الارتباط الخطي Linear Correlation

يعتبر موضوع الارتباط الخطي البسيط احد اكثر المواضيع الإحصائية تطبيقا في مجالات العلوم بسبب كونه يقيس قوة العلاقة بين المتغيرات الخاصة بالدراسة فمثلا علاقة النبات الطبيعي بالمطر، علاقة التلوث بمخلفات المصانع وغيرها من العلاقات الأخرى حيث تتراوح قيمته بين بين (1+) و(1-)

ويكون لمعامل الارتباط R الخصائص الآتية:

- ❖ لا يصلح معامل الارتباط البسيط في الا في حالة العلاقات الخطية.
- ❖ قيمته تساوي الصفر عندما تكون الظاهرتان مستقلتان تماما.
- ❖ قيمته موجبة عندما يكون الارتباط بين المتغيرين طرديا، ويكون قوياً عندما تكون قيمته قريبة من الواحد الصحيح، وضعيفاً عندما تكون قيمته قريبة من الصفر .
- ❖ قيمته سالبة عندما يكون الارتباط بين المتغيرين عكسيا، ويكون قوياً عندما تكون قيمته قريبة من -1 وضعيفاً عندما يكون قريباً من الصفر.

أي أن قيمته محصورة بين: $-1 \leq r \leq +1$

ويقسم الارتباط الخطي بصورة عامة الى ثلاثة اقسام :

1- **الارتباط الخطي البسيط**
2- **الارتباط الخطي المتعدد**
3- **الارتباط الخطي الجزئي**

ان هذه الانواع الثلاثة أعلاه تقاس في حالة كون البيانات كمية اما اذا كانت البيانات مقاسة بشكل وصفي مثل العلاقة بين ذكاء الطالب وتقديراته فاننا سوف نستخدم مقياس يدعى بمعامل ارتباط الرتب على أساس اننا سوف نعطي رتبا لهذه الصفات لكي يسهل التعامل معها بصورة عددية.

الارتباط الخطي البسيط Coefficient correlation

هو عبارة عن مقياس لدرجة قوة العلاقة بين متغيرين فقط ويرمز له بالرمز r ويحسب وفق القانون التالي:

$$r_{xy} = \frac{\sum xiyi - \frac{(\sum xi)(\sum yi)}{n}}{\sqrt{(\sum xi^2 - \frac{(\sum xi)^2}{n})(\sum yi^2 - \frac{(\sum yi)^2}{n})}}$$

او القانون التالي

$$r_{xy} = \frac{\sum(xi-\overline{x})(yi-\overline{y})}{\sqrt{\sum(xi-\overline{x})^2 \sum(yi-\overline{y})^2}}$$

وان معامل الارتباط يقع بين $-1 < r_{xy} < 1$ بحيث ان $y_{xy} > \mp 0.5$ يوجد ارتباط والعلاقة قوية و $r_{xy} < \mp 0.5$ ارتباط ضعيف

مثال// جد معامل الارتباط للبيانات التالية:

xi	*yi*
13	15
19	22
13	13
18	20
14	13
17	20
14	15
17	19
15	15
16	18

	xiyi	*Yi*^{2}	Xi^2
	195	225	169
	418	*484*	*361*
	169	*169*	*169*
	360	*400*	*324*
	182	*169*	*169*
	340	*400*	*289*
	210	*225*	*169*
	323	*361*	*289*
	225	*225*	*225*
	228	*325*	*256*
المجموع	*2710*	*2982*	*2474*

$$r_{xy} = \frac{\sum xiyi - \frac{(\sum xi)(\sum yi)}{n}}{\sqrt{(\sum xi^2 - \frac{(\sum xi)^2}{n})(\sum yi^2 - \frac{(yi)^2}{n})}} = \frac{2710 - \frac{(156)(170)}{10}}{\sqrt{(2474 - \frac{(156)^2}{10})(2982 - \frac{(170)^2}{10})}} = 0.95$$

نلاحظ ان قيمة معامل الارتباط موجبة واكبر من $r_{xy} > \mp 0.5$ لذلك فان الارتباط قوي والعلاقة طردية

معامل الارتباط المتعدد Multiple Correlation coefficient

إن جميع معاملات الارتباط التي تمت دراستها تقيس العلاقة بين متغيرين اثنين فقط ولكن في حالة وجود 3 متغيرات فأكثر نحتاج إلى صيغة أخرى هي معامل الارتباط المتعدد.

