مقيـــــــاس: إحصـــــــاء 1

المحور السابع: الارتباط والانحدار

تمهيد:

كثيرا ما تقابلنا مواقف في الحياة العملية تتضمن متغيرين (ظاهرتين) أو أكثر ويكون المطلوب معرفة ما إذا كان هناك علاقة بين هذه المتغيرات وما هو شكل هذه العلاقة، وأيضا كيفية التنبؤ بأحد هذين المتغيرين في حالة معرفتنا بالمتغير الآخر. ففي حالة متغيرين يعبر عن ما سبق بالارتباط الخطي البسيط والانحدار الخطي البسيط.

أولا. الارتباط الخطي البسيط:

1. تعريف الارتباط: الارتباط يعني وجود علاقة بين ظاهرتين او متغيرين بمعنى أن التغير في أحد المتغيرين يؤدي إلى التغير في المتغير الآخر سواء بالزيادة أو النقصان، فإذا كان المغيران يزيدان معاً أو ينقصان معاً فإن العلاقة بينهما طردية أما إذا كان أحدهما ينقص بزيادة المتغير الآخر أو يزيد بنقصان المتغير الآخر فالعلاقة بينهما عكسية. مثل العلاقة بين تكلفة الاشهار وقيمة المبيعات، العلاقة بين الكمية المطلوبة من سلعة ما وسعرها.

2. حساب معامل الارتباط: يُستخدم معامل الارتباط لقياس قوة الارتباط بين متغيرين (ظاهرتين) ويُرمز له بالرمز r، قيمته محصورة بين (-1) و (1) أي أن: $-1 \leq r \leq 1$، وتدل الاشارة الموجبة على العلاقة الطردية بينما تدل الاشارة السالبة على العلاقة العكسية. $r = -1$ ارتباط عكسي تام، $r = 0$ عدم وجود ارتباط، $r = +1$ ارتباط طردي تام تزيد قوة العلاقة الارتباطية كلما اقتربت قيمة r من -1 أو +1 وتضعف كلما اقتربت من الصفر، الجدول التالي يوضح قيم معامل الارتباط ومعانيها:

قيمة معامل الارتباط	المعنى	قيمة معامل الارتباط	المعنى
1	ارتباط طردي تام	-1	ارتباط عكسي تام
من 0,70 الى 0,99	ارتباط طردي قوي	من 0,99- الى 0,70-	ارتباط عكسي قوي
من 0,30 الى 0,69	ارتباط طردي متوسط	من 0,69- الى 0,30-	ارتباط عكسي متوسط
من 0,01 الى 0,29	ارتباط طردي ضعيف	من 0,29- الى 0,01-	ارتباط عكسي ضعيف
0	لا يوجد ارتباط	0	لا يوجد ارتباط

وتختلف طرق حساب معامل الارتباط حسب طبيعة البيانات الإحصائية.

2.2 معامل ارتباط بيرسون Pearson: يُستخدم معامل ارتباط بيرسون لقياس الارتباط بين متغيرين في حالة البيانات الكمية فقط. ولحساب معامل ارتباط بيرسون نستخدم العلاقة التالية:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}}$$

حيث: $\bar{X} = \frac{\sum X_i}{n}$ $\bar{Y} = \frac{\sum Y_i}{n}$

مثال 1: البيانات التالية تمثل علامة الرياضيات (X) وعلامة الاحصاء (Y) لعينة من عشرة طلبة.

$Etudiant$	1	2	3	4	5	6	7	8	9	10
X_i	16	16	14	12	18	18	18	14	16	18
Y_i	12	14	14	10	18	16	16	10	18	12

المطلوب: أحسب معامل الارتباط بين الدخل والاستهلاك وماذا تستنتج.

