جامعة البصرة
كلية التربية للبنات

قسم الجغرافية
المرحلة الثالثة
المادة : ألاحصاء

المحاضرة : العاشرة

تحليل

علاقة الارتباط بين المتغيرات

الكمية والوصفية للظواهر المختلفة

تأليف :

المدرس / وائل قاسم راشد

تحليل علاقة الارتباط بين المتغيرات

بشكل عام تقاس العلاقة احصائيا بين متغيرين باستخدام اسلوبين رئيسيين هما : الارتباط (correlation) والانحدار (regression) وهما اسلوبان متشابهان في نواح كثيرة جدا ولكنهما مختلفان في عدة نواح . فالارتباط يقيس درجة العلاقة واتجاهها بين المتغيرات ، بينما الانحدار يبحث في العلاقة بين المتغيرات من خلال معادلة رياضية يمكن بواسطتها تقدير التنبؤ باحد المتغيرين من خلال المتغير الاخر .

تحليل الارتباط Correlation Analysis :

الارتباط هو اداة من ادوات التحليل الوصفي لمعرفة العلاقة بين متغيرات الظواهر المختلفة. والمتغيرات على شكلين احدهما يسمى متغير مستقل (X) واخريسمى متغير تابع (Y) فاذا كانت العلاقة بين متغيرين فقط كانت العلاقة بسيطة اما العلاقة المتعددة فهي بين المتغير المعتمد (Y) ومجموعة من المتغيرات المستقلة (x_1 ......x_2) اوبين متغيرين مستقلين (x_1,x_2) ، فالارتباط اذن هو وصف درجة تاثر احد المتغيرين بالاخر وبيان مدى العلاقة الواقعة بين هذين المتغيرين اما المقياس لهذه العلاقة يدعى معامل الارتباط correlation coefficient وهو مقياس كمي لقياس قوة هذه العلاقة ويرمز له في حالة المجتمع (ρ) " رو" وفي حالة العينة بـ " r " الذي تتراوح قيمته بين (الصفر و ± ١) . وبما اننا في كثير من النواحي التطبيقية نتعامل مع بيانات عينة مسحوبة من المجتمع الاحصائي ، لذا فاننا سنركز على حساب معامل ارتباط العينة " r " كونه قيمة تقديرية لمعامل ارتباط المجتمع الماخوذة منه العينة . ان معامل الارتباط بشكل عام يركز على نقطتين اساسيتين هما :

١- نوع العلاقة التي تاخذ ثلاثة اشكال اعتمادا على اشارة معامل الارتباط وهي :

أ- العلاقة الطردية " الموجبة " بين المتغيرين أي " $r > 0$ " اذ ان أي زيادة في المتغيرين (x) يصاحبه زيادة في المتغير الثاني (y) او العكس يصاحبه نقص في المتغير الاول (x) يقابله نقص في المتغير الثاني (y) .

ب- العلاقة العكسية " السالبة " بين المتغيرين أي " $r < 0$ " اذ ان أي زيادة في المتغير الاول (x) يقابلها نقص في المتغير الثاني (y) او العكس من ذلك ، ان أي نقص في المتغير الاول (x) يقابله زيادة في المتغير الثاني (y) .

ت- انعدام العلاقة بين المتغيرين أي " $r = 0$ " اذ ان أي زيادة او نقص في المتغير الاول " x " لا يؤدي الى أي تغير في المتغير الثاني " y " .

٢- قوة العلاقة الخطية بين متغيرين يحكم عليها من خلال درجة قرب او بعد قيمة r عن (± ١) ، اذ ان قيمة معامل الارتباط تقع في المدى ($-1< r < 1$) اما اذا كان قيمة معامل الارتباط هو الصفر دلت على عدم وجود علاقة خطية بينهما ، وقد تكون هناك علاقة الا انها غير خطية أي ان التاثير بين المتغيرين يرسم بشكل منحنى . ولقد صنف بعض الاحصائيين درجات قوة العلاقة بالشكل الاتي :

