الفصل الثاني: الانــحدار الخطــي البســـيط

Simple Linear Regression

1.2 مقدمـــــــــــة:

ان دراسة العلاقة بين المتغيرات الاقتصادية يتطلب تحديد المتغيرات المؤثرة فـــي تلــك العلاقة ومن ابسط واسهل أنواع العلاقات في التقدير والتحليل الإحصائي والاقتصادي، العلاقة بين متغيرين أحدهما المتغير التابع Dependent Variable، والثاني المتغير المستقل Independent Variable، وإذا رمزنا للمتغيرين بـــ Y و X وعلى التوالي، فان العلاقة الدالية التي تجمعهما تكون كالآتي:

$$Y = F(x) \qquad \ldots(1.2)$$

حيث يشير الرمز F الى كون المتغيرالتابع Y يعتمد على المتغير المستقل X. ولتحديد شـــكل العلاقـــة هـــذه – مـــا إذا كان خطياً أم غير خطي- يمكن الاستعانة بالنظرية الاقتصـــادية، كمـــا يستعان بالاقتصاد الرياضي والإحصاء لصياغة العلاقة واختبار المتغـــيرات، كما لابد من تحديد شكل العلاقة هذه إذ تحكم العلاقة بين المتغيرات بعدد من الأشكال (الصيغ) ابسطها وأكثرها شيوعاً الصيغة الخطية، وتسمى العلاقة الخطية بيـــن متغيرين بالانحدار الخطي البسيط، Simple Linear Regression، فالعلاقة الخطية بين X ولتكن دخل الأسرة و Y ولتكن الأنفاق على سلعة معينة يمكن ان تكتب بالصيغة الرياضية الآتية:

$$Y_i = B_0 + B_1 X_i \qquad \ldots(2.2)$$

حيث :

B_1 و B_0 عــبارة عــن معلمــات مجهولة القيم وثوابت يُشرحان من وجهة النظر الرياضية كالآتي:

B_0: تمــثل تقــاطع خــط الانحدار مع المحور العمودي وهي عبارة عن القيمة التي تتخذها Y عندما تكون قيمة X مساوية للصفر.

B_1: تمثل الميل.

ومــن وجهــة الــنظر الاقتصــادية تمــثل (B_0) حالة الكفاف و(B_1) الميل الحدي للاســتهلاك، وقيمة الميل عبارة عن مقدار الزيادة المتحققة في قيمة المتغير التابع Y نتيجة زيادة المتغير المستقل بمقدار وحدة واحدة.

غــير ان العلاقة أعلاه (2.2) لا يمكن ان تشرح العلاقة بين المتغيرين بشكل دقــيق، فهــناك أسباب مهمة تجعل هذه المعادلة غير معبرة عن العلاقة بين X وY تعبــيراً كاملاً فقد يكون هناك انحراف بين العلاقة الحقيقية والمعادلة الإحصائية التي تمــثلها نتيجة أخطاء في القياسات أو في اختيار المتغير المستقل، مما يتطلب إضافة متغــير جديــد يسمى بالحد العشوائي، Random Variable ويرمز له عادةً بالرمز (U) ودوره امتصــاص العوامل غير القابلة للقياس، وكذلك أخطاء القياس، عليه فان العلاقة من الصيغة (2.2) يجب ان تعدل لكي تضم حد الخطأ العشوائي حيث يصبح:

$$Y_i = B_0 + B_1 X_i + U_i \qquad \dots(3.2)$$

2.2 الفرضيات الخاصة بالمتغير العشوائي:

1- ان المتغير العشوائي (U_i) هو متغير تعتمد قيمته في أية فترة زمنية على عامل الصـــدفة، فقد تكون اكبر أو اصغر أو مساوية إلى الصفر، إلا أنها في المتوسط تساوي صفر، أي $E(U_i)=0$، ويمكن توضيح ذلك على النحو الآتي:

$$Y_i = B_0 + B_1X_i + U_i \qquad \ldots(4.2)$$

$$U_i = Y_i - B_0 - B_1X_i \qquad \ldots(5.2)$$

وبإدخال $\sum$ على طرفي المعادلة 5.2:

$$\sum U_i = \sum(Y_i - B_0 - B_1X_i)$$

$$\sum U_i = \sum Y_i - nB_0 - B_1\sum X_i \qquad \ldots(6.2)$$

$$\because B_0 = \overline{Y} - B_1\overline{X}$$

نعوّض عن B_0 بما يساويها في المعادلة (6.2):

$$\sum U_i = \sum Y_i - n(\overline{Y} - B_1\overline{X}) - B_1\sum X_i \qquad \ldots(7.2)$$

$$\because \overline{Y} = \frac{\sum Y_i}{n} \quad , \overline{X} = \frac{\sum X_i}{n}$$

وبحاصل ضرب الطرفين في الوسطين نحصل:

$$\sum Y_i = n\overline{Y} \quad , \sum X_i = n\overline{X}$$

وبالتعويض عن ذلك في المعادلة (7.2) تكون:

$$\sum U_i = \sum Y_i - \sum Y_i + B_1 \sum X_i - B_1 \sum X_i$$
$$\sum U_i = 0$$
$$E(U_i) = 0$$

2- ان المتغـــير العشوائي (U_i) يتوزع توزيعاً طبيعياً، Normally distributed، حـــول القيمة المتوقعة أو حول الوسط الحسابي المساوي للصفر عند كل قيمة من قيم المتغير المستقل X أي بشكل جرس.

