

الفصل الثامن عشر

الانحدار والارتباط الخطي المتعدد

MULTIPLE LINEAR REGRESSION AND CORRELATION

سنعتمد في هذا الفصل على الحاسب الإلكتروني في إجراء الحسابات اللازمة لحل مشكلات الانحدار والارتباط تخففا من الأعباء الحسائية الضخمة التي يتطلبها التحليل . إلا أن هذا لا يعفينا بل يوجب علينا الإحاطة بالأسس والافتراضات والمفاهيم التي يبنى عليها التحليل تحسبا من أخطاء ومزالق التطبيق .

أولا : الانحدار الخطي المتعدد

(١٢ - ١) الانحدار الخطي المتعدد كامتداد للانحدار الخطي البسيط :

في تناولنا للانحدار الخطي البسيط في الفصل التاسع من هذا الكتاب افترضنا أن لدينا متغيرا عشوائيا v نرغب في التنبؤ بقيمه عن طريق قيم متغير غير عشوائى s على أساس أن العلاقة بين هذين المتغيرين هي علاقة خطية :

$$v = \beta + \alpha s \quad \text{حيث } \alpha, \beta \text{ بارامتران مجهولان ,}$$

وعلى أساس أنه عند أى قيمة ثابتة s يكون للمتغير v توزيع معتدل متوسطه $\beta + \alpha s$ يتوقف على قيمة s ، وتباينه عدد ثابت σ^2 لا يتوقف على s .

إلا أنه في كثير من التطبيقات يرى الباحث أن استخدام متغير واحد s لا يصلح للتنبؤ بقيم المتغير v بدقة كافية ، ويأمل أن يحصل على تنبؤات أكثر دقة إذا استخدم عدة متغيرات $s_1, s_2, s_3, \dots, s_k$ يشعر بخبرته في ميدان عمله أن لها دورا في عملية التنبؤ . وإذا تبينا افتراضات مماثلة لتلك التي تبيناها في الانحدار الخطي البسيط من أن هذه المتغيرات غير عشوائية وترتبط بالمتغير العشوائي v بعلاقة خطية :

$$(1) \quad v = \alpha + \beta_1 s_1 + \beta_2 s_2 + \dots + \beta_k s_k$$

حيث $\alpha, \beta_1, \beta_2, \dots, \beta_k$ بارامترات مجهولة ، ومن أنه عند أي مجموعة ثابتة $s_1, s_2, s_3, \dots, s_k$ يكون للمتغير v توزيع معتدل ذو تباين ثابت σ^2 ، فإن الانحدار يسمى في هذه الحال بالانحدار الخطي المتعدد أو اختصارا بالانحدار الخطي كما تسمى البارامترات $\beta_1, \beta_2, \dots, \beta_k$ بمعاملات الانحدار الجزئية **partial regression coefficients** ، وتكون دراسة هذا الانحدار مجرد تعميم للمفاهيم والأساليب والصيغ والاختبارات المستخدمة في الانحدار الخطي البسيط ولا تختلف عنها إلا في درجة المشقة في تناول البيانات ، كما تختلف بطبيعة الحال في درجات الحرية التي تعتمد على عدد الثوابت في معادلة الانحدار التي تقدر من العينة . وإذا كان القارئ قد اطمأن لمعرفة أسس وأساليب الانحدار الخطي البسيط فسوف لا يجد صعوبة تذكر في تناول الانحدار الخطي المتعدد كما يتبين من الفقرات الآتية .

(1) إيجاد معادلة الانحدار :

في الانحدار الخطي البسيط نقوم بتجربة نحصل منها على عينة من n من المشاهدات (s, v) تأخذ الشكل الآتي :

$$\begin{array}{cccc} s_1 & & s_2 & \dots & s_n \\ v_1 & & v_2 & \dots & v_n \end{array}$$

وتستخدم هذه البيانات في تقدير البارامترين المجهولين α ، β بقيمتين u ، b توطئة
لكتابة معادلة الانحدار :

$$\hat{v} = u + b s$$

التي تمثل « أحسن خط » يلائم تلك المشاهدات من وجهة نظر مبدأ المربعات
الصغرى ، الذي يستلزم كما نعلم تحديد القيمتين u ، b اللتين تجعلان الدالة :

$$S = \sum_{i=1}^n (v_i - \hat{v}_i)^2 = \sum_{i=1}^n (v_i - \alpha - \beta s_i)^2$$

نهاية صغرى . وتحقيق هذا الغرض يستلزم بدوره إجراء عملية التفاضل الجزئي
بالنسبة إلى كل من α ، β ومساواة كل من الناتجين بالصفر للحصول على معادلتين
خطيتين في هذين المجهولين تسميان بالمعادلتين المعتادتين وهما :

$$\sum_{i=1}^n v_i = \sum_{i=1}^n (\alpha + \beta s_i)$$

$$\sum_{i=1}^n v_i s_i = \sum_{i=1}^n (\alpha s_i + \beta s_i^2)$$

وحل هاتين المعادلتين معا يعطينا القيمتين u ، b المطلوبتين .

وهذه الخطوات هي بذاتها الخطوات التي تتخذ عند تناول الانحدار الخطي
المتعدد . إذ نقوم بتجربة نحصل منها على عينة من n من المشاهدات (s_i, v_i) ،
 $s_1, v_1, \dots, s_n, v_n$ حيث $s_i = 1, 2, \dots, n$ ، v_i تأخذ الشكل
المبين بالجدول (١٢ - ١) الآتي :

الجدول (١٢ - ١)

بيانات الانحدار الخطي المتعدد

المشاهدات	ص	س _١	س _٢	س _ك
(١)	ص _١	س _{١١}	س _{٢١}	...	س _{ك١}
(٢)	ص _٢	س _{١٢}	س _{٢٢}	...	س _{ك٢}
...
(س)	ص _س	س _{١س}	س _{٢س}	...	س _{كس}
...
(٧)	ص _٧	س _{١٧}	س _{٢٧}	...	س _{ك٧}

ولتقدير البارامترات المجهولة α ، β_1 ، β_2 ، ... ، β_k من هذه البيانات نستخدم مبدأ المربعات الصغرى لإيجاد القيم A ، B_1 ، B_2 ، ... ، B_k توطئة لكتابة معادلة الانحدار:

$$\hat{V} = A + B_1 S_1 + B_2 S_2 + \dots + B_k S_k \quad (2)$$

وتحقيق هذا المبدأ يستلزم أن تؤخذ القيم A ، B_1 ، ... ، B_k بحيث تكون الدالة

$$\epsilon = (\hat{V} - V)^2 = (A + B_1 S_1 + B_2 S_2 + \dots + B_k S_k - V)^2$$

نهاية صغرى ، وهذا يستلزم بدوره إجراء عملية التفاضل الجزئي بالنسبة إلى كل من α ، β_1 ، β_2 ، ... ، β_k ومساواة النواتج بالصفر للحصول على $k + 1$ من المعادلات الخطية في هذه المجاهيل تسمى بالمعادلات المعتادة . وحل هذه المعادلات معا يعطينا التقديرات المطلوبة A ، B_1 ، B_2 ، ... ، B_k وعددها $k + 1$.

ولنا أن نتصور هنا مدى المشقة التي نلقاها في حساب المجاميع ومجاميع المربعات ومجاميع حواصل الضرب المطلوبة لوضعها في هذه المعادلات ثم حل المعادلات ذاتها ، خاصة إذا كان عدد المتغيرات السينية أكبر من اثنين . وحتى في حالة وجود متغيرين s_1 ، s_2 حيث نستهدف الوصول إلى معادلة الانحدار

$$\hat{v} = 1 + s_1 b_1 + s_2 b_2$$

يكون علينا حل المعادلات المعتادة الآتية :

$$s_1 \alpha + s_2 \beta = \alpha \quad s_1 \alpha + s_2 \beta = \alpha$$

$$s_1 \alpha + s_2 \beta = \alpha \quad s_1 \alpha + s_2 \beta = \alpha \quad (3)$$

$$s_1 \alpha + s_2 \beta = \alpha \quad s_1 \alpha + s_2 \beta = \alpha$$

وهذا الحل يحتاج أولاً إلى حساب ثمانية مجاميع هي s_1 ، s_2 ، s_1^2 ، s_2^2 ، $s_1 s_2$ ، $s_1 v$ ، $s_2 v$ ، $s_1^2 v$ ، $s_2^2 v$ ، $s_1 s_2 v$ ، ولذلك نلجأ إلى الحاسب الإلكتروني ليقوم عنا بكل هذه العمليات ويعطينا التقديرات المطلوبة بكل سرعة ودقة .

