المحور 03: الارتباط والانحدار

- مفهوم الارتباط وأنواعه
- حساب معامل الارتباط لبوسن.
- حساب معامل الارتباط الرتبي لسبرمان
- مفهوم الانحدار.
- تقدير نموذج انحدار الخطي البسيط.

مفهوم الارتباط: هو تحسن طبيعة وقوة العلاقة بين متغيرين أو عدمها وهذا بالاستعانة بمعامل الارتباط وشكل الانتشار.

- أول خطوة في تحديد العلاقة هو رسم الانتشار (سحابة النقط) ثم الاستعانة بأحد معاملات الارتباط
- أنواعها: ارتباط موجب (طردي)، الارتباط السالب (عكسي)

* ارتباط موجب (طردي): العلاقة بين متغيرين في نفس الاتجاه. اذا زاد أحدهما زاد في الآخر

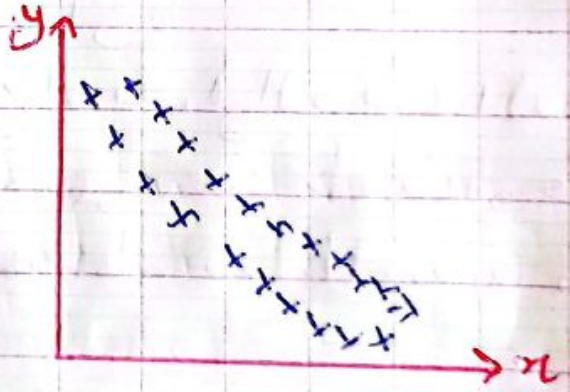
~~[illegible]~~

زاد فإن الآخر والعكس صحيح.

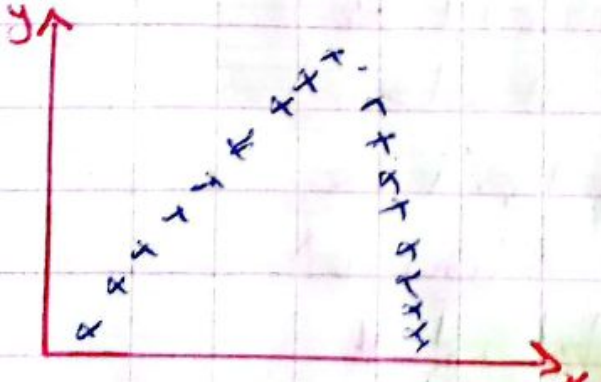
ارتباط سالب: إذا زاد x نقص y وإذا نقص x زاد y.

شكل الانتشار:

توجد علاقة طردية ذات بين المتغيرين الموجبين.



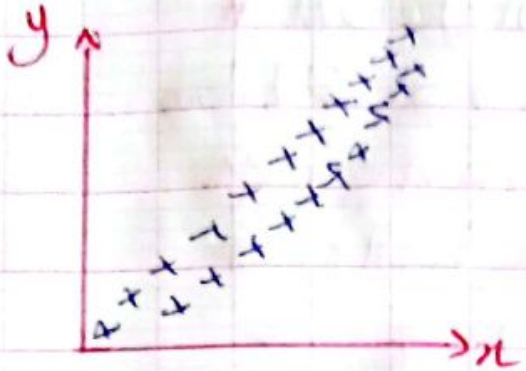
توجد علاقة عكسية ذات بين المتغيرين.



توجد علاقة ذات بين الظاهرتين

لا توجد علاقة

- يمكن تحديد قوة العلاقة من خلالها من خلال ملاحظة شكل الانتشار.

