

الافواج 4 و5

المقياس الإحصاء الوصفي

المحاضرة 3

1 - الانحراف المعياري :

مما لا شك فيه أن الانحراف المعياري يعتبر من أكثر مقاييس التشتت أهمية لأنه مفهوم جبري محدد بدقة ، وهو أقوى مقاييس التشتت حساسية وأكثرها شيوعاً . والفكرة الأساسية لهذا المقياس هي أنه بدلاً من إهمال الإشارات الجبرية عند حساب الانحراف المتوسط، نحاول التخلص من هذه الإشارات بطريقة أخرى أكثر صلاحية، وذلك بتربيع الانحرافات.

ويعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي . ومربع الانحراف المعياري يسمى **التباين** ، ويمكن حسابه بعدة طرق منها :

1.1 - الانحراف المعياري لبيانات غير مبوبة :

إذا رمزنا بـ x_i لانحراف القيمة الأولى X_1 عن الوسط الحسابي \bar{X} وللانحراف المعياري بالرمز σ و استناداً إلى تعريف الانحراف المعياري نكتب :

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}} = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}}$$

مثال : القيم التالية تمثل علامات خمسة طلاب في إمتحان المحاسبة : 10 ، 8 ، 14 ، 7 ، 16 . المطلوب : حساب الانحراف المعياري.

الحل : الجدول التالي يبين الحسابات اللازمة للحل .

X^2	$(X_i - \bar{X})^2$	$(X_i - \bar{X})$	X_i
100	1	1 -	10
64	9	3 -	8
196	9	3	14
49	16	4 -	7
256	25	5	16
$\Sigma X^2=665$	$\Sigma(X_i - \bar{X})^2= 60$	$\Sigma(X_i - \bar{X})= 0$	$\Sigma X_i=55$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}} = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N}} = \sqrt{\frac{60}{5}} = \sqrt{12} = 3,46$$

2.1- حساب الانحراف المعياري من بيانات مبوبة :

لحساب الانحراف المعياري لبيانات مبوبة يكفي أن نضرب مراكز الفئات بتكراراتها ونعوض في الصيغ السابقة التي طبقناها في البيانات غير المبوبة . فإذا كانت $X_1, X_2, X_3, \dots, X_K$ تمثل مراكز الفئات و $f_1, f_2, f_3, \dots, f_K$ تكراراتها على الترتيب ، يمكن حساب الانحراف المعياري من توزيع تكراري بالطرق التالية :

$$\sigma = \sqrt{\frac{\sum_{i=1}^K f_i (X_i - \bar{X})^2}{\sum_{i=1}^K f_i}} = \sqrt{\frac{\sum_{i=1}^K f_i X_i^2}{\sum_{i=1}^K f_i}} \quad \text{الطريقة المباشرة :}$$

مثال : أحسب الانحراف المعياري من الجدول.

الفئات	4 - 0	8 - 4	12 - 8	16 - 12	20 - 16
التكرارات	8	10	14	5	3

الحل : الجدول التالي يبين الحسابات اللازمة.

الفئات	f_i	X_i	$f_i X_i$	$(X_i - \bar{X})$	$f_i (X_i - \bar{X})^2$
4-0	8	2	16	6.5-	338
8-4	10	6	60	2.5-	62.5
12-8	14	10	140	1.5	31.5
16-12	5	14	70	5.5	151.25
20-16	3	18	54	9.5	270.75
المجموع	40	-	340	-	854

= الطريقة المباشرة : $\sigma = \sqrt{\frac{\sum_{i=1}^K f_i (X_i - \bar{X})^2}{\sum_{i=1}^K f_i}} = \sqrt{\frac{\sum_{i=1}^K f_i^2 X_i^2}{\sum_{i=1}^K f_i}}$ ، نحسب الوسط الحسابي :

ومنه : $\bar{X} = \frac{\sum fX}{\sum f} = \frac{340}{40} = 8,5$ $\sigma = \sqrt{\frac{\sum_{i=1}^K f_i (X_i - \bar{X})^2}{\sum_{i=1}^K f_i}} = \sqrt{\frac{854}{40}} = \sqrt{21,35} = 4,62$

التمرين 01

إذا كانت لديك القيم التالية : 16 ، 18 ، 16 ، 20 ، 11 ، 25 ، 14 ، أحسب كل من :

- المدى العام .
- الانحراف الربيعي .
- الانحراف المتوسط
- الانحراف المعياري

التمرين 02

الجدول التالي يبين توزيع الأراضي (بالهكتار) على 50 فلاحاً .

