المقاييس التشتت

- التباين : S^2 variance

نلاحظ أن مجموع انحراف عناصر (مفردات العينة) عن وسطها الحسابي $\sum(xi - x)$ يساوي صفر ، ذلك بسبب كون قسم من الانحرافات يكون موجباً بينما الآخر يكون سالباً ، وللتغلب على هذه المشكلة فقد تم معالجتها بأخذ القيم المطلقة للانحرافات في الانحراف المتوسط ، ويمكن معالجتها بطريقة أخرى إلا وهي تربيع قيم الانحرافات للحصول على مجموع مربعات الانحرافات ، ويمكن حساب التباين كما يلي :

1- للبيانات غير المبوبة :

$$S^2 = \frac{\sum(xi - \bar{x})^2}{n}$$

مثال :

جد التباين للقيم الآتية :

Xi = 9 , 8 , 6 , 5 , 7

الحل :

xi	$xi - \bar{x}$	$(xi - \bar{x})^2$
9	2	4
8	1	1
6	-1	1
5	-2	4
7	0	0
$\sum = 35$	0	10

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\frac{35}{5} = 7$$

$$S^2 = \frac{10}{5} = 2$$

2- للبيانات المبوبة :

$$S^2 = \frac{\sum F_i(X_i - \bar{X})^2}{\sum F_i}$$

الفئات	fiالتكرار	xi مركز الفئة	fi xi	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$F_i(X_i - \bar{X})^2$
60- 62	5	61	305	-6.45	41.6	208
63- 65	18	64	1152	-3.45	11.9	214.2
66- 68	42	67	2814	-0.45	0.2	8.4
69- 71	27	70	1890	2.55	6.5	175.5
72- 74	8	73	584	5.55	30.8	246.4
	100		6745			$\sum$ 852.5

$$\bar{X} = \frac{\sum fi\,xi}{\sum fi}$$

$$\bar{X} = \frac{6745}{100} = 67{,}45$$

$$S^2 = \frac{\sum F_i(X_i - \bar{X})^2}{\sum F_i}$$

$$S^2 = \frac{852.5}{100} =$$

$$S^2 = 8.525$$

- الانحراف المعياري S : standard daviation

عند حساب التباين قمنا بتربيع الانحرافات ، حيث تكون قيمة التباين مقاسة بمربع الوحدات المستخدمة في قياس المشاهدات ، ولا ضير في ذلك إلا أن المشكلة تظهر عندما يكون مربع الوحدات غير ذي معنى أو غير مقبول ، فعند استخدام وحدات قياس مثل عدد العمال أو الكيلو غرام أو الدينار فأن التباين يكون (عامل مربع أو كيلو غرام أو دينار مربع) وهذه كلها غير ذات معنى ، وحلاً لذلك يتم إرجاع وحدات القياس إلى أصلها وذلك بأخذ الجذر التربيعي للتباين وهو ما يسمى بالانحراف المعياري أو (القياسي) إذ يكون مقاساً بالوحدات الأصلية.

أ – الانحراف المعياري للبيانات غير المبوبة :

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

مثال: لدينا المعلومات الآتية والتي تمثل عدد الساعات الأسبوعية لدراسة أحد الطلاب خلال سبعة أسابيع: احسب الانحراف المعياري لتشتت ساعات الدراسة لهذا الطالب.

Xi	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
15	-2	4
9	- 8	64
14	-3	9
6	-11	121
23	6	36
19	2	4
33	16	256
119	0	494

الحل :

$$\bar{X} = \frac{119}{7} = 17$$

$$S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}} = \sqrt{\frac{494}{7}} = 8.4$$

$S = 8.4$ ساعة

ب : الانحراف المعياري للبيانات المبوبة : وهو عبارة عن الجذر التربيعي للتباين

$$S = \sqrt{\frac{\sum fi(xi-x)2}{\sum fi}}$$

مثال:

مدة الخدمة بالسنوات	عدد العمال
0.5-5.5	6
5.5-10.5	25
10.5-15.5	29
15.5-20.5	20
20.5-25.5	10
25.5-30.5	5
30.5-35.5	3
35.5-40.5	2
$\sum$	*100*

احسب الانحراف المعياري لمدة الخدمة لهؤلاء العمال

F_i عدد العمال	X_i وسطي الفئة	$X_i.F_i$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$F_i(X_i - \bar{X})^2$
6	3	18	−12	144	864
25	8	200	−7	49	1225
29	13	377	−2	4	116
20	18	360	+3	9	180
10	23	230	+8	64	640
5	28	140	+13	169	845
3	33	99	+18	324	972
2	38	76	+23	529	1058
∑100		1500			5900

نوجد أولا الوسط الحسابي لمدة الخدمة لهؤلاء العمال:

$$\bar{X} = \frac{\sum fi\, xi}{\sum fi}$$

$$\bar{X} = \frac{1500}{100} = 15$$

$$S = \sqrt{\frac{\sum fi(xi - \bar{x})2}{\sum fi}}$$

$$S = \sqrt{\frac{5900}{100}} = \sqrt{59} = 7.68$$ سنة

ملاحظة :

$$S = \sqrt{S^2} = \sqrt{\frac{\sum F_i(X_i - \bar{X})^2}{\sum F_i}}$$

$$S = \sqrt{S^2} = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}}$$

معامل الاختلاف : C V

تعريفه: معامل الاختلاف هو معامل نسبى يستخدم للمقارنة بين تشتت بيانات ظاهرتين أو أكثر مختلفتين في وحدات القياس أو متفقتين . أو مختلفتين في القيمة المتوسطة لهما.

