

1. مقدمة:

تناولنا فيما سبق تحليل الانحدار الخطي البسيط والمتعدد بالاعتماد على المتغيرات الكمية، وعموماً إن هذا الأسلوب الإحصائي يشترط ذلك، لكن في بعض الحالات تنشأ الحاجة إلى استخدام بعض المتغيرات من القياس الاسمي المرتبطة بالظاهرة المدروسة، وهي عبارة عن متغيرات غير قابلة للقياس الكمي كالجنس، المستوى الدراسي، المنطقة الجغرافية ... الخ. هذه المتغيرات لها تأثير على المتغير التابع إلى جانب المتغيرات المستقلة المدرجة في النموذج. وسنحاول من خلال هذا الدرس توضيح الطرق المختلفة التي تستخدم لتضمين المتغيرات الصورية في النموذج القياسي.

2. مفهوم المتغيرات الصورية وأهميتها:

المتغيرات الصورية هي متغيرات وهمية أو اصطلاحية تُستخدم في النماذج الإحصائية، مثل نماذج الانحدار، لتحويل المتغيرات النوعية (Categorical Variables) التي تأخذ قيمًا وصفية أو فئوية إلى قيم رقمية (عادةً 0 و 1)، والتي لا يمكن تمثيلها مباشرة بقيم عددية.

فمثلاً لو أردنا دراسة تأثير مستوى الطالب في الرياضيات والاحصاء على العلامة المحصلة في مقياس برامج إحصائية في شكل نموذج انحدار خطي متعدد، تنشأ الحاجة إلى التمييز بين تأثير المتغيرات المستقلة على المتغير التابع حسب متغيري الجنس وشعبة البكالوريا الخاصة بالطالب. وهنا نقوم بتضمين متغيرات مرتبطة بالجنس والشعبة تأخذ قيم 0 أو 1 بالنسبة لكل فئة من فئات المتغيرين يرمز لها عموماً بالرمز D_i ونكتب:

- الجنس يقسم إلى متغيرين:

D_1 : يأخذ القيمة 1 إذا كان الفرد المعني من الجنس ذكر وصفر لغير ذلك.

D_2 : يأخذ القيمة 1 إذا كان الفرد المعني من الجنس أنثى وصفر لغير ذلك.

بالنسبة لمتغير الشعبة يقسم إلى عدد الشعب المتوفرة، ولتكن 3: شعبة تسيير واقتصاد، شعبة العلوم التجريبية، شعبة تقني رياضي. فيكون لدينا ثلاث متغيرات:

D_3 : يأخذ القيمة 1 إذا كان الفرد المعني من شعبة تسيير واقتصاد وصفر لغير ذلك.

المناطق الجغرافية الشمال والجنوب. ومن أجل اختبار فرضية أن حجم المبيعات يختلف حسب المنطقة الجغرافية بغض النظر على التغيرات في متغيرات أخرى كالسعر ومقياس درجة الحرارة وغيرها، نقوم بإنشاء متغير صوري يعبر عن الاختلاف في المنطقة وليكن D بحيث:

$$D = \begin{cases} 1 : \text{منطقة الشمال} \\ 0 : \text{منطقة الجنوب} \end{cases}$$

تجدر الإشارة إلى أن هذا الترميز ليس له تأثير على النتائج، أي أن إعطاء القيمة 1 لمنطقة الشمال أو الجنوب ليس مهما وليس له تأثير على النتائج، لكن قراءة وتفسير النتائج يتوقف على ذلك.

فتقدير النموذج في هذه الحالة يقودنا للحصول على نموذجين:

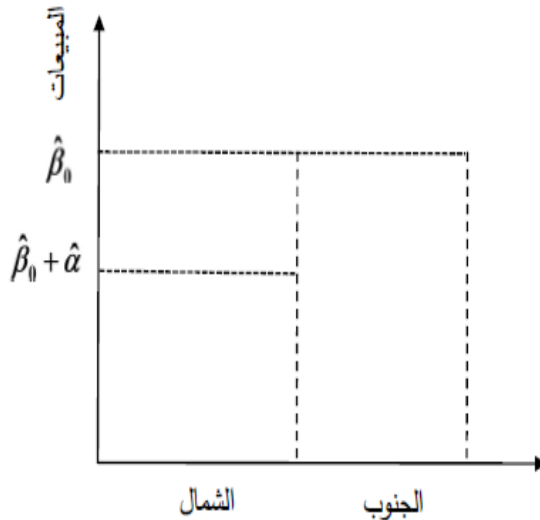
في حالة $D=0$:

$$\begin{aligned} \hat{Y}_i &= \hat{\beta}_0 + \hat{\alpha}D_i \\ \hat{Y}_i &= \hat{\beta}_0 + \hat{\alpha}(0)_i \quad (4-1) \end{aligned}$$

