المعهد العربي للتخطيط بالكويت

Arab Planning Institute - Kuwait

منظمة عربية مستقلة

الإرتباط والانحدار البسيط

سلسلة دورية تعنى بقضايا التنمية في الدول العربية

العدد السابع والأربعون - نوفمبر/تشرين الثاني 2005 - السنة الرابعة

Arab Planning Institute - Kuwait

P.O.Box : 5834 Safat 13059 State of Kuwait
Tel : (965) 4843130 - 4844061 - 4848754
Fax : 4842935

المعهد العربي للتخطيط بالكويت

ص.ب : 5834 الصفاة 13059 - دولة الكويت
هاتف : (965) - 4843130 - 4844061 - 4848754
فاكس : 4842935

E-mail ; api@api.org.kw
web site : http //www.arab-api.org

أهداف "جسر التنمية"

قائمة اصدارات "جسر التنمية"

العنوان	المؤلف	رقم العدد
مفهوم التنمية	د. محمد عدنان وديع	الأول
مؤشرات التنمية	د. محمد عدنان وديع	الثاني
السياسات الصناعية	د. أحمد الكواز	الثالث
الفقر: مؤشرات القياس والسياسات	د. علي عبدالقادر علي	الرابع
الموارد الطبيعية واقتصادات نفاذها	أ. صالح العصفور	الخامس
استهداف التضخم والسياسة النقدية	د. ناجي التوني	السادس
طرق المعاينة	أ. حسن الحاج	السابع
مؤشرات الأرقام القياسية	د. مصطفى بابكر	الثامن
تنمية المشاريع الصغيرة	أ. حسان خضر	التاسع
جداول المدخلات المخرجات	د. أحمد الكواز	العاشر
نظام الحسابات القومية	د. أحمد الكواز	الحادي عشر
إدارة المشاريع	أ. جمال حامد	الثاني عشر
الاصلاح الضريبي	د. ناجي التوني	الثالث عشر
أساليب التنبؤ	أ. جمال حامد	الرابع عشر
الادوات المالية	د. رياض دهال	الخامس عشر
مؤشرات سوق العمل	أ. حسن الحاج	السادس عشر
الاصلاح المصرفي	د. ناجي التوني	السابع عشر
خصخصة البنى التحتية	أ. حسان خضر	الثامن عشر
الأرقام القياسية	أ. صالح العصفور	التاسع عشر
التحليل الكمي	أ. جمال حامد	العشرون
السياسات الزراعية	أ. صالح العصفور	الواحد والعشرون
اقتصاديات الصحة	د. علي عبدالقادر علي	الثاني والعشرون
سياسات أسعار الصرف	د. بلقاسم العباس	الثالث والعشرون
القدرة التنافسية وقياسها	د. محمد عدنان وديع	الرابع والعشرون
السياسات البيئية	د. مصطفى بابكر	الخامس والعشرون
إقتصاديات البيئة	أ. حسن الحاج	السادس والعشرون
تحليل الأسواق المالية	أ. حسان خضر	السابع والعشرون
سياسات التنظيم والمنافسة	د. مصطفى بابكر	الثامن والعشرون
الأزمات المالية	د. ناجي التوني	التاسع والعشرون
إدارة الديون الخارجية	د. بلقاسم العباس	الثلاثون
التصحيح الهيكلي	د. بلقاسم العباس	الواحد والثلاثون
نظم البناء والتشغيل والتحويل B.O.T.	د. أمل البشبيشي	الثاني والثلاثون
الاستثمار الأجنبي المباشر : تعاريف	أ. حسان خضر	الثالث والثلاثون
محددات الاستثمار الأجنبي المباشر	د. علي عبدالقادر علي	الرابع والثلاثون
نمذجة التوازن العام	د. مصطفى بابكر	الخامس والثلاثون
النظام الجديد للتجارة العالمية	د. أحمد الكواز	السادس والثلاثون
منظمة التجارة العالمية: إنشاؤها وآلية عملها	د. عادل محمد خليل	السابع والثلاثون
منظمة التجارة العالمية: أهم الإتفاقيات	د. عادل محمد خليل	الثامن والثلاثون
منظمة التجارة العالمية: آفاق المستقبل	د. عادل محمد خليل	التاسع والثلاثون
النمذجة الإقتصادية الكلية	د. بلقاسم العباس	الأربعون
تقييم المشروعات الصناعية	د. أحمد الكواز	الواحد والأربعون
المؤسسات والتنمية	د. عماد الامام	الثاني والأربعون
التقييم البيئي للمشاريع	أ. صالح العصفور	الثالث والأربعون
مؤشرات الجدارة الإئتمانية	د. ناجي التوني	الرابع والأربعون
الدمج المصرفي	أ. حسان خضر	الخامس والأربعون
اتخاذ القرارات	أ. جمال حامد	السادس والأربعون
الإرتباط والانحدار البسيط	**أ. صالح العصفور**	**السابع والأربعون**