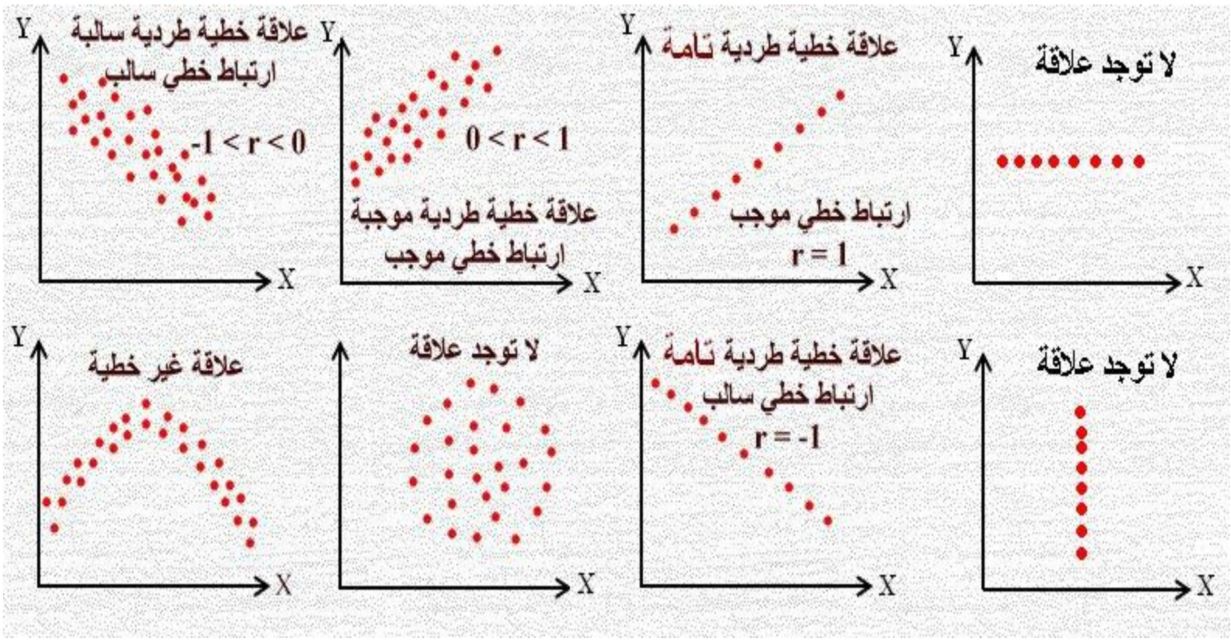
ارتباط سالب					ارتباط موجب				
قوي جدا	قوي	متوسط	ضعيف	ضعيف جدا	ضعيف جدا	ضعيف	متوسط	قوي	قوي جدا

٠.٩   -٠.٧   -٠.٥   - ٠.٣   ٠.٣   ٠.٥   ٠.٧   ٠.٩

-١ ارتباط تام   ٠ عدم وجود ارتباط   +١ ارتباط تام

شكل درجات قوة معامل الارتباط

الشكل الانتشاري لتحديد طبيعة الاتجاه العام للارتباط : يمكن اظهار العلاقة بين المتغيرات بالاشكال الانتشارية الاتية :



اولا : معامل الارتباط بين المتغيرات الكمية المقاسة :

يستخدم لقياس العلاقة بين متغيرات الظواهر القابلة للقياس الكمي " الرقمي " وهذا يشمل جميع الظواهر التي يمكن التعبير عنها بصورة رقمية ومنها الطول والوزن ، وكمية الامطار المتساقطة والانتاج الزراعي وغيرها من الظواهر التي يمكن التعبير عنها رقميا . ويقسم بحسب عدد انواع المتغيرات الى الانواع الاتية :

١- معامل الارتباط البسيط
٢- معامل الارتباط المتعدد
٣- معامل الارتباط الجزئي

١- معامل الارتباط البسيط (بيرسون) Simple Correlation Coefficient :

عندما ترصد ظواهر او مشاهدات واقعية لمتغيرات فسنطرح تساؤلات عدة منها ماهية الية العلاقة بينهما ؟ وهل ان العلاقة بينهما ناتجة عن علاقة سبب ونتيجة اوعلاقة بالصدفة ؟ فمثلا يمكن القول بان درجة سطوع ضوء الشمس هي (السبب) في تغير درجة الحرارة (كنتيجة) . ويمكن قول العكس بان تغيرات درجة الحرارة تسبب تغيرات في السطوع لضوء الشمس باعتبار ان الحرارة العالية تسبب انقشاع الغيوم وهنا تعرض ثبات او تساوي العوامل الاخرى المؤثرة على النتيجة اذن العلاقة ليست خالصة ١٠٠% كما يمكن القول ايضا ان كمية الكوليسترول بالدم تسبب السكتة القلبية ولكنها ايضا ليست خالصة لوجود عوامل اخرى .
ولقياس العلاقة بين متغيرين لبيانات مستمرة نستخدم صيغة معامل الارتباط البسيط بيرسون :

الصيغ الحسابية لمعامل الارتباط :

١-

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

٢-

$$r = \frac{\sum (x_1 - \bar{x})(y_1 - \bar{y})}{\sqrt{\sum (x_1 - \bar{x})^2}\sqrt{\sum (y_1 - \bar{y})^2}}$$

٣-

$$r = \frac{Sxy}{\sqrt{Sxx}\sqrt{Syy}}$$

$$sxy = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$sxx = \sum x^2 - \frac{(\sum x)^2}{\mathrm{n}}$$

$$syy = \sum y^2 - \frac{(\sum y)^2}{\mathrm{n}}$$

ولاختبار قوة المعنوية لـ (r) للتعرف مدى تمثيل العلاقة (r) للعينة المسحوبة الى معامل ارتباط المجتمع (P) المحسوبة منه نستخدم اختبار t لاختبار فرضية :

١- نضع الفرضيات الاتية :

H0:P=0

H1:P≠ 0

٢- نستخرج قيمة t المحسوبة وهي :

$$t=\sqrt{\frac{n-2}{1-r^2}}$$

٣- نستخرج قيمة t الجدولية من جداول t وذلك من خلال درجات الحرية التي تقاس d.f=n-2 وعند مستوى معنوية α سواء كانت ٠.٠٥ اوغيرها .

