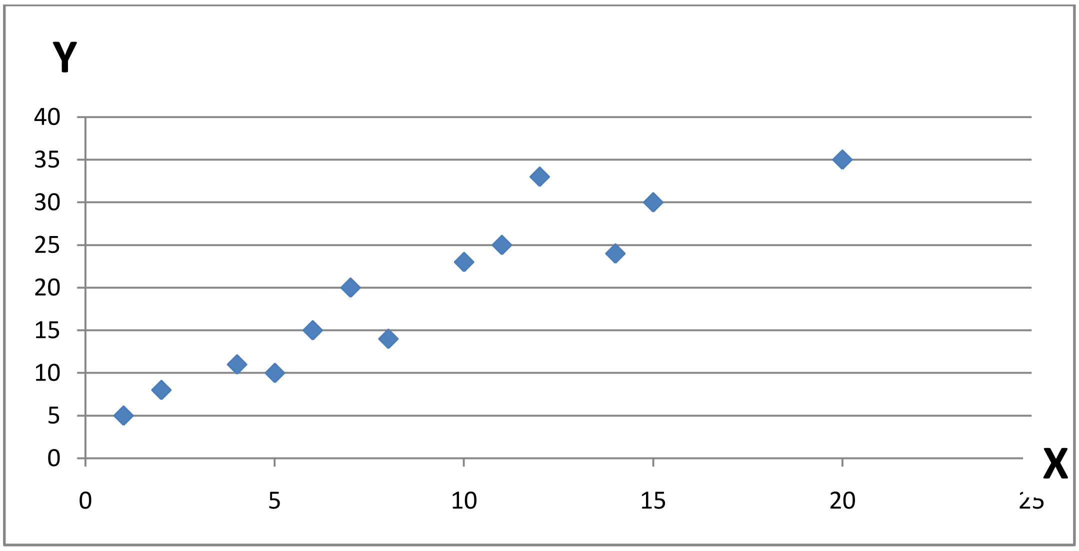
II- نظرية الارتباط البسيط:

تقوم نظرية الارتباط البسيط على دراسة العلاقة بين متغيرين أحدهما مستقل والآخر تابع، فمثلا قد نقول هل هناك علاقة بين الوزن وضغط الدم لدى الرجال ؟ أو هل هناك علاقة بين دخل الفرد الشهري وقيمة الادخار ؟ وإن وجدت علاقة فهل هي خطية أم غير خطية ؟ فوجود علاقة يعني أن أحد المتغيرين مرتبط بالآخر. فموضوع الارتباط هو في كيفية قياس قوة العلاقة بين المتغيرات.

II-1- لوحة الانتشار:

إن أول طريقة تساعدنا على معرفة وجود علاقة بين متغيرين X و Y هي لوحة الانتشار التي تمثل شكل هندسي على معلم متعامد ومتجانس محوره الأفقي يمثل المتغير المستقل (المفسر) ومحوره العمودي المتغير التابع (المفسر)، وتبين لوحة الانتشار بشكل جيد ما إذا كان هناك علاقة بين المتغيرين X و Y أو عدم وجودها

الشكل: لوحة الانتشار



II-2- معامل الارتباط الخطي البسيط:

II-2-1- معامل الارتباط البسيط بيرسون:

يستخدم لقياس العلاقة بين متغيرين ذات قيم رقمية (quantitative)، ويعتبر معامل ارتباط بيرسون من أهم الطرق المستخدمة في حالة الارتباط البسيط، ويحسب بالعلاقة التالية.

$$r_{x,y} = \frac{\mathrm{Cov}(x,y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

حيث:

$\mathrm{Cov}(x,y)$: التباين المشترك بين x وy.

$\sigma_x \sigma_y$: الانحراف المعياري لـ x والانحراف المعياري لـ y.

n: عدد المشاهدات.

بتبسيط العلاقة أعلاه نحصل على العلاقة التالية:

$$r_{x,y} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n}(x_i^2)(\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n}(y_i^2)(\sum_{i=1}^{n} y_i)^2}}$$

نستعمل x بدل x_i و y بدل y_i و $\sum x$ بدل $\sum_{i=1}^{n} x$ نحصل على العلاقة التالية:

$$r_{x,y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2}\sqrt{\sum y^2 - n\bar{y}^2}}$$

II-2-2- تفسير معامل الارتباط البسيط:

يمكن إثبات رياضيا أن معامل الارتباط يتراوح بين -1 و1

- ✓ عندما يقترب $r_{x,y}$ من 1 المتغيرات ترتبط ارتباطا إيجابيا.
- ✓ عندما يقترب $r_{x,y}$ من -1 المتغيرات ترتبط ارتباطا سلبيا أو عكسيا.
- ✓ عندما يقترب $r_{x,y}$ من 0 ليس هنا ارتباط.