ويعرف الارتباط المتعدد بأنه القيمة العددية التي تقيس العلاقة الارتباطية بين 3 متغيرات وأكثر .على سبيل المثال الارتباط بين الرصاص والخارصين والنحاس في التربة , أو الارتباط بين الكبريتات والكالسيوم والمغنيسيوم في الماء وغيرها من الأمثلة وسندرس هنا حالة ثلاثة متغيرات فقط ولغرض إيجاد معامل الارتباط المتعدد بين 3 متغيرات نتبع الخطوات التالية:

1- حساب معامل الارتباط البسيط بين المتغيران الأول والثاني
2- حساب معامل الارتباط البسيط بين المتغيران الأول والثالث
3- حساب معامل الارتباط البسيط بين المتغيران الثاني والثالث

خصائص معامل الارتباط المتعدد

1- إن قيمة معامل الارتباط المتعدد تتراوح بين (1,0) بمعنى أن قيمته موجبة دائما أي إن العلاقة بين 3 متغيرات فأكثر هي دائما طردية.
2- إن قيمته تزداد كلما ازداد عدد المتغيرات الداخلة في الدراسة بمعنى إن قيمة معامل الارتباط المتعدد ل4 متغيرات هي اكبر من قيمة معامل الارتباط المتعدد ل 3 متغيرات .

ان المعامل الذي يقيس درجة العلاقة بين متغير معين ومجموعة من المتغيرات يسمي بمعامل الارتباط المتعدد R ولنفرض انه لدينا ثلاثة متغيرات x1,x2,x3 فان معامل الارتباط المتعدد بين x1 مع x2 ,x3 ويكون وفق القانون الاتي:

$$R_{x1.x2x3} = \sqrt{\frac{(r_{x1x2})^2 + (r_{x1x3})^2 - 2(r_{x1x2})(r_{x1x3})(r_{x2x3})}{1 - (r_{x2x3})^2}}$$

وان معامل الارتباط محصور بين $0 \leq R \leq 1$

مثال // جد معامل الارتباط المتعدد للبيانات التالية : لنفرض ان x1 يمثل تلوث التربة بالرصاص وx2 يمثل التلوث بالزنك و x3 يمثل التلوث بالخارصين وعلى اعتبار ان n=20 فانه تم الحصول على المعلومات التالية:
:

$r_{x1x2} = 0.91$
$r_{x1x3} = 0.39$
$r_{x2x3} = 0.62$

المطلوب// احسب معامل الارتباط المتعدد بين($R_{x1.x2x3}$) ؟

الحل// $R_{x1.x2x3} = \sqrt{\frac{(r_{x1x2})^2+(r_{x1x3})^2-2(r_{x1x2})(r_{x1x3})(r_{x2x3})}{1-(r_{x2x3})^2}}$ =

$= \sqrt{\frac{(0.91)^2+(0.39)^2-2(0.91)(0.39)(0.62)}{1-(0.62)^2}} = 0.937$

1- احسب معامل الارتباط بين($R_{x2.x1x3}$) ؟

الحل سوف يكون بالقانون التالي $R_{x2.x1x3} = \sqrt{\frac{(r_{x1x2})^2+(r_{x2x3})^2-2(r_{x1x2})(r_{x1x3})(r_{x2x3})}{1-(r_{x1x3})^2}}$

2- احسب معامل الارتباط بين($R_{x3.x1x2}$)؟

الحل سوف يكون بالقانون التالي $R_{x3.x1x2} = \sqrt{\frac{(r_{x1x3})^2+(r_{x2x3})^2-2(r_{x1x2})(r_{x1x3})(r_{x2x3})}{1-(r_{x1x2})^2}}$

الارتباط الجزئي الخطي Partial Correlation

في بعض الظواهر والدراسات يوجد هناك عدد من المتغيرات (ثلاثة فأكثر) مرتبطة بعلاقة رياضية فيما بينها مثل تلوث الماء يكون مرتبط بالأملاح و الكبريتات وكذلك تلوث التربة يرتبط بالرصاص و والكادميوم والخارصين ففي هذه الحالة ولغرض حساب معامل الارتباط بين متغيرين اثنين في دراسة معينة مع وجود متغيرات أخرى نلجأ إلى حساب ما يسمى بالارتباط الجزئي .