ويستخدم هذا المؤشر والذي هو عبارة عن نسبة الانحراف المعياري على الوسط الحسابي للمقارنة بين عينتين إذا كانت وحدة القياس بينهما مختلفة (كالوزن والطول) أما إذا كانت وحدات القياس مشتركة فالخطأ المعياري يكفي للمقارنة ، ويحسب وفق الصيغة الآتية :

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\text{الوسط الحسابى}} x100$$

حيث : $C.V = \frac{S}{\overline{X}} x100$

مثال( ) قارن بين تشتت درجات كل من الطلاب والطالبات من واقع البيانات التالية:

المقاييس \ النوع	الطلاب	الطالبات
الوسط الحسابي	70	80
الانحراف المعياري	7	6

الحل

للمقارنة بين تشتت درجات كل من الطلاب والطالبات يتم حساب معامل الاختلاف لدرجات الطلاب ومعامل الاختلاف لدرجات الطالبات والمقارنة بينهما كما يلي:

أولا: معامل الاختلاف لدرجات الطلاب:

$$C.V = \frac{S}{\overline{X}} x100 = \frac{7}{70} x100 = 10\%$$

ثانيا: معامل الاختلاف لدرجات الطالبات:

$$C.V = \frac{S}{\bar{X}}x100 = \frac{6}{80}x100 = 7.5\%$$

من الملاحظ:

أن معامل الاختلاف لدرجات الطلاب أكبر من معامل الاختلاف لدرجات الطالبات وبالتالي درجات الطلاب أكثر تشتتا من درجات الطالبات.

مثال ([redacted]):

البيانات التالية توضح بعض المقاييس المحسوبة عن الأجر الشهري لعمال شركة سعودية وأخرى أمريكية في أحدى القطاعات والمطلوب المقارنة بين تشتت الأجور.

الشركة / المقاييس	الشركة السعودية	الشركة الأمريكية
الوسط الحسابي	300	400
الانحراف المعياري	45	80

الحل

للمقارنة بين تشتت الأجور في الشركة السعودية ،و الشركة الأمريكية يتم حساب معامل الاختلاف للأجور في الشركة السعودية، ومعامل الاختلاف للأجور في الشركة الأمريكية والمقارنة بينهما كما يلي :

أولا: معامل الاختلاف للشركة السعودية:

$$C.V = \frac{S}{\bar{X}}x100 = \frac{45}{300}x100 = 15\%$$

ثانيا: معامل الاختلاف للشركة الأمريكية:

$$C.V = \frac{S}{\bar{X}}x100 = \frac{80}{400}x100 = 20\%$$

من الملاحظ:

أن معامل الاختلاف لأجور العمال في الشركة السعودية أقل من معامل الاختلاف لأجور العمال في الشركة الأمريكية وبالتالي فان أجور العمال في الشركة السعودية أقل تشتتا من أجور العمال في الشركة الأمريكية.

مثال اخر :

أوجد أي من الدولتين توجد فيها عدالة اجتماعية أكثر في توزيع الدخل على أفراد مجتمعها

الدولة	متوسط الدخل $\bar{X}$	الانحراف المعياري S	نوعية العملة
A	*8000*	*360*	ليرة
B	*2000*	*160*	دولار

$$C.V = S\% = \frac{S}{\bar{X}} * \quad C.V_A = \frac{360}{8000} * 100 = 4.5\%$$

$$C.V_B = \frac{160}{2000} * 100 = 8\%$$

مقاييس التشتت

إن مقاييس النزعة المركزية لا تكفي لوحدها لوصف البيانات وإجراء المقارنات بين التوزيعات التكرارية، لأنها لا تعطينا فكرة عن مدى تجانس أو عدم تجانس البيانات، " فعند إجراء مقارنة بين مجموعتين من البيانات، يمكن استخدام شكل التوزيع التكراري، أوالمنحنى التكراري، وكذلك بعض مقاييس الترعة المركزية، مثل المتوسط الحسابي والوسيط، والمنوال، ولكن استخدام هذه الطرق وحدها لا يكفي عند المقارنة، فقد يكون مقياس الترعة المركزية للمجموعتين متساوي، وربما يوجد اختلاف كبير بين المجموعتين من حيث مدى تقارب وتباعد البيانات من بعضها البعض، أو مدى تباعد أو تقارب القيم عن مقياس الترعة المركزية.