II-2-3- دلالة معامل الارتباط الخطي البسيط(person):

يمكن اختبار ما إذاكان معامل الارتباط له دلالة على وجود علاقة خطية بين المتغيرين باستعمال الإحصائية t (student) للفرضية $H_0: P = 0$ مقابل الفرضية $H_0: P \neq 0$ بدرجة حرية n-2.قيمة الإحصائية المقاسية هي:

$$t_{cal} = \frac{r_{x,y}}{\sqrt{(1 - r^2)/(n - 2)}}$$

ملاحظة: وجود ارتباط خطي كامل بين متغيران لا يعني بالضرورة السببية لأنه يمكن أن نرجع هذا الارتباط إلى متغير ثالث يؤثر عليهما فيرتفعان أو ينخفضان معا.

مثال: الجدول التالي يمثل القياسات الخاصة بمتغيرين

الأداء	16	18	23	24	28	29	26	31	32	34
كمية الطاقة	20	24	28	22	32	28	32	26	41	41

المطلوب:

1- رسم لوحة الانتشار؟

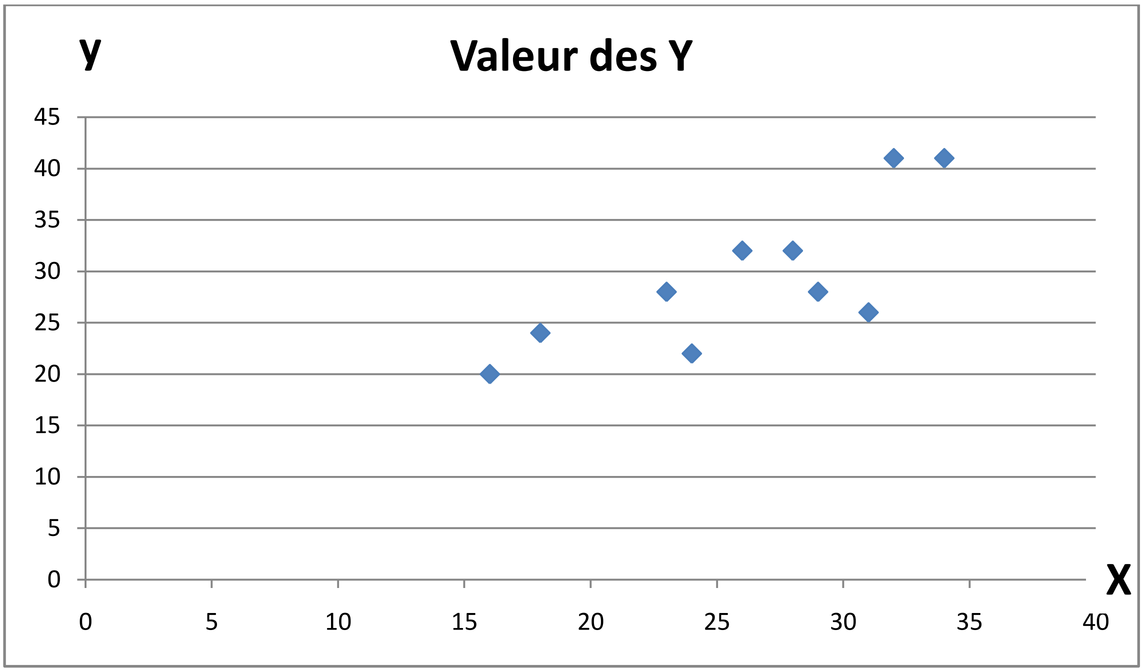
2- تحديد ما نوع العلاقة بين المتغيرين الأداء و كمية الطاقة ؟

3- حساب معامل الارتباط البسيط (person) ؟

4- اختبر دلالة معامل الارتباط عند مستوى الدلالة 5%؟

الحل:

1 - شكل الانتشار:



الشكل أدناه يبين العلاقة بين كمية الطاقة والأداء ويتضح من الشكل أنها علاقة خطية موجبة.