(ب) إيجاد الخطأ المعياري لتقدير v من خط الانحدار :

في الانحدار الخطي البسيط استخدمنا مقياساً رأينا أهميته البالغة في عمليات الاستنتاج الإحصائي هو الخطأ المعياري للتقدير الذي رمزنا له بالرمز s_v وعرفناه في البند (٩ - ٤) بأنه الانحراف المعياري للقيم المعيارية المشاهدة v حول القيم \hat{v} المقدرة من معادلة الانحدار أي عرفناه كالآتي :

$$s_v = \frac{1}{n-2} (v - \hat{v})^2$$

بدرجات حرية $n - 2$

وحيث حللنا الاختلاف الكلى فى القيم الصادية وهو ϵ (ص - ص) إلى مركبتين
 تعبر الأولى وهى ϵ (ص - ص) عن الاختلاف الناشئ عن الانحدار (الاختلاف
 المفسر) وتعبر الثانية وهى ϵ (ص - ص) عن الاختلاف المتبقى الناشئ عن
 الانحراف عن خط الانحدار (الاختلاف غير المفسر) أمكننا كتابة مربع الخطأ
 المعيارى للتقدير كالتالى :

$$\epsilon^2 \text{ ص.ص.} = \frac{1}{2-n} = \epsilon^2 (\text{ص} - \text{ص}) \times \frac{1}{2-n} = \text{الاختلاف غير المفسر}$$

بنفس المنطق نستخدم نفس التعريف فى حالة الانحدار الخطى المتعدد مع مراعاة
 درجات الحرية ، فترمز للخطأ المعيارى لتقدير ص من خط الانحدار بالرمز

ϵ ص.ص. ϵ ص... صك حيث :

$$(4) \quad \epsilon^2 \text{ ص.ص.} = \frac{1}{n - (k + 1)} = \text{الاختلاف غير المفسر}$$

حيث n حجم العينة ، k عدد المتغيرات السينية ومع ملاحظة أن $k + 1$ هو
 عدد الثوابت فى معادلة الانحدار التى تقدر من العينة .

ملاحظة (١) :

ليس من المناسب هنا تقديم الصيغة العامة للخطأ المعيارى للتقدير ويكفينا أن
 نفهم المعنى الذى يتضمنه ، أما حساب قيمته فتركه للحاسب الالىكترونى . إلا
 أنه حين يكون هناك متغيران تنبؤيان اثنان فقط ($k = 2$) فىمكن إثبات أن

$$(5) \quad \epsilon^2 \text{ ص.ص.} = \frac{1}{3-n} = (\text{محص}^2 - \text{محص} - \text{محص} - \text{محص} - \text{محص} - \text{محص} - \text{محص}) \frac{1}{3-n}$$

وهذه الصيغة هي امتداد للصيغة (١٠) بالبند (٩ - ٤) التي تتناول الحالة التي يكون لدينا فيها متغير تنبؤى واحد .

هذا مع ملاحظة أن الصيغة العامة للخطأ المعياري تكتب بأسلوب رياضى يتطلب معرفته دراسة مسبقة لموضوع المصفوفات ، والواقع أن الدراسة النظامية للانحدار المتعدد تعتمد برمتها على هذا الأسلوب غير أننا فى التطبيق العملى لا نحتاج إليه .

(ح) اختبار دلالة الانحدار ككل :

فى الانحدار الخطى البسيط اهتمامنا باختبار دلالة الانحدار أى باختبار وجود علاقة خطية بين المتغير التابع v والمتغير المستقل w وهذا يكافئ اختبار الفرض الصفرى $\beta = 0$ ضد الفرض $\beta \neq 0$. واستخدمنا لهذا الغرض اختبار t بالصيغة (١١) بالبند (٩ - ٥ - أولاً) مع وضع $\beta = 0$. أو اختبار F بالصيغة (٢١) بالبند (٩ - ٧) أو بالصيغة المكافئة (٢٤) ، بشرط توفر شروط الانحدار .

كذلك يهمننا فى حالة الانحدار الخطى المتعدد اختبار وجود علاقة خطية بين المتغير التابع v والمتغيرات المستقلة w_1, w_2, \dots, w_k ، وهذا الاختبار يكافئ اختبار الفرض الصفرى $\beta_1 = \beta_2 = \dots = \beta_k = 0$ وفى هذه الحالة نستخدم أى من الصيغتين (٢١) أو (٢٤) المذكورتين مع مراعاة استخدام درجات الحرية المناسبة . وهاتان الصيغتان هما :

$$F = \frac{\text{التباين المفسر}}{\text{التباين غير المفسر}} \quad \text{بدرجتى حرية ك ، ن - ك - ١} \quad (٦)$$

$$\text{أو } F = \frac{r^2 \div ك}{(١ - r^2) \div (ن - ك - ١)} \quad \text{بدرجتى حرية ك ، ن - ك - ١} \quad (٧)$$

حيث n حجم العينة ، k عدد المتغيرات السينية ، s^2 هو معامل التحديد الذي يعبر كالمعتاد عن نسبة الاختلاف في المتغير v التي يفسرها خط الانحدار وسنرمز له هنا بالرمز s^2 ص (٢١...٤) حيث

$$(٨) \quad s^2 = \frac{\text{الاختلاف المفسر}}{\text{الاختلاف الكلي}} \quad \text{ص (٢١...٤)}$$

(٥) اختبار دلالة معاملات الانحدار الجزئية :

في الانحدار الخطي البسيط يكون لدينا معامل انحدار واحد β ، ويمكن أن نختبر ما إذا كان هذا المعامل يأخذ قيمة ما نفترضها بواسطة الإحصاءة (١١) المقدمة بالبند (٩ - ٥ - أولاً) وهي

$$(٩) \quad t = \frac{\beta - B}{\frac{s \cdot \sqrt{C}}{n}} = \frac{\beta - B}{\frac{s \cdot \sqrt{C}}{n}}$$

التي يكون لها توزيع t بدرجات حرية $n - 2$ بشرط تحقق افتراضات الانحدار . ويلاحظ أن مقام هذا الكسر وهو $\frac{s \cdot \sqrt{C}}{n}$ هو تقدير للانحراف المعياري للمتغير العشوائى B الذى تعتبر B إحدى قيمه المشاهدة - وحيث t تقدير معامل الانحدار β الذى نجده من العينة . والمتغير B هو متغير معتدل وسطه الحسابى β وانحرافه المعياري σ (ب) يقدر من العينة بالمقدار $\frac{s \cdot \sqrt{C}}{n}$.

أما في الانحدار الخطي المتعدد فلدينا k من معاملات الانحدار الجزئية $\beta_1, \beta_2, \dots, \beta_k$ (بالإضافة إلى البارامتر α) ويهمننا أن نختبر ما إذا كانت هذه المعاملات تأخذ قيمة معينة . وبنفس منطق البند (٩ - ٥) نعتبر أن القيمة t

(و = ١ ، ٢ ، ... ، ك) التى نحصل عليها من عينة لتقدير البارامتر β_j ، هى إحدى القيم المشاهدة من متغير عشوائى B_j ذى توزيع معتدل متوسطه β_j وانحرافه المعيارى σ_j (ب_ج) نقدره بقيمة سنرمز لها بالرمز ϵ_j (ب_ج) وبالتالى يكون للإحصاءة

$$ت_j = \frac{B_j - \beta_j}{\epsilon_j} \quad \text{و = ١ ، ٢ ، ... ، ك} \quad (١٠)$$

توزيع ت بدرجات حرية $n - (ك + ١)$ بشرط تحقق افتراضات الانحدار . وهذه الإحصاءة تصلح لاختبار أن يأخذ أى معامل β_j أى قيمة نفترضها ، وبصفة خاصة لاختبار الفرض $\beta_j = ٠$ لأهمية ذلك فى بحث مدى مساهمة المتغير المناظر $س_j$ فى التنبؤ من معادلة الانحدار .

ملاحظة (٢) :

ليس من المناسب هنا أيضا تقديم الصيغة العامة التى نوجد منها التقدير ϵ_j (ب_ج) للانحراف المعيارى للمتغير B_j وسنعمد فى ذلك أيضا على الحاسب الالكترونى وكفينا أن نخطط بالدور الذى يقوم به . إلا أنه حين يكون هناك متغيران تنبؤيان اثنان فقط (ك = ٢) فيمكن إثبات أن هذا التقدير يحسب من الصيغة الآتية :

$$\epsilon_j \text{ (ب}_j\text{)} = \frac{\epsilon_{١١} \epsilon_{٢٢} \dots \epsilon_{٣٣}}{\sqrt{(١ - r_{١١}^2) \dots (١ - r_{٣٣}^2)}} \quad \text{و = ١ ، ٢} \quad (١١)$$

$$\text{حيث } \epsilon_{١١}^2 \epsilon_{٢٢}^2 \dots \epsilon_{٣٣}^2 = \frac{١}{٣ - ١} \times \text{الاختلاف غير المفسر}$$

هو مربع الخطأ المعيارى للتقدير ويمكن حسابه من الصيغة (٥) ، وحيث $r_{١١}^2$ هو معامل التحديد للمتغيرين $س_١$ ، $س_٢$ ويمكن حسابه من الصيغة :

(١٢)

$$\frac{[\sum_{i=1}^n (S_i - \bar{S})^2]}{n(n-1)} = s^2$$

كما أن $\sum_{i=1}^n (S_i - \bar{S})^2$ هو مجموع المربعات للمتغير S ، $\sum_{i=1}^n (S_i - \bar{S})^2$ هو مجموع المربعات للمتغير S .

(١٢ - ٢) استخدام الحاسب الالكترونى :

يتضح من البند السابق أن حل المشكلات العملية فى تحليل الانحدار يتطلب القيام بعمليات حسابية طويلة سواء فى إيجاد الجاميع أو حل المعادلات أو فى استخراج تقديرات للأخطاء المعيارية أو فى استخراج قيم t أو قيم F اللازمة للاستنتاج الإحصائى ، مما يستغرق جهدا شاقا ووقتا طويلا ، إضافة إلى صعوبة تلافى أخطاء هذه الحسابات إذا أجريت يدويا . ولذلك فإن معظم المشكلات التطبيقية تستعين بالحاسبات الالكترونية التى تنوب عنا فى استخراج ما نريد بكل سرعة ودقة وترك لنا فقط مهمة تفسير النتائج واتخاذ القرارات بناء على ما قدمته من معلومات .