من أجل ذلك لجأ الإحصائيون إلى استخدام مقاييس أخرى لقياس مدى تجانس البيانات، أو مدى انتشار البيانات حول مقياس الترعة المركزية، ويمكن استخدامها في المقارنة بين مجموعتين أو أكثر من البيانات، ومن هذه المقاييس ، مقاييس التشتت." [1]

يقصد بالتشتت مدى تباعد قيم المتغير الإحصائي عن بعضها البعض أو عن القيمة المركزية، فكلما ارتفعت قيم مقاييس التشتت دل ذلك على درجة كبيرة من التباعد و الاختلاف بين قيم البيانات، وكلما كانت صغيرة دل ذلك على أن الاختلاف بين قيم البيانات قليل ولذلك فإن هذه المقاييس تعطينا فكرة عن مدى تجانس أو اختلاف البيانات عن مركزها و درجة انتشارها.

أصناف مقاييس التشتت:

1- الصنف الأول ← المدى

← الانحراف الربيعي

2- الصنف الثاني ← الانحراف المتوسط

← التباين والانحراف المعياري

انواع المقاييس التشتت

A- مقاييس التشتت المطلقة

B- مقاييس التشتت النسبية

تعريف مقاييس التشتت: هي المقاييس التي تدرس تباعد المفردات والقيم عن بعضها البعض أو عن أحد مقاييس النزعة المركزية.

أصناف مقاييس التشتت:

الصنف الأول: هي المقاييس التي توضح مدى وكيفية اختلاف القيم عن بعضها البعض : المدى ، الانحراف الربيعي .

الصنف الثاني: هي المقاييس التي توضح مدى تناثر أو تجمع القيم حول أحد مقاييس النزعة المركزية : الانحراف المتوسط ، الانحراف المعياري

١- المدى : يعتبر المدى من ابسط المقاييس التشتت المطلقة ويعرف بانة الفرق بين اكبر قيمة في مجموعة البيانات واصغر قيمة .

تعريف اخر للمدى :

يعرّف المدى في البيانات بأنه (الفرق بين أعلى قيمة وأصغر قيمة في البيانات) فإذا كان المدى صغيرا كانت البيانات محصورة في فترة قصيرة، وإذا كان المدى كبيرا كانت البيانات تقع ضمن فترة طويلة.

أ – المدى في حالة البيانات غير المبوبة

$R = x_L - x_s$

R = يمثل مقياس المدى

X_L: يمثل أكبر قيمة في مجموعة البيانات.

X_s : يمثل أصغر قيمة في مجموعة البيانات.

مثال ١ :

أوجد المدى للقيم التالية :

X_i : 2 , 10 , 12 , 15 , 22

$R = x_L - x_s \rightarrow 22 - 2 = 20$

مثال اخر :

أحسب المدى لكل من المجموعتين الآتيتين من البيانات :

Y = 12 , 6 , 7 , 3 , 15 , 10 , 18 , 5

A = 9 , 12 , 12 , 12 , 22 , 22 , 22 , 24

$R_Y = 18 - 3 = 15$

$R_A = 24 - 9 = 15$

مثال اخر : اوجد مقاييس التشتت (المدى) للبيانات التالية :

5,4,16,11,4,5,9,2

الحل :

المدى $16 - 2 = 14$

ب : المدى للبيانات المبوبة

ويتم حساب المدى للبيانات المبوبة بإيجاد حاصل الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الاولى ، أي أن :

المدى = الحد الاعلى للفئة الاخيرة – الحد الأدنى للفئة الاولى :

مثال ٢ :

أوجد المدى للجدول التكراري التالي :

الحل :

$R =$ الحد الاعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

$R = 80 - 40 = 40$

الفئات	f_i
40 – 50	2
50 – 60	10
60 – 70	15
70 – 80	3

مثال اخر : اوجد مقاييس التشتت (المدى) للجدول التوزيع التكراري

فئات	26 -22	31 -27	36 -32	41 -37	46 -42	51 -47
تكرار	9	3	10	8	12	8

الحل :

المدى = الحد الاعلى للفئة الاخيره - الحد الادنى للفئة الاولى

51- 22 = 29

ملاحظة :

من خلال ما طرح ان المدى لا يعتمد على جمع البيانات بل يعتمد على أكبر قيمة وأصغر قيمة فقط. وهذا يقلل من أهميته، إذ قد يحدث أن تكون القيمتان المتطرفتان (أكبر قيمة وأصغر قيمة) قيمتين شاذتين عندئذ يكون المدى كبيرا بينما مفردات البيانات ليست متباعدة عن بعضها البعض، فمثلا إذا كانت علامات الصف الثاني الأساسي في مدرسة ما هي: 64, 72, 67, 78, 65, 74, 30, 73, 100, 70, 65, 67, 71, 65, 69, 65, 71, 70, 69 فإن المدى هو 70 = 30 - 100 بينما معظم العلامات واقعة بين 64 ، 78 أي أنها متقاربة من بعضها البعض. وأنت تلاحظ من هذا المثال أن معظم علامات الصف كانت متقاربة إلى حدّ ما إلا

أن المدى كان كبيرا. على ماذا تدل هذه الظاهرة؟ انها توضح أن المدى مقياس لا يعتمد عليه كثيرا لتحديد تشتت البيانات.

