المحـــاضـرة الحاديـة عشـر

مقاييس التشتت النسبى والدرجة المعيارية

سيتم فى هذه المحاضرة استعراض كلا من:

- مقاييس التشتت النسبى
- القيمة المعيارية

أولاً – مقاييس التشتت النسبى Coefficient of Variation

يستخدم هذا النوع من المقاييس لمقارنة تشتت مجموعتين من البيانات او ظاهرتين او توزيعين حيث يتم الاعتماد فى عملية المقارنة على مقاييس التشتت النسبى Coefficient of variations (c.v.) والتى يعبر عنها من خلال معامل الاختلاف المعيارى والذي يحسب من خلال المعادلات التالية:

تستخدم في حال وجود المتوسط الحسابي و الانحراف المعياري $c.v.=\frac{s}{\bar{x}}\times 100$

تستخدم في حال وجود الرُبيعات (الربيع الأول والثاني) $c.v.=\frac{Q_3-Q_1}{Q_3+Q_1}\times 100$

للتذكير :

معادلة حساب الربيع الأول Q1 $Q_1=\mathrm{L}_{Q_1}+\frac{\frac{n}{4}-F_a}{F_b-F_a}\times I_{Q_1}$

معادلة حساب الربيع الثالث Q1 $Q_3=\mathrm{L}_{Q_3}+\frac{\frac{3(n)}{4}-F_a}{F_b-F_a}\times I_{Q_3}$

مثال: البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الإيجار السنوى بأحد الاحياء:

الايجار بالألف ريال	٦-	١٠-	١٢-	١٤-١٨
عدد الوحدات السكنية	١٥	٢٠	١٢	١٣

المطلوب:

حساب :

- معامل الاختلاف للإيجار السنوى
- معامل الاختلاف الربيعى للإيجار السنوى

ك . م . ص	$(x-\bar{x})^2 f$	$(x-\bar{x})^2$	$x-\bar{x}$	$\bar{x}$	X f	X	التكرارات f	الفئات
15	208,65	13,91	- 3,73	11,73	120	8	15	6 -
35	10,66	0,533	-0,73	11,73	220	11	20	10 -
47	19,32	1,81	-1,27	11,73	150	13	12	12 -
60	236,99	18,23	-4,27	11,73	208	16	13	14 - 18
	475,62				$\sum xf = 704$		$\sum f = 60$	

ملاحظة : النواتج بالجدول طبقنا عليها القوانين السابقة ،، بنفس طريقة إيجاد الجداول السابقة

بالبداية نحسب التباين بهذا القانون : $S^2 = \frac{\sum f(x-\bar{x})^2}{\sum f}$

نعوض ،، $\frac{475,62}{60} = 7,93 = S^2$

الانحراف المعياري : $S = \sqrt{S^2}$ << $S = \sqrt{7,93}$ << S = 2,82

- نحسب معامل الاختلاف للإيجار السنوي .

نطبق القانون : $c.v. = \frac{S}{\bar{x}} \times 100$

$c.v = \frac{2,82}{11,73} \times 100 = 24\%$

∴ معامل الاختلاف للإيجار السنوي يبلغ ٢٤%

- معامل الاختلاف الربيعي للإيجار السنوي .

أولا نوجد الربيع الأول والثالث ،،

نوجد الربيع الأول . $K_{Q1} = \frac{n}{4} = 60 \div 4 = 10$ .. نوجد مكانها بالجدول بـ ت . م . ص ... قيمة 10 = Q1

ملاحظة مهمه :

عندما تكون قيمة التكرار المتجمع الصاعد مطابقة لترتيب الرُبيع فإن الحد الأعلى للفئة الرُبيعية تمثل قيمة الربيع المطلوب

نوجد الرُبيع الثالث . $K_{Q1} = \frac{3(n)}{4} = \frac{3(60)}{4} = 45$

نوجد مكانها بالجدول بـ ت . م . ص .. $L_{Q3} = 12$ ... $F_b = 47$ .... $F_a = 35$ ..... $I_{Q3} = 2$

نعوض بالقانون : $Q_3 = L_{Q_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3}$

$$Q_3 = 12 + \frac{45-35}{47-45} \times 2 \gg Q_3 = 13,49$$

نعوض بالقانون لنوجد معامل الاختلاف الرُبيعي .. $c.v. = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

$$C.V = \frac{13,67-15}{13,67+15} \times 100 \gg C.V = 15,49$$

ويتضح لنا من الحل السابق أن:

- معامل الاختلاف للإيجار السنوي للوحدات السكنية بلغ ٢٤%
- معامل الاختلاف الربيعي للإيجار السنوى للوحدات السكنية بلغ ١٥.٤٩٤%

ونلاحظ وجود اختلاف بين قيمتى معامل الاختلاف باستخدام كلا من المعادلة الأولى والثانية وذلك لاختلاف الأساس الرياضى فى كل من التعريفين المعادلتين. إلا أنة يفضل استخدام المعادلة الثانية فى حالة الجداول التكرارية المفتوحة أما غير ذلك فيفضل استخدام المعادلة الأولى .

ثانيا : القيمة المعيارية Standardized values .

وهي تلك القيمة التي تقيس مدى انحراف قيمة مفردة ما من مفردات الدراسة عن الوسط الحسابى لها وذلك بوحدات من الانحراف المعيارى، ويرمز للقيمة المعيارية بالرمز Z حيث أن:

X الدرجة المتحصل عليها

$\overline{X}$ المتوسط الحسابي

S الانحراف المعياري

$$z = \frac{x - \overline{x}}{S}$$

مثال: حصل أحد الطلاب فى مقرر المحاسبة على (٨٠) درجة حيث بلغ متوسط درجات الطلاب فى اختبار المحاسبة (٨٣) درجة بانحراف معيارى (٥) . بينما حصل فى اختبار مقرر الرياضيات على (٧٠) درجة حيث بلغ متوسط درجة الطلاب فى اختبار الرياضيات (٦٥) درجة بانحراف معيارى قدرة (٥) درجات .

المطلوب:

هل يمكن القول بأن درجات الطالب فى مقرر المحاسبة أفضل من درجته فى مقرر الرياضيات ؟

X المحاسبة = 80 .. $\overline{X}$ للمحاسبة = 83 .. S للمحاسبة = 5

نعوض بالقانون : $z = \frac{x - \overline{x}}{S}$

$$= \frac{80-83}{5} = \frac{-3}{5} = -0,6$$

X الرياضيات = 76 .. $\overline{X}$ للرياضيات = 65 .. S للرياضيات = 5

نعوض بالقانون : $z = \frac{x - \overline{x}}{S}$

$$= \frac{70-65}{5} = \frac{5}{5} = 1$$

يتضح لنا من الحل أن القيمة المعيارية لدرجة الطالب فى مقرر الرياضيات هى (+1) مما يعنى أن الدرجة التى حصل عليها الطالب أكبر من متوسط درجات الطالب بينما بلغت القيمة المعيارية للدرجة التى حصل عليها الطالب فى مقرر المحاسبة (-0.6) مما يدل على أن الدرجة التى حصل عليها الطلب أقل من متوسط الدرجات التى حصل عليها الطلاب .