للاطلاع على الأعداد السابقة يمكنكم الرجوع إلى العنوان الإلكتروني التالي :

http://www.arab-api.org/develop_1.htm

المحتويـــات

الارتباط والانحدار البسيط

إعداد: أ. صالح العصفور

أولا - الانحدار الخطي والارتباط

يكتسب قياس وتحديد درجة واتجاه العلاقة بين المتغيرات اهمية كبيرة في فهم الظواهر بمختلف أنواعها. وعندما يكون الأمر متعلقاً بمتغير واحد، فإن مقاييس النزعة المركزية تصف لنا القيمة التي تقع في مركز مجموعة البيانات، كما تصف لنا مقاييس التشتت درجة انتشار وتبعثر وتوزيع قيم هذه البيانات. ولكن عندما يتعلق الأمر بمتغيرين أو أكثر، فان الباحث يتطلع إلى قياس وتحديد درجة واتجاه العلاقة بين المتغيرات أو الظواهر قيد الدراسة. ويقوم بعد ذلك باستخدام العلاقات الموجودة في التنبؤ بقيمة أحد المتغيرات بدلالة المتغير (المتغيرات) الأخرى. ومحاولة التعبير عن هذه العلاقات بدالة خطية، لأن وجود علاقة أو ارتباط بين المتغيرات لا يعني بالضرورة إمكانية التعبير عن هذه العلاقة بشكل خطي. ويعبر عن هذه العلاقة بالأرقام أو القيم الكمية كما قد يعبر عنها بالوصف. فالإحصائي عندما يتحدث عن إحدى العلاقات الدالية بين متغيرين فانه يقصد بذلك أن المتغيرين يمكن ربطهما بمعادلة رياضية، بحيث إذا علمت قيمة أحدهما (المتغير المستقل) أمكن معرفة قيمة المتغير الآخر(المتغير التابع) .

يُعنى الانحدار البسيط بدراسة العلاقة بين متغيرين على هيئة علاقة دالية، بحيث يمكن الاعتماد على المعلومات المتوفرة عن أحدهما للتنبؤ عن الآخر.

لنضرب مثلاً المكالمات الهاتفية ، فإذا عرفنا أن قيمة المكالمة الواحدة هي 20 فلساً، وأن استخدام الهاتف كان لخمسة مرات، فان القيمة التي يجب دفعها هي 100 فلس. ويمكن التعبير عن ذلك بطريقة رياضية، حيث نرمز لقيمة المكالمات بالرمز Y ولعدد مرات استعمال الهاتف بالرمز X ، وبذلك تكون العلاقة الرياضية :

$$Y = 20\,X$$

وتبقى هذه الصيغة صالحة مادامت قيمة المكالمة الواحدة ثابتة أي مساوية لعشرين فلساً. واذا ما زادت قيمة المكالمة إلى 25 فلساً فان المعادلة تصبح :

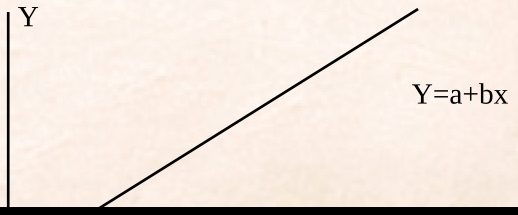
$$Y = 25\,X$$

وعليه فإنه يمكن تعميم القاعدة على كل القيم التي تبلغها المكالمة الهاتفية (b) فتصبح الصياغة الرياضية في هذه الحالة كما في المعادلة :

$$Y = b\,X$$

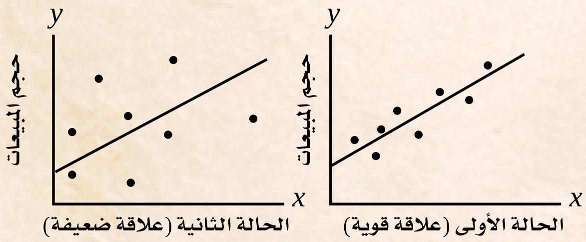
ولزيادة الإيضاح، يمكن التعبير عن هذه المعادلة بالرسم البياني، حيث يخصص محور الإحداثيات الأفقية لقيم X ومحور الإحداثيات الرأسية لما يقابلها من قيم Y وبذلك نحصل على الشكل (1)، فتبقى b ثابتة في المعادلة يمثلها الخط المعروف بخط الانحدار، وفيه نلاحظ أن كل قيمة من قيم X تقابلها قيمة من قيم Y [redacted]