فئات الأراضي	40 - 30	50 - 40	60 - 50	70 - 60	80 - 70	90 - 80	100 - 90	المجموع
عدد الفلاحين	04	06	08	12	09	07	04	50

أحسب ما يلي :

- 1 - الانحراف الربيعي .
- 2 - الانحراف المتوسط .
- 3 - الانحراف المعياري.

الافواج 4 و5

المقياس الإحصاء الوصفي

المحاضرة 4

الإرتباط الخطي

1, مفهوم الارتباط أشكاله ومقاييسه:

لقد قمنا حتى الآن في الفصول السابقة، بدراسة متغير واحد أو ظاهرة واحدة، من حيث إيجاد متوسطات هذا المتغير وكذلك تشتت تلك القيم عن الوسط المحسوب.

وفي هذا الفصل سوف نتطرق إلى دراسة العلاقة بين متغيرين أو ظاهرتين ، ومقدار التغير في إحدهما المصحوب بتغير في الآخر ، ونوع هذه العلاقة ، هل هي طردية ، بمعنى أنه كلما زاد المتغير المستقل زاد معه المتغير التابع ، وكلما نقص المتغير المستقل نقص معه المتغير التابع . أم هل هي عكسية بمعنى إذا زاد المتغير المستقل نقص المتغير التابع والعكس بالعكس . وتختلف ماهية العلاقة بين متغيرين، فقد تكون بسيطة أي مستقيمة وقد تكون غير مستقيمة. وسنبحث أبسط أشكال العلاقة بين متغيرين وهي العلاقة المستقيمة.

2- معامل الارتباط :

ويستخدم هذا المقياس لمعرفة درجة أو شدة العلاقة واتجاهها بين المتغيرين أو الظاهرتين

المدرستين . كما يمكننا التمييز بين أنواع الارتباط الرئيسية التالية :

1 - الارتباط البسيط المستقيم : ويتحقق هذا النوع من الارتباط إذا توفر الشرطين التاليين :

أ - أن تكون العلاقة موضوع الدراسة مقتصرة على ظاهرتين أو متغيرين اثنين فقط .

ب - أن تكون العلاقة موضوع الدراسة مما يمكن تمثيلها بخط مستقيم أحسن تمثيل .

2 - الارتباط البسيط غير المستقيم: ويتحقق هذا النوع من الارتباط إذا توفر الشرطين التاليين :

أ - أن تكون العلاقة موضوع الدراسة مقتصرة على ظاهرتين أو متغيرين اثنين فقط .

ب - أن تكون العلاقة موضوع الدراسة مما لا يمكن تمثيلها بخط مستقيم أي خط ذو إلتواء أو أكثر.

3 - الأرتبط المتعدد : ويتحقق هذا النوع من الارتباط عند القيام بدراسة العلاقة بين ظاهرة وظاهرتين أو أكثر في آن واحد ، حيث يعتبر في هذه الحالة أحد المتغيرات هو المتغير التابع ، وباقي المتغيرات الأخرى هي المتغيرات المستقلة .

4 - الأرتباط الجزئي : يتحقق هذا الشكل من الارتباط عند القيام بدراسة الارتباط بين ظاهرة وظاهرتين أو أكثر منفردة ، أي الارتباط بين ظاهرة وأخرى على إفتراض أن الظواهر الأخرى تعامل معاملة الثوابت .

3 . حساب معامل التحديد والارتباط :

تسمى النسبة بين التباين المفسر والتباين الكلي **بمعامل التحديد** ، وهو مقياس يبين أو يقيس مدى أو نسبة تأثير المتغير المستقل على المتغير التابع . إذا رمزنا لمعامل التحديد بـ r^2 نجد أن :

$$r^2 = \frac{\frac{\sum(\hat{Y}_i - \bar{Y})^2}{N}}{\frac{\sum(Y_i - \bar{Y})^2}{N}} \Rightarrow r^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

إذا كان التباين المفسر يساوي صفر ، فهذا يعني أن التباين الكلي جميعه غير مفسر والنسبة السابقة تصبح تساوي صفر . أما إذا كان التباين غير المفسر يساوي صفر ، أي بمعنى أن التباين الكلي جميعه مفسر ، فإن النسبة السابقة تصبح تساوي الواحد . وفي هذه الحالة نقول أن نسبة تأثير المتغير المستقل على المتغير التابع هي مئة بالمئة . أو بتعبير آخر نقول أن التغيرات التي تحصل في المتغير التابع تعود كلها إلى التغيرات التي تحصل في المتغير المستقل .
وخلاف الحالتين السابقتين فإن قيمة معامل التحديد تتراوح ما بين الصفر والواحد .