كيف يمكن التخلص من هذا النقص؟ إن إحدى الطرق للتخلص من ذلك هي حذف العلامات المتطرفة باعتبار أنها شاذة. احذف العلامتين 30 ,100 فيصبح المدى للبيانات الجديدة ؟ 14 = 64 - 78

إذا حذفت أعلى 25% من البيانات وأدنى 25% منها ثم حسبت المدى للبيانات الجديدة فإنك تحصل على المدى الربيعي.

2- المدى الربيعي (الانحراف الربيعي) : QR

الانحراف الربيعي هو أحد مقاييس التشتت المطلق ويستخدم للتغلب على العيوب الموجودة في المدى وذلك لأنه يستبعد القيم المتطرفة من الطرفين وذلك لأنه يعتمد في حسابه على الربيع الأول (Q1) ، والربيع الثالث (Q3) ويتم حسابه بالمعادلة:

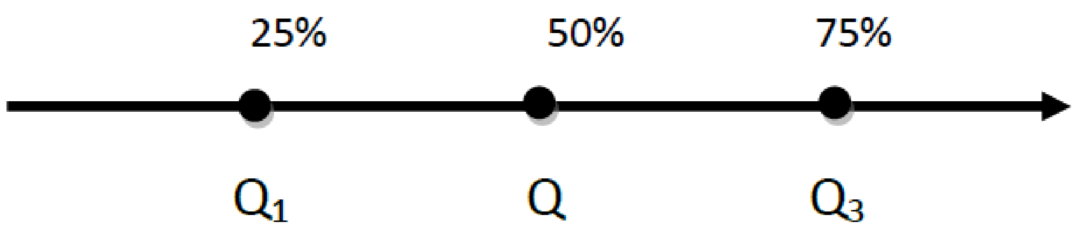
QR = Q3 – Q1

Q1 = الربيع الاول (الوسيط للربيع الادنى او الوسيط للنصف الاول)

الربيع الأول: هو القيمة التي تكون 25% من البيانات المرتبة تصاعدياً أقل منها و 75% من البيانات أكبر منها.

Q3 = الربيع الثالث (الوسيط للربيع الاعلى او الوسيط للنصف الثاني)

الربيع الثالث: هو القيمة التي تكون 75% من البيانات المرتبة تصاعدياً أقل منها و 25% من البيانات أكبر منها .



مثال : البيانات التالية تمثل درجات الطلاب قسم الحاسوب المرحلة الرابعة للعام الدراسي ٢٠٢٠-٢٠٢١ . المطلوب استخراج المدى الربيعي ؟

20,99,92,40,77,72,78,84,87,88,99,98,91,79,64

الحل : نرتب البيانات ترتيبا تصاعدياً

20,40,64,72,77,78,79,84,87,88,89,91,92,98,99

الوسيط

الربيع الاعلى Q3

Q1 الربيع الادنى

QR = Q3 – Q1 91-72= 19 QR = 19

ملاحظة :

قيمةٌ مُنخفضةٌ ← تقاربُ البياناتِ في مُنتصفِ مجموعةِ البياناتِ

قيمةٌ عاليةٌ ← تشتُّتِ البياناتِ وتباعُدِها في منتصفِ مجموعةِ البياناتِ

تمرين (١):

إذا كانت أكبر قيمة في البيانات 100 وأقل قيمة 20 وكان الربيع الأول $Q_1 = 45$ و الربيع الثالث $Q_3 = 87$ فما هو المدى وما هو المدى الربيعي.

تمرين (٢):

إذا كان الربيع الأول لمجموعة من العلامات 37 والربيع الثالث 92 وأعلى علامة 99 وأقل علامة 22. جد المدى والمدى الربيعي.

تمرين (٣) : من البيانات التالية اوجد المدى والمدى الربيعي

18,22,25 ,29,30,29,27,32,34

- الانحراف المتوسط :

وهو من مقاييس التشتت التي تقيس تشتت المفردات حول وسطها الحسابي الحقيقي ويعتمد على القيمة المطلقة للانحرافات.