واستخدام هذه الحاسبات لا يستلزم أن يكون الباحث قادرا على تشغيلها بنفسه بل يكفيه أن يتصل بأحد مراكز الحاسبات وتقديم مصفوفة البيانات التى حصل عليها من التجربة مع تحديد مجموعة الأعداد التى تخص كلا من المتغير الصادى والمتغيرات السينية ، وتحديد ما يريد إنجاده من معلومات .

وهناك كثير من الحاسبات تحتزن برامج إحصائية تقوم بكافة أنواع التحليل ، ومن هذه البرامج ما يلى :

STATPACK, MINITAB, COSAP AND SPSS

ومثل هذه البرامج من شأنها توفير الجهد والوقت مع ضمان الدقة والأمان ولذلك يعتمد هذا الفصل فى حساباته وتحليلاته على الحاسبات الالكترونية التى أصبحت

اليوم في متناول الجميع . وسنوضح ذلك بالمثال الآتي الذى يتناول حالة متغيرين
 تنبؤين S_1 ، S_2 . على أن الأسلوب المستخدم يمتد إلى الحالات التى تتناول
 أكثر من متغيرين دون أى تعديل .

مثال (١٢ - ١) :

أرادت إحدى الشركات التجارية الكبيرة أن تعد طريقة للتنبؤ بعدد الوحدات
 التى تباع فى الشهر فى أى فرع من فروعها العشرة ، على افتراض أن الوحدات
 المباعة فى أى فرع تتأثر بعاملين هما (١) عدد البائعين فى الفرع ، (٢) مقدار
 ما يصرفه الفرع شهريا على الإعلان . وتحقيقا لهذا الغرض جمعت بيانات من
 الفروع ودونت فى الجدول (١٢ - ٣) الآتى .

الجدول (١٢ - ٣)

الفرع	عدد الوحدات المباعة ص	عدد العاملين س _١	تكاليف الاعلان س _٢
(١)	٢٥	٥	١٠
(٢)	٢٠	٢	١١
(٣)	٣٠	٦	١٢
(٤)	٢٥	٤	١٣
(٥)	٢٥	٣	١٤
(٦)	٣٢	٦	١٥
(٧)	٢٥	٤	١٢
(٨)	٢١	٣	١١
(٩)	٢٠	٢	١٠
(١٠)	٢٧	٥	١٢

المطلوب إيجاد معادلة الانحدار الخطي للمتغير ص على المتغيرين س₁ ، س₂ واختبار دلالة الانحدار ككل ودلالة كل من معاملي الانحدار .

الحل :

للتوضيح سنقوم بحل هذه المسألة مرتين : باستخدام الحاسب ثم بدونه .

(أولاً) الحل باستخدام الحاسب :

أدخلت البيانات المدونة بالجدول (١٢ - ٣) في حاسب الكتروني فأخرج المعلومات المدونة بالجدول (١٢ - ٤) الآتي ، مع ملاحظة أن عدد المتغيرات التنبؤية ك = ٢ وأن حجم العينة ن = ١٠ .

الجدول (١٢ - ٤)

مخرجات الحاسب الالكتروني لبيانات المثال (١٢ - ١)

(1)	Regress. of	Y	Number of Units Sold		
(2)	on	X ₁	Number of Salespeople		
(3)		X ₂	Amount of Advertising Expenditure		
(4)	Variable Name	Regress. Coeff	S.E. of Coeff.	ت	D.F.
(5)	Constant	6.05263			
(6)	X ₁	2.10526	.17708	11.88891	7
(7)	X ₂	.87719	.16125	5.42652	7

(8) Coefficient of Determination (R²) = .974659

(9) Estimated Standard Error of Estimate = .722013

Analysis of Variance for Regression

(10)	Source of Variation	SS	D.F.	MS	F
(11)	Regression	140.351	2	70.1754	134.615
(12)	Residual	3.64912	7	.521303	

(أ) معادلة الانحدار :

بالتأمل في الجدول (١٢ - ٤) نرى أن الحاسب قد قام بإيجاد التقديرات المطلوبة للبارامترات α ، β_1 ، β_2 وهي :

$$1 = 6,05263 ، \beta_1 = 2,10526 ، \beta_2 = 0,87719$$

وهذه القيم الثلاث هي تلك المدونة بالصفوف ٥ ، ٦ ، ٧ من العمود الثاني بالجدول . وإذن معادلة الانحدار هي - التقريب إلى ٣ خانات عشرية :

$$\hat{ص} = 6,053 + 2,105 س_1 + 0,877 س_2$$

فمثلا إذا كان عدد البائعين في الفرع $س_1 = ٤$ والتكاليف الشهرية $س_2 = ١٢$ فإننا نتوقع أن يكون عدد الوحدات المباعة حوالي ٢٤ وحدة .

(ب) اختبار دلالة الانحدار الخطي :

من الجدول (١٢ - ٤) نجد أن الحاسب قد حسب لنا جدول التباين في السطور الثلاثة الأخيرة ، فقد أعطانا كلا من :

$$\text{الاختلاف المفسر (الانحدار)} = 140,351 \text{ بدرجات حرية } ك = 2$$

$$\text{الاختلاف غير المفسر (الانحراف عن خط الانحدار)} = 3,64912 \text{ بدرجات حرية } ن - ك - ١ = ٧ \text{ بل أعطانا أيضا نسبة التباين في مستخدما الصيغة (٦)}$$

وهي

$$ف_٧ = \frac{\text{التباين المفسر}}{\text{التباين غير المفسر}} = \frac{70,1754}{0,521303} = 134,615$$

الاستنتاج :

لما كانت القيمة ١٣٤,٦١٥ تزيد كثيرا عن القيمة الحرجة $F_{[٧, ٢], 0.01} = 9,55$

نرفض الفرض الصفري أن $\beta = \beta_0$ عند مستوى عالى من الدلالة بمعنى أن هناك علاقة خطية بين عدد الوحدات المباعة من ناحية وعدد البائعين وتكاليف الإعلان من ناحية أخرى .

كما أن الحاسب قد أعطانا قيمة معامل التحديد بين المتغير ص والمتغيرين S_1 ، S_2 وهو S_1^2 ص (١ ، ٢) = ٠,٩٧٤٦٥٩ وهذا يعنى أن المتغيرين التنبؤيين قد فسرا حوالى ٩٧,٥٪ من التغير في ص وهذا جزء كبير يدعم الحكم بوجود العلاقة الخطية .

(ح) اختبار دلالة كل من معاملى الانحدار :

أعطانا الحاسب فى السطرين السادس والسابع من العمود الثالث التقديرين الخاصين بالخطأ المعياري لمعاملى الانحدار الجزئيين وهما ع (ب) = ٠,١٧٧٠٨ ، ع (ب) = ٠,١٦١٢٥ ، كما أعطانا فى العمودين الأخيرين من هذين السطرين قيمتى T_1 ، T_2 مستخدما الصيغة (١٠) بعد وضع $\beta_0 = ٠$ ، $\beta_1 = ٠$ وهما $T_1 = ١١.٨٨٨٩١$ ، $T_2 = ٥,٤٢٦٥٢$ ولكل منهما سبع درجات حرية .

الاستنتاج :

نظرا لأن $T_1 = ٣,٤٩٩$ نرفض كلا من الفرضين الصفريين $\beta_0 = ٠$ ، $\beta_1 = ٠$ عند مستوى الدلالة ٠,٠١ وهذا يعنى أن كلا من المتغيرين التنبؤيين يسهم إسهاما جوهريا فى التنبؤ بعدد الوحدات المباعة .

نلخص ما وجدناه فى هذه التجربة كما يلي : بناء على البيانات المشاهدة فى العينة على أساس أن افتراضات الانحدار متحققة ، تأخذ معادلة الانحدار الشكل الآتى :

$$\hat{S} = ٦,٠٥٣ + ٢,١٠٥ S_1 + ٠,٨٧٧ S_2$$

وهذه المعادلة تعبر تعبيراً مناسباً عن العلاقة الحقيقية بين المتغير ص (عدد الوحدات المباعة) والمتغيرين S_1 ، S_2 (عدد البائعين وقيمة تكاليف الإعلان) ، وتفسر حوالى ٩٧,٥٪ من التغير في قيم ص . كما أن كلا من هذين المتغيرين يسهم إسهاماً جاداً في التنبؤ بقيم هذا المتغير .

(ثانياً) الحل بغير استخدام الحاسب :

إذا لم يكن الحاسب الآلى متوفراً وكان الانحدار ذا متغيرين تنبؤيين فقط ، يمكن إجراء العمليات الحسابية المطلوبة باستخدام حاسب الجيب دون تحمل مشقة كبيرة . وفي هذا المثال يتخذ الحل الخطوات الآتية :

(١) إيجاد معادلة الانحدار :

نبدأ بحساب المجاميع الآتية

$\text{ص} = 250$ ، $\text{مح ص}^2 = 6394$ ، $\text{مح س}_1 = 40$ ، $\text{مح س}_2 = 180$ ،
 $\text{مح س}_1 = 120$ ، $\text{مح س}_2^2 = 1464$ ، $\text{مح س}_1 \text{ س}_2 = 489$ ، $\text{مح س}_1 \text{ ص} = 1050$ ،
 $\text{مح س}_2 \text{ ص} = 3040$ ثم نعوض بهذه القيم في المعادلات المعتادة (٣) كالآتي :

$$250 = \beta_1 120 + \beta_2 40 + \alpha$$

$$1050 = \beta_1 489 + \beta_2 180 + \alpha$$

$$3040 = \beta_1 1464 + \beta_2 489 + \alpha$$

وهناك عدة طرق لحل مثل هذه المعادلات نختار منها هنا طريقة دوليتل Doolittle التي تحول مصفوفة المعاملات والثوابت في المعادلات المعتادة إلى مصفوفة مثلثية تكون جميع عناصر القطر الرئيسى فيها مساوية للواحد . وهذه الطريقة روتين خاص من التعليمات تبين من الجدول (١٢ - ٥) الآتي .

