1	0.000142	255.0093	1.052642
X2	0.028748	6.475985	1.052642