الفصل العاشر
تحليل الانحدار

Regression

تكلمنا في الفصل السابق عن مقاييس الارتباط، وهذه تظهر لنا العلاقة بين ظاهرتين أو أكثر. فإذا ما وجدت هذه العلاقة فإننا نحاول تقدير قيمة أحد المتغيرين إذا عرف المتغير الآخر وذلك عن طريق الانحدار.

فالانحدار: هو دراسة للتوزيع المشترك لمتغيرين أحدهما مثبت عند مستويات معينة، أو كما يعبر عنه أحياناً متغير يقاس دون خطأ، ويسمى عادة: متغير مستقل (Independent)، والآخر غير مقيد ويأخذ قيماً مختلفة عند كل مستوى من مستويات المتغير المستقل ويسمى هذا المتغير: بالمتغير التابع Dependent.

أهداف تحليل الانحدار:

يستخدم تحليل الانحدار كأسلوب إحصائي كمي في النواحي التالية:

١- دراسة العلاقة بين متغيرين على شكل علاقة داليّه: ص = دالة (س) والتي عن طريقها يمكن معرفة التغير في أحد المتغيرين على أساس تأثره بالمتغير الآخر. أو بعبارة أخرى توقع وتنبؤ سلوك المتغير التابع في ضوء تأثره بالمتغير أو المتغيرات المستقلة.

٢- قياس مدى الارتباط الكلي بين المتغير التابع والمتغير أو المتغيرات المستقلة.

٣- تقدير نسبة تفسير كل متغير مستقل للاختلاف في المتغير التابع .

٤- إجراء سلسلة من الاختبارات الفرضية لأي من العلاقات الثلاثة السابقة[١] .

أنواع تحليل الانحدار:

هناك ثلاثة أنواع رئيسية لتحليل الانحدار هي :

١- الانحدار البسيط Simple Regression : وهو يستخدم لدراسة العلاقة بين متغيرين فقط .

٢- الانحدار الجزئي Partial Regression : وهو يدرس العلاقة بين المتغير التابع وواحد فقط من المتغيرات المستقلة بفرض أن العوامل الأخرى ثابتة (أي بإهمال تأثير العوامل الأخرى) .

٣- الانحدار المتعدد Multiple Regression وهو يحدد مقدار العلاقة بين المتغير التابع وعدد من المتغيرات المستقلة .

تحليل الانحدار البسيط:

يعتمد تحليل الانحدار البسيط على دراسة البيانات الخاصة بمتغيرين فقط ، بهدف التوقع أو التنبؤ بتغير المتغير التابع في ضوء معرفة التغيرات في المتغير المستقل ، وهو ما يطلق عليه اسم العلاقة الدالية . ويعتمد تحليل الانحدار لدراسة العلاقة بين ظاهرتين على تكوين فكرة مبدئية عن نوع هذه العلاقة وقوتها ، وذلك باستخدام ما يعرف بشكل الانتشار .

---

(١) فتحي عبدالعزيز أبو راضي : مقدمة الأساليب الكمية في الجغرافيا ص ٦٦٠ .

تخطيط الانتشار *Scatter Diagram*:

من الأمور الهامة التي يجب البدء بها عادة عند دراسة الارتباط أو الانحدار هي معرفة شكل انتشار الظواهر المدروسة. إن معرفة شكل الانتشار تعطينا فكرة عن نوع الارتباط، ودرجة قوته، ثم تظهر لنا فيما إذا كانت نقط الانتشار قد تقع على خط مستقيم تماماً فيكون الارتباط تاماً، أو قد تنحرف عنه، أو تأخذ شكلاً آخر غير مستقيم.

إن شكل الانتشار يستعمل في العادة لتمثيل أزواج من القيم للظاهرتين المدروستين، فالنقطة تتحدد على أساس قيمة س والقيمة المناظرة لها في ص. ويستعمل الشيء نفسه في تحديد كافة قيم الظاهرة س مع ما يقابلها من قيم ص، والشكل البياني الناتج عن انتشار النقط بين المحورين المتعامدين يسمى: بالشكل الانتشاري.

وهناك صور متعددة للأشكال الانتشارية تختلف باختلاف العلاقة بين قيم س وقيم ص المناظرة لها. ولكنه يمكن القول بأن الشكل الانتشاري لمجموعة أزواج القيم سوف لا يخرج عن الصور السابقة المبينة بالشكل (٩-١) أو صورة قريبة من أي منها.

وبدراسة الشكل الانتشاري للنقط التي تمثل أزواج القيم نستطيع أن نحدد اتجاه العلاقة بين المتغيرين، كما نستطيع أن نكون فكرة عن درجة هذه العلاقة. فالشكل (٩-١-أ) يوضح لنا وجود علاقة طردية قوية، بمعنى أن التغير في أحد المتغيرين يستتبعه تغير بالدرجة نفسها تقريباً وفي الاتجاه نفسه في المتغير الآخر، ويلاحظ من الشكل أن النقط تقع قريبة من خط

يتوسط هذه النقط .

أما الشكل (٩-١-ب) فهو يعرض علاقة طردية ضعيفة بين المتغيرين، في حين أن الشكل (٩-١-ج) يعرض انتشاراً عشوائياً للنقط، وهذا يعطي فكرة عن عدم وجود علاقة بين المتغيرين. وأما الشكل (٩-١-د) فهو يعبر عن علاقة خطية عكسية، أي أن التغير في أحد المتغيرين يستتبعه تغير في الاتجاه المضاد في المتغير الآخر. وأما الشكل (٩-١-هـ) فهو يمثل علاقة غير خطية، بمعنى أن الاتجاه العام للنقط ليس مستقيماً بل يكون على شكل منحنى[١].

خط الانحدار *Regression Line*:

في الشكل السابق (٩-١) الذي يمثل أشكال الانتشار نلاحظ أن الأشكال أ، ب، د. كلها تمثل علاقة خطية، أي أنه يمكن توفيق خط مستقيم يمر بأكبر عدد ممكن من هذه النقط ويتوسط الباقي منها. هذا الخط هو ما يطلق عليه عادة خط الانحدار. إن رسم مثل هذا الخط باليد لا يؤدي المطلوب، ويختلف من شخص لآخر؛ ولهذا فمن الأفضل الحصول عليه بالطرق الجبرية. وسنستخدم الطريقة المعروفة باسم طريقة المربعات الصغرى لتوفيق خط الانحدار.

طريقة المربعات الصغرى *Mehod of Least Squares*:

هي طريقة نستخدمها لتوفيق أفضل خط (مستقيم أو منحنى) لمجموعة

---

(١) عبداللطيف عبدالفتاح وزميله - المدخل في الإحصاء ورياضياته ج١ ص ٤٠٩ - ٤١١.

من البيانات، وهذا الخط هو الخط الذي يكون مجموع مربعات انحرافات النقط عنه أصغر ما يمكن[1].

ومن الواضح أن الخط الذي نوفقه لا يمر بالنقط جميعها إلا في حالات خاصة، وعلى ذلك فتكون هناك بعض النقط التي تنحرف عن هذا الخط. وتتميز طريقة المربعات الصغرى بأنها تعطينا خطاً يكون مجموع مربعات انحراف النقط عنه أصغر ما يمكن.

نفرض أننا نريد توفيق خط مستقيم لمجموعة من القيم (س١، ص١)، (س٢، ص٢) (س٣، ص٣) . . (س ن، ص ن) والتي يوضحها شكل الانتشار (٩-١-ب) ونفرض أيضاً أن أحسن خط مستقيم هو:

ص = أ + ب س

حيث إن ص : في هذه الحالة هو المتغير التابع.

س : هو المتغير المستقل.

أ، ب : ثوابت.

و (ب): كمية ثابتة هي عبارة عن ميل المستقيم على المحور الأفقي (معامل انحدار ص / س) و (أ) هي كمة ثابتة أيضاً وهي الجزء الذي يقطعه المستقيم من المحور الرأسي (محور الصادات).

