معامل الارتباط

Coefficient of correlation

مقدمة: تدرس الجغرافية توزع الظاهرات الجغرافية على سطح الأرض، والعلاقات بين بعضها بعض، فهي تدرس العلاقات بين الظاهرات بهدف تحديد العلاقات بينها.

إن المشكلات البيئية الحالية والتي في جزء منها تعد مشكلات جغرافية تستدعي من الجغرافي الإحاطة بالأسباب مفسراً وتحليلاً. لذلك لا بد من الضروري الهتمام بدراسة الارتباط بين السبب والمسبب. فنرى الجغرافي يبحث في العلاقة بين الظاهرات الجغرافية مثل العلاقة بين الارتفاع عن سطح البحر ودرجة الحرارة، أو العلاقة بين دخل الأسرة وإنفاقها، أو بين البعد عن مركز المدينة التجاري وكثافة السكان، ..... الخ.

تقوم دراسة الارتباط على بحث العلاقة بين ظاهرتين - أو أكثر - لمعرفة احتمال وجود ترابط بينهما، فقد تكون العلاقة طردية، (مثال: ارتفاع درجة الحرارة وازدياد التبخر)، وقد تكون العلاقة عكسية مثل (الارتفاع عن سطح البحر ودرجة الحرارة).

تجدر الإشارة إلى أنه قبل إجراء أي حساب لتحليل الارتباط لابد لنا أن نتأكد من وجود علاقة منطقية بين الظاهرات المدروسة (المتغيرات المدروسة).

مثال (1): العلاقة بين كميات الهطل والإنتاج البعلي في منطقة ما هي علاقة منطقية.

مثال 2: العلاقة بين عدد رؤوس الأغنام في سورية وإنتاج أحد مصانع السيارات في مدينة ديترويت الأمريكية هي علاقة غير منطقية.

حتى نحكم بوجود علاقة بين متغيرين يجب توافر (الظاهرتين) في زمان ومكان منطقيين دون فواصل تذكر بينهما، من

قبل وجود علاقة بين غلة إنتاج زراعي واستخدام الأسمدة في موسم زراعي أقل من سنة.

وإذا تأكدنا فرضاً بوجود علاقة بين متغيرين، فهذا لا يعني بالضرورة وجود علاقة سببية بينهما، بمعنى أنه أحدهما سبب حتمي للآخر، فهناك عوامل كثيرة تتحكم بهما.

مثال: لو أخذنا ظاهرتين هما: حجم المساحة المزروعة بمحصول ما (س) والإنتاج الزراعي لهذا المحصول (ص)
(Y) (X)

لوجدنا أن هناك عوامل عديدة تؤثر في حجم المساحة المزروعة بهذا المحصول مثل إمكانيات الفلاح، قد تحدد المساحة المزروعة بهذا المحصول.

كما أن هناك عوامل عديدة غير المساحة تؤثر في الإنتاج الزراعي مثل (درجة الحرارة، الرطوبات وكميات توزعها، الصقيع ومدى فترته، الإشعاع الشمسي وكميات الأسمدة ونوع البذور المحسنة المستخدمة، ..... الخ).

- أنواع الارتباط: هناك عدة أنواع للارتباط:

أ- الارتباط البسيط: وهو دراسة العلاقة بين ظاهرتين، مثل دراسة العلاقة بين درجة الحرارة والتبخر، أو العلاقة بين دخل الأسرة وإنفاقها.

ب- الارتباط الجزئي: وهو دراسة العلاقة بين ظاهرة تابعة (ص) أو (Y) وهي متعلقة بظاهرتين مستقلتين (س1 و س2) أو (X1 و X2)، مثل دراسة العلاقة بين إنتاجية محصول القمح (كغ/هـ) (ص) أو (Y) وكمية الأمطار (س1) أو (X_1) وكمية السماد المستخدمة (س2) أو (X_2).

جـ- الارتباط المتعدد: وهو دراسة العلاقة بين متغير تابع والذي هو (ص) أو (Y) وعدة متغيرات (س1، س2، س3، ... سن) أو (X_1 و X_2 و X_3 ... X_n).