الجدول (١٢ - ٥)

طريقة دوليتل لحل المعادلات الخطية

الثوابت	β	β	α	الصف	التعليمات
٢٥٠	١٢٠	٤٠	١٠	١^M	معاملات وثابت المعادلة الأولى
١٠٥٠	٤٨٩	١٨٠	٤٠	٢^M	معاملات وثابت المعادلة الثانية
٣٠٤٠	١٤٦٤	٤٨٩	١٢٠	٣^M	معاملات وثابت المعادلة الثالثة
٢٥٠	<u>١٢٠</u>	<u>٤٠</u>	١٠	١^M	١^M
٢٥	١٢	٤	١	ص ١^M	$١٠ \div ١^M$
٥٠	<u>٩</u>	٢٠	٠	٢^M	$٢١^M - ٢٢^M$ ص ١^M
٢,٥	٠,٤٥	١	٠	ص ٢^M	$٢٠ \div ٢^M$
١٧,٥	١٩,٩٥	٠	٠	٣^M	$٣١^M - ٣٢^M$ ص ١^M
٠,٨٧٧١٩	١	٠	٠	ص ٣^M	٣٢^M ص ٢^M $١٩,٩٥ \div ٣^M$

يلاحظ في هذه التعليمات أن هناك ثلاثة مقادير ثابتة هي $١٢٠ = ٣١^M$ ، $٤٠ = ٢١^M$ ، $٩ = ٣٢^M$ ، وهذه تسمى محاور الارتكاز . أما بقية القيم التي تشتمل على الرمز s فتتغير حسب قيمة s التي تعبر عن رقم العمود ($s = ١, ٢, ٣$ ، ٤) فمثلا ٢١^M تعبر عن العنصر الذي بالصف ٢^M والعمود s وبهذا يكون

$$١٠٥٠ = ٤٢^M ، ٤٨٩ = ٣٢^M ، ١٨٠ = ٢٢^M ، ٤٠ = ١٢^M$$

تبدأ طريقة دولتيل بتدوين معاملات وثوابت المعادلات كما هو مبين بالصفوف الثلاثة الأولى المشار إليها بالرموز α_1 ، α_2 ، α_3 . ثم تحسب عناصر الصفوف التالية الواحد بعد الآخر باستخدام التعليمات المبينة أمامه .

فالصف الرابع α_4 يتكون من نفس عناصر الصف α_3 . والصف التالي α_5 يتكون بقسمة عناصر الصف السابق α_3 على معامل α وهو ١٠ ، والصف التالي α_6 تتكون عناصره باستخدام التعليمات المبينة أمامه كالتالي :

بوضع $s = 1$ نجد أن العنصر الأول في هذا الصف هو $40 - 40 \times 1 = 0$.
وبوضع $s = 2$ نجد أن العنصر الثاني في هذا الصف هو $180 - 40 \times 2 = 20$.
وبوضع $s = 3$ نجد أن العنصر الثالث في هذا الصف هو

$$9 = 40 \times 3 - 489$$

وبوضع $s = 4$ نجد أن العنصر الرابع في هذا الصف هو

$$50 = 40 \times 4 - 100$$

ويلاحظ أننا استخدمنا هنا محور الارتكاز $\alpha_3 = 40$.

بنفس الطريقة نوجد بقية الصفوف الثلاثة . وبذلك تتحول المصفوفة التي رمز لصفوفها بالرموز α_1 ، α_2 ، α_3 وهي المصفوفة الأصلية إلى المصفوفة المثلثية التي رمز لصفوفها بالرموز α_1 ، α_2 ، α_3 وبالتالي تتحول مجموعة المعادلات المعتادة إلى المجموعة المكافئة الآتية :

$$\alpha + 4\beta + 12\gamma = 25 \quad (\text{من الصف } \alpha_1)$$

$$\alpha + 0,45\beta = 2,5 \quad (\text{من الصف } \alpha_2)$$

$$\beta = 0,87719 \quad (\text{من الصف } \alpha_3)$$

من هذه المعادلات ينتج أن القيم α ، β ، γ التي تحقق هذه المعادلات معا هي :

$$ب_2 = 0,87719$$

$$\therefore ب_1 = 0,45 - 2,5 = ب_2 = 0,45 - 2,5 = 0,87719 \times 0,45 - 2,5 =$$

$$2,10526 =$$

$$أ_1 = 25 - 12 = ب_1 - 4 =$$

$$25 - 12 - 25 = 0,87719 \times 12 - 0,87719 \times 4 - 25 =$$

$$6,05268 =$$

فتكون معادلة الانحدار مقربة إلى ثلاث خانات عشرية هي :

$$\hat{ص} = 6,053 + ب_1 ص_1 + ب_2 ص_2 = 6,053 + 2,105 ص_1 + 0,877 ص_2$$

وهي نفس المعادلة التي أوجدتها لنا الحاسب الآلي .

(ب) اختبار دلالة الانحدار :

نحتاج هنا إلى تحليل الاختلاف في ص إلى مركبتين كالاتي :

$$\text{الاختلاف الكلي} = \sum (ص) = \sum (ص) - \frac{(\sum ص)^2}{ن}$$

$$= 6394 - \frac{(250)^2}{10} = 144 \text{ بدرجات حرية } 10 - 1 = 9$$

من الصيغة (5) نجد أن :

الاختلاف غير المفسر = $ص_1 - أ_1 - ب_1 ص_2 - ب_2 ص_3 - ب_3 ص_4$

$$= 6394 - 250 \times 6,05268 - 100 \times 2,10526 - 0,87719 \times 3040 =$$

$$= 3,6494 \text{ بدرجات حرية } 10 - 3 - 1 = 6$$

∴ الاختلاف المفسر = 144 - 3,6494 =

$$140,3506 = \text{بدرجات حرية ك} = 2$$

لاختبار دلالة الانحدار نستخدم الصيغة (6) كالتالي :

$$F = \frac{\text{التباين المفسر}}{\text{التباين غير المفسر}} = \frac{2 \div 140,3506}{7 \div 3,6494} = 134,6049 =$$

وهذه هي نفس القيمة التي أوجدها الحاسب (مع فارق التقريب) وهي كما ذكرنا ذات دلالة عالية وتدعوننا لرفض الفرض الصفري $\beta = 0$.

$$\text{معامل التحديد } r^2 \text{ ص (1, 2) = } \frac{\text{الاختلاف المفسر}}{\text{الاختلاف الكلي}} = \frac{140,3506}{144} = 0,97466 =$$

وهي نفس القيمة التي أوجدها الحاسب .

(ح) اختبار دلالة كل من معاملي الانحدار :

نحتاج هنا أولاً إلى تقدير الخطأ المعياري لكل من المعاملين بالصيغة (11) وهي

$$e(B_i) = \frac{e \cdot \sqrt{C_{ii}}}{\sqrt{(1 - r^2) \cdot (n - 1) \cdot \sum (y_i - \bar{y})^2}} \quad \text{و } e = 1,2$$

ثم استخدام اختبارات لكل منهما بالصيغة (10) وهي

$$t = \frac{B_i - 0}{e(B_i)} \quad \text{و } e = 1,2$$

بدرجات حرية عددها $n - k - 1 = 7$.

$$\text{لدينا: } \text{ع}^2 = \frac{1}{n - k - 1} \times \text{الاختلاف غير المفسر}$$

$$0,5212427 = 3,6494 \times \frac{1}{7} =$$

$$0,7220 = \text{ع}^2 \text{ ص. ص. ص.}$$

$$20 = \frac{240}{10} - 180 = (1) \text{ ص. ص.}$$

$$24 = \frac{1120}{10} - 1464 = (2) \text{ ص. ص.}$$

$$9 = \frac{120 \times 40}{10} - 489 = (1) \text{ ص. ص.}$$

$$0,16875 = \frac{81}{24 \times 20} = \text{ع}^2 \text{ ص. ص.}$$

$$0,83125 = 1 - \text{ع}^2 \text{ ص. ص.}$$

$$0,17707 = \frac{0,72198}{20 \times 0,83125} = \text{ع}^2 \text{ ص. ص.}$$

$$0,16164 = \frac{0,72198}{24 \times 0,83125} = \text{ع}^2 \text{ ص. ص.}$$

$$\text{بدرجات حرية } 7 \quad 11,888 = \frac{2,105}{0,17707} = \text{ت}^2 \text{ ص. ص.}$$

$$\text{بدرجات حرية } 7 \quad 5,4256 = \frac{0,877}{0,16164} = \text{ت}^2 \text{ ص. ص.}$$

وهاتان القيمتان أعطاهما الحاسب وقد رأينا أن كليهما ذو دلالة عالية وتدعوان إلى رفض كل من الفرضين $\beta_1 = 0$ ، $\beta_2 = 0$ عند مستوى الدلالة 0,01 .