الاشتراك. أي أن نقطة (a) تساوي القيمة التي يدفعها المشترك عندما لا يستخدم هاتفه .



[redacted]

إن الخطوة الأولى للبحث عن علاقة هي رسم بياني للبيانات أعلاه على الشكل أدناه (3) حيث مثلت سنوات الخبرة على الإحداثي الأفقي والمبيعات السنوية على الإحداثي الرأسي. ونكون بذلك قد حصلنا على رسم انتشاري . وقد أعطي هذا الاسم نظراً لانتشار وتبعثر النقاط على الشكل أو الرسم . وهذا الرسم الانتشاري (Scatter Diagram) المبين أدناه يساعد على تكوين استنتاج مبدئي حول إمكانية وجود علاقة بين المتغيرات، حيث يبين هذا الرسم في دراستين مختلفتين وجود علاقة قوية في الحالة الأولى، وعلاقة ضعيفة في الحالة الثانية.



الشكل (3)

وعموماً لجأ الإحصائيون إلى تصنيف المتغيرات إلى متغيرات مستقلة ومتغيرات تابعة. ويستخدم هذا التصنيف من أجل تحديد المتغير المفسّر (المستقل) والمتغير المفسَّر (التابع). وفي مثالنا هذا فانه يشار إلى سنوات الخبرة كمتغير مستقل، ويستخدم للتنبؤ بحجم المبيعات أو المتغير التابع. وتجدر الملاحظة هنا أنه جرت العادة في الرسم الانتشاري أن يكون المتغير المستقل على الإحداثي الأفقي والمتغير التابع على الإحداثي الرأسي.

ولو نظرنا إلى الشكل أعلاه فهو يعطينا لمحة عن البيانات اذ يشير إلى أن هناك فرصة لوجود علاقة بين المتغيرات. حيث أن سنوات الخبرة المنخفضة تترافق مع انخفاض حجم المبيعات، و يترافق ارتفاع سنوات الخبرة في الغالب مع ارتفاع حجم المبيعات السنوية. وتبدو هذه العلاقة بين المتغيرين خطاً مستقيماً تقريباً أو معادلة خطية. وسنبين لاحقاً كيفية تطوير مثل هذه العلاقة الخطية باستخدام الأسلوب المسمى طريقة المربعات الصغرى .

ثانيا. خط الانحدار البسيط:

سنحاول هنا التركيز على مهمة إيجاد خط مستقيم يمثل الرسم الانتشاري للبيانات أفضل تمثيل. أي أننا سنوفق معادلة على الشكل التالي:

$$\hat{y} = a + bx$$

حيث:

$\hat{y}$ = القيمة المقدرة للمتغير التابع.

x = قيمة المتغير المستقل.

a = تقاطع المحور y (أي أنها قيمة عندما تكون x=0).

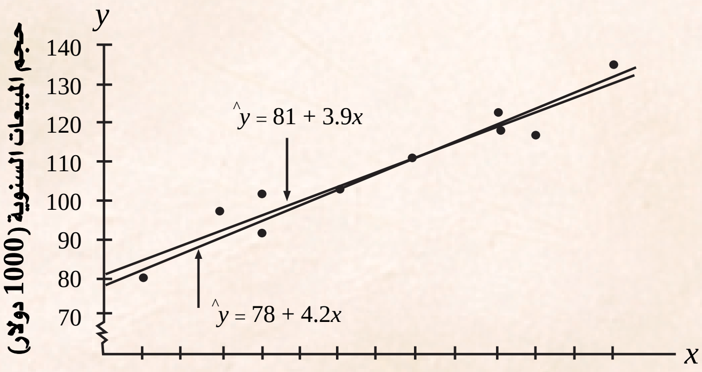
b = تغير المتغير التابع yنتيجة لتغير المتغير المستقل (ميل خط الانحدار).