3- ان تبايـــن، Variance، المتغـــير العشوائي (حد الخطأ)، حول الوسط الحسابي مقدار ثابت عند كل قيمة من قيم X أي:

$$\operatorname{var}(U_i) = E[U_i - E(U_i)]^2$$
$$\because E(U_i) = 0$$

$$\therefore \operatorname{var}(U_i) = E(U_i)^2 = 6^2$$

وإذا كان تباين الخطأ غير ثابت عندئذٍ تظهر مشكلة تسمى مشكلة عدم تجانس التباين، Heteroscedasticity، والتي سنتناولها بشيء من التفصيل لاحقاً.

الفرضيات الثلاث السابقة يمكن جمعها بشكل مختصر وتمثيلها كالآتي:

$$U_i \sim N(0, \sigma^2)$$

أي بمعـــنى ان الخطـــأ العشوائي، U_i، يتوزع، ~، توزيعاً طبيعياً، N، بوسط حسابي مساوي للصفر، 0، وتباين ثابت قيمته σ^2.

4- أن قـــيم U_i غير مرتبطة بأي من المتغيرات المستقلة، أي انعدام التباين المشترك Covariance بين U_i وX_i أي:

$$Cov(U_i X_i) = E(U_i X_i)$$
$$Cov(U_i X_i) = X_i E(U_i)$$
$$\because E(U_i) = 0$$
$$\therefore Cov(U_i X_i) = 0$$

ويمكن توضيح ذلك على النحو الآتي:

$$Y_i = B_0 + B_1 X_i + U_i \qquad ...(8.2)$$
$$U_i = Y_i - B_0 - B_1 X_i \qquad ...(9.2)$$

وبضرب طرفي المعادلة بـــ $\sum X_i$:

$$\sum X_i U_i = \sum X_i Y_i - B_0 \sum X_i - B_1 \sum X_i^2$$
$$\because B_0 = \overline{Y} - B_1 \overline{X}$$

وعند تعويض ذلك:

$$\sum X_i U_i = \sum X_i Y_i - \sum X_i (\overline{Y} - B_1 \overline{X}) - B_1 \sum X_i^2 \qquad ...(10.2)$$

$$\because \overline{Y} = \frac{\sum Y_i}{n} \quad , \quad \overline{X} = \frac{\sum X_i}{n}$$

وبتعويض ذلك يكون:

$$\sum X_i U_i = \sum X_i Y_i - \sum X_i \left(\frac{\sum Y_i}{n} - B_1 \frac{\sum X_i}{n}\right) - B_1 \sum X_i^2$$

$$\sum X_i U_i = \sum X_i Y_i - \sum X_i \left(\frac{\sum Y_i - B_1 \sum X_i}{n}\right) - B_1 \sum X_i^2$$

$$\sum X_i U_i = \sum X_i Y_i - \sum X_i Y_i + B_1 \sum X_i^2 - B_1 \sum X_i^2$$

وبعد الحذف والتبسيط يكون:

$$\sum X_i U_i = 0$$

5- القـــيم المخـــتلفة للمتغير العشوائي (U_i) تكون مستقلة عن بعضها البعض بعبارة أخـــرى التباين المشترك لــ U_i مع U_j مساوية للصفر، وعليه فان قيمة العنصر العشوائي في أي فترة لا تعتمد على قيمته في فترة أخرى أي:

$$Cov\ (U_i\ U_j) = E(U_i\ U_j) = 0 \qquad (i, j = 1,2,3,\ldots, n,\ \ i \neq j)$$

وإذا حـــدث وجـــود ارتـــباط بيـــنها تظهـــر مشكلة تسمى مشكلة الارتباط الذاتي، Autocorrelation، وسيتم شرحها لاحقاً.

6- انعـــدام العلاقـــة بين المتغيرات المستقلة وفي حالة وجود علاقة قوية بينها تظهر مشكلة تسمى مشكلة الارتباط الخطي المتعدد، Multicollinearity، والتي سيتم تناولها فيما بعد.