(١٢ - ٣) أسلوب آخر لاختبار دلالة معاملات الانحدار الجزئية :

لنبدأ بالحالة التي يكون لدينا فيها متغيران سينيان S_1 ، S_2 . إذا أوجدنا الاختلاف المفسر الناشئ عن انحدار هذين المتغيرين على المتغير ص حين يستخدمان معا (بدرجتين من درجات الحرية) ثم أوجدنا الاختلاف الناشئ عن انحدار المتغير S_1 على المتغير ص حين يستخدم وحده (بدرجة واحدة من الحرية) فإن الفرق بين هذين الاختلافين هو مقدار الاختلاف المفسر الذي ساهم به المتغير S_1 عند ضمه إلى المتغير S_2 . أي هو المساهمة الخاصة بالمتغير S_1 بعد استبعاد مساهمة S_2 ويمكن اختبار الفرض الصفرى $\beta_1 = 0$. ضد الفرض $\beta_1 \neq 0$. بواسطة اختبار ف بالصيغة (٦) كالتالي :

$$F = \frac{\text{التباين المفسر (للمتغير } S_1 \text{ بعد استبعاد } S_2)}{\text{التباين غير المفسر}}$$

$$(13) \quad F = \frac{[\text{انحدار } S_1 \text{ ، } S_2 - (\text{انحدار } S_2)] / (\text{انحدار } S_1)}{[\text{انحدار } S_1 \text{ ، } S_2 - (\text{انحدار } S_2)] / (\text{انحدار } S_1 - 1)}$$

مع ملاحظة أن $(\text{انحدار } S_1 \text{ ، } S_2 - (\text{انحدار } S_2))$ هو الاختلاف غير المفسر . وبالمثل ، إذا أوجدنا الاختلاف الذي يفسره S_1 وحده وطرحناه من الاختلاف الذي يفسره المتغيران معا نحصل على المساهمة الخاصة بالمتغير S_1 بعد استبعاد أثر S_2 ويمكن اختبار هذه المساهمة بنفس الطريقة .

ففى المثال - انظر الحل بغير الحاسب - وجدنا ما يلى :

$$٢٢ (ص) = \text{الاختلاف الكلى} = ١٤٤$$

$$٢٢ (انحدار س١ ، س٢) = \text{الاختلاف المفسر للمتغيرين معا} = ١٤٠,٣٥١٢$$

$$\text{الاختلاف غير المفسر} = ٣,٦٥ = \text{بدرجات حرية } n - ٣$$

نحسب الآن الاختلاف الذى يفسره كل من S_1 ، S_2 عندما يستخدم كل منهما على حدة .

$$\frac{[٢٢ ص (س١ ، س٢)]}{٢٢ (س٢)}$$

$$٢٥ = \frac{٢ \left(\frac{٤٠ \times ٢٥٠}{١٠} - ١٠٥٠ \right)}{٢٠} = \text{بدرجة واحدة من الحرية}$$

$$٢٢ (انحدار س١) = \frac{٢ \left(\frac{١٢٠ \times ٢٥٠}{١٠} - ٣٠٤٠ \right)}{٢٤} = ٦٦,٦٧ = \text{بدرجة واحدة من الحرية}$$

من الصيغة (١٣) ، لاختبار الفرض β ، = .

$$F_1 = \frac{٧٣,٦٨١٢}{٠,٥٢١٤} = \frac{٦٦,٦٧ - ١٤٠,٣٥١٢}{٧ \div ٣,٦٥}$$

$$\text{بدرجتى حرية } ١ ، ٧ = ١٤١,٣١٤$$

ولاختبار الفرض β ، = .

$$F_2 = \frac{١٥,٣٥١٢}{٠,٥٢١٤} = \frac{١٢٥ - ١٤٠,٣٥١٢}{٧ \div ٣,٦٥}$$

$$\text{بدرجتى حرية } ١ ، ٧ = ٢٩,٤٤٢$$

ونظراً لأن $F_{[7, 1], 0.01} = 12.2$ نرفض كلا من $\beta_1 = 0$ ، $\beta_2 = 0$

عند مستوى الدلالة 0,01

يلاحظ أن اختبار ف المستخدم هنا يكافئ اختبار ت الذي نتج عن الحل السابق ، ويتأكد هذا من ملاحظة أن $\sqrt{141,314} = 11,888 = t_{1, 0.01}$ ، $\sqrt{29,442} = 5,426 = t_{1, 0.01}$. ويمكن إذا أردنا أن ندون هذه النتائج في جدول كالآتي .

الجدول (١٢ - ٦)

اختبار كل من المتغيرين بعد استبعاد أثر المتغير الآخر

مصدر الاختلاف	د.ح	ط.م	ف
انحدار S_1 ، S_2 معا	٢		
انحدار S_1 وحده	١		
انحدار S_2 وحده	١		
انحدار S_1 بعد S_2	١	٧٣,٦٨	١٤١,٣١٤
انحدار S_2 بعد S_1	١	١٥,٣٥	٢٩,٢٤٢
الاختلاف غير المفسر	٧	٠,٥٢١٤	
الكلي	٩		

ملاحظات :

من التحليل الملخص بالجدول (١٢ - ٦) نخرج بالنتائج والملاحظات الآتية :

(١) المتغير s_1 حين يستخدم وحده للتنبؤ بقيم s_2 أفضل من المتغير s_3 حين يستخدم وحده لنفس الغرض ، وذلك لأن الاختلاف الذى يفسره s_1 منفردا وهو ١٢٥ أكبر من الاختلاف الذى يفسره s_3 منفردا وهو ٦٦,٦٧ . ونصل إلى نفس النتيجة إذا قارنا معاملي التحديد ، فمعامل التحديد للمتغير s_1 هو :

$$r_{s_1}^2 = \frac{125}{144} = 0,86806 \quad \text{بينما} \quad r_{s_3}^2 = \frac{66,67}{144} = 0,46299$$

وبهذا الأسلوب نستطيع مقارنة أى عدد من المتغيرات تستخدم فرادا .

(٢) فى تفسير الاختلاف الكلى فى s_2 تكون مساهمة أى متغير أكبر فى حالة استخدامه منفردا عنها فى حالة استخدامه بعد متغير آخر فبالنسبة للمتغير s_1 نجد أن $73,68 < 125$ وبالنسبة للمتغير s_3 نجد أن $15,35 < 66,67$ وهذه نتيجة عامة مهما كان عدد المتغيرات التنبؤية ، فالمساهمة المنفردة لأى متغير تكون أكبر دائما من مساهمته مع متغير أو عدة متغيرات أخرى .

(٣) معامل التحديد المحسوب من معادلة الانحدار للمتغيرين s_1 ، s_3 معا وهو $r_{(1,3)}^2 = \frac{140,35}{144} = 0,975$ أكبر من معامل التحديد المحسوب

من معادلة الانحدار لأى من المتغيرين وهما ٠,٨٦٨ ، ٠,٤٦٣ . وهذه نتيجة عامة ، فمعامل التحديد يقل دائما حين يستبعد واحد أو أكثر من المتغيرات التنبؤية .

(٤) فى معادلة الانحدار للمتغيرين s_1 ، s_3 معا يصعب تقدير الأهمية النسبية لهذين المتغيرين من حيث مقدار المساهمة فى التنبؤ بقيم المتغير التابع s_2 ،

حيث ٢٢ (ك) هو الاختلاف الناشئ عن الانحدار على ك من المتغيرات
 ٢٢ (ل) هو الاختلاف الناشئ عن الانحدار على ل من هذه المتغيرات
 ٢٢ (ص) هو الاختلاف في قيم المتغير التابع ص
 إذ يمكن إثبات أن كلا من بسط ومقام هذه النسبة هو تقدير مستقل للتباين
 .^{٢٥}

إن الصيغة (١٤) يمكن كتابتها بدلالة معاملات التحديد كالتالي :

$$(١٥) \quad F = \frac{[٢٢(ك) - (ل)٢٢] / (ك - ل)}{[٢٢(ك) - ١] / (١ - ك - ص)}$$

بدرجتي حرية (ك - ل) ، (١ - ك - ص) ،

مثال (١٢ - ٢) :

أجرى انحدار خطي لمتغير عشوائى ص على خمسة متغيرات تنبؤية ١ ،
 ٢ ، ٣ ، ٤ ، ٥ ، ووجد أن معامل التحديد في عينة حجمها ٤٦ هو
 ٠,٦٢. وعندما أجرى انحدار خطي لنفس المتغير على ثلاثة من هذه المتغيرات
 ١ ، ٢ ، ٣ ، وجد أن معامل التحديد ٠,٥٧. هل يمكن الاستغناء عن
 المتغيرين ٤ ، ٥ ، والاكتفاء بالمتغيرات الثلاثة الأخرى دون أن نفقد شيئاً من
 دقة التنبؤ بقيم المتغير ص ؟

الحل :

من الصيغة (١٥) :

$$F = \frac{٢ \div (٠,٥٧ - ٠,٦٢)}{٤٠ \div (٠,٦٢ - ١)} = ٢,٦٣٢$$

وهذه القيمة تقل عن القيمة الحرجة ف $F_{[40, 2], 0.05} = 3, 23$ مما يدعونا إلى قبول
الفرض الصفري أن دقة التنبؤ لم تتأثر، وبالتالي نستطيع الاكتفاء بالمتغيرات
 $\beta_1, \beta_2, \beta_3$ للتنبؤ بالمتغير β دون أن نفقد شيئا من دقة التنبؤ.