في حالة $D=1$:

$$\begin{aligned} \hat{Y}_i &= \hat{\beta}_0 + \hat{\alpha}D_i \\ \hat{Y}_i &= \hat{\beta}_0 + \hat{\alpha}(1)_i = \hat{\beta}_0 + \hat{\alpha} \quad (4-2) \end{aligned}$$

ويمكن التعبير عن ذلك بيانياً كما يلي:



يتضح من الشكل أن قيمة المعلمة المقدرة لـ α سالبة مما يبين أن المبيعات في المنطقة الشمالية أقل

- حالة دراسة تأثير المتغير النوعي على المتغير التابع الى جانب أثر متغير مستقل وليكن متوسط درجة الحرارة أو السعر أو طاقة المكيف وما الى ذلك، في هذه الحالة يكون لدينا النموذج من الشكل:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (5)$$

β_0 في المعادلة تمثل متوسط قيم المتغير التابع عندما تكون X تساوي صفر، نفترض في معادلة الانحدار أن قيمة β_0 ثابتة لكل المشاهدات في بيانات الدراسة. ولكن المعامل قد يختلف اعتمادا على اختلاف بيانات الدراسة، فمثلا الاختلاف الجغرافي في المثال السابق فمن أجل قياس هذا الاختلاف نحول متغير المنطقة الى متغير صوري واختبار صحة وجود هذا الاختلاف.

فالمتغير الصوري يمكن أن يمثل التأثير النوعي وذلك بترميز (إعطاء رموزاً رقمية) متغير المنطقة في البيانات المتوفرة بتخصيص القيمة (1) لمنطقة الشمال و (0) لمنطقة الجنوب. فيكون لدينا المتغير D على الشكل التالي:

$$D = \begin{cases} 1 : \text{منطقة الشمال} \\ 0 : \text{منطقة الجنوب} \end{cases}$$

وبإدخال المتغير الصوري في النموذج (3)، يصبح لدينا:

$$Y_i = \beta_0 + \beta_1 X_i + \alpha D_i + u_i \quad (6)$$

ومن أجل معرفة أثر المتغير الصوري D في النموذج، نأخذ بعين الاعتبار قيمتي هذا الأخير، بحيث يصبح لدينا صيغتين:

من أجل D=0:

$$Y_i = \beta_0 + \beta_1 X_i + \alpha D(0)_i + u_i$$

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (6-1)$$

من أجل D=1:

$$Y_i = \beta_0 + \beta_1 X_i + \alpha D(1)_i + u_i$$

الحالة الأولى تأثير التغيرات الموسمية فقط:

في هذه الحالة يتم تضمين 1-ل من المتغيرات الصورية الى جانب حد ثابت في النموذج، بحيث ل هو عدد المواسم أو الفصول على الشكل التالي:

$$Y_t = \beta_0 + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + u_t \quad (9)$$

فمثلا في دراسة للمبيعات لنوع معين من المشروعات خلال فترة زمنية T مقسمة الى 4 فصول، يمكن ملاحظة أن المبيعات تزيد وتنقص من فصل الى آخر . ومن أجل دراسة تأثير الفصول نستخدم المتغيرات الصورية الممثلة للفصول الأربعة ولتكن D_j بحيث $D=1$ اذا كانت المشاهدة تخص الفصل j و $D=0$ لغير ذلك. ولادراجها في النموذج نعتبر الفصل الأول وليكن الخريف مثلا كثقة أساس، ونستخدم المتغيرات الخاصة بفصل الشتاء، الربيع والصيف. ونقدر النموذج وفق الصيغة (9).

