Correlation and Regression

الفصل الخامس: (العلاقات الإحصائية) الارتباط والانحدار **Correlation &Regression**

مقدمة:

في الفصول الثلاث السابقة تم عرض بعض المقاييس الوصفية، مثل مقاييس النزعة المركزية، والتشتت، ومقاييس الالتواء وغيرها من المقاييس الأخرى والتي يمكن من خلالها وصف شكل توزيع البيانات التي تم جمعها عن متغير واحد، وننتقل من التعامل مع متغير واحد إلى التعامل مع متغيرين أو أكثر.

ويتناول هذا الفصل دراسة وتحليل العلاقة بين متغيرين، وذلك باستخدام بعض طرق التحليل الإحصائي مثل تحليل الارتباط، والانحدار الخطي البسيط، فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط، وإذا كان اهتمامه بدراسة أثر أحد المتغيرين على الآخر استخدم لذلك أسلوب تحليل الانحدار.

نستخدم تحليل الارتباط ⟸ إذا كان الغرض من التحليل هو تحديد نوع وقوة العلاقة بين متغيرين

نستخدم تحليل الانحدار ⟸ إذا كان الغرض هو دراسة وتحليل أثر أحد المتغيرين على الآخر

اولا: الارتباط:

❖ تعريف الارتباط:

يقصد بالارتباط وجود علاقة بين ظاهرتين أو أكثر ، كالعلاقة بين الطلب والسعر والعلاقة بين الدخل والاستهلاك ونحو ذلك.

❖ درجات معامل الارتباط :

– تتراوح قيمة معامل الارتباط بين (-1 ، +1)

-اذا كان معامل الارتباط : 0.00 = لا يوجد علاقة بين الظاهرتين

0.01 – 0.19 = علاقة ضعيفة جدا

0.20 – 0.29 = علاقة ضعيفة

0.30 – 0.39 = علاقة متوسطة

0.40 – 0.69 = علاقة قوية

0.70 – 1 = علاقة قوية جدا

– إذا كانت قيمة معامل الارتباط موجبة فإن العلاقة بين المجموعتين تكون طردية ، وإذا كانت سالبة فإن العلاقة تكون عكسية .

❖ مقاييس الارتباط:

١- معامل ارتباط بيرسون (Pearson) – ويرمز له: (r)

٢- معامل ارتباط الرتب (سبيرمان) (Spearman) – ويرمز له: (r_s)

اولا: الارتباط: ١- معامل ارتباط بيرسون (Pearson)

استخدامه: يتم استخدام معامل ارتباط بيرسون لقياس العلاقة في حالة البيانات الكمية

يحسب معامل الارتباط لبيرسون كالتالي:-

$$r=\frac{Cov_{xy}}{S_x.S_y}=\frac{\dfrac{\sum(x-\bar{x})(y-\bar{y})}{(n-1)}}{\sqrt{\dfrac{\sum(x-\bar{x})^2}{(n-1)}}\sqrt{\dfrac{\sum(y-\bar{y})^2}{(n-1)}}}\quad ...(1)$$

ahmed.hmood@sci.utq.edu.iq Geology dept. / Science college/ University of thi-Qar

حيث:

$Cov_{xy} = \sum(x-\bar{x})(y-\bar{y})/(n-1)$: هو التغاير المشترك بين المتغيرين (y) و (x)

$S_x = \sqrt{(x-\bar{x})^2/(n-1)}$: هو الانحراف المعياري للمتغير (x)

$S_y = \sqrt{(y-\bar{y})^2/(n-1)}$: هو الانحراف المعياري للمتغير (y)

اولا: الارتباط: ١- معامل ارتباط بيرسون (Pearson)

استخدامه: يتم استخدام معامل ارتباط بيرسون لقياس العلاقة في حالة البيانات الكمية

يمكن تبسيط الصيغة الى :

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

ويمكن ايضا تبسيط الصيغة الى:

$$\boxed{r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \ldots(3)}$$

مثال(١): اوجد معامل الارتباط بين اطوال و أوزان مجموعة من طلبة احدى الجامعات من البيانات التالية:-

الطول (x)	164	152	184	164	176	156	168	164
الوزن (y)	52	40	60	52	60	42	50	52

الحل:-

باستخدام الصيغة (3)