أ- في حالة البيانات الغير مبوبة :

إذا كان x_n , ... , x_2 , x_1 مجموعة من المشاهدات أو القيم فالانحراف المتوسط لها هو متوسط الانحرافات المطلقة (بإهمال الإشارة) عن وسطها الحسابي الحقيقي ويرمز له بالرمز M.D وعليه فإن الانحراف المتوسط هو :

$$M.D. = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

مثال ٣ :

أوجد الانحراف المتوسط للبيانات التالية :

x_i	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$
9	9 − 7 = 2	|2| = 2
8	8 − 7 = 1	1 = 1
6	6 − 7 = −1	|−1| = 1
5	5 − 7 = −2	|−2| = 2
7	7 − 7 = 0	|0| = 0
35		6

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{35}{5} = 7$$

$$M.D. = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} = \frac{6}{5} = 1.2$$

مثال اخر : احسب الانحراف المتوسط لدرجات الطالب (س) للمواد الدراسية التالية :

المادة	درجات الطالب س XI	$X_i - \bar{X}$	$\left\| X_i - \bar{X} \right\|$
1	50	-3	3
2	56	3	3
3	49	- 4	4
4	56	3	3
5	59	6	6
6	48	-5	5
المجموع	318		24

$$\bar{X} = \frac{\sum XI}{n} = \frac{318}{6} = 53$$

$$MD = \frac{\sum \left| X_i - \bar{X} \right|}{n} \qquad MD = \frac{24}{6} = 4$$

٢ – الانحراف المتوسط للبيانات المبوبة :

$$M.D. = \frac{\sum_{i=1}^{n} f_i \left| x_i - \bar{x} \right|}{\sum_{i=1}^{n} f_i}$$

مثال

أوجد الانحراف المتوسط لجدول التوزيع التكراري التالي والذي يمثل عدد الموظفين المحالين على التقاعد موزعين حسب الفئات العمرية.

الحل :

الفئات	fi	xi	fi xi	$x_i - \bar{x}$	$\left\| x_i - \bar{x} \right\|$	$fi \left\| x_i - \bar{x} \right\|$
60 – 62	5	61	(5) (61) = 305	61 – 67= -6	6	(5) (6) =30
63 – 65	18	64	1152	64 – 67= -3	3	(18) (3) = 54
66 – 68	42	67	2814	67 – 67= 0	0	0
69 – 71	27	70	1890	3	3	81
72 – 74	8	73	584	6	6	48
	100		6745			213

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{6745}{100} = 67.4 = 67$$

$$M.D. = \frac{\sum_{i=1}^{n} f_i \left| x_i - \bar{x} \right|}{\sum_{i=1}^{n} f_i} = \frac{213}{100} = 2.13$$

ملاحظة :

xi = مراكز الفئات

= <u>الحد الادنى للفئة + الحد الاعلى للفئة</u> $\qquad$ 61 = $\frac{60+62}{2}$

2

مثال اخر:

الفئات	fi	xi	fi xi	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$	$fi \lvert x_i - \bar{x} \rvert$
60 - 62	4	61	244	61-65= -4	4	16
62 - 64	10	63	630	-2	2	20
64 - 66	12	65	780	0	0	0
66 - 68	6	67	402	2	2	12
68 - 70	8	69	552	4	4	32
	40		2608			80

xi = مراكز الفئات

= <u>الحد الادنى للفئة + الحد الاعلى للفئة</u> $61 = \frac{60+62}{2}$

2

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i x_i}{\sum_{i=1}^{m} f_i} = \frac{2608}{40} = 65.2 = 65$$

$$M.D. = \frac{\sum_{i=1}^{m} f_i \lvert x_i - \bar{x} \rvert}{\sum_{i=1}^{m} f_i} = \frac{80}{40} = 2$$

الارتباط

الارتباط الخطي :

إن مفهوم الارتباط الخطي يقترن بحالة وجود متغيرين أو أكثر يرتبطان مع بعضهما بعلاقات خطية معينة ، على سبيل المثال العلاقة بين طول الشخص ووزنه ، فإذا كان التغيير في إحدى المتغيرات يؤثر في تغيير متغير آخر أو مجموعة متغيرات عندئذ يقال أن هذه المتغيرات مرتبطة فيما بينها ، وإذا كان المتغيرين يتغيران بنفس الاتجاه أي زيادة (نقصان) في أحدها تؤدي إلى زيادة (نقصان) في الآخر عندئذ يقال أن الارتباط فيما بينهما هو موجب ، أما إذا كان المتغيرين المرتبطين أو مجموعة المتغيرات يتغيران باتجاه معاكس أي زيادة (نقصان) في أحدهما يؤدي إلى نقصان (زيادة) في الآخر عندئذ يقال الارتباط هو سالب.

الارتباط: قوة العلاقة بين متغيرين وهو أحد أنواع العلاقات بين المتغير التابع والمتغير المستقل بحيث تتحدد بعض مشاهدات المتغير التابع في ضوء المتغير المستقل حيث: س: متغير مستقل ، ص : متغير تابع.

تحليل الارتباط Correlation Analysis :

يهتم تحليل الارتباط بقوة العلاقة بين متغيرين أو أكثر ، مثل العلاقة بين مهارة العاملين والإنتاجية أو بين سعر السلعة والكمية المطلوبة ، وتقاس العلاقة بمؤشر إحصائي يدعى معامل الارتباط ويرمز له بالرمز (r) وهناك عدة أنواع من معاملات الارتباط منها البسيط والجزئي ومعامل ارتباط الرتب ومعامل ارتباط فاي ... الخ.