المحـاضـرة الثانية عشرة

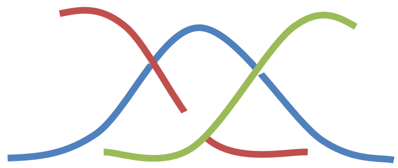
مقاييس الالتواء والتفلطح

أولا : مقاييس الالتواء Skewdness Measures

عند دراسة أشكال منحنيات التوزيعات التكرارية المختلفة نجد أن منها ما هو متماثل Symmertical ومنها الغير متماثل أى يوجد به ما يسمى بالالتواء Skewed كما يتضح من أشكال منحنيات التوزيعات التالية:

مقاييس الالتواء هي تلك الأشكال التي تأخذها التوزيعات المختلفة للبيانات عندما يكون هناك اتجاه سواء إلى الالتواء السالب او الالتواء الموجب ، بدلا من أن يكون التوزيع طبيعي أو متماثل .

وعادة يكون التوزيع طبيعي ..



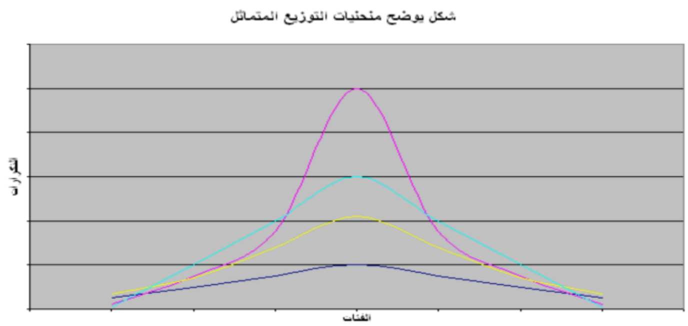
اللون الأزرق متماثل

اللون الأحمر التواء موجب

اللون الأخضر التواء سالب

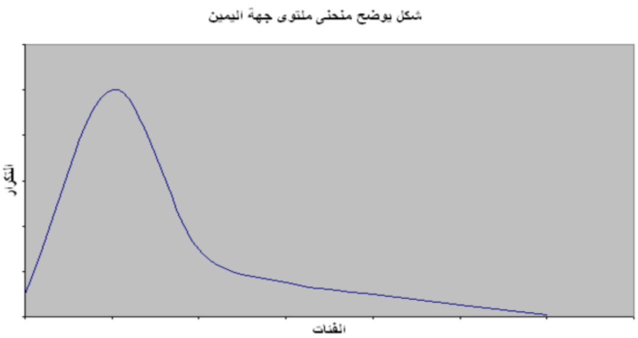
المنحنى المتماثل Symmetrical Curve

هو المنحنى الذى اذا قسمناه إلى نصفين انطبق هذان النصفان على بعضهما البعض تماما



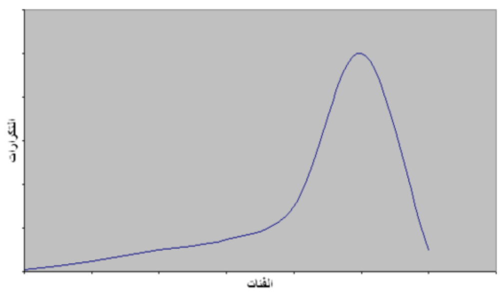
المنحنيات الملتوية Skewed

إن الكثير من التوزيعات الإحصائية تبتعد عن التماثل بتركز تكراراتها إما عند أصغر القيم فيصبح المنحنى ملتويا جهة اليمين أو التواء موجب كما يظهر فى الشكل التالي:



أما فى حالة تركز التكرارات عند عند أكبر القيم فيسمى المنحنى فى تلك الحالة منحنى ملتوي جهة اليسار (التواء سالب) كما يظهر من الشكل التالى:

شكل يوضح منحنى ملتوى جهة اليسار



ويمكن قياس الالتواء من خلال معامل الالتواء SK والذى يفيدنا فى الحكم على مدى تماثل أو التواء التوزيع

تتعدد مقاييس الالتواء إلا أن من أهمها:

معامل الالتواء لبيرسون والذى يكون فى أحد الصورتين التاليتين:

$$SK = \frac{\bar{x} - Mod}{S}$$ أو $$SK = \frac{3(\bar{x} - Med)}{S}$$

وحيث أنه لا يمكن حساب معامل الالتواء لبيرسون فى حالة المنحنيات التى تكون شديدة الالتواء أو فى حالة التوزيعات التكرارية المفتوحة .

لذلك يمكن الاعتماد على مقياس الالتواء لباولى SkB الذى يعرف كما يلي :

$$SK_B = \frac{Q_3 - 2Med + Q1}{Q_3 - Q1}$$

للتذكير :

الوسيط Med: $$Med = \mathrm{L}_{\mathrm{Med}} + \frac{k_{Med} - F_a}{F_b - F_a} \times I_{Med}$$

الربيع الأول Q1: $$Q_1 = \mathrm{L}_{Q_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1}$$

الربيع الثالث Q3: $$Q_3 = \mathrm{L}_{Q_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3}$$

مثال: البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الإيجار السنوى بأحد الاحياء في أحد المدن:

الايجار بالألف ريال	٦-	١٠-	١٢-	١٤-١٨
عدد الوحدات السكنية	١٥	٢٠	١٢	١٣

المطلوب :

حساب معامل الإلتواء لتوزيع الإيجار السنوى للوحدات السكنية.

لنفس هذا المثال بالمحاضرة السابقة حسبنا المتوسط الحسابي و الانحراف المعياري والرُبيع الأول والرُبيع الثالث و الوسيط ،، نأخذ الناتج الأخير لها ..

$$\bar{x} = 11,73 \;,,\quad Med = 11,5 \;,,\quad Q_1 = 10 \;,,$$
$$Q_3 = 13,67 \;,,,\quad S = 2,82$$

نطبق مباشرة في المعادلة من خلال استخدام الوسيط ،، $SK = \frac{3(\bar{x} - Med\ )}{S}$

نعوض .. $\frac{3(11,73-11,5)}{2,82} = 0,249$

لنطبق القانون الآخر يجب ان نعرف المنوال ..

قانون المنوال $Mod = L_{Mod} + \frac{D_1}{D_1 + D_2} \times I_{Mod}$

نجد أن الجدول غير منتظم .. لذلك نستخدم التكرار المعدل ،، نأخذ كل تكرار ونقسمه على طول الفئة ..

$15 \div 4 = 3,75$,, $20 \div 2 = 10$,, $12 \div 2 = 6$,, $13 \div 4 = 3,25$

أكبر تكرار اللي هو 10 وبكذا يكون الفئة المنواليه ..

نطلع D1 .. $D1$ يساوى تكرار فئة المنوال – تكرار الفئة السابقة

D1 = 3,75 – 10 = 6,25

نطلع D2 .. $D2$ يساوى تكرار فئة المنوال – تكرار الفئة اللاحقة

D2 = 6 – 10 = 4

نطبق بقانون المنوال ،، $Mod = 10 + \frac{6,25}{6,25+4} \times 2 \gg 11,22$

نطبق مباشرة بالمعادلة من خلال استخدام المنوال .. $SK = \frac{\bar{x} - Mod}{S}$

$$= \frac{11,73 - 11,22}{2,82} = 0,18$$

ويظهر لنا من النتيجة لجميع المعادلات الخاصة بحساب معامل الالتواء وجود التواء موجب جهة اليمين إلا أن قيمة معامل الالتواء صغيرة تقترب من الصفر مما يدل ايضا على أن التوزيع قريب من التماثل.