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \ldots(3)$$

نحتاج حساب البيانات التالية فقط :- (xy) و(y^2) و(x^2) كما في الجدول التالي:

الطول (x)	164	152	184	164	176	156	168	164	1328	$\sum x$
الوزن (y)	52	40	60	52	60	42	50	52	408	$\sum y$
(xy)	8528	6080	11040	8528	10560	6552	8400	8528	68216	$\sum xy$
(y^2)	2704	1600	3600	2704	3600	1764	2500	2704	21176	$\sum y^2$
(x^2)	26896	23104	33856	26896	30976	24336	28224	26896	221184	$\sum x^2$

الحل:-

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \ldots(3)$$

$$r = \frac{68216 - \frac{(1328)(408)}{8}}{\sqrt{\left(221184 - \frac{(1328)^2}{8}\right)\left(21176 - \frac{(408)^2}{8}\right)}} = \frac{488}{\sqrt{(736)(368)}} = \frac{488}{520.43} = 0.937$$

النتيجة ان معامل الارتباط (٠،٩٤) يدل على وجود علاقة قوية جدا وطردية بين الطول والوزن لان الاشارة موجب

ahmed.hmood@sci.utq.edu.iq Geology dept. / Science college/ University of thi-Qar

خواص معامل الارتباط:-

هناك خاصيتان يستفاد منهما لتسهيل عمليات الحساب اذا كانت الاعداد كبيرة وهما:-

الخاصية الاولى = قيمة معامل الارتباط (r) لا تتغير اذا طرحنا (او جمعنا) أي عدد ثابت من جميع قيم الظاهرة الاولى واي عدد ثابت آخر من جميع قيم الظاهرة الثانية.

الخاصية الثانية = قيمة معامل الارتباط (r) لا تتغير اذا قسمنا (او ضربنا) جميع قيم الظاهرة الاولى على عدد ثابت، وقسمنا (او ضربنا) جميع قيم الظاهرة الثانية على أي عدد ثابت آخر

مثال (٢) احسب معامل ارتباط بيرسون لبيانات الطول والوزن الواردة في المثال السابق رقم (١) باستخدام الخاصيتين:-

الحل: نكون الجدول التالي:-

لاحظ: - عمود (1) و (2) يمثلان القراءات الاصلية

- عمود (3) و (4) يمثلان الانحرافات المبسطة بعد استخدام الخاصيتين وهي الطرح والقسمة على اعداد ثابتة

1	الطول (x)	164	152	184	164	176	156	168	164		
2	الوزن (y)	52	40	60	52	60	42	50	52		
3	$x' = \frac{x-164}{4}$	0	-3	5	0	3	-2	1	0	4	$\sum x'$
4	$y' = \frac{y-50}{2}$	1	-5	5	1	5	-4	0	1	4	$\sum y'$
5	$(x'y')$	0	15	25	0	15	8	0	0	63	$\sum x'y'$
6	(x'^2)	0	9	25	0	9	4	1	0	48	$\sum x'^2$
7	(y'^2)	1	25	25	1	25	16	0	1	94	$\sum y'^2$

مثال (٢) احسب معامل ارتباط بيرسون لبيانات الطول والوزن الواردة في المثال السابق رقم (١) باستخدام الخاصيتين:-

الحل: نكون الجدول التالي:-

$$\sum x' = 4 \quad \sum y' = 4 \quad \sum x'y' = 63 \quad \sum x'^2 = 48 \quad \sum y'^2 = 94 \quad n = 8$$

$$r = \frac{\sum x'y' - \frac{\sum x' \sum y'}{n}}{\sqrt{\left(\sum x'^2 - \frac{(\sum x')^2}{n}\right)\left(\sum y'^2 - \frac{(\sum y')^2}{n}\right)}}$$

$$r = \frac{63 - \frac{(4)(4)}{8}}{\sqrt{\left(48 - \frac{(4)^2}{8}\right)\left(94 - \frac{(4)^2}{8}\right)}} = \frac{61}{\sqrt{(46)(92)}} = \frac{61}{65.05} = 0.937$$

وهي نفس النتيجة السابقة: ان معامل الارتباط (٠،٩٤) يدل على وجود علاقة قوية وطردية بين الطول والوزن لان الاشارة موجب

ahmed.hmood@sci.utq.edu.iq Geology dept. / Science college/ University of thi-Qar