٤- ثم نقارن بين المحسوبة والجدولية لتبني الفرضية المطلوبة اعتمادا على الرسم للتوزيع الملتوي للبيانات كما في اختبار كاي سكوير

مثال :

الجدول التالي يوضح العلامات النهائية لـ (٨) طلاب في مادتي الاحصاء X والرياضيات Y . جد العلاقة بينهما .

X	٨٥	٦٥	٦٠	٤٥	٨٠	٦٥	٨٥	٩٥
Y	٨٧	٨٢	٥٧	٧٢	٥٢	٦٢	٦٧	٧٧

علما بان :

$$\sum y^2=39656 \sum x^2=43950 \sum x_i=580 \sum xy=41660 \sum y_i=556 \;\; \sum x_i=580$$

$$\therefore r=\frac{(8)(41660)-(580)(556)}{\sqrt{(8)(43950)-(580)^2}\sqrt{(8)(39656)-(556)^2}}$$

097= قوية وطردية

لاختبار المعنوية نحسب tالمحسوبة كما يلي :

$$t=\sqrt{\frac{n-2}{1-r^2}}=\qquad\sqrt{\frac{8-2}{1-(0.97)^2}}$$

$$t=10.08$$

اما t الجدولية فقيمتها عند (0.05,6) = ١.٩٤

وحيث ان t المحسوبة اكبر من الجدولية تقبل فرضية H1 وهي P≠ 0

ويمكن قياس العلاقة بين حجم الانتاج والصادرات وبين مصروف الطالب ودرجات السعي الفصلية وبين اوزان الاطفال حديثي الولادة واطوالهم وبين عدد الساعات الدراسية ودرجات الامتحان وبين سطوع الشمس

ودرجات الحرارة والرطوبة وخطوط العرض وبين نمو السكان ونمو الدخل وعدد كبير من المتغيرات للظواهر المختلفة .

تمارين

تمرين ١: احسب معامل الارتباط بين حجم الاعلاف وحجم اللحوم المنتجة خلال السنوات الاتية ومن هو المتغير المستقل والتابع بحسب المثال:

السنوات	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦
الاعلاف (كغم)	١	٢	٥	٣	٢	١	٤
اللحوم (كغم)	٣	٤	٦	٣	٥	٤	٨

تمرين ٢: جد العلاقة بين عدد الساعات الدراسية للطالب ودرجات امتحان الاحصاء ومن هو المتغير المستقل والتابع بحسب المثال :

الساعات	٢	٤	٣	١	٥
الدرجات	٦٠	٧٠	٥٠	٤٠	٨

تمرين٣ : ما هي العلاقة بين خطوط العرض ودرجات الحرارة لخمس محطات جوية ومن هوالمتغير المستقل والتابع بحسب المثال ؟

X خطوط العرض	3 S	6S	21N	20 S	10N
Y درجة الحرارة	٢	٣	٤	٧	٢٢

تمرين٤ : ما هي علاقة النمو السكاني بنمو الدخل لـ (٦) دول ؟ ومن هوالمتغير المستقل والتابع بحسب المثال

نمو السكان	٢.٤	١	٠.٧	٠.٨	١.١	٣.٥
نمو الدخل القومي	٠.٣	٣.٤	٢	٤	٣.٧	٣.٤

تمرين٥ : البيانات الاتية تمثل كمية المعروض من سلع وسعرها للوحدة الواحدة ومن هوالمتغير المستقل والتابع بحسب المثال.

الكمية المعروضة Y	٣	٥	٧	٢	٩	١١	٦	٨	٦
سعر الوحدة X	٢	٢	٥	٤	٥	٦	٣	٥	٤

تمرين٦ : افترض بيانات لستة عوائل موضحة فيها مصروفاتهم الشهرية ودخلهم الشهري والعلاقة بين الدخل والمصروف ؟

٢- **معامل الارتباط المتعدد** :Multiple Correlation Coefficient

في معظم الدراسات العلمية النظرية والتطبيقية والتخطيطية لا تعتمد علاقة الارتباط على متغيرين فقط احدهما يمثل المتغير المعتمد " التابع" (Y) والاخر يمثل المتغير المستقل (X) بل يمتد ليشمل عددا من المتغيرات المستقلة التي تؤثر بشكل او باخر على المتغير المعتمد . فمثلا انتاج محصول زراعي معين يعتمد على متغيرات عدة وهي (خصوبة التربة ، وكمية المياه ، وكمية الاسمدة ، ودرجات الحرارة ، والسطوع الشمسي) وغيرها ، كما ان الانتاج الصناعي لسلعة ما يتوقف على (حجم الاستثمار ،والمواد الاولية ، والايدي العاملة ، والطاقة المستعملة) وغيرها .

ان دراسة العلاقة بين المتغير المعتمد (Y) ومجموعة المتغيرات المستقلة (X1,X2,X3....Xn) **في ان واحد** يمكن قياسه بمعامل الارتباط المتعدد الذي يقيس قوة العلاقة بين اكثر من متغيرين من المتغيرات ، ويرمز لهذا العامل بالحرف R . ان حساب قيمة هذا المعامل ما هو الا امتداد لحساب قيمة معامل الارتباط البسيط (r) باضافة متغيرات مستقلة اخرى له . وقيمة معامل الارتباط المتعدد تقع ايضا بين الصفر و ± ١.