ونتيجة لوجود اختلاف فى الاصل الرياضى لكل من المعادلات الثلاث السابقة لذا نجد أن قيمة معامل الالتواء تختلف. إلا أنة كما سبق وذكرنا بأنه يفضل استخدام معامل الالتواء لبيرسون فى أي من صيغتيه فى حالة البيانات غير المبوبة وكذلك الجداول التكرارية المغلقة أما في حالة الجداول التكرارية المفتوحة فيفضل استخدام معامل الالتواء لباولي.

لو طبقنا قانون لباولى SkB

$$SK_B = \frac{Q_3 - 2Med + Q1}{Q_3 - Q1}$$

$$= \frac{13{,}67 - 2\ (11{,}5)\ + 10}{13{,}67 - 10} = -0{,}235$$

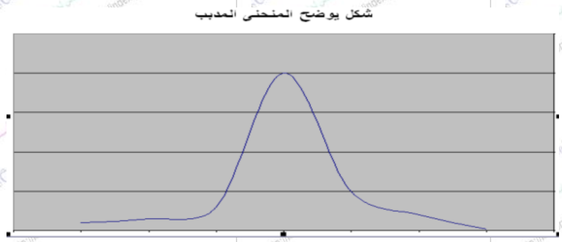
هذا القانون يستخدم للجداول التكرارية المفتوحة التي لا نستطيع أن نحسب من خلالها طول الفئة ..

ثانيا: التفلطح Kurtosis

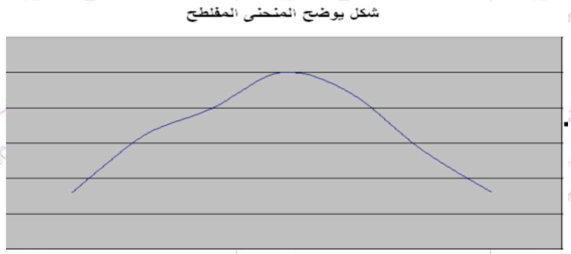
يقصد بالتفلطح مقدار التدبب (الارتفاع أو الانخفاض) فى قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي.

وتكون قيمة معامل التفلطح صفر فى حالة التوزيع الطبيعى المعياري .

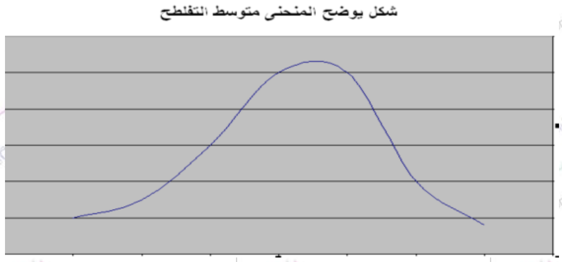
ففى حالة ما يكون معامل التفلطح للبيانات الاصلية أكبر من ٣ يكون المنحنى مدبب لأعلى كما بالشكل التالي:



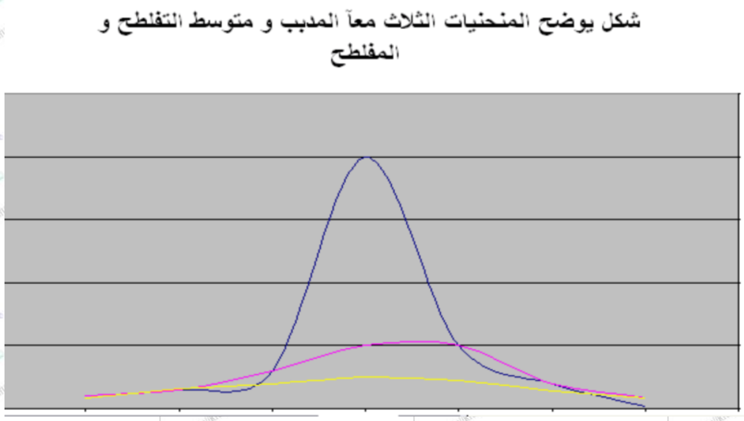
أما فى حالة ما يكون معامل التفلطح للبيانات الأصلية أقل من ٣ يعنى ذلك أن المنحنى مفلطح كما يتضح من الشكل التالي:



أما فى حالة ما يكون معامل التفلطح يساوى ثلاثة يكون المنحنى متوسط التفلطح و يكون بالشكل التالي :



وحتى يتضح الفرق بين المنحنيات الثلاث يمكن رسمها معا كما يلي:



ويتم قياس **معامل التفرطح KU** باستخدام الربيعات والمئينيات من خلال المعادلة التالية:

$$KU = \frac{Q_3 - Q_1}{2(P_{0.90} - P_{0.10})}$$

حيث يشير:

$P_{0.90}$	إلى المنين التسعين والذى يعبر عن ٩٠ % من المفردات تكون أقل منه و ١٠% منها أكبر منه
$P_{0.10}$	إلى المنين العاشر (العشير) والذى يعبر عن ١٠ % من المفردات تكون أقل منه و ٩٠% منها أكبر منه

مثال: البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الإيجار السنوى بأحد الاحياء في أحد المدن:

الايجار بالألف ريال	٦-	١٠ -	١٢-	١٤ - ١٨
عدد الوحدات السكنية	١٥	٢٠	١٢	١٣

المطلوب:

حساب معامل التفلطح لتوزيع الإيجار السنوى للوحدات السكنية.

نحسبه من خلال هذا القانون $KU = \frac{Q_3 - Q_1}{2(P_{0.90} - P_{0.10})}$

الرُبيع الأول والرُبيع الثالث طلعنا قيمها بالأمثلة السابقة $Q_1 = 10$,, $Q_3 = 13,67$

بقي لنا $P_{0,10}$ ونطبق عليه هذا القانون $P_{0,10} = L_{P0,10} + \frac{K_{P0,10} - f_a}{f_b - f_a} \times I_{P0,10}$

وأيضا بقي لنا $P_{0,90}$ ونطبق عليه هذا القانون $P_{0,90} = L_{P0,90} + \frac{K_{P0,90} - f_a}{f_b - f_a} \times I_{P0,90}$

الـ $K_{P0,10} = n \div 10$ وأيضاً الـ $K_{P0,90} = (n \times 9) \div 10$

نوجد أولا الجدول التكراري

ت . م . ص	التكرارات	الفئات
15	15	6 -
35	20	10 -
47	12	12 -
60	13	14 – 18
	$\sum f = 60$	

نعوض $K_{P0,10} = n \div 10$ >> $K_{P0,10} = 60 \div 10 = 6$

و الـ $K_{P0,90} = (n \times 9) \div 10$ >> $K_{P0,90} = (60 \times 9) \div 10 = 540 \div 10 = 54$

نبدأ نعوض بالقوانين ،،

$$P_{0,10} = L_{P0,10} + \frac{K_{P0,10} - f_a}{f_b - f_a} \times I_{P0,10}$$

$$P_{0,10} = 6 + \frac{6 - 0}{15 - 0} \times 4 = 7,6$$

$$P_{0,90} = L_{P0,90} + \frac{K_{P0,90} - f_a}{f_b - f_a} \times I_{P0,90}$$

$$P_{0,90} = 14 + \frac{54 - 47}{60 - 47} \times 4 = 16,15$$

نبدأ نعوض بقانون التفلطح $KU = \frac{Q_3 - Q_1}{2(P_{0.90} - P_{0.10})}$

$$KU = \frac{13,67 - 10}{2(16,15 - 7,6)} = \frac{3,67}{2(8,55)} = \frac{3,67}{17,1} = 0,215$$

ويتضح لنا أن معامل التفلطح أقل من ٣ مما يدل على أن المنحنى مفلطح أى أن المشاهدات (التكرارات) موزعة على الفئات المختلفة للإيجار السنوى ولا يوجد تركز بدرجة كبيرة فى أحد الفئات على حساب باقى الفئات الأخرى .