إذا لم تكن العلاقة بين المتغيرين علاقة دالية بل علاقة ترابط فقط (أي أن التغير في إحداهما لا يسبب التغير في الآخر) فإن مهمة الإحصاء هي قياس درجة العلاقة بين المتغيرين واتجاهها.

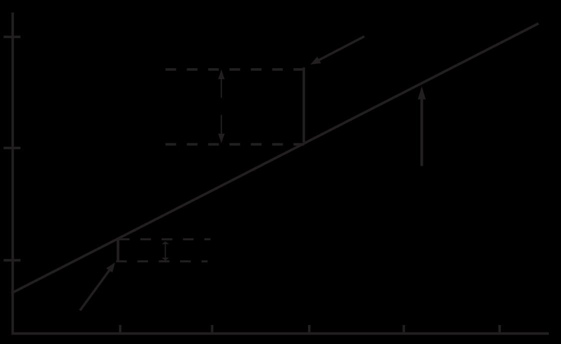
في تحليل الانحدار تكون المعادلة التي نستخدمها لوصف البيانات هي معادلة الانحدار المقدرة. وتجدر الإشارة هنا إلى أننا سنركز في هذا الجزء على معادلات خط الانحدار التي تأخذ شكل خط مستقيم. وعليه فانه يشار إلى معادلة خط الانحدار بأنها خط الانحدار المقدر. ويشار عموماً

إلى علاقة الخط المستقيم المعنية بمتغيرين أحدهما مستقل والآخر تابع على أنها انحدار خطي بسيط.

ولتوضيح فكرة خط الانحدار البسيط، نأخذ المثال المقدم في جدول (1) بخصوص حجم المبيعات السنوية وسنوات الخبرة لعشرة مندوبي مبيعات. فباستخدام الرسم الانتشاري المبين في شكل (3)، كيف يمكن اختيار



ثالثا. طريقة المربعات الصغرى :



جدول (2): حساب مجموع مربعات الفروق (الاختلافات) لخط الانحدار المقدر $\hat{y} = 78 + 4.2x$

مندوب المبيعات	سنوات الخبرة (X_i)	القيم المشاهدة لحجم المبيعات (y_i)	القيم المقدرة للمبيعات ($\hat{y} = 78 + 4.2x$)	الفروق بين القيم المشاهدة والمقدرة ($y_i - \hat{y}_i$)	مربعات الفروق $(y_i - \hat{y}_i)^2$
1	1	80	78+4.2(1)=82.2	-2.2	4.84
2	3	97	78+4.2(3)=90.6	6.4	40.96
3	4	92	78+4.2(4)=94.8	-2.8	7.84
4	4	102	78+4.2(4)=94.8	7.2	51.84
5	6	103	78+4.2(6)=103.2	-0.2	0.04
6	8	111	78+4.2(8)=111.6	-0.6	0.36
7	10	119	78+4.2(10)=120.0	-1.0	1.00
8	10	123	78+4.2(10)=120.0	3.0	9.00
9	11	117	78+4.2(11)=124.2	-7.2	51.84
10	13	136	78+4.2(13)=132.6	3.4	11.56
المجموع					179.28

وبالتالي فان مربع الفرق سيكون :

$$(y_1 - \hat{y}_1)^2 = (-2.2)^2 = 4.84$$

وسيكون الفرق للمشاهدة الثانية $= 6.4 = (y_2 - \hat{y}_2)$

ومربع الفرق $(y_2 - \hat{y}_2)^2 = 40.96$

ويبين الجدول التالي (2) بقية الحسابات للمشاهدات الثمانية الأخرى .

من جدول (2) نلاحظ أن مجموع مربعات الاختلافات = 179.28. وقد قمنا أيضاً باحتساب مجموع مربعات الاختلافات للخط المقدر الآخر: $\hat{y} = 81 + 3.9x$ والذي أعد من قبل الباحث الثاني في مثالنا أعلاه. وقد كانت النتيجة 172.32. وحيث أن مجموع المربعات لهذا الخط المقدر هي أصغر أو أقل من الخط الأول فإننا نعتبر هذا الخط أكفأ لتقدير خط الانحدار من الخط الأول. ولكن طريقة المربعات الصغرى لا تعتبر هذين الخطين هما الخطين الوحيدين. فهي التي توجد في الحقيقة خط الانحدار المقدر الذي يعطي أصغر مجموع لمربعات الفروق (الاختلافات) لكل الخطوط المختارة .