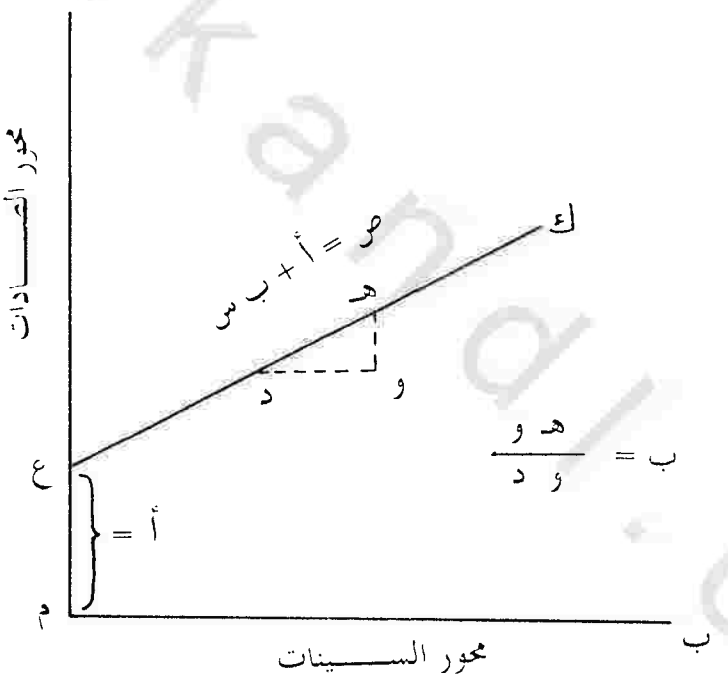
لنفترض أن المستقيم ك ع شكل رقم (١٠-١) هو خط الانحدار الذي يمثل معادلة الخط المستقيم. ففي هذه الحالة فإن الجزء المقطوع من محور

(١) أحمد عباده سرحان - طرق التحليل الإحصائي - ص ٣٥٣ .

الصادات بواسطة المستقيم ك ع هو ع م، وهذه القيمة التي تمثلها هذه المسافة هي الثابت (أ). أما الثابت (ب) فهو عبارة عن ميل المستقيم ك ع عن محور السينات، وميل المستقيم هو الانحدار، وهو عبارة عن هـ و ÷ و د ويعرف هذا بمعامل الانحدار Regression Coefficient.

شكل رقم (١٠-١)

خط الانحدار والثوابت



وعلى هذا الأساس فلتقدير قيمة ص يجب أن نحسب قيمة (ب) وقيمة (أ) ولما كانت قيمة س معروفة، فيصبح بالإمكان معرفة قيمة ص بدلالة س. وهذه القيمة لـ (ص) تمثل التوقع أو التنبؤ المطلوب.

طريقة حساب الثوابت (أ) و (ب):

هناك عدة طرق رياضية لحساب كل من أ و ب اخترنا منها المعادلتين التاليتين:

$$(١)\ ب = \frac{\text{مجـ س ص} - \text{ن } \bar{س}\ \bar{س}}{\text{مجـ س}^{٢} - \text{ن } (\bar{س})^{٢}} \quad \text{أو}$$

$$(٢)\ ب = \frac{\frac{١}{ن}\ \text{مجـ س ص} - \bar{س}\ \bar{ص}}{ع^{٢}\ س}$$

ب = معامل انحدار ص على س

ص = المتغير التابع

س = المتغير المستقل

ن = عدد أزواج القيم للمتغيرين.

$ع^{٢}$ س = التباين، الذي يساوي $\frac{\text{مجـ س}^{٢}}{ن} - (\bar{س})^{٢}$

$\bar{س}\ \bar{ص}$ = المتوسط الحسابي لكل من المتغيرين س و ص.

$أ = \bar{ص} - ب\ \bar{س}$

فإذا عرفنا قيمتي أ و ب يمكننا توقع قيم (ص) في ضوء أي تغير في قيم س، ويرمز لقيم ص المتوقعة بـ (ص٨) أي أن المعادلة المستخدمة في التوقع تصبح:

ص٨ = أ + ب س .

فنفترض أن تقديرات (ب) كانت ٤ , ٠ وتقديرات (أ) كانت ٤ وأن قيمة س تساوي ٥ فإن قيمة ص ٨ تصبح كالآتي :

ص ٨ = أ + ب س

ص ٨ = ٤ + ٤ , ٠ × ٥

ص ٨ = ٤ + ٢

ص ٨ = ٦

أي أن القيمة المتوقعة لـ (ص٨) هي ٦

مثال :

أراد باحث دراسة العلاقة بين دخل الأسرة ومقدار ما تنفقه من هذا الدخل، فما هي المعادلة الخطية للعلاقة بين الدخل والإنفاق؟ أو ما هو خط انحدار الإنفاق على الدخل ص / س، وهل يمكن التنبؤ بمقدار قيمة الدخل عندما نعرف أي قيمة من قيم الإنفاق (مثلاً قيمة الدخل عندما يكون الإنفاق ٧٥٠ ريالاً) علماً بأن مقدار الدخل والإنفاق لتسع أسر هو على النحو التالي (بالآف الريالات) :

الدخل	٣	٤	٥	٧	٩	١٠	٨	٩	١٢
الإنفاق	٣	٤	٥	٦	٨	٦	٧	٧	١١

الحل : لحساب قيمة الثوابت (أ، ب) نقوم بما يلي :

١- نرتب البيانات في جدول كالجدول رقم (١٠-١) حيث نضع قيم المتغيرات للدخل والإنفاق، ثم نربع قيم كل من الدخل والإنفاق ونضرب قيمة الدخل × الإنفاق ونحسب مجاميع القيم في نهاية الجدول.

جدول رقم (١٠-١)

حساب خط الانحدار للدخل والإنفاق

س الدخل	ص الإنفاق	س٢	س ص
٣	٣	٩	٩
٤	٤	١٦	١٦
٥	٥	٢٥	٢٥
٧	٦	٤٩	٤٢
٩	٨	٨١	٧٢
١٠	٦	١٠٠	٦٠
٨	٧	٦٤	٥٦
٩	٧	٨١	٦٣
١٢	١١	١٤٤	١٣٢
٦٧	٥٧	٥٦٩	٤٧٥

٢- إيجاد خط انحدار الإنفاق على الدخل ص / س:

$$\text{س} - = \frac{\text{مجـ س}}{\text{ن}} = \frac{٦٧}{٩} = ٧,٤٤$$

$$\text{ص} - = \frac{\text{مجـ ص}}{\text{ن}} = \frac{٥٧}{٩} = ٦,٣٣$$

$$ب = \frac{\frac{١}{ن} مجـ س ص - \bar{س}\,\bar{ص}}{ع^{٢} س} \quad حيث\ ع^{٢} س = \frac{مجـ س^{٢}}{ن} - (\bar{س})^{٢}$$

$$= \frac{٤٧٥ \div ٩ - ٦,٣٣ \times ٧,٤٤}{٥٦٩ \div ٩ - (٧,٤٤)^{٢}} = \frac{٤٧,٠٩ - ٥٢,٧٨}{٥٥,٣٥ - ٦٣,٢٢}$$

$$= \frac{٥,٦٩}{٧,٨٧} = ٠,٧٢$$

ويمكن الحصول على النتيجة نفسها باستخدام المعادلة الأولى لحساب الثابت (ب).

قيمة أ هي أ = $\bar{ص}$ - ب $\bar{س}$ = ٦,٣٣ - ٠,٧٢ × ٧,٤٤ = ٦,٣٣ - ٥,٣٦ = ٠,٩٧

إذن خط انحدار الاستهلاك على الدخل يكون:

ص = أ + ب س = ٠,٩٧ + ٠,٧٢ س

ولرسم هذا الخط يكفي أن نعين نقطتين، ونستخرج قيمتهما بواسطة المعادلة السابقة، ولما كانت قيم س هي المعلومة فإننا نأخذ أي قيمتين من قيم س في الجدول، ونستخرج قيم ص المناظرة لهما. فعلى سبيل المثال لو أخذنا القيمتين الأولى والثانية من قيم س في الجدول (١٠ : ١) نجد أن قيم ص المناظرة لها هي:

ص = أ + ب س

القيمة الأولى لـ ص = ٠,٩٧ + ٠,٧٢ × ٣

= ٣,١٣

القيمة الثانية لـ ص = ٠,٩٧ + ٤ × ٠,٧٢

= ٣,٨٥

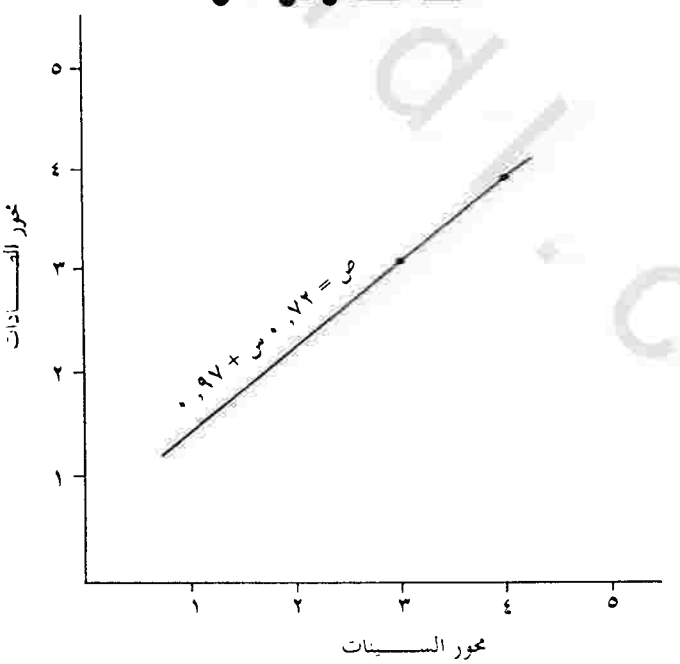
وعلى هذا عندما تكون قيمة س = ٣ فإن قيمة ص المناظرة لها تكون = ٣,١٣ وعندما تكون قيمة س = ٤ فإن قيمة ص المناظرة لها تكون = ٣,٨٥