مثال (١٢ - ٣) :

في المثال (١٢ - ١) هل نستطيع الاكتفاء بأحد المتغيرين للتنبؤ بقيم المتغير التابع
 β دون أن نفقد شيئا من دقة التنبؤ؟

الحل :

سبق أن أجبنا عن هذا السؤال حين استخدمنا الصيغة (١٣) - التي هي حالة
خاصة من الصيغة (١٤) أو (١٥) - في رفض كل من الفرضين $\beta_1 = 0, \beta_2 = 0$
لأن هذا يعنى أن وجود أى من المتغيرين له أثر ذو دلالة في عملية التنبؤ. باستخدام
الصيغة (١٥) لاختبار إمكانية الاستغناء عن المتغير β_1 لدينا :

$$F_{(1, 38)} = 0, 97466, \quad F_{(2, 37)} = 0, 46299$$

$$F_{\beta_1} = \frac{1 \div (0, 46299 - 0, 97466)}{7 \div (0, 97466 - 1)} = \frac{0, 51167}{0, 00362}$$

$$= 141, 35$$

وهذه القيمة ذات دلالة عالية وتعنى أن الاستغناء عن β_1 يقلل من دقة التنبؤ.
وبالمثل، لاختبار إمكانية الاستغناء عن β_2

$$F_{\beta_2} = \frac{0, 86806 - 0, 97466}{0, 00362} = 29, 45$$

وهذه أيضا ذات دلالة عالية وتعنى أن الاستغناء عن β_2 يقلل من دقة التنبؤ.

(١٢ - ٤) اختيار متغيرات التنبؤ :

من الصعوبات التي يلقاها الباحثون في الانحدار المتعدد كيفية اختيار المتغيرات التي تدخل في معادلة الانحدار لتعطي أعلى درجة من الدقة في التنبؤ بالمتغير التابع ص . وفي المعتاد يكون لدى الباحث مجموعة كبيرة من المتغيرات المرشحة لذلك ويكون من المفيد عمليا اختصار هذه المجموعة إلى مجموعة تتكون من أقل عدد ممكن من هذه المتغيرات بشرط أن تعطي هذه المجموعة الجزئية معادلة تنبؤ لا تقل كفاءة عن المعادلة التي تستخدم جميع المتغيرات المتاحة . إن عملية الاختصار هذه ممكنة في الغالب لأن بعض المتغيرات المتاحة لا تسهم بدرجة كافية في عملية التنبؤ كما قد يبدو لأول وهلة ، كما أنه قد توجد مجموعة (أو أكثر) من المتغيرات المتاحة ترتبط ببعضها ارتباطا عاليا وبالتالي تعطي معلومات ماثلة ويمكن حينئذ الاكتفاء بمتغير واحد فقط من هذه المجموعة ليمثلها جميعا .

ويتلخص السؤال المطروح هنا فيما يلي : إذا كان لدينا ك من المتغيرات التنبؤية المرشحة للدخول في معادلة الانحدار فكيف نختار أصغر مجموعة جزئية من هذه المجموعة بحيث تعطي نفس الدرجة من الدقة في التنبؤ بالمتغير التابع ؟ فمثلا إذا كان لدينا مجموعة من عشرة متغيرات تنبؤية فهل يمكن أن نكتفي بمجموعة من اثنين أو ثلاثة منها لتعبر عن المجموعة بكاملها (دون أن نفقد شيئا من دقة التنبؤ) ؟

الواقع أنه لا توجد إجابة واحدة شافية لهذا السؤال ولكن هناك عدة اقتراحات بمحاولات نسترشد بها في تحديد المجموعة المطلوبة ، وتتطلب هذه المحاولات النظر إلى البيانات المشاهدة عدة مرات من جوانب مختلفة ، وهنا يلعب الحاسب دورا رئيسيا لأن تنفيذ هذه المحاولات يحتاج إلى جهد حسابي شاق .

نفرض أننا حصلنا على بيانات مشاهدة في عينة حجمها n كما في الجدول

(١٢ - ١) السابق .

(أ) إن أول ما يخطر بالبال هو اختيار المتغيرات الأكثر علاقة بالمتغير التابع ص ، فنقوم بالمقارنة بينها كمتغيرات فرادا إما عن طريق مقادير الاختلافات المفسرة ٢٢ (الانحدار) أو عن طريق معاملات التحديد r^2 كما جاء بالملاحظة (١) بالبند (١٢ - ٣) السابق ، فتكون المتغيرات ذوات القيم الأكبر مرشحة مبدئيا للدخول في معادلة انحدار متعدد .

(ب) نبحث عما إذا كان هناك مجموعة (أو أكثر) من المتغيرات التنبؤية التي ترتبط ببعضها ارتباطا عاليا ، فإذا وجدت مثل هذه المجموعة فإن أحدها يمكن أن يمثلها جميعا . ويحتاج الأمر هنا إلى إيجاد معامل الارتباط بين كل زوج من هذه المتغيرات ، ويفضل وضع هذه المعاملات وأيضا معاملات الارتباط بين كل من المتغيرات والمتغير التابع في جدول لتسهيل الرجوع إليها . وعلى فرض أن هناك خمس متغيرات تنبؤية يكون الجدول كالتالي :

٥	٤	٣	٢	١	
٥١	٤١	٣١	٢١	١	١
٥٢	٤٢	٣٢	١		٢
٥٣	٤٣	١			٣
٥٤	١				٤
٥٥	٤	٣	٢	١	ص

(ج) نوجد معادلة انحدار ص على جميع ما لدينا من متغيرات ثم نفحص دلالة كل من معاملات الانحدار الجزئية b_1 ، b_2 ، b_3 ، b_4 ، b_5 عن طريق اختبار

ت بالصيغة (١٠) فتكون المتغيرات التي ترشح للدخول في معادلة الانحدار هي تلك التي تحظى بالقيم الأكبر من قيم ت ، أما المتغيرات التي تناظر القيم الأصغر من ت فتكون معرضة للاستبعاد . على أن ذلك لا ينبغي أن يتم دون تدبر فإن استخدام أو استبعاد متغير مالا يقرر لمجرد أن قيمة ت المناظرة كبيرة أو صغيرة ، لأن قيم معاملات الانحدار الجزئية وبالتالي قيم ت هي قيم شرطية وتتغير بتغير المتغيرات التي تدخل معها في الانحدار كما سبق القول بالملاحظة (٥) ، ومن ناحية أخرى قد توجد أسباب نظرية أو خبرات تحتم الاحتفاظ ببعض المتغيرات حتى ولو كان معامل الانحدار الجزئي المناظر غير ذي دلالة . فمثلا إذا كان لدينا عدة متغيرات نرغب استخدامها في التنبؤ بالمسافة التي تقطعها سيارة ما وكان من بين هذه المتغيرات متغير « مقدار الوقود المستهلك » فلا بد من الاحتفاظ بهذا المتغير حتى ولو كان معامل الانحدار المناظر له غير ذي دلالة .

(٥) على ضوء ما نجده في (أ) و(ب) و(ج) نختار مجموعة أو أكثر من المتغيرات ونحسب معادلات انحدار كل منها مع إيجاد معامل التحديد ومقارنته بمعامل التحديد الذي ينجم من الانحدار على جميع المتغيرات المتاحة واختبار دلالة الفرق بينهما في دقة التنبؤ باستخدام الصيغة (١٥) . وإذا وجدنا أن إحدى المجموعات المختارة بنفس كفاءة المجموعة الكاملة نحاول اختزالها بنفس الطريقة إلى مجموعة ذات عدد أقل من المتغيرات .

ويبنى اختيارنا النهائي لمجموعة المتغيرات التي تدخل في معادلة التنبؤ التي ننشدها على أساس أنها (أولا) تشتمل على أقل عدد من المتغيرات و(ثانيا) يمكنها أن تعبر عن مجموعة المتغيرات المتاحة بكاملها ، أى بحيث لا تقل دقة التنبؤ منها عن دقة التنبؤ من استخدام جميع المتغيرات المتاحة . هذا ويجدر الإشارة أن المجموعة التي تختار على هذا الأساس ليست فريدة ، بل يمكن أن نعثر على أكثر من مجموعة تستوفي الشرطين المطلوبين .

تمارين (١٢ - ١)

أجريت دراسة لمعرفة أثر العوامل الجغرافية على حجم نوع من الضفدع واستهدف جزء من هذه الدراسة التنبؤ بطول جسم الضفدعة عن طريق الارتفاع عن سطح البحر والمتوسط السنوي لدرجة الحرارة . وقد جمعت بيانات من ١٣ موقعا ودونت بالجدول الآتي :

الجدول (١٢ - ٧)

الموقع	طول الضفدعة ع	الارتفاع عن سطح البحر س _١	درجة الحرارة س _٢
(١)	٢٣,٧	٦٠٠	٦٠
(٢)	٢٣,٩	٨٠٠	٥٥
(٣)	٢٣,٤	١٤٥٠	٥٠
(٤)	٢٣,٤	١٤٠٠	٥٠
(٥)	٢٢,٢	١٠٠	٥٥
(٦)	١٩,٥	٤٥٠	٥٠
(٧)	٢٣,٥	٤٠٠	٥٠
(٨)	٢٤,٧	٨٠٠	٦٠
(٩)	٢٣,٣	٨٠٠	٦٠
(١٠)	٢١,٧	١١٠٠	٥٠
(١١)	٢١,٣	٥٠٠	٥٠
(١٢)	٢٠,٢	٤٠٠	٥٠
(١٣)	٢٢,١	٤٠٠	٥٠

$$\begin{aligned} 8495000 &= \sum_{i=1}^n Y_i^2 & 9200 &= \sum_{i=1}^n X_1 & 6627,27 &= \sum_{i=1}^n X_2 \\ 486500 &= \sum_{i=1}^n Y_i X_1 & 36850 &= \sum_{i=1}^n X_1 X_2 & 690 &= \sum_{i=1}^n X_2 \\ 10092,5 &= \sum_{i=1}^n Y_i & 210,265 &= \sum_{i=1}^n X_1 & &= \sum_{i=1}^n X_2 \end{aligned}$$

(أولاً) استخدم مخرجات الحاسب الالكتروني المبينة بالجدول (١٢ - ٨) الآتي لإيجاد معادلة الانحدار الخطي للمتغير ص على المتغيرين س_١ ، س_٢ .
 (ثانياً) اختبر دلالة الانحدار ككل ودلالة كل من معاملي الانحدار . هل من الممكن الاستغناء عن أحد المتغيرين التنبؤيين دون أن نفقد شيئاً من دقة التنبؤ ؟
 (ثالثاً) أجب عن الجزئين (أولاً) و(ثانياً) دون الاستعانة بمخرجات الحاسب .