1- تقدير خط الانحدار بطريقة المربعات الصغرى :

عند استعمال طريقة المربعات الصغرى لتقدير معالم خط الانحدار المقدر ، فان الإحصائيين يرون بأن أفضل قيم لكل من a (ثابت الانحدار) و b (معامل الانحدار) يمكن إيجادها باستعمال المعادلات التالية :

$$b = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \quad \text{.............(1)}$$

$$a = \bar{y} - b_1 \bar{x} \quad \text{.............(2)}$$

حيث أن :

x_i = قيم المتغير المستقل لـ (i) من المشاهدات .

y_i = قيم المتغير التابع لـ (i) من المشاهدات .

$\bar{x}$ = قيمة متوسط المتغير المستقل .

$\bar{y}$ = قيمة متوسط المتغير التابع .

(2) احتساب الميل (b):

$$b = \frac{n\sum x_i \sum y_i - \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$= \frac{10(8,128) - 70(1080)}{10(632) - (70)^2}$$

$$= \frac{5680}{\quad} = 4$$

[redacted]

في حالة العلاقة الدالية فإن علاقة الارتباط حتمية، ولكن وجود الارتباط بين متغيرين لا يعني بالضرورة وجود علاقة دالية.

عدم وجود أي خط آخر يمكن أن يؤمن مجموع مربعات أصغر من 170 في مثالنا الحالي . لذلك نعتبر أن طريقة المربعات الصغرى هي التي تضمن أكفأ تقدير لخط الانحدار.

2- استخدام معادلة الانحدار الخطي البسيط في التنبؤ:

مع الاعتقاد بأن معادلة خط الانحدار المقدرة بطريقة المربعات الصغرى تصف العلاقة بين x و y بشكل دقيق ، فانه يبدو أنه من المعقول استخدام هذه المعادلة الرياضية في تقدير قيم y عند معرفة قيم x. وفي هذه الحالة فإن معادلة خط الانحدار المقدرة تسمى معادلة خط انحدار y على x . أما إذا كان المتغير التابع هو x والمتغير المستقل هو y فإن المعادلة المقدرة ستكون معادلة انحدار x على y، وبالتالي فإننا سنتنبأ من خلالها بقيم x بعد معرفة قيم y .

في مثالنا الحالي المتعلق بمسألة حجم المبيعات السنوية وسنوات الخبرة ، فانه يمكننا استخدام معادلة خط الانحدار المقدرة بطريقة المربعات الصغرى $\hat{y} = 80 + 4x$ في تقدير حجم المبيعات السنوية المتوقعة لثلاثة متقدمين لوظيفة مندوب مبيعات . ويبين الجدول (4) نتائج عملية التقدير .

من الجدول السابق نرى أن تحليل الانحدار السابق يعطي دعماً لقرار أن السيد رشوان محمد هو الشخص الأفضل كمندوب مبيعات ، حيث يمكن تقدير مبيعاته خلال السنة الأولى من التحاقه بالشركة بحوالي 100,000 دولار ، ويفوق بذلك المبيعات المقدره لعلي مرتجى المتقدم الثاني للوظيفة بحوالي 12,000 دولار. في هذا المثال لمسنا أحد أهم استخدامات تحليل الانحدار : بتزويد صيغة رياضية تدعم بالمعلومات صانعي القرار في المجال التجاري والاقتصادي .

هناك الكثير من العلاقات الزائفة بين بعض المتغيرات، فبالرغم من وجود علاقة ارتباط قوية، إلاّ أن هذه العلاقة لا تعني بالضرورة وجود متغير تابع ومتغير مستقل.

3- جودة التقدير :

الخطوة التالية في دراسة الانحدار هي قياس جودة التقدير في نموذج خط الانحدار . ولكن قبل ذلك لا بد من وضع مجموعة من الفرضيات (في قيم البيانات) تمكننا من إنجاز تحليلات إحصائية إضافية .

جدول (4): استخدام معادلة خط الانحدار في تقدير حجم المبيعات السنوية لثلاثة متقدمين

المتقدم للوظيفة	سنوات الخبرة	حجم المبيعات السنوية المقدرة (بالألف دولار)
رشوان محمد	5	100 = (5) 4 + 80
علي مرتجى	2	88 = (2) 4 +80
يحيى الربيعي	0	80 = (0) 4 + 80

يفترض النموذج أن قيم المتغير المستقل x يمكن أن تقابلها قيم للمتغير التابع y . وبالتالي فان تغير y ناتج عن (1) التغير في x (2) التغير الباقي e وهو تغير عشوائي غير مفسر من قبل النموذج.