فينتج لدينا النتائج على النحو التالي:

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\alpha}_1(0) + \hat{\alpha}_2(0) + \hat{\alpha}_3(0) = \hat{\beta}_0 \quad (9-1) \text{ بالنسبة لفصل الخريف}$$

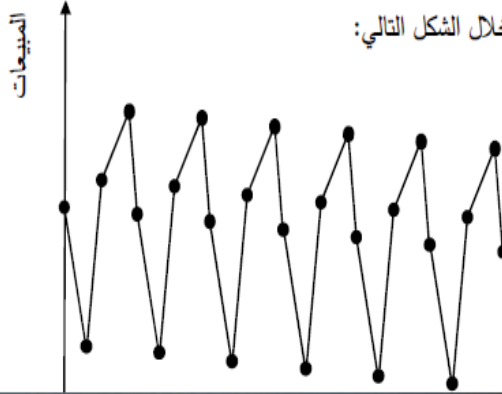
$$\hat{Y}_t = \hat{\beta}_0 + \hat{\alpha}_1(1) + \hat{\alpha}_2(0) + \hat{\alpha}_3(0) = \hat{\beta}_0 + \hat{\alpha}_1 \quad (9-1) \text{ بالنسبة لفصل الشتاء}$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\alpha}_1(0) + \hat{\alpha}_2(1) + \hat{\alpha}_3(0) = \hat{\beta}_0 + \hat{\alpha}_2 \quad (9-1) \text{ بالنسبة لفصل الربيع}$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\alpha}_1(0) + \hat{\alpha}_2(0) + \hat{\alpha}_3(1) = \hat{\beta}_0 + \hat{\alpha}_3 \quad (9-1) \text{ بالنسبة لفصل الصيف}$$

ومن هنا يتضح لنا أنه تم قياس تأثير كل فصل على المبيعات من المشروعات بالنسبة الى فصل الخريف، فهو يزيد أو ينقص حسب قيمة معاملات المتغيرات الصورية المقدرة وإشارتها.

ويمكن توضيح ذلك بيانيا من خلال الشكل التالي:



الحالة الثانية تأثير التغيرات الموسمية وتأثير الزمن:

في هذه الحالة يتم تحليل تأثير التغيرات الموسمية وتأثير الزمن الذي يعرف بالاتجاه العام، فكما هو معروف فإن المبيعات تأخذ منحى متزايد عبر الزمن t مع تسجيل زيادة وانخفاض موسمي، ويتم ترجمة ذلك في شكل نموذج يعرف بالنموذج الخطي الموسمي، وهو نموذج يمزج بين أثر الاتجاه والأثر الموسمي، ويكتب على الشكل التالي:

$$Y_t = \beta_0 + \beta_1 t + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \alpha_4 D_4 + u_t \quad (10)$$

بحث t يمثل متغير الزمن يأخذ قيم من 1 إلى T .

مثال ذلك، تحليل المبيعات من منتج معين عبر الزمن. ففي هذه الحالة يكون المتغير التابع Y (المبيعات) يفسر بالزمن، وتأثير الزمن هنا يتم تقسيمه الى قسمين: تأثير الفترة الزمنية ويعرف بالاتجاه العام، وتأثير التغيرات الموسمية. ويتقدير النموذج نتحصل على النتائج التالية:

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_2(t) + \hat{\alpha}_1(0) + \hat{\alpha}_2(0) + \hat{\alpha}_3(0) = \hat{\beta}_0 + \hat{\beta}_2(t) \quad (10-1) \text{ بالنسبة لفصل الخريف:}$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_2(t) + \hat{\alpha}_1(1) + \hat{\alpha}_2(0) + \hat{\alpha}_3(0) = \hat{\beta}_0 + \hat{\alpha}_1 + \hat{\beta}_2(t) \quad (10-1) \text{ بالنسبة لفصل الشتاء:}$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_2(t) + \hat{\alpha}_1(0) + \hat{\alpha}_2(1) + \hat{\alpha}_3(0) = \hat{\beta}_0 + \hat{\alpha}_2 + \hat{\beta}_2(t) \quad (10-1) \text{ بالنسبة لفصل الربيع:}$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_2(t) + \hat{\alpha}_1(0) + \hat{\alpha}_2(0) + \hat{\alpha}_3(1) = \hat{\beta}_0 + \hat{\alpha}_3 + \hat{\beta}_2(t) \quad (10-1) \text{ بالنسبة لفصل الصيف:}$$

ويمكن توضيح ذلك بيانيا كما يلي:

