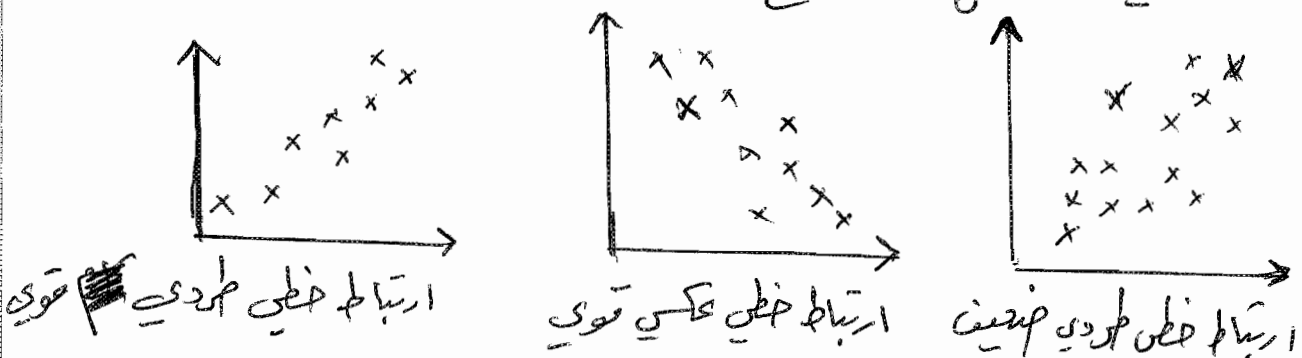
البند 7 : العلاقة بين المتغيرات

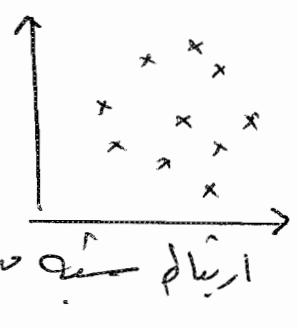
لدراسة العلاقة بين المتغيرين X و Y

$X : x_1, x_2, \dots, x_n$

$Y : y_1, y_2, \dots, y_n$

نرسم شكل الانتشار للعلاقة بين X على المحور الأفقي و Y على المحور العمودي ونمثل الأزواج $(x_1, y_1), \dots, (x_n, y_n)$





(1) معامل الارتباط الخطي البسيط (معامل بيرسون)

$$S_{XY} = \sum XY - n\bar{X}\bar{Y}$$

$$S_{XX} = \sum X^2 - n\bar{X}^2$$

$$S_{YY} = \sum Y^2 - n\bar{Y}^2$$

ويكون معامل الارتباط بيرسون $r = \dfrac{S_{XY}}{\sqrt{S_{XX} S_{YY}}}$

- قيمة r تقع بين -1 و 1
- $r = 1 \Leftarrow$ ارتباط تام طردي
- $r = -1 \Leftarrow$ = = عكسي
- $r = 0 \Leftarrow$ لا توجد علاقة

مثال : X : درجة اللغة الانجليزية

Y : درجة علم النفس

الدرجات مدرجة في الجدول التالي

X	Y	X^2	Y^2	XY
8	9	64	81	72
7	7	49	49	49
6	8	36	64	48
9	10	81	100	90
6	8	36	64	48
36	42	266	358	306

$$\overline{X} = \frac{36}{5} = \boxed{7.2} \qquad \overline{Y} = \frac{42}{5} = \boxed{8.4}$$

$$S_{XX} = 266 - 5(7.2)^2 = 266 - 5(51.84) = \boxed{6.8}$$

$$S_{YY} = 358 - 5(8.4)^2 = 358 - 5(70.56) = \boxed{5.2}$$

$$S_{XY} = 306 - 5(7.2)(8.4) = \boxed{3.6}$$

$$\Rightarrow r = \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}} = \frac{3.6}{\sqrt{6.8}\sqrt{5.2}} = \frac{3.6}{5.95} = \boxed{0.61}$$

دلالة طردية متوسطة بين X و Y

مثال : لمعرفة بين المساحة المقطوعة في الحالة X وكمية الوقود X أخذت البيانات التالية.

n	ΣX	ΣY	ΣX^2	ΣY^2	ΣXY
30	2444	960	204232	31514	79800

المطلوب حساب معامل ارتباط بيرسون بين X و Y .

(47)

$$\bar{X} = \frac{2444}{30} = \boxed{81.47} \qquad \bar{Y} = \frac{960}{30} = \boxed{32}$$

$$S_{xx} = 204232 - 30(6637.36) = \boxed{5111.17}$$

$$S_{yy} = 31514 - 30(1024) = \boxed{794}$$

$$S_{xy} = \cancel{79800}\ \cancel{6664} - 30(81.47)(32) = \boxed{1588.8}$$

$$\Rightarrow r = \frac{1588.8}{\sqrt{(5111.7)(794)}} = \frac{1588.8}{2014.62} = \boxed{0.79}$$ طردي قوي

(2) معامل ارتباط الرتب

يمكن استخدام مقياس لمعامل الارتباط بديل للمقياس السابق وهو مقياس يعتمد فقط على الرتب وحسب حجم العينة