الـمحـاضـرة الثـالـثـة عشر

تحليل الإرتباط - ١

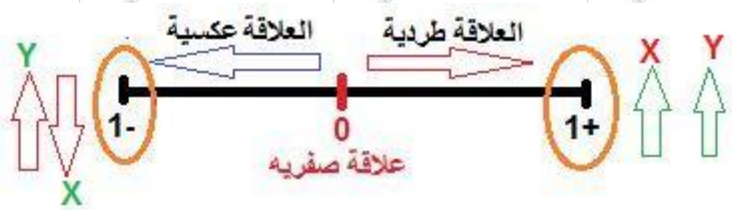
بالرغم من أن مقاييس العلاقة تختلف عما سبقها من مقاييس ، فهي تتعلق بدراسة العلاقة بين متغيرين (الإنتاجية والجودة) مثلا، بينما المقاييس السابقة فتهتم بدراسة الفروق بين المتغيرات .

وعندما نقول مقاييس العلاقة نعني بذلك تلك المقاييس التي تبين درجة العلاقة والارتباط بين متغيرين أو أكثر مثلا، كأن يكون الهدف معرفة هل هناك علاقة بين مستوى الإنتاجية وجودة المنتج في مصنع ما؟، أي هل كلما زادت الإنتاجية تقل جودة المنتج أو العكس .

معامل الارتباط: هو تعبير يشير إلى المقياس الإحصائي الذي يدل على مقدار العلاقة بين المتغيرات سلبية كانت أم إيجابية، وتتراوح قيمته بين الارتباط الموجب التام (+١) وبين الارتباط السالب التام (-١) .

معظم الأحيان يكون معامل الارتباط كسر بمعنى الحصول على ارتباط تام نادر جدا ،، إذا كان تام وصل إلى +1 و -1 أي يكون عدد صحيح .. هذا يعطينا مؤشر أنه عندما يكون ناتج معامل الارتباط أكبر من +1 أو أصغر من –1 يكون الحل خطأ ..

العلاقة الطردية بين المتغيرات : هو تعبير يشير إلى تزايد المتغيرين المستقل والتابع معا ، فإذا كانت الإنتاجية مرتفعة ، ومستوى الجودة مرتفع ، يقال حينئذ أن بينهما ارتباط موجب ، وأعلى درجة تمثله هي (+١) .



العلاقة العكسية بين المتغيرات : هو تعبير يشير إلى تزايد في متغير يقابله تناقص في المتغير الآخر ، فإذا كانت الإنتاجية منخفضة ومستوى الجودة مرتفع ، يقال حينئذ أن بينهما ارتباط سالب ، وأعلى درجة تمثله هي (-١) .

ومن الطبيعي ملاحظة أن الارتباط الكامل لا وجود له أو نادر الوجود في الظواهر الطبيعية ، وأن معامل الارتباط الناتج في الأبحاث والدراسات الإنسانية والاجتماعية يكون عادة كسرا موجبا أو سالبا. والجدول التالي يوضح أنواع العلاقات بين المتغيرات كما يصفها معامل الارتباط:

نوع العلاقة	قيمة معامل الارتباط
طرديه كاملة	+١
طرديه ناقصة	+ كسر(قيمة موجبة)
صفرية	صفر
عكسية ناقصة	- كسر (قيمة سالبة)
عكسية كاملة	-١

إن معامل الارتباط التام الموجب (+١) يعنى التغير في اتجاه واحد في كلا الظاهرتين مع بقاء الأوضاع النسبية لوحدات الظاهرة ثابتة، سواء كان هذا التغير في اتجاه الزيادة (أي زيادة قيم الظاهرة الأولى تتبعها زيادة في قيم الظاهرة الأخرى)، أو في اتجاه النقص (أي نقص قيم الظاهرة الأولى يتبعها نقص في قيم الظاهرة الأخرى) .

طرق التعرف على العلاقة بين متغيرين وحسابها

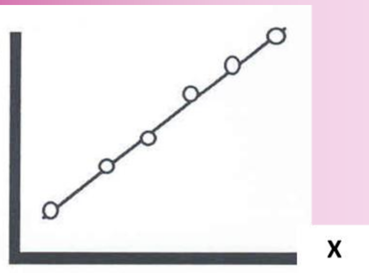
أولا: طريقة شكل الانتشار Scatter Diagram :

هناك وسيلة مبدئية يعرف الباحث من خلالها نوع الارتباط بين المتغيرين وما إذا كان الارتباط قوياً وضعيفاً أو منعدماً ، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة. هذه الوسيلة هي " شكل الانتشار " والتي تصلح إذا كان المتغيران كميين. وجدير بالذكر أن هذه وسيلة مبدئية تساعد فقط في معرفة نوع الارتباط ولا تعتبر بديلاً عن الطرق الإحصائية التي سوف نتناولها بالتفصيل في هذه المحاضرة.

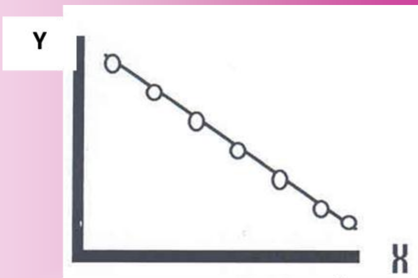
والمقصود بشكل الانتشار هو تمثيل قيم الظاهرتين بيانياً على المحورين، المتغير الأول X على المحور الأفقي والمتغير الثاني Y على المحور الرأسي، حيث يتم تمثيل كل زوج Pair من القيم بنقطه ، فنحصل على شكل يمثل كيفية انتشار القيم على المستوى، وهو الذي يسمى شكل الانتشار. وطريقة انتشار القيم تدل على وجود أو عدم وجود علاقة بين المتغيرين ومدى قوتها ونوعها. فإذا كانت تتوزع بشكل منتظم دل ذلك على وجود علاقة (يمكن استنتاجها)، أما إذا كانت النقط مبعثرة ولا تنتشر حسب نظام معين دل ذلك على عدم وجود علاقة بين المتغيرين أو أن العلاقة بينهما ضعيفة. والأشكال التالية تظهر بعض أشكال الانتشار المعروفة :

الشكل الأول :

إذا وقعت جميع النقاط على خط مستقيم، دل ذلك على أن العلاقة بينهما خطية وأنها ثابتة أو تامة. وهذه تمثل أقوى أنواع الارتباط بين المتغيرين " ارتباط تام ". فإذا كانت العلاقة طردية فإن " الارتباط طردي تام " كما في الشكل الأول (أ). ومثاله العلاقة بين الكمية المشتراة من سلعة والمبلغ المدفوع لشراء هذه الكمية. أما إذا كانت العلاقة عكسية (وجميع النقاط تقع على خط مستقيم واحد فإن "الارتباط عكسي تام" كما في الشكل الأول (ب). ومثال على ذلك العلاقة بين السرعة والزمن.