سوف نستعرض هنا في هذا الجزء طريقة إحصائية لقياس أو وصف جودة خط الانحدار المقدر . وقد كنا قد

[redacted]

المتوسط (108) هو 2,442 . وهذه القيمة تمثل مجموع المربعات الكلي للفروق قبل تحليل الانحدار ، ويشار إليها عموماً على أنها المجموع الكلي لمربعات الاختلاف حول المتوسط . وسوف نرمز لها بـ SST . وفي مثالنا السابق سيكون المجموع الكلي لمربعات الاختلاف :

$$SST = \sum(Y_i - \overline{Y})^2 = 2442$$

[redacted]

تسمى النسبة بين الاختلافات المفسَّرة والاختلاف الكلي بمعامل التحديد، فإذا كانت تساوي صفراً فإن ذلك يعني أن الاختلاف الكلي جميعه غير مفسَّر.

الاختلاف (التباين) المفسرة بواسطة معادلة خط الانحدار . وعموماً يدعى مجموع المربعات هذا بمجموع مربعات الانحدار ويرمز له بـ SSR وهو الاختلاف المفسر . ورغم أن الطريقة أعلاه هي الطريقة المستخدمة عادة في احتساب الاختلاف المفسر (SSR)، إلا أنه يمكن تبيان طريقة مباشرة لاحتساب SSR باستخدام الصيغة التالية :

$$SSR = \sum(\hat{Y}_i - Y)^2$$

واحتساب (SSR) باستخدام الصيغة أعلاه مبين في جدول (6)، وقد تم الحصول على قيمة (2272=SSR) وهي قيمة مطابقة لما تم الحصول عليه أعلاه . والعلاقة بين SSE، SST و SSR تشكل أساساً لواحدة من أهم النظريات في الإحصاء التطبيقي . وتقول هذه النظرية ، بشكل عام : أن مجموع مربعات المشاهدات حول متوسطها (SST) يمكن تجزئتها إلى جزئين : SSE و SSR بحيث : SST = SSE + SSR

والآن ، لنرى كيف يمكن استعمال العلاقة المبينة أعلاه في تطوير مقياس لجودة التوفيق لمعادلة خط الانحدار المقدرة .

يكون التوفيق كاملاً لمعادلة خط الانحدار إذا كانت جميع المشاهدات واقعة على خط مستقيم، وعندها يمر خط الانحدار المقدر بطريقة المربعات الصغرى في جميع النقاط.

سيكون لدينا توفيق كامل لخط الانحدار المقدر لو كانت كل المشاهدات تقع على خط مستقيم . وفي هذه الحالة فان خط الانحدار المقدر بطريقة المربعات الصغرى سيمر في جميع النقاط . ولذلك ستكون SSE=0. وسيكون SSR = SST، وبالتالي SST/SSR = 1. ومن جانب آخر فان توفيقاً ضعيفاً للبيانات المشاهدة ينتج عن SSE كبيرة . وحيث أن SST = SSE + SSR، فان أكبر

جدول (6): الاحتساب المباشر لمجموع مربعات الانحدار (SSR)

مندوب المبيعات	سنوات الخبرة x_i	حجم المبيعات y_i	حجم المبيعات المقدرة ($\hat{y} = 80 + 4x_i$)	$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$
1	1	80	84	-24	576
2	3	97	92	-16	256
3	4	92	96	-12	144
4	4	102	96	-12	144
5	6	103	104	-4	16
6	8	111	112	4	16
7	10	119	120	12	144
8	10	123	120	12	144
9	11	117	124	16	256
10	13	136	132	24	576
$\sum$					2272

SSE (وبالتـالي أضعف جودة توفيق) سيـحدث عنـدما تكون SST = SSE ، وفي هذه الحالة فان SSR = صفر ، وعندها لن يكون لخط الانحـــدار المقـــدر أي دور في المساعدة بتخمين قيم y . وهكذا فان أسوأ توفيق ممكن لخط الانحدار يكون عندما تكون SSR = صفر وعندها تكون نسبة SST/SSR = صفر .

