التعدد الخطي (Multicollinearity)

توجد في نماذج الانحدار الخطي المتعدد حصراً ولا تعاني منها مشكلات الانحدار البسيط ، فهي تنشأ بسبب تداخل (تشابك) بين المتغيرات التوضيحية (التفسيرية) ذاتها. كما ان الفرضية التي ترتبط بوجود قدر كافٍ من التغيرات في قيم المتغيرات التوضيحية وان عدد المشاهدات يكون اكبر من عدد المتغيرات التوضيحية هي فرضيات متممة للتعدد الخطي .

ويمكن تحديدها بالفرضية :

{ رتبة مصفوفة المعلومات(X) تكون تامة من ناحية الاعمدة.}

وبالرموز :

$\rho(X) = k + 1 < n$

حيث ان n : حجم العينة المستخدمة في الانحدار.

k : عدد المتغيرات التوضيحية في المعادلة .

(k+ 1): عدد المعلمات في المعادلة بضمنها المقطع الثابت.

والمفهوم يتضمن " التعدد الخطي التام "(perfect multicolinearity)" كما يتضمن التعدد الخطي شبه التام (semi- perfect multicolinearity)

ان التعدد الخطي التام يتحقق اذا كان واحد أو اكثر من المتغيرات التوضيحية (المستقلة) توليفة خطية تامة من المتغيرات الأخرى كالآتي:

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_k X_k = 0$$

λ: ثوابت ليست جميعها أصفاراً .

اما التعدد الخطي شبه التام فيتحقق اذا ارتبطت المتغيرات التوضيحية أو المستقلة ببعضها ارتباطاً قوياً ولكن غير تام :

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_k X_k + e_i = 0$$

e_i: متغير عشوائي.

وفي حالة انعدام مشكلة التعدد الخطي فان المتغيرات التوضيحية تسمى متغيرات متعامدة (orthogonal) أي ان كل متغير توضيحي يؤثر في المتغير التابع بطريقه منفصلة تماماً. وبذلك تكون نتائج التقدير في الانحدار المتعدد متطابقة تماماً مع نتائج التقدير في الانحدار البسيط حيث ان مصفوفة فيشر(X´X) تكون قطرية لان معامل الارتباط بين المتغيرات التوضيحية معدوم $(r_{X_i X_j} = 0)$ أو ان الارتباط المشترك بين المتغيرات التوضيحية يسأوي صفراً .

$$\sum_{t=1}^{n} (X_{ti} - \overline{X}_i)(X_{tj} - \overline{X}_j) = 0 \qquad \forall \qquad i \neq j = 1,2.\cdots,k$$