ولرسم خط الانحدار نقوم بتوقيع النقطة الأولى الناجمة من تقاطع س مع ص، وكذلك النقطة الثانية، ونصل بينهما بخط مستقيم يكون هو خط انحدار الاستهلاك على الدخل (انظر شكل ١٠-٢) ومعادلة هذا الخط هي:

ص = ٠,٧٢ س + ٠,٩٧ أو ص = ٠,٩٧ + ٠,٧٢ س

شكل (١٠-٢)

خط انحدار ص / س



مثال آخر:

البيانات التالية في الجدول رقم (١٠-٢) تمثل مساحة حوض النهر وطوله لعينة من عشرة أنهار لأحد نظم التصريف النهري في منطقة ما. والمطلوب إيجاد المعادلة الخطية للعلاقة بين المتغيرين[١].

جدول رقم (١٠-٢)

مساحة حوض النهر وطوله

المتغير المستقل س (طول النهر الكلي)	المتغير التابع ص (مساحة حوض النهر)	س٢	س ص
٩,٤٠	٤,٠٨	٨٨,٤	٣٨,٤
١٠,٨٥	٤,٥٣	١١٧,٧	٤٩,٢
١٠,٦٩	٤,٩٢	١١٤,٣	٥٢,٦
١٠,٩٧	٦,١٨	١٢٠,٣	٦٧,٨
١١,٩٤	٦,٢٢	١٤٢,٦	٧٤,٣
١٨,٦٩	٦,٥٤	٣٤٩,٣	١٢٢,٢
١٧,٢٣	٧,٥٥	٢٩٦٠٩	١٣٠,١
١٨,٤١	٨,٩٢	٣٣٩,٩	١٦٤,٢
٢٠,١٤	١٠,١٦	٤٠٥,٦	٢٠٤,٦
٢٠,٦٨	١٢,٥٠	٤٢٧,٧	٢٥٨,٥
١٤٩٠,٠٠	٧١,٦	٢٤٠٢,٧	١١٦١,٩

(١) فتحي أبو راضي : مرجع سابق ص (٦٦٧).

١- لحساب الثوابت (أ، ب) يتم ترتيب الجدول السابق بحيث تظهر فيه مربعات قيم المتغير المستقل (س^{٢})، ثم حاصل ضرب المتغير التابع × المتغير المستقل(س×ص)، ثم نستخدم المعادلة الأولى أو المعادلة الثانية من معادلات خط الانحدار لحساب العلاقة الرياضية.

٢- باستخدام المعادلة الأولى وهي:

$$ب = \frac{مجـ س ص - ن \bar{س} \bar{ص}}{مجـ س^{٢} - ن (\bar{س})^{٢}} \text{ فإن}$$

$$ب = \frac{١١٦١,٩ - ١٠ \times ٧,١٦ \times ١٤,٩}{٢٤٠٢,٧ - ١٠ (١٤,٩)^{٢}} = \frac{٩٥,٠٦}{١٨٢,٦} = ٠,٥٢$$

وهذه قيمة معامل الانحدار، ويمكن استخراج القيمة نفسها بواسطة المعادلة الثانية.

ولإيجاد قيمة أ فإن: أ = $\bar{ص}$ - ب $\bar{س}$

أ = ٧,١٦ - ٠,٥٢ × ١٤,٩ = - ٠,٥٨٨

وبهذا تكون معادلة خط الانحدار هي ص = ٠,٥٢ س - ٠,٥٨٨

ومن خلال هذه المعادلة نستطيع أن نتنبأ بأي قيمة من قيم ص، وهي في هذه الحالة مساحة حوض النهر إذا عرفنا طول النهر الكلي. فعلى سبيل المثال كم ستبلغ مساحة حوض النهر إذا كان طول هذا النهر ٣٠ كم؟

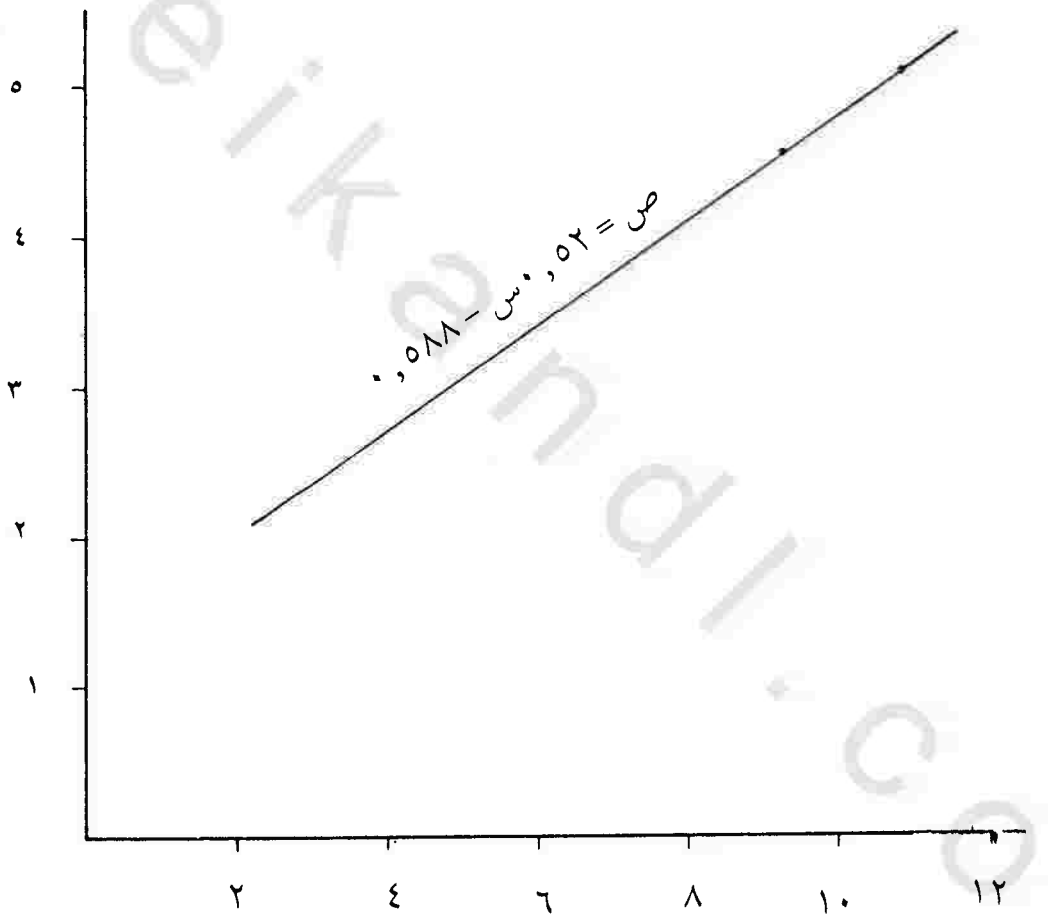
وباستخدام المعادلة السابقة هي:

ص = ٠,٥٢ س - ٠,٥٨٨

ص = ٠,٥٢ × ٣٠ - ٠,٥٨٨

ص = ١٥,٦٠ - ٠,٥٨٨

ص = ١٥,٠١٢ كم٢ مساحة حوض النهر.



شكل رقم (١٠-٣)

خط انحدار ص / س

obeikandl.com

٣- يوضح الشكل رقم (١٠-٣) طريقة رسم خط الانحدار وذلك بتطبيق معادلة الانحدار التي حصلنا عليها وذلك لأي قيمتين من قيم س وعن طريقها نحدد قيمتي (ص) المناظرتين لهما. فمثلاً تكون قيم ص المتوقعة من خلال المعادلة إذا كانت قيمة س هي ٩,٤ على النحو التالي:

ص = أ + ب س

ص٨ = -٠,٥٨٨ + ٠,٥٢ × ٩,٤ = ٤,٣

و بالكيفية نفسها نستخرج قيمة أخرى ص ٨ إذا كانت قيمة س = ١٠,٨٥

ص٨ = -٠,٥٨٨ + ٠,٥٢ × ١٠,٨٥ = ٥,٠٥٤

وعلى شكل بياني نمثل نقطة أ من تقاطع ٩,٤ مع ٤,٣ وكذلك نقطة ب من تقاطع ١٠,٨٥ مع ٥,٠٥٤ ونوصل بين النقطتين أ، ب ونمد هذا الخط من الاتجاهين. وهذا الخط هو خط الانحدار الذي معادلته (انظر شكل ١٠: ٣).