الارتباط الجزئي هو العلاقة الرياضية الصافية بين متغيرين اثنين فقط مع وجود متغيرات أخرى قيد الدراسة ويمكن حساب هذه العلاقة الرياضية من خلال معامل الارتباط الجزئي.

أن الفرق بينه وبين معامل الارتباط البسيط هو أن معامل بيرسون يستخرج العلاقة بين متغيرين اثنين لأي ظاهرة بدون يأخذ بنظر الاعتبار وجود متغيرات أخرى تؤثر في الظاهرة أو لا بينما معامل الارتباط الجزئي لا يأخذ بنظر الاعتبار وجود متغيرات أخرى تؤثر في الظاهرة فحسب وإنما يقوم باستبعاد أثرها لكي يستخرج الارتباط الصافي بين أي متغيرين.

خطوات حساب معامل الارتباط الجزئي (ثلاث متغيرات فقط)

إذا كان لدينا ثلاث متغيرات داخلة في الدراسة وأردنا حساب معامل الارتباط الجزئي بين كل متغيرين اثنين منها نتبع الخطوات التالية:

1- حساب معامل الارتباط البسيط بين المتغيران الأول والثاني
2- حساب معامل الارتباط البسيط بين المتغيران الأول والثالث
3- حساب معامل الارتباط البسيط بين المتغيران الثاني والثالث

خصائص معامل الارتباط الجزئي

1- إن قيمة معامل الارتباط الجزئي تتراوح بين (1,-1)
2- تفسر قيمته كما تفسر قيمة معامل الارتباط البسيط.

إن معامل الارتباط الجزئي لأي متغيرين تكون إشارته مماثلة لإشارة معامل الارتباط البسيط بينهما.

معامل ارتباط الصفات(الرتب)

Rank correlation Coefficient

ان الصيغ السابقة الخاصة بمعامل الارتباط البسيط والجزئي والمتعدد كانت على أساس كون المتغيرات كمية ولكن في كثير من الأحيان يكون لدينا متغيرات غير قابلة للقياس تسمى بالمتغيرات الوصفية والمعامل الذي يقيس درجة العلاقة بين هذه المتغيرات يسمى بمعامل ارتباط الرتب وسوف ناخذ معامل ارتباط الرتب لسبيرمان كمثال على ذلك:

معامل ارتباط الرتب لسبيرمان

لنفرض انه لدينا متغيرين هما (x ,y) وهما من النوع الوصفي مترابطين ويمكن ترتيبها تصاعديا او تنازليا وفق معيار معين ويحسب وفق القانون الاتي :

$$r_s = 1 - \frac{6\sum di^2}{n(n^2-1)}$$

حيث ان معامل ارتباط الرتب r_s= والفرق بين القيم المتناظرة di=xi-yi= وعدد المشاهدات n= وقيمة هذا المعامل **$-1 < r_s < 1$**

مثال // اذا كانت تقديرات (6) طلاب في مادة الإحصاء(x) ومادة ميكانيك الموائع (y) بالشكل التالي والمطلوب حساب معامل الارتباط بينهما؟

x	y
جيد	متوسط
متوسط	جيد
ضعيف	مقبول
مقبول	ضعيف
جيدجدا	امتياز
امتياز	جيدجدا

الحل/ نرتب التقديرات أعلاه تصاعديا او تنازليا واعطائهم ارقام معينة وكالاتي: (ضعيف=1 , مقبول= 2 , متوسط=3 ,جيد=4 , جيدجدا=5 ,امتياز=6) ثم نعود الى التقديرات الاصلية بالسؤال واعطائها هذه الأرقام حسب ترتيبها بالسؤال

x	y	x	y	di=xi-yi	di
جيد	متوسط	4	3	1	1
متوسط	جيد	3	4	1-	1
ضعيف	مقبول	1	2	1-	1
مقبول	ضعيف	2	1	1	1
جيدجدا	امتياز	5	6	1-	1
امتياز	جيدجدا	6	5	1	1

$$r_s = \frac{6\sum di^2}{n(n^2-1)} = 1 - \frac{6(6)}{6(6^2-1)} = 1 - \frac{6}{35} = 0.829$$