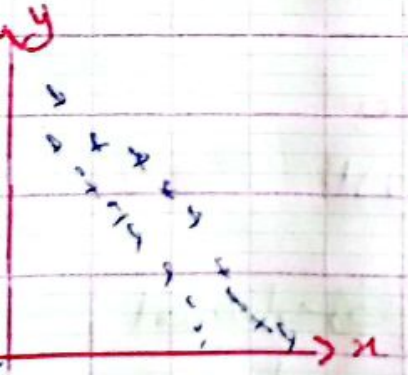


ارتباط قوي

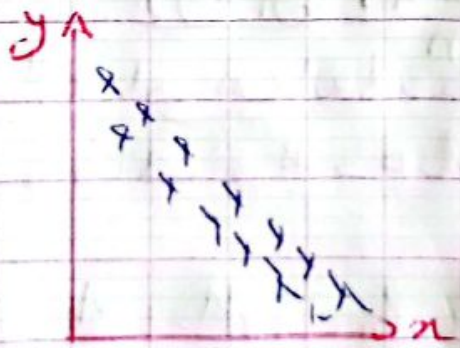
ارتباط قوي تام

ارتباط قوي

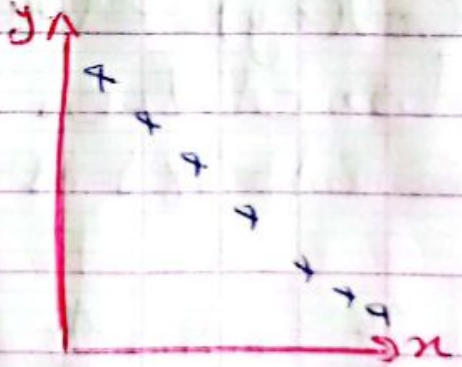
* في 03 منحنيات هناك ارتباط خطي سالب موجب.



ارتباط عكسي



ارتباط عكسي قوي



ارتباط عكسي تام

في 02 منحنيات هناك ارتباط خطي سالب.

Scanné avec CamScanner

قياس الارتباط: نستخدم عدة معاملات لقياس درجة الارتباط بين متغيرين

* المعامل الذي يقيس الارتباط بالرسم هو من أكثر معاملات الارتباط استخداما نتراوح قيمتها بين

$-1 \leq r \leq 1$

تدل الإشارة (+) على العلاقة الطردية
و (-) على العكسية.

نستعمل في البيانات الكمية فقط

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

كما يمكن كتابته:

$$rp = \frac{Cov(xi,yi)}{\sigma x \, \sigma y}$$

$Cov(x,y)$ التباين المشترك بين x و y.

$$Cov(x,y) = \frac{\sum (xi - \bar{x})(yi - \bar{y})}{n}$$

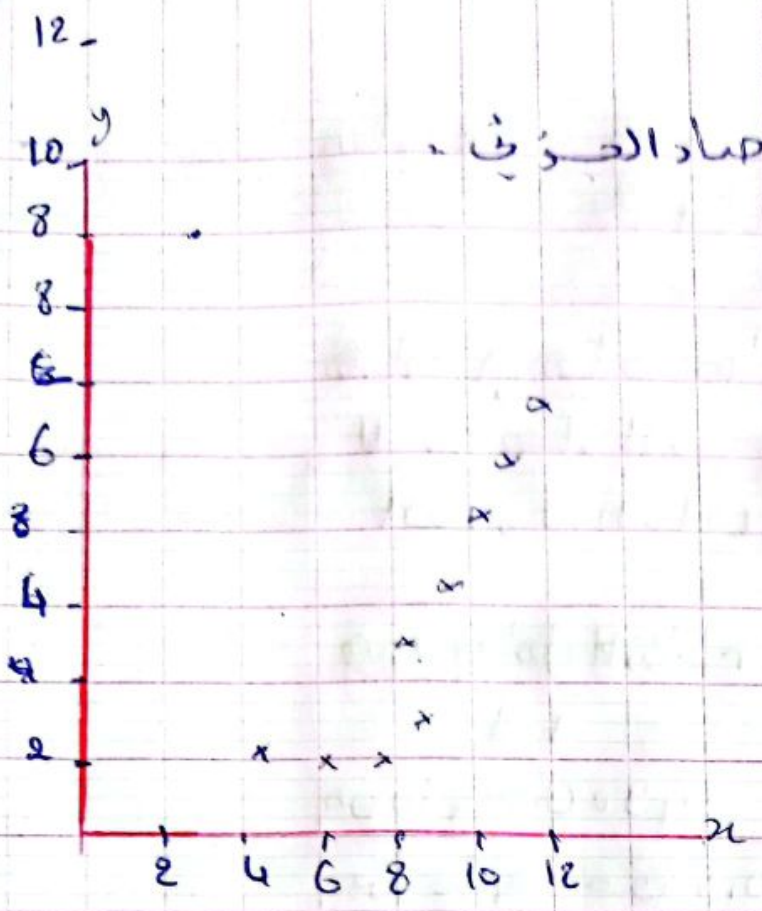
مثال: أدرس العلاقة بين علامات مقياس اقتصاد الجزء في 10 و اقتصاد 10 لـ 8 طلبة سنة أولى.