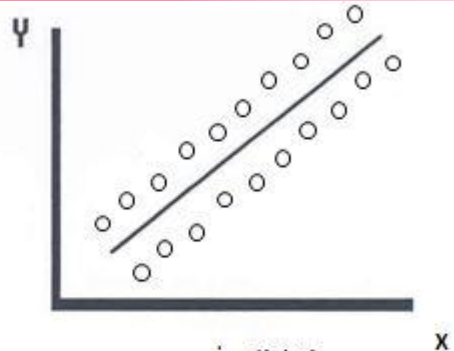
الشكل الأول (أ) ارتباط طردي تام (موجب)



الشكل الأول (ب) ارتباط عكسي تام (سالب)

الشكل الثاني :

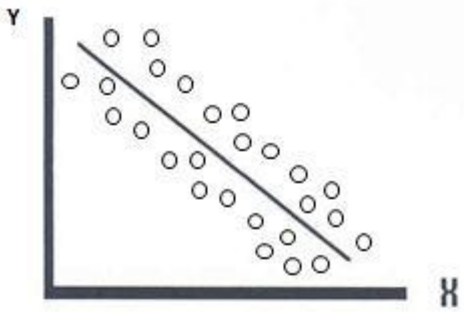
أما إذا كانت النقاط تأخذ شكل خط مستقيم ولكن لا تقع جميعها على الخط قيل أن العلاقة خطية (موجبة أو سالبة) كما في الشكل الثاني أ، ب.



الشكل الثاني (أ)

ارتباط موجب قوي

(ارتباط خطي طردي)



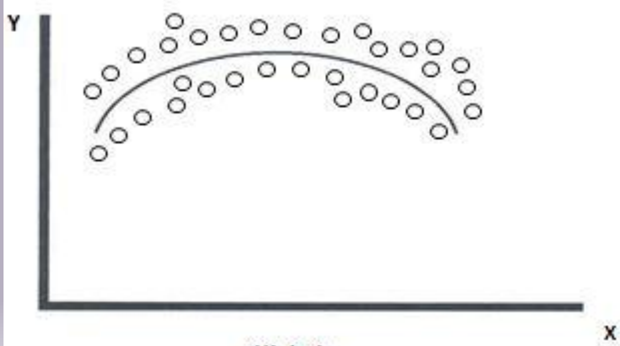
الشكل الثاني (ب)

ارتباط سالب قوي

(ارتباط خطي عكسي)

الشكل الثالث :

وإذا كانت العلاقة تأخذ شكل منحنى فإن الارتباط لا يكون خطياً "ارتباط غير خطي" Non Linear Correlation كما في الشكل الثالث :

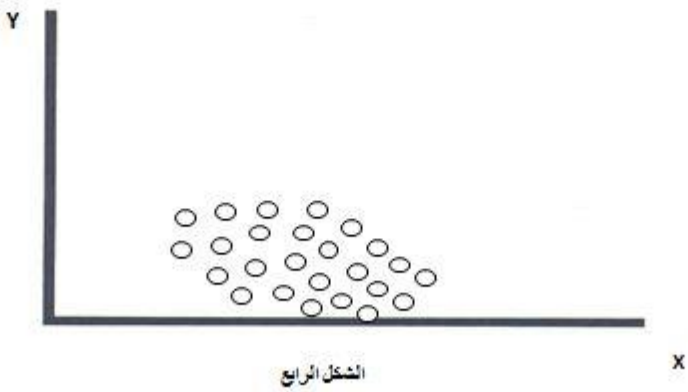


الشكل الثالث

(ارتباط غير خطي)

الشكل الرابع :

أما إذا كانت النقاط تتبعثر بدون نظام معين فإن ذلك يدل على عدم وجود علاقة بين المتغيرين (أو أن العلاقة بينهما ضعيفة جداً) كالعلاقة مثلاً بين دخل الشخص وطوله كما في الشكل الرابع :



الشكل الرابع

(لا توجد علاقة)

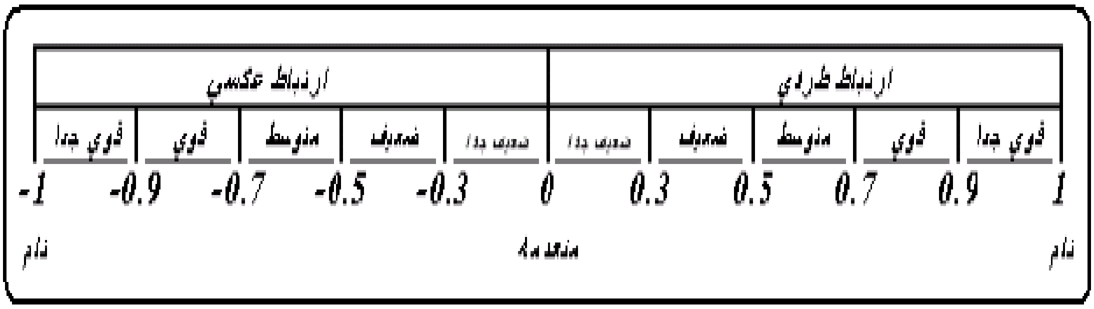
ثانيا: معامل الارتباط Correlation Coefficient :

يقاس الارتباط بين متغرين بمقياس إحصائي يسمى " معامل الارتباط " ويعكس هذا المقياس درجة أو قوة العلاقة بين المتغيرين واتجاه هذه العلاقة. وتنحصر قيمة معامل الارتباط بين + ١، - ١.

- فإذا كانت قيمة معامل الارتباط تساوي + ١ فمعنى ذلك أن الارتباط بين المتغيرين طردي تام، وهو أقوى أنواع الارتباط الطردي بين متغيرين.
- وإذا كانت قيمة معامل الارتباط تساوي – ١ فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام، وهو أقوى أنواع الارتباط العكسي بين متغيرين.
- وإذا كانت قيمة معامل الارتباط تساوي صفر، فمعنى ذلك أنه لا يوجد ارتباط بين المتغيرين.
- وكلما اقتربت قيمة معامل الارتباط من + ١ أو – ١ كلما كان الارتباط قوياً، وكلما اقترب من الصفر كلما كان الارتباط ضعيفاً.

