المحور الرابع: الارتباط الجزئي، الازدواج الخطي وطرق اختيار المتغيرات التفسيرية:

معامل الارتباط

يعرف معامل الارتباط والذي يرمز له بالرمز r بأنه مقياس رقمي يقيس قوة الارتباط بين متغيرين، حيث تتراوح قيمته ما بين $(+1)$ و (-1) أي أن: $-1 \leq r \leq +1$

وتدل إشارة المعامل الموجبة على العلاقة الطردية، بينما تدل إشارة المعامل السالبة على العلاقة العكسية. ويمكن حسابه بالقانون التالي:

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

مثال:

سجلت قراءة تقريبية لحجم الإنتاج وحجم صادرات بالنفط الخام في أحد الدول العربية (بالمليار برميل) خلال عدة سنوات كما يلي:

	2000	2001	2002	2003	2004	2005
حجم الانتاج	3	4	2	2	2	2
حجم الصادرات	2	2	2	1	1	1

أدرس وجود علاقة ارتباط خطية بين حجم الإنتاج وحجم صادرات النفط الخام.

الحل:

Y^2	X^2	XY	Y	X
4	9	6	2	3
4	16	8	2	4
4	4	4	2	2
1	4	2	1	2
1	4	2	1	2
1	4	2	1	2
15	41	24	9	15

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

$$r_p = \frac{6(24) - (15)(9)}{\sqrt{((6 \times 41) - 15^2)((6 \times 15) - 9^2)}} = \frac{144 - 135}{\sqrt{(246 - 225)(90 - 81)}}$$

$$r_p = \frac{9}{\sqrt{189}} = \frac{9}{13.75} = 0.65$$

الارتباط الجزئي

يقيس الارتباط الجزئي العلاقة بين متغيرين Y,X بعد ثبات أثر أي متغيرات أخرى, أي تحديد أثر المتغيرات الاخرى التي قد تؤثر على أحد المتغيرين Y,X واللذان نريد قياس العلاقة بينهما.

السؤال الآن.....لماذا نستخدم الارتباط الجزئي؟؟؟؟

إن العلاقة بين المتغيرين Y,X قد تكون علاقة كاذبة أو غير حقيقية Spurious وذلك عندما يكون هناك متغير ثالث خارجي يؤثر في كل منهما و في نفس الوقت لا يؤثر أحدهما في الآخر. فقد تكون هناك علاقة بين زيادة مبيعات الآيس كريم وانخفاض الانتاجية ولكن هذه العلاقة غير ناتجة عن تأثير الايس كريم في الانتاجية أو تأثير الانتاجية في الآيس كريم. انها علاقة غير حقيقية لأنه قد يكون هناك متغيراً ثالثاً مثل درجة الحرارة المرتفعة يؤثر في كلا المتغيرين, وهذا المتغير الثالث يسمى المتغير العرضي Extraneous .

ويستخدم تحليل الارتباط الجزئي لاختبار قوة واتجاه العلاقة الخطية بين متغيرين بعد تثبيت أثر متغير آخر غير حقيقي قد يتسبب وجوده في نتائج غير دقيقة.

فإذا كان $r_{12.3}$ يعبر عن الارتباط الجزئي بين X_1 و X_2 مع تثبيت X_3 فإننا نجد:

$$r_{12.3} = \frac{r_{12} - (r_{13} \times r_{23})}{\sqrt{[1 - r_{13}^2][1 - r_{23}^2]}}$$

وبصورة مماثلة، إذا كانت $r_{12.34}$ هي معامل الارتباط الجزئي بين بين X_1 و X_2 مع تثبيت X_3 و X_4 فإننا نجد:

$$r_{12.34} = \frac{r_{12.4} - (r_{13.4} \times r_{23.4})}{\sqrt{[1 - r_{13.4}^2][1 - r_{23.4}^2]}} = \frac{r_{12.3} - (r_{14.3} \times r_{24.3})}{\sqrt{[1 - r_{14.3}^2][1 - r_{24.3}^2]}}$$

العلاقة بين معاملات الارتباط المتعددة والجزئية:

يمكن الحصول على نتائج ذات أهمية تربط بين معاملات الارتباط المتعددة ومعاملات الارتباط الجزئية المختلفة كما يلي:

$$1 - r_{1.23}^2 = (1 - r_{12}^2)(1 - r_{13.2}^2)$$

$$1 - r_{1.234}^2 = (1 - r_{12}^2)(1 - r_{13.2}^2)(1 - r_{14.23}^2)$$

مثال عن معامل الارتباط الجزئي

مثال

قام باحث بدراسة إحصائية على 100 مؤسسة لمعرفة العلاقة بين أسعار الفائدة والضرائب المدفوعة وحجم الاستثمارات، عند مستوى دلالة (0.05) علماً أن معامل الارتباط بين أسعار الفائدة والضرائب (0.40) وعلاقة أسعار الفائدة والاستثمارات (0.30) وعلاقة الضرائب والاستثمارات (0.60)

علاقة أسعار الفائدة بالضرائب بعد استبعاد حجم الاستثمارات

$$r_{12.3} = \frac{r_{12} - (r_{13} \times r_{23})}{\sqrt{[1 - r_{13}^2][1 - r_{23}^2]}} = \frac{0.40 - (0.30 \times 0.60)}{\sqrt{[1 - (0.30)^2][1 - (0.0.60)^2]}}$$
$$= 0.288$$

علاقة أسعار الفائدة والاستثمارات بعد استبعاد الضرائب

$$r_{13.2} = \frac{r_{13} - (r_{12}.r_{23})}{\sqrt{[1 - r_{12}^2][1 - r_{23}^2]}} = \frac{0.30 - (0.40 \times 0.60)}{\sqrt{[1 - (0.40)^2][1 - (0.60)^2]}}$$
$$= 0.081$$

علاقة الضرائب بالاستثمارات بعد استبعاد أسعار الفائدة

$$r_{23.1} = \frac{r_{23} - (r_{12}.r_{13})}{\sqrt{[1 - r_{12}^2][1 - r_{13}^2]}} = \frac{0.60 - (0.40 \times 0.30)}{\sqrt{[1 - (0.40)^2][1 - (0.30)^2]}}$$
$$= 0.544$$

ويتم اختبار الدلالة الاحصائية لمعامل الارتباط الجزئي بالاختبار $(Student)$ وفق الصيغة الاتية علما ان درجة الحرية هي $(n - w - 1)$.

$$H_0 : r_{ij} = 0$$

$$H_1 : r_{ij} \neq 0$$

$$t = r_{ij} \times \sqrt{\frac{n-w-2}{1-(r_{ij})^2}} = r_{23} \times \sqrt{\frac{n-1-2}{1-(r_{23})^2}}$$

$$= 0.544 \sqrt{\frac{100-3}{1-(0.544)^2}} = 0.544 \sqrt{\frac{97}{1-0.295}}$$

$$= 0.544 \times 11.729 = 6.38$$

وبعد مقارنتها بالقيمة الجدولية عند مستوى دلاله $\alpha = 0.05$ درجة حرية (97) فأن العلاقة دالة مقارنة مع القيمة الجدولية 1,98

وحيث أن قيمة معيار الاختبار المستخرجة هي أكبر من قيمته الجدولية، عليه ترفض فرضية العدم (H_0) أي أن هنالك علاقة معنوية بين دخل الأسرة الشهري وأنفاقها باستبعاد أثر عدد أفراد الأسرة، استناداً إلى