يسمى الجذر التربيعي لمعامل التحديد (r^2) **بمعامل الارتباط** ، ويعرف بمعامل ارتباط

بيرسون ، وصيغته هي :

$$r = \pm \sqrt{\frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y - \bar{Y})^2}}$$

وتتراوح قيمة الجذر السابق ما بين (- 1 ، + 1) . أما إشارة الجذر فهي من إشارة معامل الانحدار (α_1) .

ويجب التأكيد على أن قيمة (r) المحسوبة في أية حالة تقيس درجة أو شدة العلاقة بالنسبة إلى نوع المعادلة المفترضة ، فإذا افترضنا معادلة خطية ، ونتج عن العلاقة (1) قيمة لـ r تقترب من الصفر (أو تساوي صفر) فهذا يدل على أنه لا توجد تقريباً علاقة خطية بين المتغيرين . ولكن هذا لا يعني أنه لا توجد علاقة بين المتغيرين على الإطلاق ، فقد تكون هناك بالفعل علاقة كبيرة غير خطية بين المتغيرين .

إذاً قرارنا على وجود علاقة أو عدم وجودها يتوقف على نوع المعادلة المفترضة فقط ، ومن هنا فإن مصطلح معامل الارتباط يستخدم ليعني الارتباط الخطي ، ما لم يذكر خلاف ذلك .

إن وجود معامل ارتباط مرتفع (يقترب من +1 أو -1) لا يعني دائماً وجود علاقة تبعية مباشرة بين المتغيرات . فقد يكون هناك معامل ارتباط مرتفع بين عدد الكتب المنشورة في كل سنة وعدد مباريات الكرة الملعبوبة في كل سنة ، فمن الواضح أنه لا توجد علاقة سببية بين هذين المتغيرين . في مثل هذه الحالات يشار إلى العلاقة بين هذين المتغيرين بأنها ذات ارتباط لا معنى له ، ويقال عنه ارتباط زائف . يمكن أن نستخرج عدة طرق لحساب معامل الارتباط

حساب معامل ارتباط بيرسون

يمكن حساب معامل الارتباط بالطريقة المباشرة ، كما يلي :

$$r = \frac{N\sum XY - \sum X.\sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

مثال :

من الجدول أحسب معامل الارتباط وأشرحه ، ثم أحسب معامل التحديد وأشرحه .
الجدول التالي يبين الدخل (X) والإنفاق (Y) لسبعة أسر بآلاف الدينارات الجزائرية :

الجدول رقم (01)

9	11	14	10	6	8	12	الدخل (X)
8	9	12	10	7	7	10	الإنفاق (Y)

الحل : من الجدول أدناه :

Y ²	XY	X ²	Y	X
100	120	144	10	12
49	56	64	7	8
49	42	36	7	6
100	100	100	10	10
144	168	196	12	14
81	99	121	9	11
64	72	81	8	9
ΣY^2	ΣXY	ΣX^2	ΣY	ΣX
=	=	=	=	=
587	657	742	63	70

$$r = \frac{N \Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{[N \Sigma X^2 - (\Sigma X)^2][N \Sigma Y^2 - (\Sigma Y)^2]}}$$

من العلاقة :

$$r = \frac{7(657) - 70(63)}{\sqrt{[7(742) - (70)^2][7(587) - (63)^2]}} = \frac{189}{202.19} \approx 0.93$$

الشرح :

- بالنسبة لمعامل الارتباط هناك علاقة طردية قوية بين الدخل والإنفاق .