وسوف نقتصر في دراستنا لهذا المعامل على العلاقة الخطية بين ثلاثة متغيرات Y و X1 و X2 لنحصل على الصيغ الاتية :

$$r_{YX_1} = \frac{n\sum YX_1 - \sum Y \sum X_1}{\sqrt{N\sum Y_1^2 - \left(\sum Y_1\right)^2}\sqrt{n\sum X_1^2 - \left(\sum X_1\right)^2}}$$

$$r_{YX_2} = \frac{n\sum YX_2 - \sum Y \sum X_2}{\sqrt{n\sum Y_1^2 - \left(\sum Y_1\right)^2}\sqrt{n\sum X_2^2 - \left(\sum X_2\right)^2}}$$

$$r_{X_1X_2} = \frac{n\sum X_1X_2 - \sum X_1 \sum X_2}{\sqrt{n\sum X_1^2 - \left(\sum X_1\right)^2}\sqrt{n\sum X_2^2 - \left(\sum X_2\right)^2}}$$

لذا فان الصيغة النهائية لحساب معامل الارتباط المتعدد هي :

$$R_{YX_1X_2} = \sqrt{\frac{r^2_{YX_1} + r^2_{YX_2} - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1 - r^2_{X_1X_2}}}$$

مثال : جد قيمة معامل الارتباط المتعدد من الجدول الاتي :

Y	X1	X2
2	1	1
3	6	8
2	5	2
1	7	6
4	10	8

الحل :

ماهي العمليات اللازمة للحصول على معامل الارتباط المتعدد بين Y و المتغيرين X1 و X2 ؟

Y	X_1	X_2	YX_1	YX_2	X_1X_2	Y_1^2	X_1^2	X_2^2
2	1	1	2	2	1	4	1	1
3	6	8	18	24	48	9	36	64
2	5	2	10	4	10	4	25	4
1	7	6	7	6	42	1	49	36
4	10	8	40	32	80	16	100	64
12	٢٩	25	77	68	181	34	211	169

$$r_{YX_1} = \frac{n\sum YX_1 - \sum Y \sum X_1}{\sqrt{n\sum Y_i^2 - \left(\sum Y_i\right)^2}\sqrt{n\sum X_1^2 - \left(\sum X_1\right)^2}}$$

$$= \frac{5(77)-(12)(29)}{\sqrt{5(34)-(12^2)}\sqrt{5(211)-(29)^2}}$$

$$= \frac{383-348}{\sqrt{(170-144)}\sqrt{(1055-841)}} = \frac{37}{\sqrt{26}\sqrt{214}} = \frac{37}{74.5} = 0.49$$

$$r_{YX_2} = \frac{n\sum YX_2 - \sum Y \sum X_2}{\sqrt{n\sum Y_i^2 - \left(\sum Y_i\right)^2}\sqrt{n\sum X_2^2 - \left(\sum X_2\right)^2}}$$

$$= \frac{5(68)-(12)(25)}{\sqrt{5(34)-(12)^2}\sqrt{5(169)-(25)^2}} = \frac{40}{}$$

$$\frac{340-300}{\sqrt{(170-144)}\sqrt{(845-625)}} = \frac{40}{\sqrt{26}\sqrt{220}} = \frac{40}{75.6} = 0.53$$

$$r_{X_1X_2} = \frac{n\sum X_1X_2 - \sum X_1 \sum X_2}{\sqrt{n\sum X_1^2 - \left(\sum X_1\right)^2}\sqrt{n\sum X_2^2 - \left(\sum X_2\right)^2}}$$

$$r_{X_1X_2} = \frac{5(181)-(29)(25)}{\sqrt{\sum 5(211)-(29)^2}\sqrt{5(169)-(25)^2}}$$

$$= \frac{905-725}{\sqrt{(1055-841)}\sqrt{842-625}} = \frac{180}{\sqrt{214}\sqrt{220}} = \frac{180}{217} = 0.83$$

$$R_{YX_1X_2} = \sqrt{\frac{r^2{}_{YX_1} + r^2{}_{YX_2} - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1-r^2{}_{X_1X_2}}}$$

$$= \sqrt{\frac{(0.49)^2 + (0.53)^2 - 2(0.49)(0.53)(0.83)}{1-(0.83)^2}}$$

$$= \sqrt{\frac{0.240 + 0.280 - 2(0.431)}{0.312}}$$

$$= \sqrt{\frac{0.52 - 0.431}{0.312}}$$

$$= \sqrt{\frac{0.089}{0.312}} = \sqrt{0.285} = 0.53$$

من قيمة معامل الارتباط المتعدد اعلاه البالغة ٠.٥٣ يمكن القول بانه توجد علاقة موجبة بين المتغير المعتمد (Y) والمتغيرات المستقلة (X1 و X2) ولكنها علاقة ضعيفة .

٣- معامل الارتباط الجزئي Partial correlation Coefficient :

لقد عرفنا سابقا ان الارتباط المتعدد يدرس العلاقة بين المتغير المعتمد " التابع" ومجموعة من المتغيرات المستقلة في ان واحد . اما الارتباط الجزئي فانه يدرس العلاقة بين متغيرين اثنين فقط من تلك المتغيرات بعد عزل او تثبيت باقي المتغيرات الاخرى او استبعاد اثرها احصائيا .