حيث أن الباحثين في الدراسات الاقتصادية والتجارية يشعرون بنتائج جيدة إذا حصلوا على R^2 مساوية لـ 0.60 وأكثر .

تكمن مساوي معامل ارتباط بيرسون بأنه لا يطبق إلاّ على المتغيرات الرقمية أو الكمية ويفقد مفعوله بالنسبة للمتغيرات الوصفية.

رابعا . الارتباط :

ومن الجدير بالذكر أن معامل الارتباط لا يتناول موضوع العلاقة الدالية، بمعنى أنه يقيم مقدار العلاقة بين المتغيرين واتجاههما دون التعرض إلى موضوع أي منهما متغير مستقل وأيهما متغير تابع.

سنستعرض فيما يلي أهم الصيغ المستخدمة لاستنباط معامل الارتباط :

1- تحديد معامل الارتباط من خلال تحليل الانحدار :

من خلال مناقشتنا السابقة للانحدار الخطي كنا قد افترضنا أن معادلة الانحدار الخطي بطريقة المربعات الصغرى هي $\hat{y} = a + bx$. وفي هذه الحالة فان معامل الارتباط يمكن احتسابه من معامل التحديد (r^2) كما يلي :

$$R = \pm\sqrt{r^2}$$

$$= \pm\sqrt{\text{معامل التحديد}}$$

وتتحدد إشارة معامل الارتباط من خلال إشارة الميل (b) في معادلة خط الانحدار . وفي مثالنا السابق فان معامل الارتباط هو

$$r = \pm\sqrt{0.93}$$

$$= +0.96$$

وحيث أن الميل (b = 4) كان موجباً فان معامل الارتباط هو إيجابي أيضاً.

2- معامل الارتباط البسيط (بيرسون):

هو معامل يمكن اللجوء إليه عندما لا يكون صانع القرار معنياً بالعلاقة الدالية بين المتغيرين x و y ، بل يكون معني فقط العلاقة بين المتغيرين من عدمها . فيمكن احتساب معامل الارتباط دون اللجوء إلى إنجاز تحليل خط الانحدار. وفي هذه الحالة فان الصيغة المستخدمة هي :

$$R = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

وفي مثالنا السابق حول حجم المبيعات السنوية فان الحسابات الضرورية لاستخدام الصيغة أعلاه موضحة في الجدول (7) التالي :

جدول (7): الحسابات الضرورية لاستخراج معامل الارتباط

مندوب المبيعات	سنوات الخبرة x_i	حجم المبيعات y_i	$x_i y_i$	x_i^2	y_i^2
1	1	80	80	1	6400
2	3	97	291	9	9409
3	4	92	368	16	8464
4	4	102	408	16	10404
5	6	103	618	36	10609
6	8	111	888	64	12321
7	10	119	1190	100	14161
8	10	123	1230	100	15129
9	11	117	1287	121	13689
10	13	136	1768	169	18496
المجموع	70	1080	128,8	632	082,119

وبتطبيق الصيغة الرياضية أعلاه فان :

$$R = \frac{10(8,128) - 70(1,080)}{\sqrt{10(632) - (70)^2}\sqrt{10(119,082) - (1,080)^2}}$$

ويلاحظ أن قيمة معامل الارتباط التي حصلنا عليها هي نفس القيمة التي احتسبت كجذر تربيعي لمعامل التحديد .

وهناك خصائص مميزة يتمتع بها معامل بيرسون ، تسهل من عملية احتسابه وهي :

- أنه لا يتأثر بإضافة أو طرح أي ثابت على/من جميع قيم أي من المتغيرين .
- كذلك فانه لا يتأثر بضرب جميع قيم أي من المتغيرين في أية كمية ثابتة .

هذا يعني أن العلاقة بين نفقات الدعاية وحجم المبيعات طردية ، فكلما زادت نفقات الدعاية كلما زادت المبيعات .