تعريف تحليل الارتباط: هو دراسة العلاقة بين متغيرين أحدهما يسمى المتغير المستقل والأخر المتغير التابع وتحديد قوة العلاقة وجهتها (طردية ، عكسية) ويتم ذلك من خلال حساب معامل الارتباط.

درجة الارتباط: تقاس بعدد يتراوح مقداره بين (–1 ، 1) مروراً بالصفر

خصائص معامل الارتباط :

1- تتراوح قيمة معاملة الارتباط بين سالب واحد والواحد الصحيح :

$$-1 \leq r \leq 1$$

إذ يمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن (∓ 1).

2- تكون قيمته تساوي صفراً عندما يكون المتغيران مستقلان عن بعضهما تماماً ، ويكون مساوٍ للواحد الصحيح عندما يكون الارتباط تاماً.

3- تكون قيمته موجبة عندما يكون الارتباط بين المتغيرين طردياً ويكون قوياً عندما يكون معامل الارتباط قريباً من الواحد الصحيح وضعيفاً عندما يكون قريباً من الصفر.

4- تكون قيمته سالبة عندما يكون الارتباط بين المتغيرين عكسياً ويكون قوياً عندما يكون المقدار السالب قريباً من (-1) ، وضعيفاً عندما يكون المقدار السالب قريباً من الصفر.

وبالإمكان تقسيم الارتباط الخطي إلى :

أ- الارتباط الخطي البسيط : وهو الارتباط بين الظاهرتين [redacted].

ب- الارتباط الجزئي : وهو الارتباط بين ظاهرتين بثبوت الظاهرة الثالثة.

ج- الارتباط المتعدد : وهو الارتباط بين ثلاث ظواهر.

درجة الارتباط: تقاس بعدد يتراوح مقداره بين (−1 ، 1) مروراً بالصفر

مقياس درجة الارتباط [معامل الارتباط]

1	صفر	1-

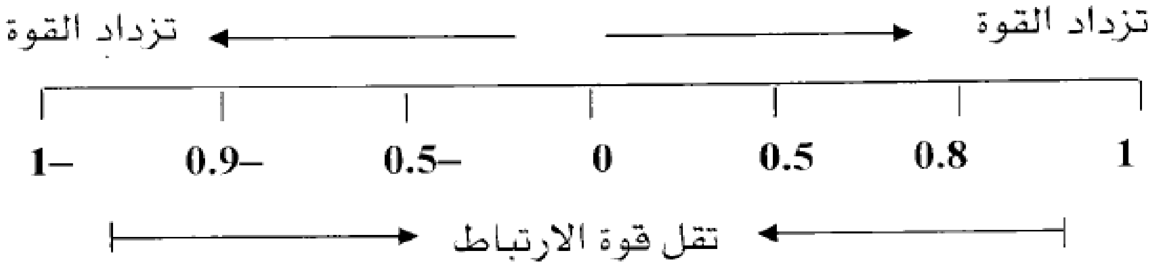
وبناء على هذا المقياس صنفت علاقة الارتباط إلى ما يلي

ارتباط طردي تام (درجة الارتباط = 1) (موجب)	لا يوجد ارتباط (درجة الارتباط =	ارتباط عكسي تام (درجة الارتباط = −1) (سالب)

أما قوة درجة الارتباط فإنها تصنف وفق الأصناف التالية

قوي طردي (بين $\frac{1}{2}$ و 1)	ضعيف بين ($-\frac{1}{2}$ و $\frac{1}{2}$)	قوي عكسي بين (−1، $-\frac{1}{2}$)

ملاحظة هامة: تزداد قوة الارتباط كلما اقترينا من الأطراف وتقل كلما ابتعدنا عن الأطراف.



قانون الارتباط

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

مثال

البيانات التالية تمثل الكمية المعروضة من سلعة معينة وسعر الوحدة الواحدة من هذه السلعة.

المطلوب : حساب معامل الارتباط البسيط بين الكمية المعروضة والسعر بالصيغتين.

[redacted] :

السعر x	الكمية المعروضة y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
2	3	$(2-4) = -2$	$(3-7) = -4$	$(-2)^2 = 4$	$(-4)^2 = 16$	$(-2)(-4) = 8$
2	5	$(2-4) = -2$	$(5-7) = -2$	$(-2)^2 = 4$	$(-2)^2 = 4$	$(-2)(-2) = 4$
5	7	$(5-4) = 1$	$(7-7) = 0$	$(-1)^2 = 1$	0	$(1)(0) = 0$
4	8	0	1	0	1	0
5	9	1	2	1	4	2
6	11	2	4	4	16	8
3	6	-1	-1	1	1	1
5	8	1	1	1	1	1
4	6	0	-1	0	1	0
36	63	0	0	16	44	24

$$\bar{x} = \frac{\sum_{i=1}^{n}(x_i)}{n} = \frac{36}{9} = 4 \text{ ,,,, } \bar{y} = \frac{\sum_{i=1}^{n}(y_i)}{n} = \frac{63}{9} = 7$$