فمثلا اذا كانت لدينا ظاهرة معينة (Y) وهناك مجموعة من المتغيرات المستقلة (X4,X3,X2,X1) تؤثر بشكل متفاوت على هذه الظاهرة فايجاد معامل الارتباط الجزئي بين المتغيرين X3,X1 سيتم مع استبعاد اثر المتغيرين X4,X2 والغرض الاساس من هذا الاستبعاد هو للتعرف على المتغير المستقل الاكثر اهمية في التاثير على المتغير التابع اوللتعرف على متغير قد يتسبب وجوده في نتائج غير دقيقة .

ان حساب قيمة معامل الارتباط الجزئي يتم على وفق الخطوات الاتية :

١- حساب معاملات الارتباط البسيطة بين المتغيرات الثلاث ، وهذا يقتضي تكوين الجدول الاتي :
٢- ايجاد معامل الارتباط الجزئي من الصيغة الاتية :

ا- صيغة معامل الارتباط الجزئي بين Y و X_1 بثبات X_2

ب- صيغة معامل الارتباط الجزئي بين Y و X_2 بثبات X1 .

مثال : قام باحث باختيار عينة من ستة عمال يعملون في احد مصانع الاغذية وقد جمع معلومات عن مقدار الانتاج الشهري لهؤلاء العمال الستة ، ومدة الخدمة الفعلية لهم في المصنع ،وكانت بيانات العينة بحسب الاتي :

مقدار الانتاج (بالطن) لكل عامل Y	1	4	4	5	8	9
الطاقة التصميمية للانتاج (بالطن) X_1	3	5	8	12	12	15
مدة الخدمة الفعلية لكل عامل (بالسنة) X_2	2	4	6	8	9	9

المطلوب : ايجاد معاملات الارتباط الجزئي وبيان ايهما اكثر تاثيرا على الانتاج .

١- بعد استخراجنا قيم معامل الارتباط البسيطة للمتغيرات الثلاث بنفس طريقة المتعدد نحصل على النتائج الاتية $r_{yx1}=0,91$ ، $r_{yx2}=0,92$ ، $r_{x1x2}=0,97$

٢- نبدا بتطبيق قانون الارتباط الجزئي $YX_1.X_2$ بتثبيت X_2 ثم ايجاد $YX_2.X_1$ بتثبيت X_1 :

$$r_{YX_1X_2} = \frac{r_{YX_1} - r_{YX_2}\, r_{X_1X_2}}{\sqrt{1-(r_{X_1X_2})^2}\,\sqrt{1-(r_{YX_2})^2}}$$

$$= \frac{0.91-(0.92)(0.97)}{\sqrt{1-(0.97)^2}\,\sqrt{1-(0.92)^2}} = \frac{0.02}{0.097} = 0.20$$

$$R_{YX_2X_1} = \frac{r_{YX_2} - r_{YX_1}\, r_{X_1X_2}}{\sqrt{1-(r_{X_1X_2})^2}\,\sqrt{1-(r_{YX_1})^2}}$$

$$= \frac{0.92-(0.91)(0.97)}{\sqrt{1-(0.97)^2}\,\sqrt{1-(0.91)^2}} = \frac{0.04}{0.100} = 0.40$$

X_2

من نتائج الارتباط الجزئي اعلاه يمكن القول بان المتغير الثاني " مدة الخدمة الفعلية للعمال " اكثر اهمية في التاثير على كمية الانتاج (Y) من متغير الطاقة التصميمية للمصنع .

تمرين : ارادت احدى المؤسسات للدعاية والاعلان معرفة العلاقة بين عدد المستجيبين لاعلانها Y وبين حجم الاعلان المنشور في الصحيفة X_1 وعدد الصحف الموزعة X_2 التي يتم نشر الاعلان فيها ، واستطاعت المؤسسة الحصول على المعطيات التالية :

جد : (١) الارتباط المتعدد (٢) الارتباط الجزئي

X_2 عدد الصحف الموزعة (بالالف)	X_1 حجم الاعلان (بالانج)	Yi عدد المستجيبين (بالمئات)
٢	١	١
٨	٨	٤
١	٣	١
٧	٥	٣
٤	٦	٢
٦	١٠	٤

الحل :

$$\sum y = 15 \quad , \quad \sum X_1 = 33 \quad , \quad \sum X_2 = 28$$
$$\sum y^2 = 47 \quad , \quad \sum X_1^2 = 235 \quad , \quad \sum X_2^2 = 170$$
$$\sum yX_1 = 103 \quad , \quad \sum yX_2 = 88 \quad , \quad \sum X_1X_2 = 188$$