ص٨ = ٠,٥٢ س - ٠,٥٨٨

٤- من الأهمية بمكان حساب مستوى دلالة معامل الانحدار أي الثابت (ب) في المعادلة السابقة ص = أ + ب س، حيث إن مستوى الدلالة يعطي الباحث فكرة عن مدى ترابط العلاقة بين المتغيرين (س، ص) أو بمعنى آخر توضيح ما إذا كان الارتباط قد تم بتأثير الصدفة أم من خلال عوامل أخرى مسببة له.

جدول رقم (١٠-٣)

القيم المتوقعة لمساحة حوض النهر

ص المشاهدة	ص٨ المتوقعة	(ص-ص٨)^{٢}
٤,٠٨	٤,٣	٠,٠٤٨
٤,٥٣	٥,٠٥	٠,٢٧
٤,٩٣	٤,٩٧	٠,٠٠٢
٦,١٨	٥,١٢	١,١٢٤
٦,٢٢	٥,٦٢	٠,٣٦
٦,٥٤	٩,١٣	٦,٧٠٨
٧,٥٥	٨,٣٧	٠,٦٧٢
٨,٩٢	٨,٩٨	٠,٠٠٤
١٠,١٦	٩,٨٨	٠,٠٧٨
١٢,٥	١٠,١٦	٥,٤٧٦
		١٤,٧٤

ومن المعروف في تحليل الانحدار أن هناك نوعان من الانحرافات : النوع الأول انحرافات القيم الأصلية عن القيم المتوقعة بواسطة معادلة الانحدار، والنوع الثاني وهو الانحرافات العشوائية التي لا يستطيع الباحث التكهن بها، وهي ترجع إلى الأخطاء العشوائية.

٥- وللتخلص من الأخطاء العشوائية السابقة لا بد من حساب الخطأ المعياري للتقدير. وحساب الخطأ المعياري يقتضي إجراء ما يلي :

أ - حساب القيم المتوقعة للمتغير التابع ص ٨ من خلال استعمال معادلة خط الانحدار على كل قيمة من قيم المتغير ص على النحو الوارد في الجدول رقم (١٠-٣)

ب- طرح كل قيمة من القيم المتوقعة المحسوبة (ص٨) من قيم (ص) الفعلية المشاهدة على النحو الوارد في الجدول.

جـ- حساب تباين الانحدار من خلال المعادلة التالية:

$$\text{تباين الانحدار} = \sqrt{\frac{\text{مربع مجموع انحرافات الانحدار}}{\text{عدد المفردات} - ٢}}$$

$$= \sqrt{\frac{\text{مجـ}(\text{ص} - \text{ص}\,٨)^{٢}}{\text{ن} - ٢}} = \sqrt{\frac{١٤,٧٤٢}{٨}} = \sqrt{١,٨٤٢٨} = ١,٣٥٧٥$$

د- حساب الخطأ المعياري للتقدير (خ) من خلال المعادلة التالية:

$$\text{خ} = \frac{\sqrt{\frac{\text{مربع مجموع انحرفات الانحدار}}{\text{عدد المفردات} - ٢}}}{\text{مجموع مربعات انحرفات المتغير المستقل عن الوسط الحسابي}}$$

$$\text{خ} = \frac{\sqrt{\frac{\text{مجـ}(\text{ص} - \text{ص}٨)^{٢}}{\text{ن} - ٢}}}{\text{مجـ}(\text{س} - \text{س})^{٢}}$$

ولما كانت (س - س)^{٢} = ١٧٩,٦٤ (ونحصل عليها من طرح الوسط الحسابي من كل قيمة من قيم س ثم نربع هذه القيمة ونجمع جميع القيم معاً)

$$\text{فإن خ} = \frac{١,٣٥٧٥}{١٧٩,٦٤} = ٠,٠٠٧٦$$

هـ - لإجراء اختبار مستوى الدلالة ت لمعامل الانحدار (ب) نستخدم الصيغة التالية :

$$\text{ت} = \frac{\text{ب}}{\text{خ}}$$

حيث ب = معامل الانحدار = ٠,٥٢

خ = الخطأ = ٠,٠٠٧٦

$$\therefore \text{ت} = \frac{٠,٥٢}{٠,٠٠٧٦} = ٦٧,٩٤$$

و - لما كانت قيمة ت في جدول توزيع (ت) بمستوى دلالة ٠,٠١ بدرجات حرية (ن - ٢) أو ١٠ - ٢ = ٨ هي ٣,٣٦ وحيث إن قيمة (ت) المحسوبة ٦٧,٩٤ أكبر من قيمة (ت) في الجدول؛ لذلك نرفض الفرضية القائلة بأنه لا توجد أية علاقة بين المتغيرين (س) و (ص)، أو بعبارة أخرى نقبل الفرض البديل وهو أن هناك علاقة بين المتغيرين في مستوى دلالة ٠,٠١ أي أن هناك احتمال قدرة ٩٩٪ أن لا تكون العلاقة قد حدثت بفعل الصدفة أو العشوائية.

٦- لتحديد قوة العلاقة بين المتغير التابع والمتغير المستقل، لا بد لنا من معرفة معامل التحديد الذي هو مربع معامل الارتباط، فإذا كان معامل الارتباط = ٠,٨ فإن معامل التحديد تساوي $(٠,٨)^{٢}$ = ٠,٦٤ وهذا يعني أن ٦٤٪ من الاختلافات في قيم ص يمكن تفسيرها من خلال المتغير س. وبناء عليه إذا كانت العلاقة بين س، ص قوية فإن ذلك يعني ارتفاع قيمة معامل التحديد. وتحسب قيمة معامل التحديد بالمعادلة التالية :

$$ر^{٢} = \frac{أ\,[مجـ (س - \bar{س})(ص - \bar{ص})]}{مجـ (ص - \bar{ص})^{٢}}$$

وإذا رجعنا إلى المثال السابق الخاص بالعلاقة بين مساحة حوض النهر وطوله (انظر جدول ١٠ : ٤) وحسبنا معامل التحديد لهذين المتغيرين وجدنا أن :

$$ر٢ = \frac{٩٩,١٣ \times ٠,٥٢}{٦٧,٢٣} = \frac{٥١,٥٥}{٦٧,٢٣} = ٠,٧٦٧$$

أي أن ٠,٧٧ (٧٧٪) من قيمة الاختلاف في قيم (ص) يمكن تفسيرها بالاختلاف في قيم (س) وأن ٠,٢٣ (٢٣٪) من هذه الاختلافات ترجع إلى عامل الصدفة أو إلى أخطاء عشوائية أو إلى عوامل أخرى غير ظاهرة .

جدول رقم (١٠-٤) طريقة حساب معامل الارتباط بين طول النهر (س) ومساحة حوض (ص) وكذلك حساب معامل التحديد

س	ص	س-$\bar{س}$	ص-$\bar{ص}$	(س-$\bar{س}$)^{٢}	(ص-$\bar{ص}$)^{٢}	(س-$\bar{س}$)(ص-$\bar{ص}$)
٩,٤٠	٤,٠٨	-٥,٥٠	-٣,٠٨	٣٠,٢٥	٩,٤٩	١٦,٩٤
١٠,٨٥	٤,٥٣	-٤,٠٥	-٢,٦٣	١٦,٤٠	٦,٩٢	١٠,٦٥
١٠,٦٩	٤,٩٢	-٤,٢١	-٢,٢٤	١٧,٧٢	٥,٠٢	٩,٤٣
١٠,٩٧	٦,١٨	-٣,٩٣	-٠,٩٨	١٥,٤٤	٠,٩٦	٣,٨٥
١١,٩٤	٦,٢٢	-٢,٩٦	-٠٩٢	٨,٧٦	٣,٦٩	٢,٧٢
١٨,٦٩	٦,٥٤	٣,٧٩	-٠,٦٢	١٤,٣٦	٠,٣٨	,٢٣٥
١٧,٢٣	٧,٥٥	٢,٣٣	٠,٧٩	٥,٣٤	٠٣,١٥	٠,٩٠
١٨,٤١	٨,٩٢	٣,٢٤	١,٧٦	١٠,٥٠	٣,١٠	٥,٧٠
٢٠,١٤	١٠,١٦	٥,٢٤	٣,٠٠	٢٧,٤٦	٩,٠٠	١٥,٧٢
٢٠,٦٨	١٢,٥٠	٥,٧٨	٥,٣٤	٣٣,٤١	٢٨,٥٢	٣٠,٨٧
١٤,٠٠	٧٠,٦٠	-	-	١٧٩,٦٤	٦٧,٢٣	٩٩,١٣

الانحدار الخطي المتعدد *Multiple Linear Regression*:

لقد تم حتى الآن مناقشة الانحدار الخطي البسيط لمتغير تابع وآخر مستقل؛ ونظراً لأن واقع الحياة، يقوم على تأثر أي ظاهرة بأكثر من متغير مستقل؛ لذلك لا بد من توسيع أسلوب الانحدار البسيط ليشتمل على انحدار المتغير التابع مع أكثر من متغير مستقل.