- بالنسبة لمعامل التحديد يمكن القول أن الإنفاق يعتمد على الدخل بنسبة 86.49 % ، أما النسبة الباقية وهي 13.51 % فهي تعود إلى عوامل أخرى (مداخل أخرى) نجهلها

معامل ارتباط الرتب :

يستخدم هذا المعامل لدراسة الارتباط بين البيانات النوعية أي تلك التي لا يمكن قياسها كميًا ، وتعتمد هذه الطريقة على إعطاء المتغيرات رتبًا لتحل محل القياس العددي ، فإذا رتبنا قيم المتغير X ترتيبًا تصاعديًا ووجدنا أن قيم المتغير Y المناظرة لها مرتبة ترتيبًا تصاعديًا أيضًا نستنتج وجود ارتباط طردي تام بين المتغيرين X و Y . أما إذا رتبنا قيم المتغير X ترتيبًا تصاعديًا

ووجدنا أن قيم المتغير Y المناظرة لها مرتبة ترتيباً تنازلياً نستنتج وجود ارتباط عكسي تام بين المتغيرين X و Y ، غير أن هذا النوع من الارتباط التام ، نادرًا ما يصادفنا في الدراسات الاجتماعية والاقتصادية ، وفي الحالات الأخرى يتراوح معامل الارتباط كما رأينا في معامل بيرسون بين (1+) و (1-).

إذا لقياس الارتباط بين المتغيرين X و Y نرتب كل منهما حسب أفضليته ثم نحسب الفرق (D) بين كل رتبتين متقابلتين ، فنجد أن مجموع الفروق يساوي صفر ($\sum D = 0$) ، وبحساب مربعات هذه الفروق (D^2) يمكن إيجاد معامل ارتباط الرتب باستخدام العلاقة :

$$r_s = 1 - \frac{6\sum D^2}{N(N^2 - 1)}$$

حيث : r_s هو معامل ارتباط سبيرمان

D الفرق بين رتب القيم

N عدد القيم

مثال 1 :

الجدول التالي يبين تقديرات خمس طلبة في المقياسين X و Y . أحسب معامل ارتباط الرتب .

الجدول رقم (06)

تقديرات المقياس X	جيد	ممتاز	مقبول	ضعيف	جيد جدًا
تقديرات المقياس Y	جيد جدًا	جيد	ضعيف	ضعيف جدًا	ممتاز

الحل :

لدينا أحسن تقدير بالنسبة للمقياس X هو ممتاز ويأخذ الرتبة الأولى أي رقم 1 ، ثم يليه من حيث الأهمية التقدير جيد جدًا الذي يأخذ الرتبة الثانية أي رقم 2 ، ثم يأتي التقدير جيد الذي يأخذ الرقم 3 ، ثم التقدير مقبول ويأخذ الرقم 4 ، وأخيرًا التقدير ضعيف الذي يأخذ الرقم 5 .

أما بالنسبة لتقديرات المقياس Y فنجد أن أحسن تقدير هو ممتاز ويأخذ الرقم 1 ، ثم يليه جيد جدًا برقم 2 ، وجيد يأخذ الرقم 3 ، وضعيف يأخذ الرقم 4 ، وأخيرًا التقدير ضعيف جدًا ويأخذ الرقم 5.

كما يمكن أن يكون الترتيب تصاعدياً من حيث الأهمية نبدأ بأقل تقدير فنعطيه رقم 1 ، ثم الذي أحسن منه وهكذا حتى نصل إلى أحسن تقدير ضمن التقديرات ونعطيه آخر رقم في الترتيب

، فالترتيب قد يكون تصاعديًا أو تنازليًا . والجدول التالي يبين الرتب المقابلة للتقديرات وكذلك الفروق بين هذه الرتب .

D ²	D	رتب Y	رتب X	تقديرات Y	تقديرات X
1	1	2	3	جيد جدًا	جيد
4	2-	3	1	جيد	ممتاز
0	0	4	4	ضعيف	مقبول
0	0	5	5	ضعيف جدًا	ضعيف
1	1	1	2	ممتاز	جيد جدًا
$\sum D^2 = 6$	$\sum D = 0$	المجموع			

$$r_s = 1 - \frac{6\sum D^2}{N(N^2 - 1)} = 1 - \frac{6(6)}{5(25 - 1)} = 1 - 0.3 = 0.7$$

بتطبيق قانون سبيرمان نجد :

مثال 2 :

الجدول التالي يبين علامات ثمانية (08) طلاب في امتحان مقياسي الرياضيات (Y) والإحصاء (X) .

المطلوب :

1 - حساب معامل الارتباط لبيرسون .

2 - حساب معامل الارتباط لسبيرمان (معامل ارتباط الرتب) .

1 - معامل ارتباط بيرسون : الجدول التالي يبين الحسابات اللازمة .