مثل دراسة العلاقة بين كمية الإنتاج الزراعي لمحصول ما (ص) أو (Y) وحجم مساحة المزرعة (س١) أو (X_1) وكمية الرطل (س٢) أو (X_2) وكمية السماد (س٣) أو (X_3) ونوعية التربة (س٤) أو (X_4) وغيرها من العوامل.

وقد تكون العلاقة بين المتغيرات خطية، أي يمكن تمثيلها بخط مستقيم، في حين أنها قد تكون علاقة لا خطية، أي يمكن تمثيلها بخط منحني.

وفي دراستنا لمعامل الارتباط سنكتفي بالعلاقة الخطية.

أ- معامل الارتباط لكارل بيرسون

ويعرف باسم معامل ارتباط بيرسون نسبة إلى واضعه كارل بيرسون، وهو يستخدم من أجل قياس درجة الترابط بين ظاهرتين كميتين.

يتراوح قيمته ما بين (-١) و(+١)، وإذا كانت قيمة معامل الارتباط سالبة مثلاً يعني أن الارتباط عكسي، وإذا كانت موجبة فالارتباط طردي.

تجدر الإشارة إلى أنه يوجد عدة طرق لحسابه، واختيار الطريقة الأنسب يعتمد على حسابه.

مثال: يبين الجدول الآتي إنتاجية أحد المحاصيل والارتفاعات عن سطح البحر.

(الهدف: معرفة هل يوجد علاقة بين الارتفاع والإنتاجية؟ إذاً لدينا هنا ظاهرتين.

الارتفاع: مقدّر بالقدم = ٣٠,٤٨ م = ٣٠,٤٨ سم.

الإنتاجية: مقدّرة بالبوشل/أكر

البوشل: مكيال أمريكي للحبوب يعادل ٣٥,٢٣٩٠٧ ليتر

الأكر: وحدة مساحة تعادل ٤٠٤٦,٨٥٦٤٢ م٢

رقم الحقل	عمود (١)	عمود (٢)	عمود (٣)	عمود (٤)	عمود (٥)
الرقم	الارتفاع (قدم) (س) (X)	الإنتاجية (بوشل/أكر) (ص) (Y)	($س^2$) (X^2)	($ص^2$) (Y^2)	(س.ص) (X.Y)
١	١٠	٣٠	١٠٠	٩٠٠	٣٠٠
٢	٢٠	٣٠	٤٠٠	٩٠٠	٦٠٠
٣	٥٠	٣١	٢٥٠٠	٩٦١	١٥٥٠
٤	٧٠	٢٤	٤٩٠٠	٥٧٦	١٦٨٠
٥	٨٠	٢٦	٦٤٠٠	٦٧٦	٢٠٨٠
٦	١٠٠	٢٣	١٠٠٠٠	٥٢٩	٢٣٠٠
٧	١٤٠	١٣	١٩٦٠٠	١٦٩	١٨٢٠
٨	١٥٠	١٧	٢٢٥٠٠	٢٨٩	٢٥٥٠
٩	١٨٠	١٤	٣٢٤٠٠	١٩٦	٢٥٢٠
١٠	٢٠٠	١٢	٤٠٠٠٠	١٤٤	٢٤٠٠
المجموع (مج) (Σ)	١٠٠٠	٢٢٠	١٣٨٨٠٠	٥٣٤٠	١٧٨٠٠

المطلوب: احسب معامل الارتباط حسب طريقة كارل بيرسون.

الحل: ١- نحدد الظاهرة التابعة (في مثالنا الإنتاجية ونرمز لها بالرمز (ص) أو (Y)، أما الظاهرة المستقلة فنرمز لها بالرمز (س) أو (X).

٢- تربيع قيم الظاهرة المستقلة (س) أو (X) وتضعها في العمود رقم (٣)، والذي سيكون عنوانه (س٢) أو (X^2).

٣- تربيع قيم الظاهرة التابعة (ص) أو (Y) وتضعها في العمود رقم (٤)، والذي عنوانه (ص٢) أو (Y^2).