تعرف معاملات الانحدار في تحليل الانحدار المتعدد باسم: معاملات الانحدار الجزئية Partial Regression Coefficients، وذلك تمييزاً لها عن معامل الانحدار البسيط الذي سبق ذكره.

وكما ذكرنا من قبل فإن معامل الانحدار البسيط يقيس معدل التغير المتوقع في المتغير التابع نتيجة تأثير المتغير المستقل. أما معامل الانحدار الجزئي في تحليل الانحدار المتعدد فيقيس معدل التغير المتوقع في المتغير التابع نتيجة تأثير المتغير المستقل مع بقاء أثر بقية المتغيرات المستقلة الأخرى ثابتاً.

وهناك طرق لحساب الانحدار المتعدد أبرزها طريقة الانحدار التدريجي Stepwise Regression التي تعطي نسبة تفسير كل متغير مرتبة حسب أهمية المتغير في التحليل (أي إن التحليل يبدأ بتحديد أهم متغير وينتهي بالمتغير الأقل أهمية في تفسير الاختلاف الذي يحدث في المتغير التابع).

ورغم وجود بعض الاختلاف في العمليات الحسابية لكل نوع من أنواع التحليل السابقة إلا أنها تسير في الاتجاه نفسه تقريباً. وهو تحديد معادلة خط الانحدار للعلاقة بين عدة متغيرات، وهذه تحتاج إلى عمليات حسابية

معقدة وطويلة وخاصة إذا زاد عدد الحالات Caes والمتغيرات Variables ، وقد ظهرت أهمية الحاسوب كعامل مساعد وهام للقيام بمثل هذه التحليلات . وكان لتقدم الدراسات الخاصة بتنظيم معالجة البيانات أن تعددت البرامج الجاهزة التي تستخدم في مثل هذا النوع من التحليل الإحصائي ، غير أن أشهرها ما يعرف باسم حزمة البرامج الإحصائية للعلوم الاجتماعية Spss التي تستخدم في هذا المجال .

وتستند طريقة حساب الانحدار المتدرج من خلال حزمة البرامج الإحصائية للعلوم الاجتماعية Spss ، على أنه في البدء يتم حساب ارتباط ثنائي بين أحد المتغيرات المستقلة والمتغير التابع . ولا يكون اختيار هذا المتغير المستقل عشوائياً بل يتم الاختيار على أساس أن المتغير المستقل هو أقوى المتغيرات ارتباطاً مع المتغير التابع ويفسر أكبر قدر منه ، ثم يدخل متغير مستقل آخر ، وهذا المتغير المستقل الآخر يكون بدوره أقوى المتغيرات بعد المتغير الأول من حيث علاقته بالمتغير التابع ، ثم يدخل المتغير الثالث الذي يفسر أكبر كمية ممكنة مما تبقى من تباين في المتغير التابع . وهكذا تدخل المتغيرات تباعاً حسب أهميتها في تفسير التباين في المتغير التابع . ويمكننا هذا التحليل من معرفة أهمية المتغيرات الداخلة إلى التحليل ومدى قوتها في تفسير النتائج المتحصلة من معادلة الانحدار ، حيث يتم استبعاد المتغيرات غير المهمة وتبقى المتغيرات الهامة في نموذج الانحدار . وتكون هذه المعادلة على النحو التالي :

$$س = أ + ب_{٢} س_{٢} + ب_{٣} س_{٣} + ب_{٤} س_{٤} + \ldots ب_{ن} س_{ن}$$

حيث (أ) هي قيمةالقاطع، و (ب) هي معاملات الانحدار للمتغيرات المستقلة، و($\hat{س}_1$) المتغير التابع س٢، س٣، س٤ . . . المتغيرات المستقلة.

فعلى سبيل المثال: إذا كان لدينا ثلاثة متغيرات هما س١، س٢، س٣ فإن معادلة الانحدار تكون على النحو التالي:

$$\hat{س}_8 = أ + ب_2 س_2 + ب_3 س_3$$

وكما هو الحال في الانحدار الثنائي البسيط، يستخدم في حساب معاملات الانحدار، وفي تقدير القيم المتوقعة للمتغير التابع، أسلوب المربعات الصغرى، إلا أن العمليات الحسابية اللازمة لذلك عمليات طويلة تستخدم فيها المصفوفات الرياضية. ولا مجال لتوضيحها هنا. وسنكتفي بتوضيح المفاهيم الأساسية لأسلوب الانحدار المتعدد معتمدين على الحاسوب في تنفيذ العمليات الحسابية.

مثال:

أجريت دراسة لمعرفة أثر بعض العوامل الاقتصادية والاجتماعية على مستويات خصوبة المرأة في مدينة الرياض[1]. وقد اختيرت المتغيرات التالية لدراسة أثرها على مستوى الخصوبة الفعلية مقدرة بعدد المواليد كمتغير تابع، وقد استخدمت المتغيرات التالية كمتغيرات مفسرة لمستويات الخصوبة أو متغيرات فعلية وهي حجم الأسرة (س١)، دخل الأسرة (س٢)، المستوى التعليمي للمرأة (س٣)، الحالة العملية للمرأة(س٤)، مكان النشأة

(١) بدرية الجندان: مستويات الخصوبة واتجاهاتها في مدينة الرياض - دراسة في جغرافية السكان. رسالة دكتوراه غير منشورة مقدمة إلى كلية التربية للبنات، بالرياض، ١٤١١هـ.

(س٥)، عـدد وفـيـات الأطفـال للمـرأة (س٦)، عـمـر المرأة عند الزواج (س٧)، مدة الحياة الزوجية (س٨)، مدة الرضاعة (س٩)، ممارسة وسائل منع الحـمل (س.١)، العـمل الحـالي للزوجـة (س١١)، ونوع الرضـاعـة (س١٢).

اختارت الباحثة أسلوب تحليل الانحدار المتدرج وذلك لتوفير العدالة في المعاملة الإحصائية حيث تدخل إلى هذا التحليل الإحصائي المتغيرات التي تفسر أكبر قيمة من التباين تباعاً حسب أهميتها، ثم تستبعد المتغيرات التي ليس لها دلالة إحصائية، وقد أدخلت البيانات إلى الحاسوب حيث جاء متغير حجم الأسرة (س١) في مقدمة المتغيرات الداخلة في تحليل الانحدار المتدرج حيث فسر وحده مربع معامل ارتباط مقداره ٦٥١٩ , ٠ ثم تلاه متغير مدة الحياة الزوجية (س٨) حيث فسرا معاً ما مقداره ٦٨٢١ , ٠ من التباين، ثم دخل إلى التحليل متغير عمر الزوجة الحالي (س١١) حيث فسر الثلاثة معاً ٦٨٧٣ , ٠ من التباين، ثم دخل متغير العمر عند الزواج (س٧)، ثم تلاه متغير الرضاعة (س١٢) حيث فسرت هذه المتغيرات جميعاً ٦٩٢٤ , ٠ وبذا أصبح لدينا أفضل معـدالة للانحـدار المتـدرج على النحـو التالي:

س٨= -٨ , ١ + ٦ , ٠ س١ + ٠٢ , ٠ س٨ - ٠٧ , ٠ س١١ - ٠٦ , ٠ س٧ + ١٤ , ٠ س١٢

باعتبار أن س٨ هي قيمة المتغير التابع المتنبأ به وهو عدد المواليد.

أما معامل التحديد R2 فكان ٦٩ , ٠

أما ثابت الانحدار (أ) فيساوي - ٨ , ١

ومن خلال المعادلة السابقة نستطيع أن نتنبأ بعدد المواليد استناداً إلى متغيرات حجم الأسرة ومدة الحياة الزوجية، ثم العمر الحالي للزوجة، ثم عمر الزوجة عند الزواج، ثم نوع الرضاعة. لقد فسرت هذه المتغيرات ما يقرب من ٦٩٪ من التباين في المتغير التابع (معامل التحديد). أما ما تبقى من التباين وهو ٣١٪ فقد بقي غير مفسر. ولم يكن لبقية المتغيرات أثر في عملية الخصوبة فقد استبعدت من التحليل لأنها غير مهمة.