وباشتقاق الصيغة يكون لدينا :

$$r_{y1} = \frac{n\sum y X_1 - \sum y \sum X_1}{\sqrt{\left[n\sum y^2 - (\sum y)^2\right]\left[n\sum X_1^2 - (\sum X_1)^2\right]}}$$
$$= \frac{6(103) - (15)(33)}{\sqrt{\left[6(47) - (15)^2\right]\left[6(232) - (33)^2\right]}}$$
$$= \frac{123}{131.42} = 0.936$$
$$r_{y2} = \frac{n\sum y X_2 - \sum y \sum X_2}{\sqrt{\left[n\sum y^2 - (\sum y)^2\right]\left[n\sum X_2^2 - (\sum X_2)^2\right]}}$$

$$= \frac{6(88)-(15)(28)}{\sqrt{[6(47)-(15)^2][6(170)-(28)^2]}}$$

$$= \frac{108}{115.983} = 0.931$$

$$r_{12} = \frac{n\sum X_1X_2 - \sum X_1 \sum X_2}{\sqrt{[n\sum X_1^2 - (\sum X_1)^2][n\sum X_2^2 - (\sum X_2)^2]}}$$

$$= \frac{6(188)-(15)(28)}{\sqrt{[6(232)-(33)^2][6(170)-(28)^2]}}$$

$$= \frac{204}{267.41} = 0.763$$

وبتطبيق الصيغة نحصل على :

$$R_{y.12} = \sqrt{\frac{r_{y.1}^2 + r_{y.2}^2 - 2r_{y.1}\ r_{y.2}\ r_{1.2}}{1-r_{1.2}^2}}$$

$$= \sqrt{\frac{0.876+0.866-2(0.936)(0.931)(0.763)}{1-0.582}}$$

$$= \sqrt{\frac{.412}{.418}} = 0.99$$

اما صيغة اختبار فرضية العدم فهي :

H0:Py.12…K=0

في حين ان F المحسوبة فتتمثل بالصيغة التالية :

$$F = \frac{R_{y.12\ldots k}^2}{1-R_{y.12\ldots k}^2}\ \frac{n-k-1}{k}$$

حيث ان K تمثل عدد المتغيرات .

ويتم مقارنة قيمة F المحتسبة مع قيمة F الجدولية الموجودة بجداول F بدرجات حرية عددها (K و n-k-1) وعند مستوى معنوية α ، عندما تكون القيمة المحتسبة اقل من تلك الجدولية ، تقبل H_0 ويستدل على عدم معنوية ارتباط المتغير المعتمد مع المتغيرات المستقلة X_i للمجتمع المسحوبة منه العينة .

ثانيا ـ معامل الارتباط بين المتغيرات الوصفية (غير الكمية) :

توجد بعض الظواهر لا يمكن قياسها كميا " رقميا " وقد تكون على شكل صفات او على شكل رتب ومن هذه الظواهر ، الحالة الاجتماعية ، والحالة الصحية لافراد المجتمع والتدخين والاشباع والذكاء وغيرها ، وهذه الظواهر لا يوجد مقياس رقمي لقياسها وكل ما نستطيع ان نقوم به هو تصنيف افراد المجتمع من حيث الحالة الصحية مثلا الى اصناف متدرجة ، اما من الاعلى الى الادنى او بالعكس ، فمثلا نبتدأ من الحالة الصحية الممتازة وننتهي بالحالة الصحية السيئة ، وهكذا نطبقه على بقية الظواهر المماثلة الاخرى ، ومن اهم تلك المقاييس هي :

١- معامل ارتباط الرتب لسبيرمان
٢- معامل ارتباط الرتب لكيندل

١- معامل ارتباط الرتب لسبيرمان Spearmiain's Rank correlation coefficient

وهو من المقاييس المهمة والشائعة الاستعمال في الحالات الاتية :

١- اذا كان كلا المتغيرين او احدهما من النوع الترتيبي الوصفي ordinal varibles القابل للترتيب التصاعدي اوالتنازلي .
٢- عندما تكون قيم المتغيرات الكمية كبيره وبدلالة الرتب يمكن استخدامه باعتباره معامل بديل لمعامل ارتباط بيرسون.
٣- اذا كان كلا المتغيرين او احدهما لا يتبع التوزيع الطبيعي ، او في حالة البيانات اللامعلمية

. وقد اكتسب معامل سبيرمان في الدراسات الجغرافية مكانة ممتازة نظرا لكون كثير من البيانات الجغرافية هي بيانات غير كمية ، وقد امتد استعماله حتى في البيانات الكمية ، وان قوته في قياس الارتباط لا تقل عن ٠,٩١ من قوة معامل بيرسون . وقد استنبط سبيرمان معادلة لحساب معامل ارتباط الرتب الذي يرمز له بالرمز (r_s) وعلى وفق ما ياتي :

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$

اذ ان :

n تمثل عدد ازواج الرتب

d : تمثل الفرق بين رتب مستويات المتغير الاول (y) ورتب مستويات المتغير الثاني .

ويمكن اختبار معنوية معامل سبيرمان باعتبارها تمثل المجتمع باختبار t كما في الخطوات الاربعة الاتية :

١) نضع الفرضيات :

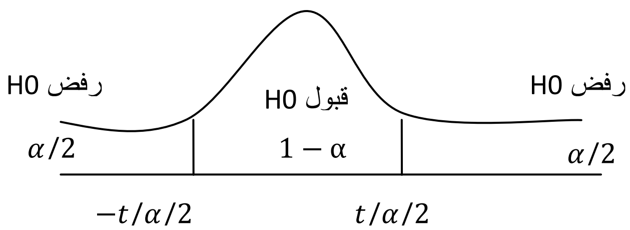
$H0{:}Ps=0$

$H1{:}Ps\neq 0$

٢) $t = rs\sqrt{\frac{n-2}{1-r_s^2}}$ المحسوبة

٣) من جداول (t) وبدرجة حرية (n-2) وبمستوى معنوية α معينة نستخرج t الجدولية

٤) نقارن بين t المحسوبة و t الجدولية لنثبت الفرضية والجدولية



مثال (١) : الجدول الاتي يمثل تقديرات لكفاءة اداء خمسة من العاملين في مصنع ما وتحصيلهم الدراسي :