الحل: لتطبيق صيغة معامل ارتباط بيرسون (حالة البيانات الكمية) يجب أولا حساب الوسط الحسابي لكل من X_i و Y_i ثم الانحرافات $X_i - \bar{X}$ و $Y_i - \bar{Y}$ ومربعاتهم وحاصل ضربهم كما في الجدول التالي:

$Etudiant$	X_i	Y_i	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$
1	16	12	0	-2	0	4	0
2	16	14	0	0	0	0	0
3	14	14	-2	0	4	0	0
4	12	10	-4	-4	16	16	16
5	18	18	2	4	4	16	8
6	18	16	2	2	4	4	4
7	18	16	2	2	4	4	4
8	14	10	-2	-4	4	16	8
9	16	18	0	4	0	16	0
10	18	12	2	-2	4	4	-4
$\sum$	**160**	**140**	**0**	**0**	**40**	**80**	**36**

$$\bar{X} = \frac{\sum X_i}{n} = \frac{160}{10} = 16 \qquad \bar{Y} = \frac{\sum Y_i}{n} = \frac{140}{10} = 14$$

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} = \frac{36}{\sqrt{40 \times 80}} = \frac{36}{\sqrt{3200}} = \frac{36}{56{,}57} = \mathbf{0{,}64}$$

الاستنتاج: نستنتج أن هناك ارتباط طردي متوسط بين المتغيرين.

2.2 معامل ارتباط الرتب سبيرمان Spearman: هو مقياس تقريبي يُستخدم غالباً في حالة البيانات النوعية كما يُستخدم أيضاً في حالة البيانات الكمية. ولحساب معامل ارتباط الرتب سبيرمان نستخدم العلاقة التالية:

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

مثال 2: البيانات التالية تمثل رتبة مادتي الفلسفة (X) والتاريخ (Y) لعينة من ستة تلاميذ.

$Etudiant$	1	2	3	4	5	6	7
X	السادسة	الخامسة	السابعة	الثانية	الأولى	الرابعة	الثالثة
Y	السابعة	السادسة	الخامسة	الثانية	الثالثة	الخامسة	الرابعة

المطلوب: أحسب معامل الارتباط بين تقدير مادتي الفلسفة والتاريخ وماذا تستنتج.

الحل: بداية نرتب الرتب ترتيبا تصاعديا: الأولى = 1؛ الثانية = 2؛ الثالثة = 3؛ الرابعة = 4؛ الخامسة = 5؛ السادسة = 6؛ السابعة = 7.

ولتطبيق صيغة معامل ارتباط سبيرمان للرتب (حالة البيانات النوعية الترتيبية) يجب أولا ترتيب رتب المادتين ثم نقوم بحساب فروقات الرتب $d_i = Rang\ X_i - Rang\ Y_i$ ونقوم بتربيعها d_i^2 كما في الجدول التالي:

$Etudiant$	1	2	3	4	5	6	7
X_i	السادسة	الخامسة	السابعة	الثانية	الأولى	الرابعة	الثالثة
Y_i	السابعة	السادسة	الخامسة	الثانية	الثالثة	الخامسة	الرابعة
$Rang\ X_i$	6	5	7	2	1	4	3
$Rang\ Y_i$	7	6	5	2	3	5	4
d_i	-1	-1	2	0	-2	-1	-1
d_i^2	1	1	4	0	4	1	1

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6 \times 12}{7(49-1)} = 1 - \frac{72}{336} = 1 - 0{,}21 = \mathbf{0,79}$$

الاستنتاج: نستنتج أن هناك ارتباط طردي قوي بين المتغيرين.