اقتصاد جزئي	10	11	8	16	7	13	6	9
اقتصاد كلي	8	7	2	12	1	9	2	4

1. سحابة النقط

على سحابة النقط هناك علاقة موجبة بين النقط

الإحصاء و نقاط الإقتصاد الجزئي .



x_i^2	y_i^2	$x_i y_i$	y_i	x_i
100	9	30	3	10
121	49	77	7	11
64	4	16	2	8
256	144	192	12	16
49	1	7	1	7
169	81	117	9	13
36	4	12	2	6
81	16	36	4	9
876	308	487	40	80

$n = 8$.

Scanné avec CamScanner

$$r = \frac{8(487) - 80 \times 40}{\sqrt{[8(816) - (80)^2][8(308) - (40)^2]}}$$

$r = 0,96$

هناك ارتباط طردي قوي بين علامات الحساب والهندسة والجزئي.

الجدول التالي يوضح درجة الارتباط حسب قيمة r

قيمة معامل الارتباط r	المعنى
+1	ارتباط طردي تام
من 0,70 حتى 0,90	" " متوسط
من 0,50 حتى 0,70	" " ضعيف
0	لا يوجد ارتباط

وما قل عن الضعيف الطردي يقال عن الارتباط العكسي مع وضع اشارة (-)

معامل سبيرمان للارتباط الرتبي:

نستخدم معامل سبيرمان إذا أردنا قياس درجة الارتباط بين متغيرين أحدهما أو كلاهما كيفي (رتبي)، أو في حالة ما إذا كان قيم أحد المتغيرات كبيرة جدا

طريقة حساب معامل سبيرمان:

إذا افترضنا أن المتغير x له رتبة Rx والمتغير y له الرتبة Ry نفرق "d" نرمز لفرق الرتبتين

أي: $d = Ry - Rx$.

Scanné avec CamScanner

يُحسب معامل سبيرمان بالعلاقة التالية:

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

حيث n ... الأزواج المدروسة.

مثال:

- لدراسة علاقة الارتباط بين الكتابة والإملاء لـ 5 أخذنا 5 تلاميذ من قسم وسجلنا تغيراتهم في الكتابة والإملاء.

الكتابة	الإملاء	رتبة x	رتبة y	d	d^2
ضعيف	ضعيف	1	2	-1	1
ممتاز	متوسط	3	3	0	0
متوسط	جيد	3	4	-1	1
ضعيف	ضعيف جدا	2	1	1	1
جيد	ممتاز	4	5	-1	1
✓	/	/	/	/	04

$$\rho_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \qquad n = 5$$

$$= 1 - \frac{6\sum (8)^2}{5(25-1)} = 0,6$$

$\rho_s = 0,6$

~~الحكم: علاقة طردية قوية جدا~~

مثال 01: في دراسة معرفة العلاقة بين عدد الحقول المكتشفة وطول الأنابيب بالكيلومتر الناقلة للنفط الخام، خلال أن سنوات سجلنا البيانات التالية.

عدد الحقول x_i	طول الأنابيب y_i	الرتبة R_x	الرتبة R_y	d	d^2
55	21960	1	2	1	1
54	22027	2	1	-1	1
56	23006	3	3	0	0
61	23008	4	4	0	0
62	23020	5	5	0	0
63	23125	7	6	-1	1
67	23120	6	7	1	1
					04

$$\rho_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \qquad n = 7$$

$$1 - \frac{6(4)^2}{7(49-1)} =$$

$$\rho_s = 0,92$$

نلاحظ وجود علاقة طردية قوية جدا بين عدد الحقول المكتشفة وطول الأنابيب الناقلة للنفط

Scanné avec CamScanner

الإنحدار:

الإنحدار هو أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلات الإنحدار مثلا، هناك بعض التطبيقات في الهاتف، يمكننا معرفة الوزن المثالي بمجرد ادخال الطول، هنا متغير الطول (x) هو المتغير المستقل (المؤنس) ومتغير الوزن (y) هو المتغير التابع

هناك عدة أنواع:

أ- الإنحدار الخطي البسيط: تعني كلمة بسيط أن هناك المتغير التابع y يعتمد على متغير مستقل واحد x وكلمة خطي تعني أن العلاقة بين متغيرين خطية.

ب- الإنحدار الخطي المتعدد: المتغير التابع في هذه الحالة يعتمد على عدة متغيرات مستقلة ($x_1, x_2, x_3 \dots$)

ج- الإنحدار الغير خطي: هنا تكون العلاقة بين x و y غير خطية (منحنى).

يمكن كتابة معادلة الإنحدار الخطي البسيط بالشكل التالي:

$$y_i = a x_i + b + e_i$$

y_i: المتغير التابع x_i: المتغير المستقل.

b: معلمات التقدير e_i: الخطأ العشوائي.

Scanné avec CamScanner

الانحدار الخطي البسيط:

معادلة خط الانحدار:

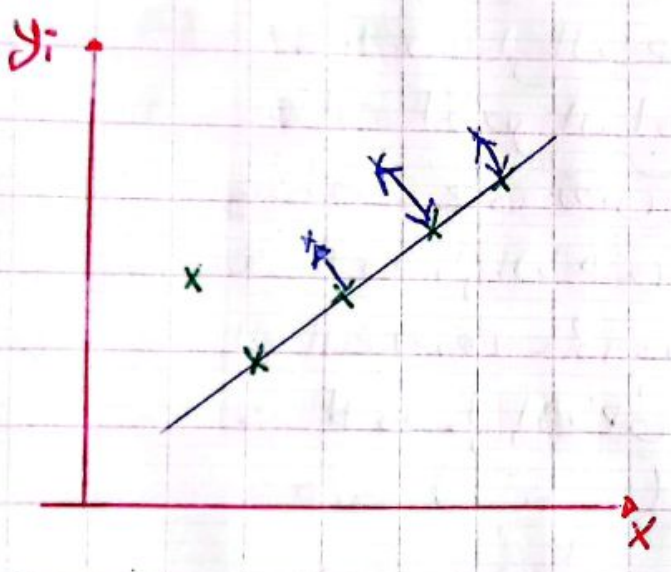
$$y_i = a x_i + b + e_i$$

a و b معاملات سيتم تقديرها لاحقاً

e_i: الخطأ العشوائي.

$$e_i = y_i - \hat{y}_i$$

e_i = الفرق بين القيمة الحقيقية لـ y_i والقيمة المقدرة لـ $\hat{y}_i$.



(ei) مثلا: البعد بين النقطة الحقيقية والساقط عليها على المستقيم.

تحدد العلاقة بين المتغير التابع والمتغير المستقل بعدة طرق أشهرها طريقة المربعات الصغرى حيث تعتمد على مبدأ تقدير المعلمات a, b بشرط

أ ن يكون الخطأ العشوائي عند حده الأدنى (مجموع مربعات الفروقات بين المشاهدات والقيم المقدرة أصغر ما يمكن)

Min $\hat{\varepsilon}$

يمكن تقدير المعلمين (a, b) باستخدام طريقة المربعات الصغرى، لا يمكن للطالب أن يلجأ إلى مراجع أخرى عن تفضيل الحساب إضافة إلى شروط توظيف النموذج). بحل جملة المعادلات

$$\begin{cases} \sum y_i = a \sum x_i + n b \\ \sum y_i x_i = a \sum x_i^2 + b \sum x \end{cases}$$

نحصل على المقدرات التالية:

$$\hat{a} = \frac{\sum x_i y_i - n \overline{X} \overline{y}}{\sum x_i^2 - n \overline{X}^2}$$

$$\hat{b} = \overline{y} - \hat{a} \overline{x}$$

مثال: يبين الجدول التالي انتاج محصول الذرة لـ من المساحة المزروعة X

المطلوب: رسم شكل الانتشار

أوجد معادلة انحدار لـ على X

- ما هو انتاج المتوقع عندما تكون المساحة المزروعة هي 300 هكتار.

Scanné avec CamScanner

الملاحظة	X المساحة المزروعة H	Y: انتاج الذرة	$X_i Y_i$	X_i^2
1	50	140	7000	2500
2	200	500	100000	40000
3	110	400	44000	12100
4	80	300	24000	6400
5	120	356	42720	14400
6	745	940,5	700972,5	555025
7	889	200,6	178333,4	790321
8	5,7	33,5	190,95	32,49
9	11	69,8	767,8	121
10	3,2	18,7	59,84	10,24
المجموع	743,3	2259,1	254418,18	889017,19

1. تشكل الإنتشار:

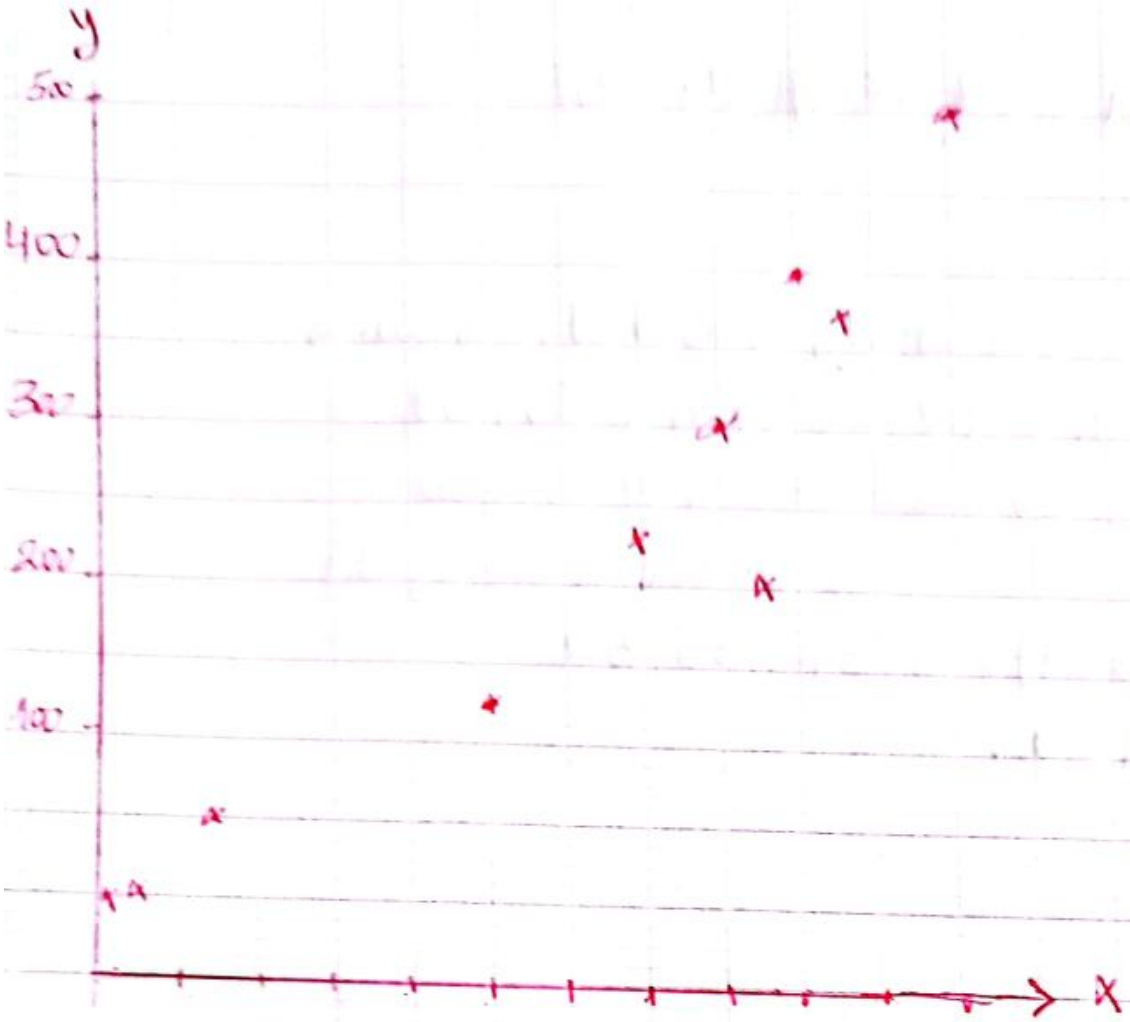
نلاحظ أن العلاقة بين المساحة المزروعة وكمية الإنتاج علاقة خطية طردية موجبة

• وجدنا العلاقة بين X و Y خطية

$$y = ax + b + e_i$$

ولتقدير المعلمين (a, b)

Scanné avec CamScanner



$$\bar{x} = \frac{\sum x_i}{N_i} = \frac{743,3}{10} = 74,33$$

$$\bar{y} = \frac{2259,1}{10} = 225,91.$$

$$\hat{a} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{254489,18 - (10)(74,33)(225,91}{(89017,19) - 10(74,33)^2}$$

$$\hat{a} = 2,564.$$

$$\hat{b} = \bar{y} - a\bar{x}$$

Scanné avec CamScanner

$$\hat{b} = 225,91 - (2,564)(14,33)$$
$$\hat{b} = 35,32.$$

$\hat{a} = 2,564.$ ومنه معادلة الانحدار تأخذ الشكل:

$$y_i = (2,564)x_i + 35,32$$

مستوى الإنتاج المتوقع عندما تكون المساحة المزروعة = 300 هكتار.

$$y_{i+1} = (2,564) \times 300 + 35,32$$
$$y_{i+1} = 804,52.$$

Scanné avec CamScanner