كفاءة الاداء (y)	جيد جدا	ضعيف	ممتاز	جيد	متوسط	مقبول
التحصيل الدراسي (X)	دبلوم	متوسط	بكالوريوس	يقرا ويكتب	ثانوية	ابتدائية

المطلوب :

احسب قيمة معامل الارتباط بين كفاءة الاداء والتحصيل الدراسي وما هي مدلولاته؟

خطوات الحل :

١- القيام بترتيب التقديرات تصاعديا اوتنازليا لكل متغير ومنحها رتبا متسلسلة كما في الشكل الاتي :

كفاءة الاداء (y)	ضعيف	مقبول	متوسط	جيد	جيد جدا	ممتاز
الرتب	١	٢	٣	٤	٥	٦
التحصيل الدراسي (X)	يقرا ويكتب	ابتدائية	متوسطة	ثانوية	دبلوم	بكالوريوس
الرتب	١	٢	٣	٤	٥	٦

٢- نقوم باعادة توزيع قيم الرتب بحسب الجدول الاصلي ، وتجمع الرتب المتشابهة اذا تواجدت وتقسم على عددها ثم نطرح قيم الرتب للمتغيرات d ثم نربعها d2 .

di^2	$d_i=y_i-x_i$	رتب x	رتب y	X	y
صفر	صفر	٥	٥	دبلوم	جيد جدا
٤	٢-	٣	١	متوسط	ضعيف
صفر	صفر	٦	٦	بكالوريوس	ممتاز
٩	٣	١	٤	يقرا ويكتب	جيد
١٠	١-	٤	٣	ثانوية	متوسط
صفر	صفر	٢	٢	ابتدائية	مقبول
١٤	صفر				Σ

٣- نستخدم الصيغة rs :

$$r_s = 1 - \frac{6\sum di^2}{\mathrm{n}(\mathrm{n}^2-1)} = 1 - \frac{6(14)}{6(36-1)} = 1 - \frac{84}{210} = 1 - 0.4 = 0.6$$

نستنتج بوجود ارتباط موجب متوسط بين كفاءة الاداء والتحصيل الدراسي .

مثال (2) : البيانات الاتية تمثل معدل انتاج الدونم من محصول الحنطة في المناطق شبه المطرية وكمية الامطار المتساقطة للمدة من ٢٠٠٥ – ٢٠٠٩ .

السنة	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨	٢٠٠٩
معدل انتاج الدونم / كغم	١٥٠	٢٠٠	١٢٠	١٨٠	٩٠
كمية الامطار المتساقطة /ملم	٢٥	٣٠	٢٠	٢٨	٢٦

المطلوب : اوجد معامل ارتباط الرتب (سبيرمان) بين معدل انتاج الدونم من محصول الحنطة وكمية الامطار المتساقطة .

الحل : قمنا باعداد الجدول بنفس الخطوات السابقة بعد ترتيب القيم تصاعديا :

معامل الانتاج y	كمية الامطار x	ترتيب y	ترتيب x	di=yi-yi	d_i^2
١٥٠	٢٥	٣	٤	١-	١
٢٠٠	٣٠	١	١	صفر	صفر
١٢٠	٢٠	٤	٥	١-	١
١٨٠	٢٨	٢	٢	صفر	صفر
٩٠	٢٦	٥	٣	٢	٤
				صفر	٦

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

$$= 1 - \frac{6(6)}{5(25-1)}$$

$$= 1 - \frac{36}{120}$$

$$= 1 - 0.3 = 0.7$$

تمرين : مستوى التقدير لمادتي الموارد والجغرافية الزراعية عند طالبة في قسم الجغرافيا كما يلي فهل توجد علاقة بينهما :

ضعيف	جيد جدا	جيد	جيد جدا	ممتاز	جيد جدا	مقبول	x
مقبول	جيد	مقبول	ممتاز	جيد جدا	جيد	ضعيف	y

٢- معامل ارتباط الرتب Kendal Tau كندل تاو :

يستخدم على البيانات الصغيرة التي تقل عن (١٠) وحدات وبشكل رتب أي (البيانات الرتيبة) اوالتي يمكن ان تحول الى رتب وهي مشابهة لسبيرمان في قياس العلاقات درجتها واتجاهها الا انها الافضل استخداما في حالة العينات الكبيرة ، وهي تستخدم لاختبار استقلال متغيرين واختبار وجود اتجاه عام عندما يمثل احد المتغيرين بالزمن وقيمته بين (١ ، -١) والملاحظ بان معامل سبيرمان يميل ان يكون اكبر قيمة من معامل كندل . ومن اجل التعرف على خطوات حسابه ناخذ المثال الاتي :