الجدول (١٢ - ٨)

Regress. of	Y	Length of Frog		
On	X ₁	Altitude		
	X ₂	Temperature		
Variable Name	Regress. Coeff.	S.E. of Coeff.	t	D.F.
Constant	9.785579	4.129044	2.37	
X ₁	.1700912 E-02.	.81158094—03	2.10	10
X ₂	.2174479	.7589070 E—1	2.87	10

Multiple Corr. Coeff. (R) = .73379

Coeff. of Determination (R²) = .5384

Estimated Standard Error of Estimate = 1.1390588

Analysis of Variance for Regression

Source of Variation	SS	D.F.	M S	F
Regression	15.13315	2	7.566575	5.8319
Residual	12.97455	10	1.297455	

ملاحظة :

الرمز 01 - E يعني ضرب العدد الذى على يسار هذا الرمز فى 10^{-1}
الرمز 02 - E يعني ضرب العدد الذى على يسار هذا الرمز فى 10^{-2}
وهكذا ...

الرمز 02 + E يعني ضرب العدد الذى على يسار هذا الرمز فى 10^{+2}
وهكذا ...

فمثلا العدد 02 - E 1700912 . يعنى العدد 001700912 .
هذا ما يسمى بالرمزية العلمية scientific notation

ثانيا - الارتباط الخطى المتعدد

سنفترض هنا أن لدينا $k + 1$ من المتغيرات العشوائية $x_1, x_2, \dots, x_k, x_{k+1}$ ،
كما سنفترض أن توزيع الاحتمال المشترك لهذه المتغيرات هو توزيع معتدل
متعدد المتغيرات multivariate normal distribution إن هذا الفرض يتضمن
أمرين : الأول أن توزيع أى من هذه المتغيرات (على حدة) هو توزيع معتدل ،
والثانى أن العلاقة بين أى من هذه المتغيرات وأى متغير آخر أو مجموعة من المتغيرات
الأخرى هى علاقة خطية .

وامتدادا لدراساتنا فى الارتباط الخطى البسيط فى الفصل العاشر نتناول فى
البندين الآتين معاملين رئيسيين هما معامل الارتباط الخطى المتعدد ومعامل الارتباط
الجزئى ونرى كيف نختبر دلالة كل منهما .

(١٢ - ٥) معامل الارتباط المتعدد

MULTIPLE CORRELATION COEFFICIENT

إذا أردنا تقدير درجة العلاقة الخطية بين أحد هذه المتغيرات وليكن x_1 وبقية
المتغيرات ، فإننا نعرّف معامل الارتباط المتعدد بين x_1 والمتغيرات الأخرى بأنه

معامل الارتباط البسيط بين المتغير v_1 والمتغير العشوائى الذى يتكون من التركيب الخطى $\alpha + \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_k v_k$ وسنرمز لهذا المعامل بالرمز $\rho_{(1+e \dots k)}$.

ولتقدير هذا المعامل نأخذ عينة عشوائية من n من وحدات المجتمع ونقوم بقياس قيمة كل من هذه المتغيرات لكل وحدة فنحصل على n من المشاهدات كل منها على الصورة $(v_{1e}, v_{2e}, \dots, v_{ke}, v_{1+e})$ حيث $e = 1, 2, \dots, n$ وتوضع هذه القياسات عادة كما فى الجدول الآتى .

جدول (١٢ - أ)

المشاهدات	v_1	v_2	v_k	\dots	v_{1+e}
(١)	v_{11}	v_{21}	v_{k1}	\dots	$v_{1+e, 1}$
(٢)	v_{12}	v_{22}	v_{k2}	\dots	$v_{1+e, 2}$
\dots	\dots	\dots	\dots	\dots	\dots
(٣)	v_{1e}	v_{2e}	v_{ke}	\dots	$v_{1+e, e}$
\dots	\dots	\dots	\dots	\dots	\dots
(٥)	v_{1n}	v_{2n}	v_{kn}	\dots	$v_{1+e, n}$

من هذه المشاهدات نحصل على التقدير المطلوب بعدد سنرمز له بالرمز $\rho_{(1+e \dots k)}$ ونحسبه كما فى حالة الارتباط الخطى البسيط من الصيغة

$$(16) \quad \frac{\text{الاختلاف المفسر}}{\text{الاختلاف الكلى}} = \frac{\rho_{(1+e \dots k)}^2}{\rho_{(1+e \dots k)}}$$

أى أن التقدير المطلوب هو الجذر التربيعى لنسبة الاختلاف الذى يفسره الانحدار الخطى المتعدد إلى الاختلاف الكلى فى المتغير v_1 . ويمكن إثبات أنه فى هذه الحالة .

$$(17) \quad 1 \leq r_{12} \leq (1 + k \dots r_2) \leq 0$$

أى أن هذا المعامل لا يمكن أن يكون سالبا وهو يختلف فى هذه الصفة عن معامل الارتباط البسيط الذى نعلم أنه يمكن أن يأخذ قيما سالبة .

وكما فى الانحدار الخطى المتعدد نعلم فى حساب هذا المعامل على الحاسب الالكترونى توفيراً للجهد والوقت . وفى الحالة البسيطة التى يكون لدينا فيها ثلاثة متغيرات عشوائية v_1 ، v_2 ، v_3 يمكننا حساب معامل الارتباط المتعدد r_{12} بين المتغير v_1 والمتغيرين v_2 ، v_3 من الصيغة الآتية :

$$(18) \quad r_{12} = \frac{r_{12} - r_{13}r_{23}}{r_{22} - r_{23}r_{33}}$$

ويتطلب حساب هذه القيمة إيجاد كل من :

- r_{12} وهو معامل الارتباط البسيط بين المتغيرين v_1 ، v_2
- r_{13} ، وهو معامل الارتباط البسيط بين المتغيرين v_1 ، v_3
- r_{23} ، وهو معامل الارتباط البسيط بين المتغيرين v_2 ، v_3

ومن الواضح أن قيمة معامل الارتباط المتعدد تعتمد على قيم معاملات الارتباط البسيط بين أزواج المتغيرات المستخدمة ، وهذه الملاحظة صحيحة مهما كان عدد المتغيرات .

اختبار دلالة معامل الارتباط المتعدد :

لاختبار الفرض الصفري $H_0: \rho = 0$ ضد الفرض ص $H_1: \rho > 0$ (٣٢... ك + ١) \neq .
 نستخدم نفس الصيغة (٦) التي وردت بالبند (١٢ - ١) وهي :

$$F = \frac{\text{التباين المفسر}}{\text{التباين غير المفسر}}$$

$$(19) \quad \frac{r^2 / (1 + \dots + 32) \text{ ك}}{(1 - r^2) / (1 + \dots + 32) \text{ ك} - 1)} =$$

إذا توفرت الشروط المذكور في مستهل الجزء الثاني من هذا الفصل فإن هذه الإحصاءة يكون لها توزيع ف بدرجتي حرية ك ، $1 - ك - ٧$ حيث $ك + ١$ عدد جميع المتغيرات المستخدمة . يلاحظ أن كلا من الصيغتين (٦) ، (١٩) تختبر وجود أو عدم وجود العلاقة الخطية .

مثال (١٢ - ٤) :

في عينة عشوائية من ٢٣ مجموعة من القيم من مجتمع معتدل ذي ثلاثة متغيرات

عشوائية وجدت معاملات الارتباط البسيطة الآتية : $r_{11} = ٠,٣٨$ ، $r_{12} = ٠,٤٦$ ، $r_{22} = ٠,٥١$.
 أوجد معامل الارتباط المتعدد $r_{1(2)}$ واختبر دلالته .

الحل :

من الصيغة (١٨) :

$$\frac{٠,٤٦ \times ٠,٣٨ \times ٢ - ٠,٢١١٦ + ٠,١٤٤٤}{٠,٢٦٠١ - ١} = \frac{٠,١٤٤٤}{٠,٢٤٠} = r_{1(2)}$$

∴ معامل الارتباط المتعدد المطلوب $= \sqrt{٠,٢٤٠} = ٠,٤٩$ تقريبا .

من الصيغة (١٩) :

$$F_y = \frac{2 \div 0,240}{20 \div (0,24 - 1)} = 3,1579 \quad \text{بدرجتي حرية } 20, 2$$

وهذه القيمة ليست بذات دلالة عند المستوى ٠,٠٥ وتشير إلى عدم وجود ارتباط خطي بين المتغير V_1 والمتغيرين V_2 ، V_3 .