2 – حساب معامل الارتباط

	x	y			xy
	16	20	256	400	320
	18	24	324	576	432
	23	28	529	784	644
	24	22	576	484	528
	28	32	784	1024	896
	29	28	841	784	812
	26	32	676	1024	832
	31	36	961	1296	1116
	32	41	1024	1681	1312
	34	41	1156	1681	1394
المجموع	**261**	**304**	**7127**	**9734**	**8286**

$$r_{x,y} = \frac{8286 - (10 \times 26.1 \times 30.4)}{\sqrt{7127 - (10 \times 26.1^2)}\sqrt{9734 - (10 \times 30.4^2)}} = 0.89$$

$$\boldsymbol{r_{x,y} = 0.89}$$

اختبار الدلالة لمعامل الارتباط البسيط:

القيمة الجدولية:

$$t_{tab}\,(n-2,\alpha/2)=1.85$$

القيمة الحسابية:

$$t_{cal}\ \frac{0.89}{\sqrt{(1-0.89^2)/(10-2)}}=5.52$$

القيمة الجدولية أكبر من القيمة الحسابية إذن نرفض الفرضية العدمية ونقبل الفرضية البديلة وبالتالي معامل الارتباط يدل على وجود علاقة خطية بين المتغيرين.

II-3- معامل الارتباط المتعدد:

II-3-1- تعريف معامل الخطي المتعدد

ويستخدم لقياس العلاقة بين أكثر من متغيرين، إلا أن إشارة معامل الارتباط هنا لا تدل على اتجاه العلاقة لان هذا الاتجاه لا يكون موحدا لجميع المتغيرات، وأن عملية التحليل تقوم على فرض أن المتغيرات عشوائية متصلة ويدعى توزيعها بمتعدد المتغيرات، وصيغة حسابه هي امتداد لمعامل الارتباط البسيط، ففي حال 3 متغيرات مثلا لإيجاد العلاقة بين x_2 و x_1 x_3 و فإن صيغة الحساب هي[9]:

$$r_{2.13}=\sqrt{\frac{r_{21}^2+r_{23}^2-(2)r_{21}r_{23}r_{13}}{1-r_{13}^2}}$$

ولإيجاد العلاقة بين x_1 و x_2 x_3 و فإن صيغة الحساب هي:

$$r_{1.23}=\sqrt{\frac{r_{12}^2+r_{13}^2-(2)r_{12}r_{13}r_{23}}{1-r_{23}^2}}$$

[9] عبد الحميد عبد المجيد البلداوي، **أساليب الإحصاء للعلوم الاقتصادية وإدارة الأعمال مع استخدام برنامج SPSS**، دار وائل للنشر والتوزيع، الطبعة الأولى، الأردن، 2009، ص.201.

حيث أن r_{12}, r_{13}, r_{23} هي معاملات ارتباط يتم حسابها صيغة الارتباط البسيط لبيرسون المبينة أعلاه

مثال: لدينا المعطيات التالية:

$$\sum y = 15, \quad \sum x_1 = 33, \quad \sum x_2 = 28,$$

$$\sum y^2 = 47, \quad \sum x_1^2 = 235, \quad \sum x_2^2 = 170,$$

$$\sum yx_1 = 103, \quad \sum yx_2 = 88, \quad \sum x_1 x_2 = 188,$$

$$\boldsymbol{r_{12}} = 0.763 \quad \boldsymbol{r_{y1}} = 0.936 \quad \boldsymbol{r_{y2}} = 0.931$$

المطلوب: حساب معامل الارتباط المتعدد ؟

الحل:

$$r_{y.12} = \sqrt{\frac{r_{y1}^2 + r_{y2}^2 - (2)r_{y1}r_{y2}r_{12}}{1 - r_{12}^2}}$$

$$r_{y.12} = \sqrt{\frac{(0.876) + (0.866) - (2)(0.936)(0.931)(0.763)}{1 - (0.613)}}$$

$$r_{y.12} = \sqrt{\frac{(0.876) + (0.866) - (2)(0.936)(0.931)(0.763)}{1 - (0.582)}}$$

$$r_{y.12} = \sqrt{\frac{0.412}{0.418}}$$

$$r_{y.12} = 0.99$$

II-3-2- دلالة معامل الارتباط المتعدد:

يمكن اختبار ما إذاكان معامل الارتباط المتعدد له دلالة على وجود علاقة خطية بين المتغيرات باستعمال الإحصائية F (ficher) للفرضية $H_0: P_{y,12} = 0$ مقابل الفرضية $H_0: P_{y,12} \neq 0$ بدرجة حرية n-k-1 للبسط و k للمقام .قيمة الإحصائية المقاسية هي:

$$f_{\text{cal}} = \frac{r_{y,12}^2}{1 - r_{y,12}^2} \cdot \frac{n-k-1}{k}$$

مثال: في المثال السابق اختبر دلالة معامل الارتباط المتعدد

القيمة الجدولية:

$$f_{tab}\,(n-k-1, k\,, \alpha/2) = 16.04$$

القيمة الحسابية:

$$f_{\text{cal}} = \frac{0.98}{1 - 0.98} \cdot \frac{6-3-1}{3}$$

$$f_{\text{cal}} = 32.667$$

قيمة f_{cal} المحسوبة أكبر من قيمة f_{tab} الجدولية ومنه نرفض الفرضية العدمية ونقبل الفرضية البديلة وبالتالي معامل الارتباط يدل على وجود علاقة خطية بين المتغيرات.

II-4- معامل الارتباط للرتب:

يهتم الباحث أحيانا بإيجاد معامل الارتباط بين رتب البيانات والمشاهدات وليس بين المشاهدات نفسها. وتبرز أهمية هذا المعامل عندما مع متغيرات في مستوى القياس الرتبي. إن أشهر العلاقات التي وضعت لهذا الغرض هي العلاقة التي وضعها الإحصائي البريطاني تشارلز سبيرمان عام 1904م بإسم معامل ارتباط سبيرمان للرتب والتي تعطى بالعلاقة التالية[10]:

---

[10] سالم عيسى بدر، عماد غصاب عبابنة، **مباديء الإحصاء الوصفي والاستدلالي**، دار المسيرة للنشر والتوزيع والطباعة، الطبعة الثانية، 2010، ص.158.

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

حيث:

n: هي عدد المشاهدات

d: هي الفرق بين رتبة x و y

ملاحظة: يكون هذا المعامل كفؤ إذا كان عدد المشاهدات أقل من 30، وأكثر كفاءة عندما تكون عدد المشاهدات بين 25 و30.

مثال:

الجدول التالي يبين العلاقة بين معدلات الطلبة في مقياسي المحاسبة (x) والرياضيات (y) لعشرة طلاب:

x	12	9	15	13	5	6	18	14	3	11
y	14	12	16	11	8	10	17	13	9	5

المطلوب:

- أحسب معامل الارتباط للرتب سبيرمان ؟

الحل:

نقوم بترتيب المعدلات من الأكبر إلى الأصغر

x	y	رتبة x	رتبة y	الفرق بين الرتب d	مربع الفرق d^2
12	14	**5**	**3**	**2**	**4**
9	12	**7**	**5**	**2**	**4**
15	16	**2**	**2**	**0**	**0**
13	11	**4**	**6**	**−2**	**4**
5	8	**9**	**9**	**0**	**0**
6	10	**8**	**7**	**1**	**1**
18	17	**1**	**1**	**0**	**0**
14	13	**3**	**4**	**−1**	**1**
3	9	**10**	**8**	**2**	**4**
11	5	**6**	**10**	**4**	**16**
المجموع					**34**

$$r_s = 1 - \frac{6(34)}{10(100-1)}$$

$$\boldsymbol{r_s = 0.80}$$

عندما تكون لدينا بيانات متساوية تأخذ نفس الرتبة فإننا نأخذ الوسط الحسابي للرتب

II-5- الارتباط النوعي:

II-5-1- معامل الاقتران:

هو وصف درجة العلاقة بين ظواهر غير كمية لا يمكن التعبير عنها بالأرقام بل بالوصف فقط مثل الارتباط بين الجنسين (الذكر والأنثى)، أو بين الحالة التعليمية لمجموعة من الأفراد، أو دراسة التطعيم بمصل واقي والإصابة بمرض معين وغير ذلك من الأمثلة ولدراسة هذا النوع من العلاقات نستعمل معامل الاقتران[11]. ويعرض وفق الجدول التالي لمتغيرين X و Y:

[11] وليد إسماعيل السيفو، عيد أحمد أبو بكر، غالب عوض الرفاعي، **أساسيات الأساليب الإحصائية للأعمال وتطبيقاتها في العلوم المالية والإدارية والاقتصادية**، زمزم ناشرون وموزعون، الطبعة الأولى، 2010، ص.2016.