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{24}{\sqrt{(16)(44)}} = 0.95$$

مثال اخر : سجلت ستة قراءات لحجم الانتاج وحجم الصادرات النفط الخام في جمهورية العراق (بالمليار برميل) خلال عدة سنوات كما يلي :

حجم الانتاج x	حجم الصادرات y
3	2
3	3
4	3
3	2
3	2
6	4

المطلوب : اوجد علاقة ارتباط خطية بين حجم الانتاج وحجم الصادرات لنفط الخام

حجم الانتاج x	حجم الصادرات y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
3	2	-1	-1	1	1	1
3	3	-1	0	1	0	0
4	3	0	0	0	0	0
3	2	-1	-1	1	1	1
3	2	-1	-1	1	1	1
6	4	2	1	4	1	2
22	16			8	4	5

$\bar{x} = 4$

$\bar{y} = 3$

$$R = \frac{5}{\sqrt{32}} = \frac{5}{5.65} = 0.88$$

معامل الارتباط بيرسون : rp

ويرمز له بالرمز rp وهو معامل ارتباط خطي بسيط يقيس قوة واتجاه العلاقة بين متغيرين فقط ، وهذان المتغيران هما متغيران كميان أي يعبر عنهما بالأرقام ، ويحسب المعامل وفق القانون الآتي :

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

مثال : جد قوة العلاقة بين كفاءة الانتاج التي تمثل xi وجودة المنتوج التي تمثل yi باستخدام معامل الارتباط بيرسون :

Xi	Yi	$(X_i - \bar{X})$	$(y_i - \bar{y})$	$(\bar{x_i} - \bar{x})(\bar{y_i} - \bar{y})$	$(X_i - \bar{X})^2$	$(y_i - \bar{y})^2$
5	3	-5	-2	10	25	4
7	5	-3	0	0	9	0
10	5	0	0	0	0	0
20	7	10	2	20	100	4
8	5	-2	0	0	4	0
$\sum$ 50	25			30	138	8

$$\bar{X} = \frac{50}{5} = 10 \quad , \quad \bar{y} = \frac{25}{5} = 5$$

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}} = \sqrt{\frac{138}{5}} = \sqrt{27.6} = 5.25$$

$$S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n}} = \sqrt{\frac{8}{5}} = \sqrt{1.6} = 1.26$$

$$rp = \frac{\sum (xi - \bar{X})(yi - \bar{y})}{n\, S_x\, S_y}$$

$$rp = \frac{30}{5\,(5.25)(1.26)} = 0.907$$

مثال : البيانات التالية توضح عدد ساعات دراسة الطلاب (xi) ودرجاتهم النهائية (yi) لقسم الحاسوب . احسب معامل ارتباط بيرسون :

Xi عدد ساعات الدراسة	Yi الدرجة	$(X_i - \bar{X})$	$(y_i - \bar{y})$	$(\overline{xi} - \bar{x})(\overline{yi} - \bar{y})$	$(X_i - \bar{X})^2$	$(y_i - \bar{y})^2$
150	200	-17	-45	765	289	2025
162	250	-5	5	-25	25	25
180	300	13	55	715	169	3025
160	200	-7	-45	315	49	2025
170	240	3	-5	-15	9	25
180	280	13	35	455	169	1225
$\sum$ 1002	1470			2210	710	8350

$$\bar{x} = \frac{1002}{6} = 167$$

$$\bar{y} = \frac{1470}{6} = 245$$

$$S_x = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} = \sqrt{\frac{710}{6}} = 10.88$$

$$S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n}} = \sqrt{\frac{8350}{6}} = 37.31$$

$$rp = \frac{\sum (xi - \bar{X})(yi - \bar{y})}{n\, S_x\, S_y}$$

$$rp = \frac{2210}{6\,(10.88)(37.31)} = \frac{2210}{2435.6} = 0.91$$

معامل ارتباط الرتب سبيرمان rs

معامل ارتباط الرتب لسبيرمان Spearman Rank correlation coff. :

أفرض أن x , y متغيرين من النوع الوصفي وأفرض أن البيانات المستحصل عليها على أساس عينة عشوائية من المفردات قوامها n هي بهيئة صفات غير قابلة للقياس الكمي ويسمى الارتباط لهذه الحالة بارتباط الرتب والصيغة المستخدمة لإيجاد معامل الارتباط هي :

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

حيث ان : rs تمثل معامل (ارتباط الرتب)

n : عدد القيم لكلا الظاهرتين

$6\sum di^2$: عبارة عن مجموع مربعات الفروق بين الرتب الظاهرة الاولى ورتب الظاهر الثانية المناظرة لها

أما الخطوات الواجب اتباعها لتطبيق هذه الصيغة هي كما يلي :

١- نعطي قيم من الأعداد الطبيعية على قيم الظاهرتين مباشرة بترتيب تصاعدي.