مثال (١) :
قام مدربان بتقييم فريق رياضي من (٧) اشخاص برتب حسب افضلية كل لاعب بنظر كل منها ، ما هو معامل الارتباط بين رتب تقديرات المدربين ؟

٧	٦	٥	٤	٣	٢	١	رقم اللاعب
٤	٦	٢	١	٥	٣	٧	X المدرب الاول
٣	٧	١	٢	٥	٤	٨	Y المدرب الثاني

الحل :

١- نقوم باعادة ترتيب x تصاعديا ، ثم نعيد ترتيب y على اساس وضع قيم y مقابل قيم الـ x المرتبة تصاعديا فيصبح كما يلي :

Rank X	١	٢	٣	٤	٥	٦	٧
Rank Y	٢	١	٤	٣	٥	٧	٨

نبدأ برتب y من اليسار وهي الرتبة ٢ حيث نضيف (+١) لكل رتبة من y تقع على يمين ٢ وتزيد على 2 كما نضيف (-١) لكل رتبة من y تقع على يمين٢ وتقل عن ٢ وهكذا نستمر ونكرر العملية بنفس الطريقة

لباقي الرتب التي تقع يمين ٢ وهي ٦ رتب في هذا المثال ونكرر العملية لجميع القيم التي تقع ثم نجمع درجات كل رتبة للحصول على الدرجة الكلية و كما يلي :

١) -١ + ١ + ١ + ١ + ١ + ١   -١ + ٥ = ٤
٢) +١ +١ +١ +١ +١   ٥
٣) -١ + ١ + ١ +١   -١ + ٣ = ٢
٤) +١ +١ +١   ٣
٥) + ١ + ١   ٢
٦) +١   ١

المجموع   +١٥ = الدرجة الكلية T

٢- نطبق صيغة معامل كندل

$$w = \frac{2T}{n(n-1)} = \frac{2-15}{7(7-1)} = \frac{30}{42} = 0.71$$

مما يدل على ان العلاقة طردية متوسطة

مثال (٢) : نفرض لدينا الرتب الاتية للمتغيرين y و x وكما يلي :

رتب x	٤	٣	١	٢
رتب y	٤	٢	٣	١

خطوات الحل :

١- نعيد ترتيب رتب x تصاعديا وكما يلي ، ثم نربطه بما يقابل من y وكما يلي :

X= 1 2 3
Y= 3 1 2 4

٢- -١ – ١ + ١ ⇐ (-٢) + ١ = -١
+ ١ + ١ ⇐ ٢
+ ١ ⇐ ١

T = ٢

٣- $w = \frac{2\times T}{n(n-1)} = \frac{2\times 2}{4(4-)} = \frac{4}{12} = 0.33$ العلاقة ضعيفة

تمارين

تمرين ١ : الجدول يمثل تقارير موظفين في احدى الشركات ولسنتين متتاليتين ،كيف نقيس العلاقة بينهما:

السنة الاولى	جيد	جيد جدا	مقبول	جيد جدا	جيد	ممتاز	ممتاز
السنة الثانية	مقبول	ممتاز	ضعيف	ممتاز	جيد جدا	جيد جدا	جيد

تمرين٢ : قام احد مدربي الرياضة بتقييم عينة من (٨) لاعبين في لعبتي الطائرة والسلة وكانت نتائج التقييم كما في الجدول ، كيف نقيس العلاقة بينهما ؟

ت	١	٢	٣	٤	٥	٦	٧	٨
الطائرة	جيد	ضعيف	مقبول	جيد	ممتاز	مقبول	ضعيف جدا	جيد جدا
السلة	ضعيف جدا	ممتاز	ممتاز	جيد	مقبول	جيد جدا	مقبول	ضعيف

تمرين ٣: يمثل الجدول التالي الضغط العالي والسفلي لدم عينة من (٦) افراد ، ما هي قوة واتجاه العلاقة بينهما حسب سبيرمان :

ضغط الدم العالي x	١٢٠	١٣٠	١٢٠	١٥٠	١٠٠	١٢٠
ضغط الدم السفلي y	٨٠	٨٠	٩٠	٧٠	٨٠	٧٠

تمرين ٤: يمثل الجدول الترتيب الذي قدمته ممرضتان عن صعوبات الشفاء الممكنة لثمانية مرضى ، احسب معامل سبيرمان وكندل .

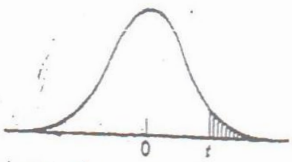
المريض	علي	احمد	سلمان	صادق	عدنان	محمد	انور	احمد
ترتيب المريض حسب الممرضة ١	١	٢	٣	٤	٥	٦	٧	٨
ترتيب المريض حسب الممرضة٢	٢	١	٥	٣	٧	٦	٤	١٠

تمرين ٥ : ما هي علاقة الارتباط بين نتيجة الامتحان ونتيجة المقابلة لستة طلبة مقدمين على الدراسات العليا

المتقدم	١	٢	٣	٤	٥	٦
نتيجة الامتحان التنافسي	الاول	الثاني	الثالث	الرابع	الخامس	السادس
نتيجة المقابلة	الثاني	الاول	الخامس	الرابع	السادس	الثالث

جدول II : قيم t الجدولية

Values of t for Given Probability Levels



α / ν	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576