٤- تضرب قيم الظاهرة المستقلة بقيم الظاهرة التابعة، أي (س.ص) وتضع الناتج في العمود رقم (٥)، والذي عنوانه (س.ص) أو (X.Y).

٥- تجمع القيم الواردة في العمود الأول فتحصل على (مجـ س) أو (ΣX)، وفي مثالنا يعادل (١٠٠٠).

٦- تجمع القيم الواردة في العمود رقم (٢) فتحصل على (مجـ ص) أو (ΣY)، وفي مثالنا يعادل (٢٢٠).

٧- تجمع القيم الواردة في العمود رقم (٣) فتحصل على (مجـ س٢) أو (ΣX^2)، وفي مثالنا يعادل (١٣٨٨٠٠).

٨- تجمع القيم الواردة في العمود رقم (٤) فتحصل على (مجـ ص٢) أو (ΣY^2)، وفي مثالنا يعادل (٥٣٤٠).

٩- تجمع القيم الواردة في العمود رقم (٥)، فتحصل على مجـ (س.ص) أو (ΣXY)، وفي مثالنا يعادل (١٧٨٠٠).

١٠- تربيع مجموع (س) أو (مجـ س)٢ = (١٠٠٠)٢ = ١٠٠٠٠٠٠

١١- تربيع مجموع (ص) أو (مجـ ص)٢ = (٢٢٠)٢ = ٤٨٤٠٠

١٢- تطبيق العلاقة التالية:

$$ر = \frac{ن\ مجـ (س.ص) - (مجـ س)(مجـ ص)}{\sqrt{[ن\ (مجـ س^2) - (مجـ س)^2][ن\ (مجـ ص^2) - (مجـ ص)^2]}}$$

إذ أن: ن = عدد أزواج القيم، وفي مثالنا ن = ١٠

[لاحظ عند استخدام علاقة معامل الارتباط فإن ن لا تمثل عدد المفردات كما ورد في المحاضرات السابقة، بل تمثل عدد أزواج القيم].

مجـ س = ١٠٠٠ ، مجـ (ص)^{٢} = ٥٣٤٠

مجـ ص = ٢٢٠ ، مجـ (ص)^{٢} = (٢٢٠)^{٢} = ٤٨٤٠٠

مجـ (س)^{٢} = ١٣٨٨٠٠ ، ر = معامل الارتباط

(مجـ س)^{٢} = (١٠٠٠)^{٢} = ١٠٠٠٠٠٠

أو نستخدم العلاقة التالية بالرموز الأجنبية:

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

وبالتعويض:

$$ر = \frac{١٠(١٧٨٨٠٠) - (١٠٠٠)(٢٢٠)}{\sqrt{\left[١٠(١٣٨٨٠٠) - (١٠٠٠٠٠٠)\right] \cdot \left[١٠(٥٣٤٠) - (٤٨٤٠٠)\right]}}$$

$$ر = \frac{٢٢٠٠٠٠ - ١٧٨٨٠٠}{\sqrt{\left[(١٣٨٨٠٠٠ - ١٠٠٠٠٠٠)\right] \cdot \left[(٥٣٤٠٠ - (٤٨٤٠٠)\right]}}$$

$$ر = \frac{٤٢٠٠٠}{\sqrt{٣٨٨٠٠٠ \times ٥٠٠}}$$

$$ر = \frac{-٤٢٠٠٠}{\sqrt{١٩٤٠٠٠٠٠٠ \times ٤٤٠٤٥٩٤٣١}} = \frac{-٤٢٠٠٠٠٠٠}{٤٤٠٤٥٩٤٣١} = -٠٫٩٥٣٥$$

أي أن قيمة معامل الارتباط هي **-٠٫٩٥٣٥**

١- لاحظ أن قيمة معامل الارتباط سالبة هنا، أي أنه هناك علاقة عكسية بين الارتفاع والإنتاجية. أي كلما ارتفعنا عن سطح البحر قلّت الإنتاجية، (نبحث عن السبب فقد يكون السبب هو غسل التربة بواسطة الأمطار).