YX	X ²	Y ²	X	Y
120	100	144	10	12
208	169	256	13	16
168	144	196	12	14
80	64	100	8	10
165	121	225	11	15

144	144	144	12	12
182	196	169	14	13
96	64	144	8	12
$\Sigma YX = 1163$	$\Sigma X^2 = 1002$	$\Sigma Y^2 = 1378$	$\Sigma X = 88$	$\Sigma Y = 104$

$$r = \frac{N \Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{[N \Sigma X^2 - (\Sigma X)^2][N \Sigma Y^2 - (\Sigma Y)^2]}} = \frac{8(1163) - 88(104)}{\sqrt{[8(1002) - (88)^2][8(1378) - (104)^2]}}$$

2 - معامل ارتباط سبيرمان : الجدول التالي يبين رتب القيم .

D ²	D	رتب X	رتب Y	X	Y
0	0	6	6	10	12
1	1-	2	1	13	16
0.25	0.5-	3.5	3	12	14
0.25	0.5	7.5	8	8	10
9	3-	5	2	11	15
6.25	2.5	3.5	6	12	12
9	3	1	4	14	13
2.25	1.5-	7.5	6	8	12
$\Sigma D^2 = 28$	$\Sigma D = 0$	المجموع			

$$r_s = 1 - \frac{6 \Sigma D^2}{N(N^2 - 1)} = 1 - \frac{6(28)}{8(64 - 1)} = 1 - 0.33 \approx 0.67$$

مثال 3:

إذا كانت لديك التقديرات التالية لعشرة طلاب في مقياسي المحاسبة (X) والرياضيات (Y) أحسب معامل ارتباط سبيرمان (معامل ارتباط الرتب) .

تقديرات الرياضيات (Y)	تقديرات المحاسبة (X)
جيد جداً	جيد

مقبول	مقبول
جيد جدًا	ممتاز
جيد	مقبول
مقبول	ضعيف
ممتاز	جيد جدًا
مقبول	جيد
مقبول	ضعيف
جيد	مقبول
مقبول	ضعيف جدًا

الحل :

نرتب كل من تقديرات المحاسبة والرياضيات ترتيبًا تنازليًا كما يلي :

ترتيب تقديرات الرياضيات	ترتيب تقديرات المحاسبة
1 ممتاز	1 ممتاز
2.5 = $2 \div (3+2)$ { 2 جيد جدًا 3 جيد جدًا	2 جيد جدًا
4.5 = $2 \div (5+4)$ { 4 جيد 5 جيد	3 جيد 4 جيد
8 = $5 \div (10+9+8+7+6)$ { 6 مقبول 7 مقبول 8 مقبول 9 مقبول 10 مقبول	5 مقبول 6 مقبول 7 مقبول 8 ضعيف 9 ضعيف
	10 ضعيف جدًا

فإذا وضعنا أمام كل تقدير الرتبة التي نالها في الشكل أعلاه ، نحصل على الجدول التالي :

D ²	D	رتب Y	رتب X	تقدير Y	تقدير X
1	1	2.5	3.5	جيد جدًا	جيد
4	2-	8	6	مقبول	مقبول
2.25	1.5-	2.5	1	جيد جدًا	ممتاز
2.25	1.5	4.5	6	جيد	مقبول
0.25	0.5	8	8.5	مقبول	ضعيف
1	1	1	2	ممتاز	جيد جدًا
20.25	4.5-	8	3.5	مقبول	جيد
0.25	0.5	8	8.5	مقبول	ضعيف
2.25	1.5	4.5	6	جيد	مقبول
4	2	8	10	مقبول	ضعيف جدًا
$\sum D^2 = 37.5$	$\sum D = 0$	المجموع			

$$r_s = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} = 1 - \frac{6(37.5)}{10(100 - 1)} = 1 - 0.23 = 0.77$$

معامل ارتباط بيرسون : 0.77 = 1 - 0.23 = 1 - $\frac{6(37.5)}{10(100 - 1)}$

التمرين 01

لدينا المتغيرين X و Y حيث أن X يمثل الدخل العائلي الشهري بمئات الدنانير و Y يمثل النفقات الغذائية الشهرية بمئات الدنانير . فإذا كان لدينا 10 عائلات حسب الجدول التالي :

ترتيب العائلة	1	2	3	4	5	6	7	8	9	10
X	38	22	7	15	33	17	19	24	16	12
Y	22	14	7	12	20	13	12	17	12	10

المطلوب : (1) أحسب معامل الارتباط بين X و Y

التمرين 02

أحسب معامل الارتباط بطريقتي (بيرسون) و (سبيرمان) بين قيم X و Y من البيانات الآتية :

13	12	16	11	15	14	13	X
17	15	14	12	15	16	13	Y