حساب الانحدار البسيط عن طريق الحاسوب:

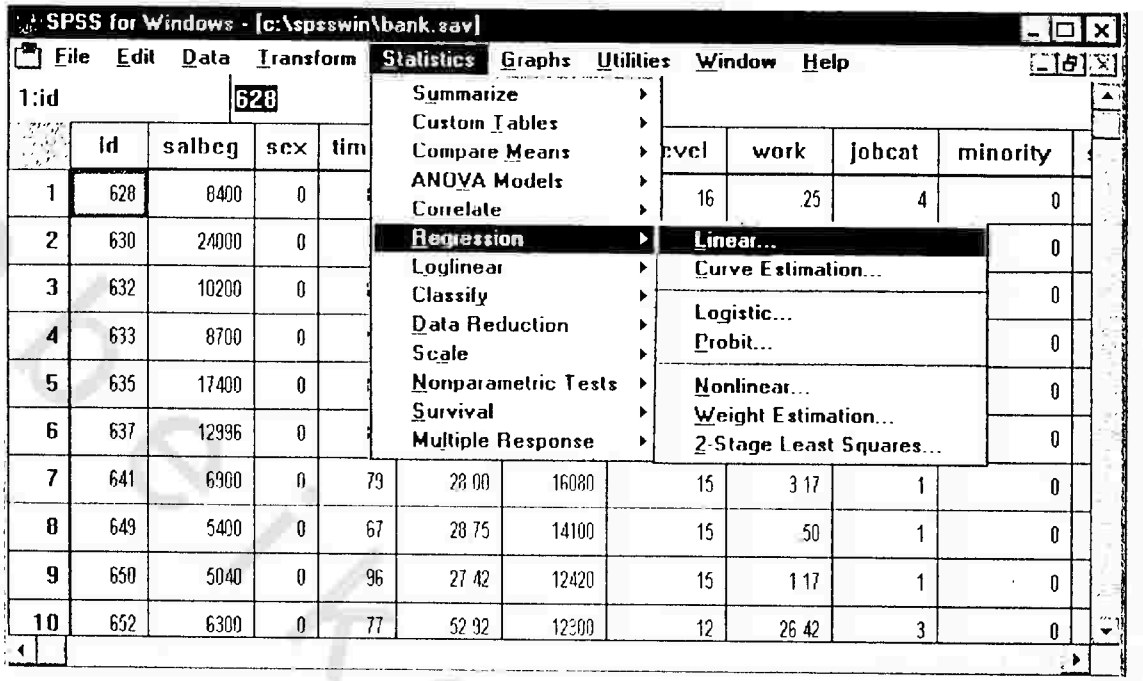
سنوضح استخدام الحاسوب في تنفيذ الانحدار الخطي البسيط على أحد متغيرات الملف pop. sav الخاص بسكان بعض البلاد العربية السابق. وسيكون المتغير التابع هو معدل الوفيات (Death) أما المتغير المستقل فهو نسبة سكان المدن (Urban) ولعمل التحليل نقوم بما يلي:

١- نطلب الملف pop. sav من قائمة الملفات المخزنة في برنامج Spss بالطريقة التي تحدثنا عنها سابقاً.

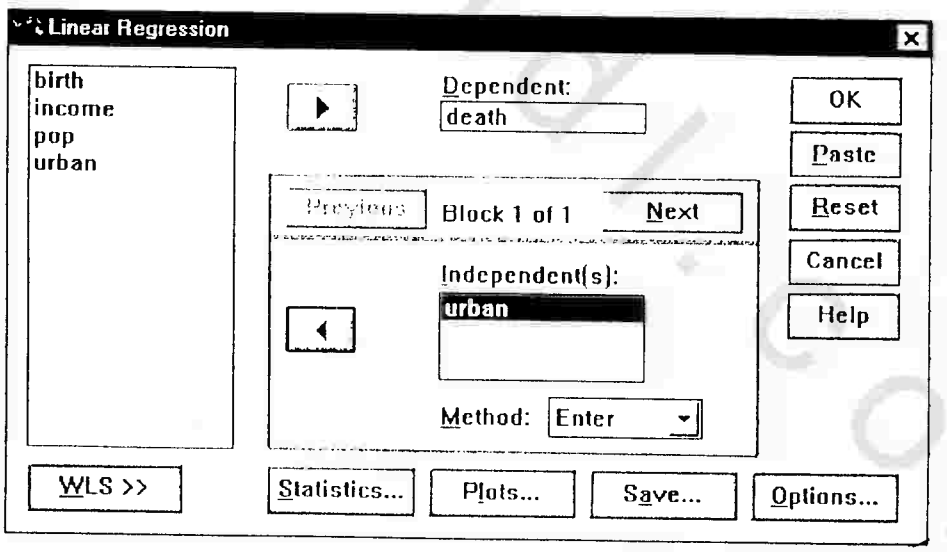
٢- يمكن طلب أسلوب الانحدار الثنائي البسيط من statistics في قائمة شريط القوائم، واختيار موضوع انحدار Regression فتظهر شاشة ثانوية خاصة تتضمن أساليب متعددة لتحليل الانحدار كالانحدار الخطي (Linear) وغير الخطي Nonlinear واللوجستيكي logistic وغيرها (انظر شكل رقم ١٠-٤) نختار الأسلوب (Linear) فتظهر شاشة جديدة تتضمن المتغيرات التي يتضمنها الملف pop. sav جميعها (شكل رقم ١٠-٥).

٤- ننقل المتغير التابع الوفيات (Death) إلى المكان المخصص للمتغير التابع وعنوانه (Dependent) وننقل المتغير المستقل (Urban) إلى المكان المخصص للمتيغرات المستقلة (Independent).

٥- ننفذ البرنامج فور الانتهاء من الخطوة السابقة بالأمر (OK).



شكل (٤:١٠) الأوامر الخاصة بطلب الانحدار البسيط



شكل (٥:١٠) البيانات الخاصة بإجراء تحليل الانحدار البسيط

تفسير نتائج الانحدار من خلال الحاسوب: يظهر الجدول التالي رقم (١٠-٥) نتائج تحليل الانحدار لمتغيري نسبة الوفيات ونسبة سكان المدن. ويشتمل هذا الجدول على ثلاثة حقول هي:

١- الحقل الأول ويشمل: معامل الارتباط الكلي (Multiple R) وهو يقيس الارتباط بين المتغير التابع (الوفيات) والمتغير المستقل (نسبة سكان المدن). ويبلغ في هذا المثال ٨٣, ٠ بينما يبلغ معامل التحديد (R Square) ٧٠, ٠ وهو مربع معامل الارتباط ٨٣, مما يدل على أن ٧٠٪ من تباين الوفيات يمكن تفسيره من خلال متغير (سكان المدن)، وهناك تعديل بسيط يجريه الحاسب لمزيد من الدقة في معامل التحديد فيقوم بحساب معامل التحديد المعدل (Adjusted R Square) وهو في حدود ٦٥٪ وذلك بحذف الأخطاء العشوائية؛ ولذلك يكون معامل التحديد المعدل أكثر دقة من معامل التحديد.

وبجانب معامل التحديد المعدل يتم حساب الخطأ المعياري للتقدير، وهو كما ذكرنا سابقاً يساعدنا على حساب مستوى الدلالة، وتبلغ قيمة الخطأ المعياري Standard Error ١, ٧٣ .

٢- الحقل الثاني: ويضم حساب التباين، وذلك لاختبار الدلالة الإحصائية لنموذج الانحدار، وتظهر في هذا الحقل درجات الحرية لمجموع مربعات الانحدار، وكذلك درجات الحرية لمربعات البواقي، وكذلك تظهر القيمة ف (F) ومستوى الدلالة الإحصائية لها.

٣- الحقل الثالث: ويضم مكونات معادلة الانحدار والمتغيرات المكونة لها، ويشمل هذا الحقل خمس خانات:

أ- الخانة الأولى وتضم اسم المتغير المستقل الذي رمزنا له بمعادلة الانحدار (س) كما تضم الثابت (constant) (أ).

ب- الخانة الثانية: تضم معاملات الانحدار. حيث نجد قيمة ب = ١٠٥ , ٠ وقيمة الثابت (أ) ١٢ , ٨١ وعلى هذا فمعادلة خط الانحدار هي:

ص٨ = أ + ب س

جدول (١٠-٥)

نتائج تحليل الانحدار

```
* * * *  M U L T I P L E   R E G R E S S I O N  * * * *

Listwise Deletion of Missing Data

Equation Number 1     Dependent Variable..   DEATH

Block Number  1.  Method:  Enter      URBAN


Variable(s) Entered on Step Number
   1..     URBAN


Multiple R            .83671
R Square              .70008
Adjusted R Square     .65009
Standard Error       1.73002

Analysis of Variance
                 DF        Sum of Squares      Mean Square
Regression        1             41.91729         41.91729
Residual          6             17.95771          2.99295

F =       14.00533       Signif F =  .0096


------------------ Variables in the Equation ------------------
   (أ)                  (ب)         (حـ)         (ى)          (هـ)     (و)
Variable                B          SE B         Beta            T     Sig T

URBAN           -.105585      .028213    -.836708      -3.742    .0096
(Constant)     12.815691     1.826479                    7.017    .0004


End Block Number   1   All requested variables entered.
```

ص٨ = ١٢,٨١ - ١٠٥و . س

ج - الخانة الثالثة وتضم الأخطاء المعيارية لمعاملات الارتباط.