الخلاصة :

أنه كلما كانت العلاقة قوية بين المتغيرين كلما اقترب معامل الارتباط من + ١ أو – ١ فإذا وصلت قيمة المعامل إلى + ١ أو – ١ كان الارتباط تاماً بين المتغيرين. وأنه كلما كانت العلاقة ضعيفة بين المتغيرين كلما اقترب معامل الارتباط من الصفر، فإذا وصلت قيمة المعامل إلى الصفر كان الارتباط منعدماً بين المتغيرين.
ومعنى ذلك أيضاً أنه لا يوجد ارتباط بين متغيرين تكون قيمة المعامل فيه أكبر من + ١ ولا أصغر من – ١.
ويمكن تمثيل قوة العلاقة بالشكل التالي:



معامل بيرسون للارتباط الخطى البسيط

Person's Correlation Coefficient

يعتبر معامل الارتباط الخطى البسيط لبيرسون Person's Correlation Coefficient والذى سنرمز له بالرمز r_p من أكثر الأدوات الإحصائية استخداما فى تحديد قوة العلاقة بين متغيرين كما يستعمل لتحديد مدى وجود علاقة خطية بين متغيرين.

وهناك أكثر من صيغة يمكن الاعتماد عليها فى حساب معامل الارتباط الخطى البسيط لبيرسون منها:

تستخدم هذه المعادلة من خلال إيجاد المتوسطات

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

وكذلك المعادلة الرياضية التالية والتي تعتبر اسهل وابسط:

تستخدم هذه المعادلة من خلال الدرجات الخام

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

وتتراوح قيمة معامل الارتباط بين الواحد الصحيح الموجب و الواحد الصحيح السالب أى أن قيمة معامل تكون كالتالي: $1 \geq r_p \geq -1$

والارتباط غالبا قيمته كسر أي اقل من الواحد الصحيح

ولتحديد نوع العلاقة نعتمد على اشارة معامل الارتباط فإذا كانت الإشارة:

- موجبة فإن العلاقة تكون طرديه
- سالبة فإن العلاقة تكون عكسية

ولتحديد قوة العلاقة نعتمد على قيمة معامل الارتباط فإذا كانت القيمة:

- من أكبر من صفر إلى أقل من 0.3 فتكون علاقة ضعيفة جدا
- من أكبر من 0.3 إلى أقل من 0.5 تكون علاقة ضعيفة
- من أكبر من 0.5 إلى أقل من 0.7 تكون علاقة متوسطة
- من أكبر من 0.7 إلى أقل من 0.9 تكون علاقة قوية
- من أكبر من 0.9 إلى أقل من 1.00 تكون علاقة قوية جدا
- الواحد الصحيح تكون علاقة تامة
- إما إذا كانت قيمة معامل الارتباط تساوى صفر فلا توجد علاقة خطيه او ارتباط بينهما أى يكون المتغيرين مستقلين عن بعضهما البعض وتكون العلاقة منعدمة

فمثلا إذا كانت قيمة معامل الارتباط r_p كالتالي فإن تفسيرة يكون:

تفسير معامل الارتباط	قيمة
أرتباط طردى قوى جدا	0.91
ارتباط عكسى قوى	-0.87
ارتباط عكسى ضعيف جدا	-0.21
ارتباط طردى ضعيف	0.43
ارتباط طردى تام	1 +
ارتباط عكسى متوسط	-0.51

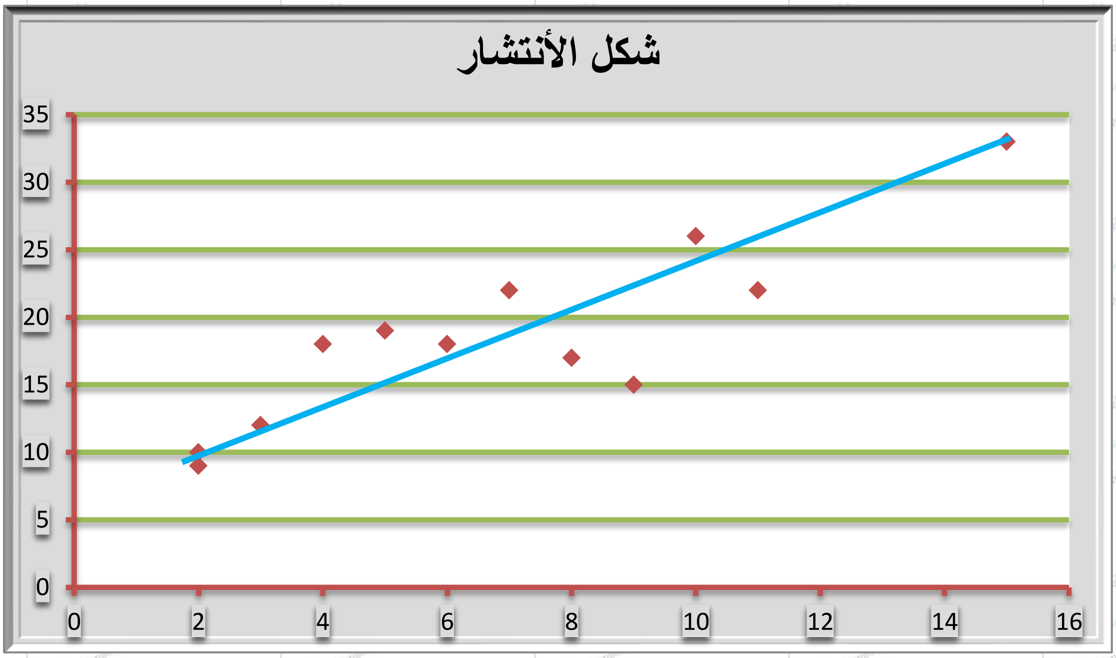
مثال: فيما يلى بيان بالمنفق على الاعلان والمبيعات لأحد المنتجات فكانت بالمليون ريال كمايلى:

المنفق على الاعلان	2	3	2	7	6	5	10	15	4	11	9	8
المبيعات	10	12	9	22	18	19	26	33	18	22	15	17

المطلوب:

ارسم شكل الانتشار يوضح العلاقة بين المنفق على الاعلان و المبيعات ؟

احسب معامل الارتباط الخطى البسيط (بيرسون)، مع التعليق



١- نستنتج بأن المؤشر يتجه تصاعدياً جهة اليمين مما يدل على وجود علاقة طرديه وقوية إلى حد ما ..

٢- نحسب معامل الارتباط ،، بداية نصلح الجدول ..

بيانات X المنفق	بيانات Y المبيعات	X Y	X^2	Y^2
2	10	20	4	100
3	12	36	9	144
2	9	18	4	81
7	22	154	49	484
6	18	108	36	324
5	19	95	25	361
10	26	260	100	676
15	33	495	225	1089
4	18	72	16	324
11	22	242	121	484
9	15	135	81	225
8	17	136	64	289
82	221	1771	734	4581

نطبق بالقانون :
$$r_p = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

$$r_P = \frac{12\,(1771) - [(82)(221)]}{\sqrt{12\,(734) - (82)^2}\,\sqrt{12\,(4581) - (221)^2}}$$
$$= \frac{21252 - 18122}{\sqrt{8808 - 6724}\,\sqrt{54972 - 48841}} = \frac{3130}{\sqrt{2084}\,\sqrt{6131}}$$
$$= \frac{3130}{(45,65)\,(78,30)} = \frac{3130}{3574,395} = 0,87567$$

التعليق على النتيجة : توجد علاقة طرديه بين المنفق على الإعلان وبين المبيعات

ومن أهم خصائص معامل الارتباط الخطى البسيط لبيرسون أنه لا يعتمد على قيم المتغيران نفسها عند حساب قيمته وإنما يعتمد على مقدار التباعد بين هذه القيم بعضها البعض.

