الارتباط الجزئي، الارتباط الخطي المتعدد (الازدواج الخطي) Multicollinearity

- ما طبيعة المشكلة.
- ماهي نتائج المشكلة.
- كيف يمكن تشخيص المشكلة.
- ما هو علاج المشكلة.

1.**طبيعة المشكلة:** وجود ارتباط قوي بين متغيرين مستقلين في نموذج الانحدار الخطي. يوجد نوعين من الارتباط الخطي.

1.1.الارتباط الخطي التام (Perfect Multicollinearity) : وهنا نجد ارتباط تام بين إحدى المتغيرين المستقلين. حيث من الممكن التعبير عن واحد أو أكثر من المتغيرات كمزيج خطي للمتغير الآخر، فمثلا نجد تعدد خطي تام بين X1, X2 إذا كانت مثلا $X_1=2X_2$ وهذا الأمر نادر الحدود في الواقع العملي. وهنا من غير الممكن تقدير معامل OLS.

2.1.**الارتباط الخطي غير التام (Imperfect Multicollinearity)**: يحدث عندما تكون المتغيرات التفسيرية في النموذج الخطي مرتبطة ارتباطا مرتفعا. ويجعل هذا من الصعب أو من غير الممكن عزل تأثير كل متغير مفسر من بين المتغيرات ذات الارتباط المرتفع فيما بينها على المتغير التابع.

مثال:

X1	X2	X2*
10	50	52
15	75	75
18	90	97
24	120	129
30	150	152

Correlation

	X1	X2	X2_
X1	1.000000	1.000000	0.995882
X2	1.000000	1.000000	0.995882
X2_	0.995882	0.995882	1.000000

ملاحظاتك

- نلاحظ أن معامل الارتباط بين (X1, X2) يساوي 1 وهو ارتباط تام بين المتغيرين، كما هو واضح أن $X_2=5X_1$.

- المتغير X_2^* تم تشكيله من X2 وهذا بإضافة أرقام عشوائية (2، 0، 7، 9 و2)، حيث لم يعد هناك ارتباط تام بين X_1 و X_2^* ، حيث أن معامل الارتباط بينهما يساوي 0.9958

السؤال الجوهري: كيف يمكن اكتشاف تعدد العلاقات الخطية أو الازدواج الخطي في النموذج؟

- الحالة الكلاسيكية للتعدد الخطي تحدث عندما لا يكون أي من المتغيرات المفسرة في النموذج معنويا احصائيا (وبعضها قد يأخذ الإشارة الخطأ)، بالرغم من أن R2 قد يكون مرتفعا (مثلا بين 0.7 و1.00).
- في الحالات الأقل وضوحا، قد يكون اكتشاف التعدد الخطي أكثر صعوبة، أحيانا يستخدم المعامل المرتفع للارتباط البسيط أو للارتباط الجزئي بين المتغيرات المفسرة كمقياس للتعدد الخطي. كما أنه قد يوجد تعدد خطي حتى لو كان الارتباط البسيط أو الجزئي منخفضا نسبيا (أقل من 0.50 مثلا).

2.ما هي نتائج الارتباط المتعدد.

- ✓ تكون إحصائية t منخفضة؛
- ✓ R^2 مرتفعة جدا؛
- ✓ معاملات الانحدار غير معنوية.

3.طرق الكشف عن الارتباط الخطي:

1.3. معامل الارتباط البسيط: لما يكون أكبر من 0.80 هذا يدل على وجود مشكلة الازدواج الخطي في النموذج.

$$r_{YX_1} = \frac{\sum (Y-\bar{Y})(X_1-\bar{X})}{\sum (X_1-\bar{X})^2 \sum (Y-\bar{Y})^2}$$

2.3. معاملات الارتباط الجزئي: يقيس صافي الارتباط بين المتغير التابع ومتغير مستقل بعد حذف التأثير المشترك، أي مع تثبيت للمتغيرات المستقلة الأخرى في النموذج. فمثلا $r_{YX_1}.x_2$ هو الارتباط الجزئي بين y و X_1 بعد حذف تأثير X_2 من كل y و X_1. وعلى هذا الأساس يمكن حساب معاملات الارتباط الجزئي كما يلي:

- معامل الارتباط الجزئي بين y و X_1 بعد حذف تأثير X_2 من كل y و X_1:

$$r_{YX_1}.x_2 = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{1-r^2_{X_1X_2}} \cdot \sqrt{1-r^2_{YX_2}}}$$

- معامل الارتباط الجزئي بين y وX_2 بعد حذف تأثير X_1 من كل y وX_2:

$$r_{YX_2}.x_1 = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{1-r^2_{X_1X_2}} \cdot \sqrt{1-r^2_{YX_1}}}$$

حيث r_{YX_1} معامل الارتباط البسيط بين y و$x1$، ويعرف r_{YX_2} و $r_{X_1X_2}$ على نفس النمط، وتتراوح معاملات الارتباط الجزئية بين ± 1 كما هو الحال بالنسبة لمعاملات الارتباط البسيطة. وتستخدم لتحديد الأهمية النسبية للمتغيرات المفسرة المختلفة في الانحدار المتعدد. مفهوم الارتباط الجزئي مهم جدا لأنه يسمح بمعرفة مدى ملائمة أدراج المتغيرات التفسيرية في النموذج، بحيث كلما ارتفع معامل الارتباط الجزئي للمتغير كلما دل ذلك على أهميته في المساهمة في التفسير العام للنموذج.

4.العلاج: يمكن علاج هده المشكلة من خلال:

- ✓ اسقاط المتغيرات الزائدة؛
- ✓ تحويل المتغيرات من خلال مزج المتغيرات أو تحويل متغيرات المعادلة إلى الفرق الأول؛
- ✓ زيادة حجم العينة.