المشاكل القياسية وطرق الكشف عنها في تحليل الانحدار الخطي المتعدد

الغرض من هذه الاختبارات هو التحقق من صلاحية النموذج الإحصائي، حيث أن انتهاك الافتراضات الأساسية قد يؤدي إلى نتائج مضللة أو تقديرات غير دقيقة للمعاملات. لذا، يجب أن تُجرى هذه الاختبارات لضمان الثقة في النتائج الناتجة عن التحليل الإحصائي.

تعريف مبسط للاختبارات الموجودة في الجدول أدناه:

1. **الارتباط الذاتي:(Autocorrelation)**

تخيل أن لديك سلسلة من البيانات، مثل درجات الحرارة اليومية على مدار العام. الارتباط الذاتي يخبرنا بمدى ارتباط كل نقطة بيانات بالنقطة التي تسبقها. فمثلاً، في حالة درجات الحرارة، من المرجح أن تكون درجة الحرارة اليوم مشابهة لدرجة حرارة الأمس، وهذا يعني وجود ارتباط ذاتي.

2. **عدم تجانس الانتشار:(Heteroscedasticity)**

عندما ترسم بياناتك على شكل نقاط، فإن عدم تجانس الانتشار يعني أن انتشار النقاط (توزيعها) حول خط الانحدار ليس ثابتاً. ففي بعض الأحيان تكون النقاط قريبة جداً من الخط، وفي أحيان أخرى تكون بعيدة جداً.

3. **التوزيع الطبيعي (Normality)**

الطبيعية تشير إلى توزيع البيانات بشكل يشبه الجرس. بمعنى أن معظم البيانات تتجمع حول المتوسط، وهناك عدد قليل جداً من القيم البعيدة جداً عن المتوسط.

4. **الارتباط المتعدد:(Multicollinearity)**

عندما يكون لديك متغيرات متعددة في نموذجك الإحصائي، فإن الارتباط المتعدد يعني أن بعض هذه المتغيرات مرتبطة جداً ببعضها البعض. وهذا يمكن أن يسبب مشاكل في تفسير نتائج النموذج.

الاختبار	الفرضية الصفرية (H_0)	الفرضية البديلة (H_1)	الحكم على الفرضية	الدوروالغرض من الاختبار
الارتباط الذاتي (Autocorrelation)	لا يوجد ارتباط ذاتي في الأخطاء	يوجد ارتباط ذاتي في الأخطاء	p-value صغيرة: رفض H_0 ، قبول H_1	يُستخدم لاكتشاف إذا ما كانت قيم الخطأ مرتبطة زمنيًا بعضها ببعض، مما يؤثر على الثقة في النموذج.
عدم تجانس الانتشار (Heteroscedasticity)	الأخطاء لديها تباين ثابت	الأخطاء لديها تباين غير ثابت	p-value صغيرة: رفض H_0 ، قبول H_1	يُستخدم لفحص ثبات التباين في الأخطاء، وتجانس الانتشار مطلوب لكفاءة تقديرات الأقل تربيعات.
الطبيعية (Normality)	الأخطاء تتبع التوزيع الطبيعي	الأخطاء لا تتبع التوزيع الطبيعي	p-value صغيرة: رفض H_0 ، قبول H_1	يُستخدم للتأكد من أن الأخطاء تتوزع بشكل طبيعي، وهو شرط لصحة العديد من الاختبارات الإحصائية.
الارتباط المتعدد (Multicollinearity)	لا يوجد ارتباط متعدد بين المتغيرات المستقلة	يوجد ارتباط متعدد بين المتغيرات المستقلة	رفض H_0 : VIF > 10 ، قبول H_1	يُستخدم للكشف عن الارتباط القوي بين المتغيرات المستقلة، الذي يمكن أن يؤدي إلى مشاكل في تفسير النموذج.

1. **الارتباط الذاتي :(Autocorrelation)** يمكن استخدام اختبار Durbin-Watson في Stata باستخدام الأمر **dwstat** قبل ذلك يجب تأكيد السلسلة الزمنية بأنها سنوية أو غير ذلك من خلال الأمر tsset years بعد تنفيذ النموذج الانحداري. القيم القريبة من 2 تعني عدم وجود ارتباط ذاتي، بينما القيم القريبة من 0 أو 4 تشير إلى وجود ارتباط ذاتي . بعد قراءة جدول Durbin-Watson إذا كانت القيمة بين 1.5 و 2.5 فانه لا يوجد ارتباط ذاتي وإذا كانت اقل من 1.5 أو أكبر من 2.5 فانه يوجد مشكل الارتباط الذاتي .

2. **عدم تجانس الانتشار :(Heteroscedasticity)** يمكن القيام بذلك باستخدام اختبار Breusch-Pagan أو White باستخدام الأوامر **hettest** إذا كانت chi-2 أقل أو تساوي 5 فإن هناك مشكل و العكس.

3. **الطبيعية :(Normality)** يتم التحقق من طبيعية الأخطاء باستخدام اختبار Shapiro-Wilk باستخدام الأمر **swilk** بعد تشغيل النموذج. إذا كانت prob>zأكبر من 5% فإن قيم المتغير ذات توزيع طبيعي و العكس.

4. **الارتباط المتعدد :(Multicollinearity)** يمكن التحقق من الارتباط المتعدد باستخدام مؤشر التضخم في التباين VIFعن طريق الأمر **vif** بعد تقدير النموذج. قيم VIF التي تزيد عن 10 تشير عادة إلى وجود مشكلة في الارتباط المتعدد و 1/ vif يجب أن يكون أكبر من 0.1 .

د . عبد الحق لفيلف