ثانيا. الانحدار الخطي البسيط:

تعريف الانحدار: يقيس درجة تأثير متغير مستقل على متغير تابع، ويستخدم لتقدير القيمة المستقبلية لمتغير بناءا على معرفة قيم متغير آخر، وتُكتب معادلة خط الانحدار البسيط كما يلي: $Y = a + bX$ حيث: Y المتغير التابع، X المتغير المستقل، a ثابت الانحدار، b ميل الانحدار. ويتم حساب المَعلَمتين a و b كما يلي:

- **حساب ميل الانحدار b:** يتم حساب ميل الانحدار بطريقتين:

طريقة الانحرافات:

$$b = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

طريقة المربعات الصغرى:

$$b = \frac{n\sum(X_i \times Y_i) - \sum X \times \sum Y}{n\sum X_i^2 - (\sum X_i)^2}$$

- **حساب ثابت الانحدار a:** يتم حساب ثابت الانحدار كما يلي:

$$a = \bar{Y} - b\bar{X}$$

حيث: $\bar{X} = \frac{\sum X_i}{n}$ $\bar{Y} = \frac{\sum Y_i}{n}$

مثال 3: الجدول التالي يوضح الانتاج (X) والاستهلاك المحلي (Y) لمادة الاسفلت (مليون برميل) في احدى الدول خلال عشرة سنوات.

$Année$	1	2	3	4	5	6	7	8	9	10
X_i	10	13	15	14	9	7	6	6	5	5
Y_i	6	8	9	8	7	6	5	6	5	5

المطلوب: أوجد معادلة خط الانحدار البسيط للاستهلاك المحلي بدلالة الانتاج، ثم ما هو تقدير قيمة الاستهلاك المحلي عندما يصل الانتاج إلى 16 مليون برميل؟

الحل: تكتب معادلة الانحدار الخطي البسيط على الشكل التالي: $Y = a + bX$ ولحساب المَعلَمتين a و b نشكل الجدول التالي:

$Annee$	1	2	3	4	5	6	7	8	9	10	$\sum$
X_i	10	13	15	14	9	7	6	6	5	5	**90**
Y_i	6	8	9	8	7	6	5	6	5	5	**65**
$X_i \times Y_i$	60	104	135	112	63	42	30	36	25	25	**632**
X_i^2	100	169	225	196	81	49	36	36	25	25	**942**
$X_i - \bar{X}$	1	4	6	5	0	-2	-3	-3	-4	-4	-
$Y_i - \bar{Y}$	-0,5	1,5	2,5	1,5	0,5	-0,5	-1,5	-0,5	-1,5	-1,5	-
$(X_i - \bar{X})^2$	1	16	36	25	0	4	9	9	16	16	**132**
$(X_i - \bar{X})(Y_i - \bar{Y})$	-0,5	6	15	7,5	0	1	4,5	1,5	6	6	**47**

1. تحديد قيمة الوسط الحسابي:

$$\bar{X} = \frac{\sum X_i}{n} = \frac{90}{10} = 9 \qquad \bar{Y} = \frac{\sum Y_i}{n} = \frac{65}{10} = 6{,}5$$

2. تحديد قيمة المَعلمة b:

1.2 طريقة المربعات الصغرى:

$$b = \frac{n\sum(X_i \times Y_i) - \sum X \times \sum Y}{n\sum X_i^2 - (\sum X_i)^2} = \frac{10 \times 632 - 90 \times 65}{10 \times 942 - 8100} = \frac{470}{1320} = \mathbf{0,36}$$

2.2 طريقة الانحرافات:

$$b = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{47}{132} = \mathbf{0,36}$$

3. تحديد قيمة المَعلمة a:

$$a = \bar{Y} - b\bar{X} = 6{,}5 - 0{,}36 \times 9 = \mathbf{3,26}$$

ومنه معادلة خط الانحدار البسيط تكتب كما يلي:

$$\boldsymbol{Y = 3,26 + 0,36X}$$

تقدير قيمة الاستهلاك المحلي عندما يصل الانتاج إلى 16 مليون برميل

$$Y = 3{,}26 + 0{,}36X \Leftrightarrow Y = 3{,}26 + 0{,}36(16) = \mathbf{9,02}$$

أي أن حجم الاستهلاك المحلي المتوقع يساوي 9,02 مليون برميل لما الانتاج يصل إلى 16 مليون برميل.