٢- لاحظ أن قيمة معامل الارتباط قريبة من (-١) أي أن الارتباط عكسي قوي جدًا.

إذا سُئلنا ما هي دلالة، يكون الجواب عكسي قوي جدًا.

- مزايا ومساوئ معامل الارتباط:

أ- المزايا: ١- قيمته تتراوح ما بين (+١) و (-١)، وهذين الحدين يدلان على ارتباط تام، أما الصفر فيدل على انعدام الارتباط.

٢- تدل الإشارة الموجبة على علاقة طردية، والإشارة السالبة على علاقة عكسية.

٣- ليس له وحدة قياس.

ب- المساوئ: ١- لا يُعبّر عن متانة العلاقة (شدة العلاقة) بشكل صحيح إلا إذا كان الارتباط خطيًا، وفي الحالات الأخرى نستعمل غيره.

٢- لا يمكن استخدامه للمتغيرات النوعية. (فقط يمكن استخدامه في حالة كانت المتغيرات كمية).

- تفسير قيمة معامل الارتباط (دلالة):

١- نقول عن معامل الارتباط أنه ضعيف إذا كان ر < ٣٠

٢- نقول عن معامل الارتباط أنه متوسط الشدة إذا كان ٣٠ ≤ ر < ٧٠

٣- نقول عن معامل الارتباط أنه قوي إذا كان ٧٠ ≤ ر

- اختبار قيمة معامل الارتباط عن طريق حساب الخطأ المعياري لمعامل الارتباط

من أجل معرفتنا فيما إذا كانت قيمة معامل الارتباط أهمية إحصائية (لا يخضع لعامل الصدفة)، يجب أن تكون قيمة معامل الارتباط التي حصلنا عليها أكبر بثلاث مرات من قيمة الخطأ المعياري له والذي نحسب من العلاقة:

$$خ م ر = \sqrt{\frac{1 - ر^2}{ن - 2}}$$

إذ أن: خ م ر = الخطأ المعياري لمعامل الارتباط،

ر = قيمة معامل الارتباط،

ن = عدد أزواج القيم.

أو يمكن كتابة العلاقة السابقة بالرموز الأجنبية:

$$SE_r = \sqrt{\frac{1 - r2}{n - 2}}$$

يجب أن تكون قيمة (ر) أكبر من (٣ × خ م ر) عندئذ تكون لمعامل الارتباط أهمية إحصائية (ذو دلالة)، أما إذا كانت قيمة معامل الارتباط أقل من (٣ × خ م ر) فعندئذ ليس لمعامل الارتباط أهمية إحصائية، ونقول هنا بعدم وجود علاقة بين المتغيرين (خاضع لعامل الصدفة)، بالنسبة لمثالنا السابق كانت قيمة ر = ٩٥٣٥٠٠

نحسب الخطأ المعياري لمعامل الارتباط من العلاقة:

$خ_{م_ر} = \sqrt{\frac{1-ر^2}{ن-2}}$ إذا أن : ن = عدد أزواج القيم

بالتعويض : $خ_{م_ر} = \sqrt{\frac{1-(-0.935)^2}{10-2}} = \sqrt{\frac{1-0.991}{8}}$

$خ_{م_ر} = \sqrt{\frac{0.9 \times 0.1}{8}} = \sqrt{0.1126} = 0.1658$

$3 خ_{م_ر} = 3 \times 0.1658 = 0.3197$

وبما أنه $ر > 3 خ_{م_ر}$ فهذا يعني أن قيمة معامل الارتباط لا تخضع لعامل الصدفة، ونقول أن هناك ارتباط عكسي قوي بين الارتفاع والإنتاجية.

ملاحظة : هناك طريقة ثانية لحساب معامل الارتباط حسب طريقة بيرسون وذلك عن طريق حساب الانحراف المعياري، وهي طريقة طويلة وغير مطلوبة من الطلاب، فقط المطلوب الطريقة المشروحة أعلاه.

انتهت المحاضرة