	X	
Y	A	B
	C	D

$$T_n = \frac{AD - BC}{AD + BC}$$

- تتراوح قيمة هذا المعامل بين -1 و1 أي $-1 \leq T_n \leq 1$

مثال:

في دراسة أجرية لمعرفة العلاقة بين الدروس الإضافية ونجاح الطلبة أجريت دراسة على عينة من 400 طالب من الطلبة المقبلين على امتحان البكالوريا فكانت النتائج كما في الجدول التالي:

	قاموا بدروس إضافية	لم يقوموا بدروس إضافية	**المجموع**
نجحوا في امتحان البكالوريا	220	80	300
لم ينجحوا في امتحان البكالوريا	60	40	100
المجموع	280	120	**400**

المطلوب:

- حساب معامل الاقتران ؟

الحل:

$$T_n = \frac{AD - BC}{AD + BC}$$

$$T_n = \frac{(220)(40) - (80)(60)}{(220)(40) + (80)(60)}$$

$$T_n = 0.29$$

معنى هذا أن هناك علاقة طردية ضعيفة بين الدروس الإضافية ومستوى النجاح في البكالوريا.

II-5-2- معامل ارتباط فاي:

له نفس خصائص معامل الاقتران حيث يتعامل مع متغيرات منفصلة ثنائية التصنيف (dichotomous) تم تطويره من قبل يدني يول (Udny yule) في عام 1912، وهو معطى بالعلاقة التالية:

$$r_f = \frac{AD - BC}{\sqrt{(A+B)(A+C)(A+D)(C+D)}}$$

مثال: في المثال السابق أحسب معامل الارتباط لفاي

الحل:

$$r_f = \frac{(220)(40) - (80)(60)}{\sqrt{(220+80)(220+60)(220+40)(60+40)}}$$

$$r_f = 0.08$$

معنى هذا أن هناك علاقة ضعيفة جدا بين الدروس الإضافية ومستوى النجاح في البكالوريا.

II-5-3- معامل التوافق لبيرسون:

يدرس هذا المعامل العلاقة بين الظواهر الوصفية التي تنقسم إلى أكثر من نوعين. حيث لا يساعد معامل الاقتران السالف الذكر بل يستخدم معامل التوافق الذي يصلح أيضا لقياس العلاقة بين ظواهر كمية قابلة للقياس وأخرى وصفية لا يمكن قياسها[12]. ويحسب بموجب العلاقة التالية:

$$T_p = \sqrt{\frac{G-1}{G}}$$

[12] وليد إسماعيل السيفو، عيد أحمد أبو بكر، غالب عوض الرفاعي، نفس المرجع السابق، ص 218.

G: مجموع خارج قسمة مربع كل تكرار في الجدول التكراري على حاصل ضرب مجموع بيانات الصف في مجموع بيانات العمود التي يقع فيها التكرار. حيث نقوم بتربيع كل تكرار وارد في الجدول التكراري. ثم نقوم بقسمته على حاصل ضرب التكرار الكلي الرأسي في التكرار الكلي الأفقي.

مثال:

أحسب درجة إرتباط بين تقدير الطالب في المادة (X) والمادة (Y) من البيانات التالية:

X Y	مقبول	جيد	جيد جدا	ممتاز	المجموع
مقبول	10	18	23	14	65
جيد	45	103	52	50	280
جيد جدا	27	50	30	21	128
ممتاز	82	171	135	85	473

الحل:

$$G = \frac{(10)^2}{82 \times 65} + \frac{(45)^2}{82 \times 280} + \frac{(27)^2}{28 \times 128} + \frac{(18)^2}{171 \times 65}$$

$$+ \frac{(103)^2}{171 \times 280} + \frac{(50)^2}{171 \times 128} + \frac{(23)^2}{135 \times 65} + \frac{(52)^2}{135 \times 280}$$

$$+ \frac{(30)^2}{135 \times 128} + \frac{(14)^2}{85 \times 65} + \frac{(50)^2}{85 \times 280} + \frac{(21)^2}{85 \times 128}$$

$$G = 1.0066$$

$$T_p = \sqrt{\frac{1.0066 - 1}{1.0066}}$$

$$\boldsymbol{T_p = 0.08}$$