د- الخانة الرابعة: تمثل معاملات الانحدار القياسية (Beta). وهي تمثل نسبة التغير في التغير التابع عندما نستخدم في حساب الانحدار القيم المعيارية للمشاهدات بدلاً من المشاهدات الفعلية. وتمتاز معاملات الانحدار القياسية (بيتا B) عن معاملات الانحادر العادية (B) بأنها متحررة من وحدات القياس؛ ولهذا فإنها تستخدم لمقارنة معدلات التغير الناتجة عن المتغيرات المستقلة.

هـ- الخانة الخامسة: تظهر قيمة ت (T) لمعامل الانحدار (ب).

وكذلك لثابت Constant (أ) وهما - ٣,٧٤ و ٧,٠١ على التوالي.

و - الخانة السادسة: تظهر مستويات المعنوية الخاصة بمعاملات الانحدار، فمستوى المعنوية لقيمة الثابت (ب) هي ٠,٠٠٩ وهي ذات دلالة أكيدة بين الوفيات وبين سكان المدن، وكذلك أيضاً نقطة التقاطع (أ) ذات مستوى دلالة ٠,٠٠٠٤ مما يدل على وجود علاقة قوية وينفي فرضية العدم، ويثبت وجود علاقة بين المتغيرات.

افتراضات نموذج تحليل الانحدار(*): يعد نموذج الانحدار من أكثر الأساليب الإحصائية شيوعاً بين الجغرافيين، وذلك لسهولته وملاءمته لتحليل كثير من المسائل الجغرافية، غير أن لهذا التحليل العديد من الافتراضات التي يجب تحقيقها حتى تكون نتائجه دقيقة وصحيحه، ونظراً لأهمية هذه الافتراضات نورد ملخصاً بأهمها وهي:

١- طبيعة المتغير التابع:

يجب أن يكون المتغير التابع متغيراً كمياً، إذ لا يصلح أسلوب المربعات

(*) للتوسع في هذا الموضوع انظر نعمان شحادة: الأساليب الكمية في الجغرافيا باستخدام الأسلوب، دار صفاء للنشر والتوزيع عمان، الأردن، ١٩٩٧م، ص ٣٩٠ - ٤٠٢.

الصغرى Least Square Method في اشتقاق معاملات الانحدار عندما يكون المتغير التابع نوعياً، ويستخدم في هذه الحالة أسلوب آخر، وهو أسلوب الاحتمالات الكبرى Maximum Liklihood Method ويعرف أسلوب الانحدار عندئذ بالانحدار اللوجيستيكي Logistic Regression، وهو أسلوب يعمل على اشتقاق معاملات انحدار مختلفة عن معاملات الانحدار التي يتم اشتقاقها بأسلوب المربعات الصغرى، فالميزة الرئيسة لمعاملات المربعات الصغرى هي أنها تجعل تشتت النقاط حول خط الانحدار أقل منه عن أي خط آخر. أما معاملات الاحتمالات الكبرى فتعمل على أن تكون احتمالات القيم الموقعة للمتغير التابع التي يتم اشتقاقها بهذا الأسلوب، أكبر من احتمالات أية قيم، يتم اشتقاقها بأي أسلوب آخر.

يفترض، إذن، أن يكون المتغير التابع في الانحدار الخطي متغيراً كمياً ولا يجوز غير ذلك. أما المتغيرات المستقلة فيجوز أن يكون بعضها كمياً والبعض الآخر نوعياً[1].

٢- العلاقة الخطية Linearity:

يشترط نموذج الانحدار الخطي أن تكون العلاقة بين المتغير التابع والمتغير المستقل علاقة خطية، أي أن تكون نسبة التغير في (ص) ثابتة، مقارنة مع نسبة التغير في (س). أما إذا استخدمناه لتمثيل علاقة غير خطية، فإننا نحصل على تقديرات غير دقيقة لقيم (ص) المتوقعة كما أن مستوى معنوية بعض معاملات الانحدار المتعدد ينخفض كثيراً بحيث لا يصبح لها دلالة إحصائية بالرغم من أهميتها من الناحية العملية، وإزاء ذلك لا بد من تحويل

(١) نعمان شحاده: مرجع سابق ص ,٣٩٢

العلاقات غير الخطية إلى علاقات خطية، ويكون ذلك باستخدام لوغرتم أحد المتغيرين أو كليهما. ومن عيوب هذه الطريقة لعمليات التحويل اللوغارتمي Logarethmic Transformation بالرغم من مزاياها الكثيرة أنها تجعل فهم وتفسير معاملات الانحدار صعباً وغير مباشر؛ ولهذا يقترح بعض الباحثين. استخدام اللوغرتمات الطبيعية في عملية التحويل بدلاً من اللوغاريتمات العشرية، إذ إن استخدام ذلك النوع من اللوغريتمات يجعل تفسير معامل الانحدار مباشراً ويعيد إليه مدلوله الأصلي[1].

٣- التنبؤ بالمتغير التابع:

يصلح نموذج الانحدار الخطي للتنبؤ بالمتغير التابع من خلال متغير مستقل واحد كما هو الحال في الانحدار الثنائي، أو من خلال مجموعة من المتغيرات المستقلة، كما هو الحال، في نموذج الانحدار المتعدد، إلا أن مجال التنبؤ يجب أن يكون محصوراً في نطاق المتغيرات المستقلة فقط، فإذا كان المدى المطلق للمتغير المستقل يمتد على سبيل المثال من (١) إلى (١٠٠) فلا يجوز لنا أن نتوقع قيمة (ص) التي تقابل أية قيمة للمتغير المستقل تقل عن واحد، أو تزيد عن المائة، إذ إننا لا نعرف شيئاً عن طبيعة العلاقة بين المتغير التابع والمتغير المستقل خارج ذلك المدى، فربما تكون قد تغيرت وأصبحت غير خطية أو حدث بها انقطاع.

٤- القيم المتطرفة Outliers:

يعطي نموذج الانحدار أفضل النتائج عندما تخلو البيانات الخاصة بأي

(١) نعمان شحادة: مرجع سابق ص ٣٩١.

من المتغيرين المستقل والتابع من القيم المتطرفة، إذ إن تأثير تلك القيم على تقدير معاملات الانحدار كبير، كما أنها تضاعف الخطأ المعياري لخط الانحدار، وفي بعض الحالات الشاذة تتمكن نقطة متطرفة واحدة من تغيير اتجاه خط الانحدار تغييراً كلياً.

٥- التوزيع الطبيعي للبواقي Normality:

يستحسن قبل أن نبدأ بمناقشة هذا الافتراض أن نوضح المقصود بالتوزيع الطبيعي للبواقي، فمن الممارسات الشائعة في الدراسات الجغرافية أن يختبر الباحثون مدى توافق التوزيع التكراري لبياناتهم مع التوزيع الطبيعي باعتبار ذلك شرطاً من الشروط الرئيسية لتطبيق نموذج الانحدار، علماً بأن الافتراض الفعلي لذلك النموذج هو أن يُكوّن التوزيع التكراري للبواقي الخاصة بكل واحدة من قيم المتغير المستقل توزيعاً طبيعياً، فلو تحقق هذا الشرط فإن التوزيع الطبيعي للبيانات الأصلية يتحقق دائماً، لكن توافق التوزيع التكراري للمشاهدات الأصلية مع التوزيع الطبيعي لا يعني بالضرورة أن التوزيع التكراري للبواقي توزيعاً طبيعياً.

ولفهم المقصود بالتوزيع التكراري للبواقي تصور متغيراً مستقلاً (س) تتكرر فيه قيم معينة كالقيم ٥ ,٢و٠ ,٣ و٥ ,٣وغيرها، أما القيم الفعلية للمتغير التابع التي تقابل هذه القيم المكررة من (س) فمختلفة، وبما أن نموذج الانحدار يعطي قيمة متوقعة واحدة لكل قيمة من (س) مهما كان تكرارها، فإن تلك القيمة تمثل في الواقع متوسطاً للقيم المتوقعة، التي تقابل كل مشاهدة من (س) مهما كان تكرارها، ويمثل الفرق بين القيم الفعلية (ص) التي تقابل كل مشاهدة من (س) والقيمة المتوقعة لها توزيعاً من

البواقي Destributions of Errors والشرط الرئيسي لنموذج الانحدار أن يكون ذلك التوزيع توزيعاً طبيعياً.