(١٢ - ٦) معامل الارتباط الجزئي

PARTIAL CORRELATION COEFFICIENT

إن معامل الارتباط البسيط r_{12} بين متغيرين عشوائيين V_1 ، V_2 يفترض أن هذين المتغيرين يتأثران ببعضهما فقط ولا يتأثران بأي متغير آخر . إلا أنه في كثير من الحالات يكون كل من هذين المتغيرين متأثراً تأثراً خطياً بمتغير عشوائي ثالث V_3 ، وفي هذه الحالة يقتضى القياس الصحيح للارتباط بين V_1 ، V_2 استبعاد أثر هذا المتغير الثالث من كل منهما . إن المعامل الذى يقيس مثل هذا الارتباط يسمى بمعامل الارتباط الجزئي بين V_1 ، V_2 وسنرمز له بالرمز $r_{12 \cdot 3}$. ويعرف هذا المعامل أيضا بأنه معامل الارتباط البسيط بين V_1 ، V_2 عند قيمة ثابتة للمتغير V_3 أى في قطاع الوحدات التى تأخذ نفس القيمة للمتغير V_3 . ويتفق هذان التعريفان طالما كان توزيع الاحتمال المشترك للمتغيرات الثلاثة معتدلاً *trivariate normal distribution* إذ أنه في هذه الحالة تظل قيمة $r_{12 \cdot 3}$ كما هى مهما كانت القيمة الثابتة للمتغير V_3 ، وهذه هى الحالة التى نتناولها هنا والتى تتمشى مع ما افترضناه في مستهل هذا الجزء .

ويقدر هذا المعامل من العينة بمقدار سنرمز له بالرمز $r_{12 \cdot 3}$ حيث

$$(20) \quad r_{12 \cdot 3} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{(r_{11}^2 - 1)(r_{22}^2 - 1)}}$$

وحيث r_{11} ، r_{12} ، r_{22} هي معاملات الارتباط البسيط بين أزواج المتغيرات .
وبنفس الطريقة نحسب كلا من r_{2-31} ، r_{1-32} مع تذكر أننا هنا لا نميز بين
متغير مستقل ومتغير تابع .

أما دلالة هذا المعامل فنختبره عن طريق الإحصاءة :

$$(21) \quad \text{بدرجات حرية } 3 - n \quad t = \frac{\sqrt{3-n} r_{2-31}}{\sqrt{r_{2-31}^2 - 1}}$$

أو عن طريق الإحصاءة المكافئة:

$$(22) \quad \text{بدرجات حرية } 1 ، 3 - n \quad f = \frac{r_{2-31}^2}{(3-n) / (r_{2-31}^2 - 1)}$$

مثال (١٢ - ٥) :

في دراسة للعلاقة بين ضغط الدم وكلوسترول الدم أخذت عينة عشوائية من
١٤٢ من السيدات كبار السن وقيس كل من ضغط الدم r_1 ودرجة تركيز
الكلوسترول r_2 والعمر r_3 ، ثم حسبت معاملات الارتباط البسيط لكل زوج
من هذه المتغيرات فوجدت كما يلي : $r_{11} = 0,2495$ ، $r_{21} = 0,3332$ ،
 $r_{32} = 0,5029$. ابحث دلالة معامل الارتباط بين ضغط الدم ودرجة تركيز
الكلوسترول (أولا) دون استبعاد متغير العمر ، (ثانيا) بعد استبعاد متغير
العمر .

الحل :

(أولا) معامل الارتباط بين ضغط الدم ودرجة تركيز الكلوسترول دون

استبعاد متغير العمر هو $r_{11} = 0,2495$ ونختبر دلالة هذا المعامل باستخدام الصيغة (٨) بالبند (٥ - ١٠) وهي :

$$t = \frac{r_{11} \sqrt{2 - n}}{\sqrt{r_{11}^2 - 1}}$$

بدرجات حرية $n - 2$

أو الصيغة المكافئة ف = $\frac{r_{11}}{(2 - n) / (r_{11}^2 - 1)}$ بدرجتي حرية ١ ، $n - 2$

لدينا : $t = \frac{140 \sqrt{0,2495}}{\sqrt{(0,2495) - 1}}$ = ٣,٠٤٨٦ بدرجات حرية ١٤٠

وهذه القيمة تزيد عن القيمة الحرجة $t_{0,01} [140] = 2,076$ فهي ذات دلالة عالية وتشير إلى وجود ارتباط خطي بين المتغيرين .

(ثانيا) معامل الارتباط بين ضغط الدم ودرجة تركيز الكلوسترول بعد استبعاد متغير العمر هو معامل الارتباط الجزئي r_{3-21} . من الصيغة (٢٠) نجد أن

$$r_{3-21} = \frac{0,2495 - 0,3332 \times 0,5029}{\sqrt{(0,3332 - 1)(0,5029 - 1)}} = 0,1233$$

ونختبر دلالة هذا المعامل من الصيغة (٢١) كالآتي :

$$t = \frac{139 \sqrt{0,1233}}{\sqrt{0,1233 - 1}} = 1,465$$

بدرجات حرية ١٣٩

وهذه القيمة تقل عن القيمة الحرجة ت = 1.1980 ، فهي ليست بذات دلالة عند المستوى 0.05 ، وتشير إلى أنه حين نستبعد عامل العمر لا نستطيع الاستدلال على وجود ارتباط خطى بين ضغط الدم وكلوسترول الدم . هذا مع ملاحظة أن كلا من ضغط الدم ومقدار الكلوسترول في الدم يزداد بزيادة العمر ، ومن هنا نرى أن النتيجة التى حصلنا عليها فى (أولا) كانت نتيجة مضللة .

معاملات الارتباط الجزئى من مراتب أعلى :

إن معامل الارتباط الجزئى من النوع r_{r-s} يوصف بأنه معامل ارتباط جزئى من المرتبة الأولى of the first order لأنه يعطى معامل الارتباط بين متغيرين بعد استبعاد أثر متغير واحد من كل منهما . وبالمثل إذا كان لدينا أربعة متغيرات عشوائية v_1 ، v_2 ، v_3 ، v_4 ، فإننا نرمز بالرمز r_{3-21} لمعامل الارتباط الجزئى بين المتغيرين v_3 ، v_4 بعد استبعاد أثر المتغيرين v_1 ، v_2 من كل منهما ، ونصف هذا المعامل حينئذ بأنه معامل ارتباط جزئى من المرتبة الثانية . ويحسب هذا المعامل من معاملات الارتباط الجزئية من المرتبة الأولى كالآتى :

$$(23) \quad \frac{r_{4-32} \cdot r_{4-31} - r_{4-21}}{\sqrt{(r_{4-32}^2 - 1)(r_{4-31}^2 - 1)}} = r_{4-21}$$

أو كالآتى :

$$(24) \quad \frac{r_{3-42} \cdot r_{3-41} - r_{3-21}}{\sqrt{(r_{3-42}^2 - 1)(r_{3-41}^2 - 1)}} = r_{3-21}$$

وتختبر دلالة هذا المعامل من الصيغة :

$$T = \frac{\sqrt{4 - n} \sqrt{r_{32} - r_{21}}}{\sqrt{r_{32} - r_{21}}} \quad \text{بدرجات حرية } n - 4 \quad (25)$$

ويتمد مفهوم معاملات الارتباط الجزئية لأي عدد منتهى ك من المتغيرات العشوائية التي تشترك في توزيع معتدل متعدد المتغيرات . ومعامل الارتباط الجزئي من المرتبة 2 (حيث $2 > k - 2$) هو معامل الارتباط بين اثنين من المتغيرات بعد استبعاد أثر 2 من المتغيرات العشوائية الأخرى . ويعتمد حساب معامل الارتباط الجزئي من المرتبة 2 على معاملات الارتباط الجزئية من المرتبة السابقة عليها أى من المرتبة 2 - 1 . هذا ويجدر بنا ملاحظة التماثل في صيغ هذه المعاملات . أما اختبار دلالة معامل الارتباط الجزئي من المرتبة 2 فهو تعميم للصيغتين (21) ، (25) ، ويأخذ الصيغة الآتية :

$$T = \frac{\sqrt{2 - 2 - n} \sqrt{r_{32} - r_{21}}}{\sqrt{r_{32} - r_{21}}} \quad \text{بدرجات حرية } n - 2 - 2 \quad (26)$$

حيث r_{32} هو معامل الارتباط الجزئي من المرتبة 2 .

تمارين (12 - 2)

(1) في عينة عشوائية من 30 وحدة من مجتمع معتدل ذي ثلاثة متغيرات وجد أن معامل الارتباط المتعدد $r_{(321)} = 0,5$. بين أن هذه القيمة تدل على وجود ارتباط في المجتمع بين المتغير x_1 والمتغيرين x_2 ، x_3 مستخدما مستوى الدلالة 0,05

(٢) فى عينة عشوائية حجمها ٣٩ من مجتمع معتدل ذى ثلاثة متغيرات وجد أن معاملات الارتباط البسيط بين أزواج هذه المتغيرات هى $r_{١٢} = ٠,٢٨$ ، $r_{١٣} = ٠,٤٩$ ، $r_{٢٣} = ٠,٥١$ أثبت أن $r_{١٢} = ٠,٤٤٥$ ، $r_{١٣} = ٠,٤٢$ واختبر دلالة كل منهما .

(٣) بين أن معامل الارتباط الجزئى من المرتبة الثانية والذى قيمته $r_{٢٣-١} = ٠,٥$ المحسوب من عينة عشوائية حجمها ٢٠ من مجتمع معتدل ذى أربعة متغيرات يكون ذا دلالة عند المستوى $٠,٠٥$

(٤) فى عينة عشوائية من ٥٤ من السيدات حسبت المقادير المستهلكة فى فترة ما من نوعين من الغذاء هما البروتين (ص١) والدهون (ص٢) كما حسب العمر ص٣ وقد وجدت معاملات الارتباط البسيط الآتية :

$r_{١٢} = ٠,٥٧٨٤$ ، $r_{١٣} = -٠,٤٨٦٥$ ، $r_{٢٣} = -٠,٥٢٩٦$ أوجد معامل الارتباط الجزئى بين استهلاك البروتين واستهلاك الدهون مستقلا عن أثر العمر ، واختبر دلالاته عند المستوى $٠,٠١$