٢- نعطي القيم المتكررة معدل من الرتب التي تناظرها.

٣- نرتب قيم الظاهرة الأولى بتسلسل الأعداد الطبيعية وتذكر الرتب التي تناظرها من رتب الظاهرة الثانية.

٤- تحسب الفروق بين الرتب والرتب المناظرة فنحصل على di ثم نربع هذه الفروق ثم نجمعها والخطوة الأخيرة نطبق الصيغة المذكورة أعلاه.

ملاحظة : يمكن استخدام صيغة (سبيرمان) لإيجاد معامل ارتباط الرتب بين ظاهرتين غير مقاسة كمياً ، وذلك بذكر تسلسل الظاهرتين تصاعدياً وإعطاءهما رتب من الاعداد الطبيعية ثم نكمل الحل بإتباع الخطوات المذكورة أعلاه.

مثال : لديك البيانات التالية والمطلوب معرفة العلاقة بين الظاهرتين (x ,y) ، علماً ان البيانات الحالة الانتاجية yi لاحدى المصانع وكمية الانتاج xi لاحدى السلع في هذا المصنع .

[redacted]

x_i	y_i	الرتب x_i والرتب المناظرة y_i	d_i	d_i^2
$^{3}76$	متوسط ^{2}	3 - 2	1	1
$^{4}82$	جيد ^{3}	4 - 3	1	1
$^{1}63$	ضعيف ^{1}	1 - 1	0	0
$^{2}70$	جيد جداً ^{4}	2 - 4	-2	4
				$6 \rightarrow \sum_{i=1}^{n} d_i^2$

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

$$rs = 1- \frac{6(6)}{4\ (\ 16-1\ )} = 1- \frac{36}{4(15)} = 1- \frac{36}{60} = \frac{60-36}{60} = \frac{24}{60} = 0.4$$

مثال :

أجاب عشرة موظفين عن حالتهم العلمية ومقدار دخلهم الشهري بالدينار فكانت إجاباتهم كما في الجدول أدناه ، المطلوب إيجاد معامل ارتباط الرتب باستخدام صيغة (سبيرمان) بين الدخل الشهري والحالة العلمية وتفسير النتائج.

الحالة العلمية x_i	الدخل الشهري y_i	الرتب x_i والرتب المناظرة إلى y_i	d^i	d^{i2}
يقرأ ويكتب ^{4}	$^{7}43$	4 − 7	−3	9
متوسطة ^{7}	$^{6}39$	7 − 6	1	1
أمي $^{1.5}$	$^{2}11$	1.5 − 2	−0.5	0.25
يقرأ ويكتب ^{4}	$^{3}16$	4 − 3	1	1
عليا $^{9.5}$	$^{1}10$	9.5 − 1	8.5	72.25
عليا $^{9.5}$	$^{9}78$	9.5 − 9	0.5	0.25
متوسطة ^{7}	$^{10}87$	7 − 10	-3	9
يقرأ ويكتب ^{4}	$^{5}32$	4 5	-1	1
أمي $^{1.5}$	$^{4}19$	15 4	-2.5	6.25
متوسطة ^{7}	$^{8}50$	7 8	1	1
المجموع				101

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

$$rs=1- \frac{6(101)}{10\ (\ 100-1\ )} = 1- \frac{606}{10(99)} = 1- \frac{606}{990} = \frac{990-606}{990} = \frac{384}{990} = 0.388 \cong 0.39$$

مثال / اوجد معامل ارتباط رتب سبير مان للمتغيرين (xi) و(yi) :

Xi	8	10	6	4	12	13	5	11	9
Yi	150	160	150	130	165	180	120	160	150

الحل :

نرتب البيانات ترتيباً تصاعدياً

xi	4	5	6	8	9	10	11	12	13
رتب xi	1	2	3	4	5	6	7	8	9
Yi	120	130	150	150	150	160	160	165	180
رتب yi	1	2	$\frac{3+4+5}{3}$ 4	$\frac{3+4+5}{3}$ 4	$\frac{3+4+5}{3}$ 4	$\frac{6+7}{2}$ 6.5	$\frac{6+7}{2}$ 6.5	8	9

Xi	Yi	رتب xi	رتب yi	Xi –yi	di	di^2
8	150	4	4	4 -4	0	0
10	160	6	6.5	6 - 6.5	-0.5	0.25
6	150	3	4	3 -4	-1	1
4	130	1	2	1-2	-1	1
12	165	8	8	8-8	0	0
13	180	9	9	9 -9	0	0
5	120	2	1	2-1	1	1
11	160	7	6.5	7-6.5	0.5	0.25
9	150	5	4	5-4	1	1
						$\sum 4.5$

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

$$rs = 1 - \frac{6(4.5)}{9\ (\ 81-1\ )} = 1 - \frac{27}{9(80)} = 1 - \frac{27}{720} = 1 - 0.04 = 0.96$$