والحقيقة أن تكرار قيم المتغير التابع في العينات الجغرافية أمر ليس شائعاً، ولهذا فإن هذا الافتراض لا يتم التحقق منه دائماً، ويعد من الافتراضات الصعبة التي يمكن تجاوزها ضمن حدود معقولة.

٦- ثبات التباين *Homoscedasticity*:

يشترط نموذج الانحدار أن يكون تباين التوزيعات التكرارية للبواقي التي تحدثنا عنها في الفقرة السابقة متساوياً، إذ إن اختلافه سواء بالزيادة أم بالنقصان يجرد معاملات الانحدار الجزئي من ميزتها الرئيسة، وهي تحقيق الحد الأدنى من مجموع المربعات، كما أنه يزيد الخطأ المعياري لخط الانحدار، ويقلل من مستوى معنوية معاملات الانحدار نفسها، ويمكننا التحقق فيما إذا كان التباين ثابتاً أم لا بإعداد شكل انتشار يبين العلاقة بين القيم المتوقعة من (ص) والبواقي، فإذا كان انتشار النقاط عشوائياً، فإن شرط ثبات التباين قد تحقق، أما إذا وجدنا تغيراً منهجياً في نمط الانتشار، فإن التباين لا يكون ثابتاً.

وبالرغم من الأهمية الكبيرة لهذا الافتراض فإنه لا يتحقق في كثير من البيانات الجغرافية خاصة في العينات التي تقل فيها نسبة البيانات الصغيرة مقارنة مع البيانات الكبيرة. كالبيانات الخاصة بعدد سكان المدن، مقارنة مع عدد المشتغلين منهم بمهنة معينة، وعدد الأنهار من رتب مختلفة في الأحواض النهرية وغيرها[1].

(١) نعمان شحادة: مرجع سابق ص ٣٩٩ - ٤٠٢.

٧- الارتباط الذاتي *Autocorrelation*:

المقصود بالارتباط الذاتي هو ارتباط بيانات المتغير المستقبل بعضها ببعض. ولنضرب على ذلك الأمثلة التالية: فعدد سكان دولة الإمارات عام ١٩٩٢م الذي يقدر بحوالي ١,٢٦٥,٠٠٠ نسمة غير مستقل عن عدد سكانها في العام الماضي أو العام الذي سبقه، إذ إن عدد سكان الأعوام السابقة متضمن في عدد سكان هذا العام، وهذا النوع من البيانات بيانات بها قدر كبير من الإرتباط الذاتي. ومن البيانات المناخية التي يوجد بها قدر كبير من الارتباط الذاتي درجة الحرارة خلال كل ساعة من اليوم، والمعدل الشهري لدرجة الحرارة وغير ذلك.

يشترط نموذج الانحدار عدم وجود ارتباط ذاتي في بيانات المتغير المستقل، أي أن تكون البيانات مستقلة عن بعضها، ولكن البيانات الجغرافية بالذات يكثر فيها وجود الارتباط الذاتي.

٨- عدم ارتباط المتغيرات المستقلة *(Collinearity)*:

يختص هذا الشرط بنموذج الانحدار المتعدد، إذ يفترض أن نراعي عند اختيار المتغيرات المستقلة في أسلوب الانحدار المتعدد ألا تكون تلك المتغيرات مترابطة مع بعضها، إذ إن وجود ذلك الارتباط يجرد معاملات انحدار بعض تلك المتغيرات من دلالته الإحصائية، ويعكس اتجاهها أحياناً، فتتحول بعض المعاملات من طردية إلى عكسية، ومن عكسية إلى طردية.

ويمكن التغلب على مشكلة ارتباط المتغيرات المستقلة في نموذج الانحدار المتعدد باستخدام أسلوب التحليل العاملي Principal components

analysis حيث يتم استبدال المتغيرات بعدد من العوامل التي لا يوجد بينها ارتباط، كما يمكن التغلب على المشكلة باستثناء بعض المتغيرات ذات الارتباط القوي من التحليل، إلا أن لكل من الحالين المذكورين عيوبه التي لا يتسع المجال هنا لبحثها والتي لا بد من الحذر منها عند تفسير النتائج[1]

(١) نعمان شحادة: مرجع سابق ص ٣٩٩ - ٤٠٢ .

obeikandl.com

أسئلة وتطبيقات

س١ : قام أحد الجغرافيين بدراسة العلاقة بين عدد ساعات العمل، وعدد القطع المنتجة من أحد المصانع، وقد سجل الباحث البيانات التالية للمتغير (س) والذي يعبر عن عدد ساعات العمل، والمتغير (ص) الذي يعبر عن عدد القطع المنتجة في مقابل قيم (س) المعطاة وذلك لعينة من سبعة عمال من عمال المصنع المذكور فكانت القراءات كالآتي :

(س)	١	٢	٣	٤	٥	٦	٧
(ص)	٧	١٤	١٧	٢٤	٣٣	٣٩	٤١

والمطلوب :

١) إيجاد معامل انحدار ص على س واستنتاج معادلة الانحدار .

٢) رسم شكل الانتشار

٣) من خلال معادلة الانحدار تحسب قيم ص المتنبأ بها (ص٨) لكافة قيم (س) المعطاة .

٤) أجري تحليل الانحدار بواسطة الحاسوب باعتبار س المتغير المستقل (ص) هو المتغير التابع .

س٢ : يقوم أحد الباحثين بدراسة تأثير أحد أنواع المغذيات على وزن نوع من الحيوانات، وإذا كانت (س) تمثل الوزن المبدئي للحيوان وكانت (ص) تمثل الزيادة في وزنه . وبإجراء التجربة على ستة من هذه الحيوانات حصلنا على النتائج التالية :

(س)	٣	٣	٥	٤	٦	٧
(ص)	١٤	١٢	٢٣	٣١	٢٢	٢٥

١- ارسم شكل الانتشار.

٢- أوجد معادلة انحدار ص / س بعد حساب قيم الثوابت (أ) و(ب).

٣- استخدم المعادلة لتقدير الزيادة في وزن حيوان وزنه المبدئي ٢ كغم.

٤- احسب معامل التحديد.

٥- احسب الخطأ المعياري للتقدير، وما هي دلالة هذا المؤشر الإحصائي.

٦- احسب قيمة ت واستخرج القيمة الجدولية ثم اختبر الفرضية الخاصة بالعلاقة بين المتغيرين في مستوى دلالة ٠,٠١ و ٠,٠٥.

س٣: المعادلة التالية ص = ٠,٣ س + ٠,٩٥ هي معادلة خط الانحدار التي تمثل العلاقة بين البعد عن مركز المدينة (س) وعدد الأسواق المركزية (ص). احسب عدد الأسواق المركزية حينما يكون البعد عن مركز المدينة (٤) كم، (٦) كم، (٨) كم، (١٠)كم.

س٤: فيما يلي البيانات الخاصة بستة أشخاص تم اختيارهم عشوائياً:

ضغط الدم	٨٩	٨١	١٠٠	٨٥	٩٠	١٠٠
الوزن	٦٠	٧٠	٨٠	٧٥	٧٧	٩٠
الطول	١٧٠	١٦٨	١٧٠	١٧٥	١٧٤	١٧٣

اعتبر أن ضغط الدم هو المتغير التابع وبقية المتغيرات هي مستقلة، ثم احسب:

١- معامل الارتباط الجزئي للوزن وضغط الدم مع ثبات الطول.

٢- معامل الارتباط الجزئي للطول وضغط الدم مع ثبات الوزن.

٣- احسب معادلة الانحدار واستخرج قيم ثوابتها.

٤- احسب معامل الارتباط المتعدد ر ١/ ٣٠٢.

س٥: استخدم الحاسوب لاستخراج قيم معاملات الارتباط البسيط والجزئي والمتعدد وكذلك الانحدار المتعدد للمتغيرات التالية:

البعد عن مركز المدينة (كم)	كثافة السكان نسمة/ كم	أسعار السلع دولار	المبيعات اليومية بالدولار	إيجارات الأراضي دولار/ م
٠,٠١	٦	١٣,٨	٣,٤	١٩٥
٠,٠٥	٨	١٤,٢	٣,٢	١٨٠
١,٦	١٠	١٤,٧	٣,٠	١٢٠
٣,٠	١٢	١٤,٩	٢,٦	١٠٠
٣,٥	١٠	١٥,١	١,٩	١٤٠
٤,٨	٨	١٥,٣	١,٩	٩٠
٦,٠	٧,٦	١٥,٨	١,٥	٨٥
٧	٥	١٦,٠	١,٢	٨٤

obeikandl.com