لذلك لا يتأثر معامل الارتباط الخطى البسيط بأى عمليات جبرية يتم إجراءها على بيانات اى من المتغيرين أو أحدهما من جمع أو طرح أو ضرب أو قسمة .

مثال : فى بيانات المثال السابق إذا اكتشفت إدارة الشركة أن البيانات تم تجميعها وحسابها بطريقة خاطئة حيث يجب إضافة ٥ مليون ريال إلى جميع قيم المنفق على الإعلان. كما أن المبيعات يجب مضاعفة قيمتها لجميع القيم.

المطلوب : أحسب معامل الارتباط فى هذه الحالة بين المنفق على الإعلان والمبيعات.

أولا نصلح الجدول نفس السابق ولكن نضيف 5 لجميع القيم المنفق .. أما المبيعات مضاعفة القيم ..

بيانات X المنفق	بيانات Y المبيعات	X Y	X^2	Y^2
7	20	140	49	400
8	24	192	64	576
7	18	126	49	324
12	44	528	144	1936
11	63	396	121	1296
10	38	380	100	1444
15	52	780	225	2704
20	66	1320	400	4356
9	36	324	81	1296
16	44	704	256	1936
14	30	420	196	900
13	34	442	169	1156
142	442	5752	1854	18324

نبدأ نطبق بالقانون :

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

$$r_p = \frac{12\,(5752) - (142 \times 442)}{\sqrt{12\,(1854) - (142)^2}\,\sqrt{12\,(18324) - (442)^2}}$$

$$= \frac{69024 - 62764}{\sqrt{22248 - 20164}\,\sqrt{219888 - 195364}} = \frac{6260}{\sqrt{2084}\,\sqrt{24524}}$$

$$= \frac{6260}{45,65 \times 156,60} = \frac{6260}{7148,79} = 0,87567$$

معامل التحديد Determination Coefficient

وهو مربع معامل الارتباط لذلك يرمز له بالرمز R^2 أي R^2 أو R-Square و هو يشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغير فى المتغير التابع .

فمثلا:

نجد أن المنفق على الاعلان يفسر نسبة (0.8756^2) أى 76.675 % من التغير فى قيمة المبيعات بينما 23.32 % من التغير فى المبيعات ترجع إلى عوامل أخرى منها الخطاء العشوائى .

المــحاضــرة الرابـعة عشــر

تحليل الإرتباط – ٢

معامل ارتباط الرتب لسبيرمان Spearman's Rank Correlation Coefficient

معامل الارتباط لبيرسون r_s لا يمكن استخدامه فى حساب قوة العلاقة بين متغيرين الا اذا كانت البيانات المتوافرة عنهما فى صورة كمية فقط ، أما اذا كانت البيانات فى صورة وصفية فلا يمكن تطبيق معامل ارتباط بيرسون وحساب الارتباط بين المتغيرين محل الدراسة.

أما في حالة المتغيرات الوصفية فنستخدم معامل ارتباط الرتب لـ سبيرمان ، والذى يتم استخدامه فى قياس الارتباط خاصة فى حالة البيانات الوصفية الترتيبية مثل تقديرات الطلاب (ممتاز – جيد جدآ – جيد – مقبول – ضعيف) وكذلك قوة المركز المالى (جيد - متوسط - ضعيف) ودرجة الموافقة على الرأى فى اسئلة الإستبانه (موافق تماما – موافق – محايد – غير موافق – غير موافق على الاطلاق).

ويتم حساب معامل الارتباط الرتب لسبيرمان r_s باستخدام المعادلة التالية:

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

حيث أن:

d الفرق بين رتبة المتغير

n عدد المشاهدات

ملاحظات يجب مراعاتها عند ترتيب المتغيرات:

- **يتم ترتيب قيم مشاهدات المتغير x وتسمى القيم الترتيبية للمتغير x "رتب x" وكذلك الامر للمتغير y تسمى بـ "رتب y" . والترتيب يكون تصاعديا أو تنازليا ولكن أهم شىء هو اذا كان ترتيب x تصاعدى لابد ان يكون ترتيب y تصاعدى ايضا والعكس صحيح.**
- **فى حالة الترتيب التصاعدى مثلا يتم اعطاء أقل قيمة الرتبة ١ والقيمة التى هى أكبر منها الرتبة ٢ وهكذا .**
- **فى حالة تكرار أو تساوى بعض القيم لأي متغير تعطى كل منهم رتبة كما لو كانت القيم غير متساوية ثم نحسب الوسط الحسابى (مجموع الرتب ÷ عددها) لتلك الرتب ويعطى الوسط الحسابى كرتبة تلك القيم المتساوية .**

مثال: فيما يلى بيان بالمنفق على الاعلان والمبيعات لأحد المنتجات فكانت بالمليون ريال كمايلى:

المنفق على الاعلان	2	3	2	7	6	5	10	15	4	11	9	8
المبيعات	10	12	9	22	18	19	26	33	18	22	15	17

المطلوب:

أحسب معامل الارتباط لسبيرمان بين المنفق على الاعلان و المبيعات ؟

ملاحظه : تم تعديل الحل وفقاً للمثال لأن الدكتور أخطأ في أحد الأعداد واستبدله بعدد آخر . فغير مسار الحل جميعه ..

X	Y	رتبت X	رتبت Y	d	d^2
2 ← ١	10 ← ٢	1,5	2	-0,5	0,25
3 ← ٣	12 ← ٣	3	3	0	0
2 ← ٢	9 ← ١	1,5	1	+0,5	0,25
7 ← ٧	22 ← ٩	7	0,86	6,14	37,69
6 ← ٦	18 ← ٦	6	0,72	5,28	27,88
5 ← ٥	19 ← ٨	5	8	-3	9
10 ← ١٠	26 ← ١١	10	11	-1	1
15 ← ١٢	33 ← ١٢	12	12	0	0
4 ← ٤	18 ← ٧	4	0,72	3,28	10,76
11 ← ١١	22 ← ١٠	11	0,86	10,14	102,81
9 ← ٩	15 ← ٤	9	4	5	25
8 ← ٨	17 ← ٥	8	5	+3	9
					223,69

ملاحظه : الأعداد اللي باللون الأحمر طريقة تحديدنا للرتب ..

الفئات المتشابهه في المنفق والمبيعات طبقنا عليها هذا القانون مجموع الرتب ÷ عددها

مثلاً 3 ÷ 2 = 1,5 الناتج نضعه ترتيب لتلك الأعداد المتشابهه ..