وإذا ما استخدمنا أحد الخاصيتين المذكورتين بقسمة (Y) على 10 فإن النتائج ستكون كالتالي :

جدول (9): الحسابات الضرورية لاحتساب معامل الارتباط بعد قسمة أحد المتغيرين (y) على 10

[redacted]

جدول (2): حساب مجموع مربعات الفروق (الاختلافات) لخط الانحدار المقدر $\hat{y} = 78 + 4.2\,x$

الشهر	نفقات الدعاية X_i	ترتيب قيم X_i	حجم المبيعات Y_i	ترتيب قيم X_i	الفرق (d)	D^2
1	2	5	60	5	0	0
2	5	2	100	1	1	1
3	4	3	70	4	1	1
4	6	1	90	2	1	1
5	3	4	80	3	1	1
المجموع						4

(d = فروق الرتب) ويحسب معامل ارتباط الرتب حسب المعادلة التالية :

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

بتطبيق هذه المعادلة على المثال السابق (العلاقة بين الإنفاق على الدعاية وبين المبيعات) .

$$r = 1 - \frac{6 \times 4}{125 - 5}$$

$$= 1 - \frac{24}{120}$$

$$= 1 - 0.2$$

$$= 0.8$$

وعادة ما يعطي بيرسون وسبيرمان قيماً متقاربة . ولو أن التطابق في هذا المثال لا يعدو أن يكون مجرد مصادفة . ويمكن التحقق بسهولة من أن معامل سبيرمان يتمتع كما يتمتع بيرسون بالخواص التي تم ذكرها لهذا المعامل وهي أنه لا يتأثر بطرح أو إضافة أو حتى قسمة أي من المتغيرين على أي ثابت.

مثال : عند احتساب العلاقة بين عدد الأطباء ومعدل الوفيات في دولة الكويت حسب طريقة سبيرمان كانت النتائج كما هي مبينة أدناه :

السنوات	عددالأطباء X_i	معدل الوفيات لكل ألف من السكان الكويتيين Y_i	ترتيب X_i	ترتيب Y_i	الفرق (d)	D^2
1973	990	6.7	16	1	15	225
1974	1087	6.6	15	2	13	169
1975	1178	6.1	14	4	10	100
1976	1341	5.6	13	5	8	64
1977	1567	6.2	12	3	9	81
1978	1703	5.4	11	6.5	4.5	20.25
1979	1957	5.4	10	6.5	3.5	12.25
1989	2341	5.1	9	8	1	1
1981	2580	4.5	8	10	2	4
1982	2734	4.6	7	9	2	4
1983	2834	4.2	6	11	5	25
1984	2983	3.9	5	13	8	64
1985	3095	4.0	4	12	8	64
1986	3129	3.5	2	14	12	144
1987	3128	3.4	3	15.5	12.5	156.25
1988	3256	3.4	1	15.5	14.5	210.25
المجموع						1344

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

$$= 1 - \frac{6 \times 1344}{(16)^3 - 16} = 1 - \frac{8064}{4096 - 16}$$

$$= 1 - \frac{8064}{4080} = -0.9764$$

وتدل هذه النتيجة على أن العلاقة قوية وعكسية بين [redacted]

هناك من علاقة جوهرية بين الاستهلاك والدخل، أي هل يمكن الاعتماد على العلاقة الخطية بالتنبؤ بقيمة الاستهلاك الشخصي إذا ما عرفنا قيمة الدخل (وذلك بمستوى 0.05) ؟.

Year	Income	Cons
1981	80	70
1982	100	65
1983	120	90

[redacted]

مثال تطبيقي على الحاسب الآلي:

المصادر العربية

- د. علي أبو القاسم محمد، مقدمة في علم الإحصاء التطبيقي - المعهد العربي للتخطيط بالكويت.
- د. محمد عبدالحميد طه، مقدمة في الإحصاء - الهيئة العامة للتعليم التطبيقي والتدريب، 1985.
- د. علي أبو القاسم محمد، أساليب الإحصاء التطبيقي - المعهد العربي للتخطيط بالكويت، دار الشباب للنشر والترجمة، 1987.
- د. رمضان حسن عبدالرحيم، مبادىء في الإحصاء الوصفي، مكتبة عين شمس.
- د. أحمد عبادة سرحان ود. صلاح الدين طلبه، مقدمة الإحصاء الاجتماعي، دار الكتب الجامعية، 1973.
- جوردن بانكروفت وجورج أوسليفان، الرياضيات والإحصاء لدراسات المحاسبة والأعمال، ترجمة الدكتور سامي مقدسي، دار ماكجرو هيل للنشر، 1981.

المصادر الانجليزية

- David Anderson, Dennis J. Sweeney and Thomas Williams, Introduction to Statistics, West Publishing Co., 1981.
- W.M. Harper, Statistics, Pitman Publishing, 1989.
- William L. Hays, Statistics for the Social Sciences, 1979.
- Ya-Lan Chau, Statistical Analysis for Business Economics, Elsevier Science Publishing Co., 1989.