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

d: الفرق في الرتب

مثال السابق

X	Y	رتب X	رتب Y	d	d^2
8 (4)	9 (4)	4	4	0	0
7 (3)	7 (1)	3	1	2	4
6 (1.5)	8 (2.5)	1.5	2.5	-1	1
9 (5)	10 (5)	5	5	0	0
6 (1.5)	8 (2.5)	1.5	2.5	-1	1
					6

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{36}{5(24)} = 1 - \frac{36}{120} = 1 - 0.3 = \boxed{0.7}$$

ارتباط طردي قوي

(3) اختبار إستقلال الصفات

لإختبار استقلالية متغيرين نستخدم الفرضيات التالية

H_0: المتغيرين او العاملين مستقلين

H_1: المتغيرين او العاملين غير مستقلين

		B			
		B_1	B_2	---	B_c
A_1	A_1	O_{11}	O_{12}	---	O_{1c}
	A_2	O_{21}	O_{22}	---	O_{2c}
	⋮				
	A_r	O_{r1}	O_{r2}	---	O_{rc}

$O_{1.}$					
$O_{2.}$					
⋮					
$O_{r.}$					
n	$O_{.1}$	$O_{.2}$	---	$O_{.c}$	

O : عدد التكرارات او المشاهدات

E : القيمة المتوقعة عند افتراض صحة H_0

$E_{ij} = \frac{O_{i.} * O_{j}}{n}$ (تحسب لكل خلية)

احصاء الاختبار $\chi^2 = \sum \frac{(O-E)^2}{E}$ ويتبع تقريباً

توزيع χ^2 بدرجات حرية $(r-1)(c-1)$

مثال: لدراسة فيما اذا كانت الحالة الاجتماعية تعتمد على الجنسية، لديك جدول مأخوذ من SPSS. اختبر فيما اذا كانت الحالة الاجتماعية تعتمد على الجنسية عند مستوى معنوية $\alpha = 0.05$

(49)

Chi-Square Test			
	Value	df	Asympt. Sig. (2-tailed)
Pearson Chi-Square	5.73	3	0.126
Likelihood ratio	6.815	3	0.078
Linear by Linear...	3.219	1	0.073
N of Valid Cases	290		

(1) اكتب فرضية العدم والفرضية البديلة

H_0: الحالة الاجتماعية والجنس مستقلين

H_1: الحالة الاجتماعية والجنس غير مستقلين

(2) ما هي قيمة اختبار ~~الطلب~~ ؟

$$\chi^2 = 5.73$$

(3) ما هي قيمة مستوى المعنوية المقدر ؟ P-Value

$$P\text{-value} = \boxed{0.126}$$

(4) القرار: نقبل H_0 [الحالة الاجتماعية تقتل مع الجنس]

السبب: لأن مستوى المعنوية المقدر $= 0.126 >$

قيمة مستوى المعنوية $= 0.05$

(4) الانحدار الخطي البسيط

لتحديد العلاقة الخطية بين X و Y نكتب بالشكل التالي

$$Y = \beta_0 + \beta_1 X_i + \epsilon_i \quad , \quad i = 1, 2, \ldots, n$$

نبدأ مع البيانات بحساب تقدير β_0 و β_1 بواسطة b_0 و b_1 كما يلي :

تقدير معامل الانحدار $b_1 = \dfrac{S_{xy}}{S_{xx}}$

تقدير الثابت $b_0 = \bar{Y} - b_1 \bar{X}$

وعليه فان تقدير معادلة الانحدار كما يلي :

$$\boxed{\hat{Y} = b_0 + b_1 \bar{X}}$$ *

* خطأ التقدير : الفرق بين القيمة الحقيقية والقيمة المقدرة

$$e_i = Y_i - \hat{Y}_i$$

وكلما قل خطأ التقدير كلما زادت حصانة خط الانحدار

* معامل التحديد R^2

نسبة الاختلافات او التشتت بين المتغير التابع Y والمتغير المستقل X والتي يمكن الاستدلال عليها من معادلة الانحدار

$R^2 =$ مربع معامل الارتباط $= r^2$

وكلما زادت R^2 كان النموذج جيداً .

مثال: استخدم SPSS لإيجاد معادلة الانحدار الخطي البسيط لعلامة الثانوية X. والمعدل العام في الجامعة Y على معدل الثانوية X.

Model Summary				
Model	R	R-square	Adjusted R-Square	Std. Error
1	.801	0.642	0.64	0.2039

	Coefficients				
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-2.5	0.255		-9.813	0.00
High Sch. Grade	0.065	0.003	0.801	22.74	0.00

(1) اوجد معادلة الانحدار التي تربط معدل الجامعة Y بمعدل الثانوية X

$$\hat{Y} = -2.5 + 0.065X$$

(2) ما هي درجة الطالبة في الجامعة اذا كانت درجتها في الثانوية 80؟

$$\hat{Y} = -2.5 + (0.065)(80) = \boxed{2.7}$$

(3) معامل التحديد $R^2 = 0.64$

64% من التغير في معدل الطالبة في الجامعة سببه معدل الثانوية

(4) معامل الارتباط بين X و Y = 0.8 [علاقة طردية قوية]