وكذلك 19 ÷ 22 = 0,86

وأيضاً 13 ÷ 18 = 0,72

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

$$r_s = 1 - \frac{6\,(223,69)}{12\,(12^2-1)} = 1 - \frac{1342,14}{1716} = 0,782$$

قيمة عامل ارتباط سبيرمان (لو لا حظتوا الناتج لا يختلف عن الناتج اللي حسبه الدكتور)

مثال: البيانات التالية تمثل التقديرات التى حصل عليها عشر طلاب فى مقررى المحاسبة والقانون:

المحاسبة	ممتاز	جيدجدا	جيد	مقبول	ضعيف	جيد	مقبول	جيدجدا	جيد	مقبول
القانون	جيد	جيد	مقبول	جيد	جيدجدا	جيدجدا	ممتاز	مقبول	جيد	جيدجدا

المطلوب:

أحسب معامل الارتباط المناسب.

نرسم الجدول ونوزع الدرجات من 0 إلى 10 (الحل من الكتاب)

المحاسبة	القانون	رتب المحاسبة	رتب القانون	d	d^2
ممتاز	جيد	10	4,5	5,5	30,25
جيد جيدا	جيد	8,5	4,5	4	16
جيد	مقبول	6	1,5	4,5	20,25
مقبول	جيد	3	4,5	-1,5	2,25
ضعيف	جيد جدا	1	8	-7	49
جيد	جيد جدا	6	8	-2	4
مقبول	ممتاز	3	10	-7	49
جيد جدا	مقبول	8,5	1,5	7	49
جيد	جيد	6	4,5	1,5	2,25
مقبول	جيد جدا	3	8	-5	25
				0	247

- نلاحظ عند ترتيب تقديرات مقرر المحاسبه أن التقدير مقبول اخذ الرتب ٢ و ٣ و ٤ لذلك تم وضع ٣ أمام التقدير مقبول في مقرر المحاسبة .

- كما أن تقدير جيد في مقرر القانون أخذ الرتب ٣ و ٤ و ٥ و ٦ لذلك تم وضع الرتبه ٤.٥ امام التقدير جيد في مقرر القانون .

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

$$r_s = 1 - \frac{6\,(247)}{10\,(10^2-1)} = 1 - \frac{1482}{10\,(100-1)} = 1 - \frac{1482}{990} = 1 - 1,4969$$
$$= -0,4969$$

معامل الاقتران Conjunction Coefficient

ويستخدم معامل الاقتران فى حساب العلاقة الإرتباطية بين المتغيرات الوصفية التى ليس فى طبيعتها صفة الترتيب أى الوصفية الأسمية التى يكون لها زوج من الصفات مثل:

النوع (ذكر – انثى) ، والحالة التعليمية (متعلم - غير متعلم)

وعلى ذلك إذا كان لدينا متغيران لدي كلاً منهما زوج من الصفات فيكون جدول تكرارات الصفات المشتركة بينهما على الصورة التالية:

Y / X	الصفة الأولى لـ y	الصفة الثانية لـ y
الصفة الأولى لـ x	A	B
الصفة الثانية لـ x	C	D

حيث أن A , B , C , D تشير إلى التكرارات المشتركة بين صفات المتغيرين، ويمكن حساب معامل الاقتران فى هذه الحالة كما يلي:

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

مثال: فى دراسة اجريت لمعرفة هل هناك علاقة بين العمل والتعليم تم سوال ٢٠٠ شخص سؤالين هما:

هل انت متعلم ؟ نعم لا

هل انت ملتحق بأى عمل ؟ نعم لا

وبتجميع الاجابات تم عمل جدول الاقتران التالي:

العمل \ التعليم	متعلم	أمي
يعمل	113	23
لا يعمل	49	15

المطلوب:

أحسب معامل الاقتران ؟

نطبق بالقانون : $r_C = \dfrac{AD - BC}{AD + BC}$

$$r_c = \frac{[(113)\ (15)] - [(23)\ (49)]}{[(113)\ (15)] + [(23)\ (49)]} = \frac{1695 - 1127}{1695 + 1127} = \frac{568}{2822} = +0,20$$

يظهر لنا في حل المثال أن العلاقة ضعيفة جداً بين العمل والتعليم ..

معامل التوافق Concordance coefficient

ويستخدم معامل التوافق لحساب الارتباط بين المتغيرات الوصفية الاسمية والتى يكون لصفاتها قيم أكثر من ٢، مثل الحالة الاجتماعية (اعزب - متزوج – متزوج ويعول – أرمل – مطلق)

وحتى يمكن حسابه يتم إعداد الجدول المزدوج بين صفات المتغيرين ومنه يتضح لنا التكرارات المشتركة بين الصفات التى نعتمد عليها فى حساب مقدار يطلق علية " M "

ويتم حساب معامل التوافق من خلال المعادلة التالية:

$$M = \sum \frac{(f_{ij})^2}{f_{i.} f_{.j}}$$

حيث أن:

f_{ij}	التكرار المشرك بين الصفة i والصفة j
$f_{i.}$	مجموع صف الصفة i
$f_{.j}$	مجموع عمود الصفة j

أي يتم إيجاد: مربع تكرار كل خلية مشتركة

مجموع الصف × مجموع العمود

ثم نجمعهم كلهم

وعلى ذلك يتم حساب معامل التوافق كما يلي : $r_T = \sqrt{\dfrac{M-1}{M}}$

مثال : أوجد معامل التوافق بين تخصص الطالب ودرجة الرضا عن الدراسة بالكلية الملتحق بها إذا كانت البيانات كما يلي :

التخصص / الرضا	لغة عربية	جغرافيا	تربية خاصة	المجموع
عالي	30	15	45	90
متوسط	20	30	20	70
منخفض	10	5	5	20
المجموع	60	50	70	180

نطبق القانون : $M = \sum \frac{(f_{ij})^2}{f_{i.} f_{.j}}$

$$M = \frac{30^2}{60 \times 90} + \frac{15^2}{50 \times 90} + \frac{45^2}{70 \times 90} + \frac{20^2}{60 \times 70} + \frac{30^2}{50 \times 70} + \frac{20^2}{70 \times 70} + \frac{10^2}{60 \times 20} + \frac{5^2}{50 \times 20} + \frac{5^2}{70 \times 20}$$

$$M = \frac{900}{5400} + \frac{225}{4500} + \frac{2025}{6300} + \frac{400}{4200} + \frac{900}{3500} + \frac{400}{4900} + \frac{100}{1200} + \frac{25}{1000} + \frac{25}{1400}$$

$$M = 0,166 + 0,05 + 0,32 + 0,095 + 0,257 + 0,081 + 0,083 + 0,025 + 0,017$$

$$M = 1,094$$

أوجدنا قيمة M نبدأ نعوض بالقانون : $r_T = \sqrt{\frac{M-1}{M}}$

$$r_T = \sqrt{\frac{1,094 - 1}{1,094}} = \sqrt{\frac{0,094}{1,094}} = \sqrt{0,0859} = +\ 0,293$$

نجد ان هناك علاقة ضعيفة إلى حد ما بين التخصص و الرضا عن الدراسة ..

تـمــت بـحـمـد الله

نعتذر عن التقصير والأخطاء ،، وأي ملاحظه نسعد بتقبلها ،، فهذا جهد شخصي قد لا يخلو من الخطأ ..

نتمنى لكم التوفيق و الاستفادة من هذه الملزمه ،، و لا نريد منكم غير الدعاء ..

شرح وتنسيق .. oOo al maznei oOo .. حلم المشاعر ..