3.2 طريقة المربعات الصغرى،

:The Ordinary Least Squares (OLS)

بالرجوع إلى العلاقة الخطية بين دخل الأسرة X وانفاقها على سلعة معينة Y.

$$Y_i = B_0 + B_1 X_i + U_i \qquad \ldots(11.2)$$

يتبيـــن لـــنا بان تأثير الدخل في الأنفاق على السلعة موضوع البحث يتحدد من خلال العلاقة المنتظمة $(B_0 + B_1 X_1)$، أما تأثير العوامل الأخرى فانه متجسد في (U_i).

وعلـــيه فانه لمعرفة العلاقة الحقيقية بين دخل الأسرة وانفاقها على السلعة في القطر يتطلب احتساب B_0 و B_1، إلا أن احتساب المعالم المذكورة لا يمكن ان يتم إلا في حالة الحصول على دخل وانفاق جميع الأسر في ذلك القطر وهذا أمر غير ممكن

بسبب صعوبة العملية الإحصائية اللازمة ولتسهيل العمل تسحب عينة من أسر القطر، ومن ثم تقدر قيم المعالم ويتم التقدير بواسطة المعادلة:

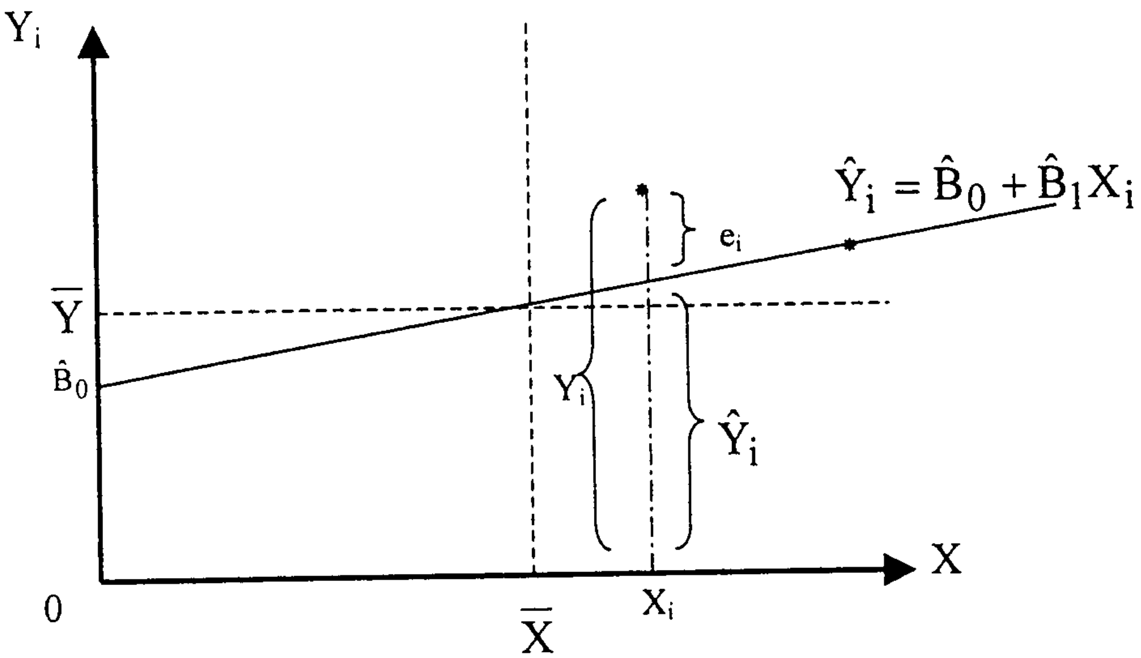
$$Y_i = \hat{B}_0 + \hat{B}_1 X_i + e_i \qquad \ldots(12.2)$$

ولتقدير تأثير الدخل بصورة مستقلة في الإنفاق فانه يتم بواسطة المعادلة التالية:

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i \qquad \ldots(13.2)$$

تسمى المعادلة (13.2) بمعادلة خط الانحدار، وتشير العلامة (^) إلى كون القيم تقديرية وليست حقيقية وكل نقطة من نقاطه ($\hat{Y}_i$) تمثل القيمة التقديرية لمتوسط إنفاق جميع العوائل ذات الدخل البالغ X. ويتبين من المعادلتين (12.2) و(13.2) بأن قيم المشاهدات الفعلية Y_i تنحرف عن القيم التقديرية ($\hat{Y}_i$) بمقدار e_i وكما مبين في الشكل الآتي:

شكل (1.2)



من الشكل يتبين ان:

$$e_i = Y_i - \hat{Y}_i$$

حيـــث يمكن للبواقي، e_i، ان تكون سالبة أو موجبة حسب موضع نقطة المشاهدة من الخـــط المقدر. ولإيجاد افضل خط مستقيم لعينة مشاهدات Y ، X من بين خطوط لا نهائـــية العـــدد تصف المعادلة الخطية، تستخدم طريقة المربعات الصغرى (OLS)، ويتضـــمن ذلـــك في محاولة جعل مجموع مربع انحرافات القيم الحقيقية Y_i عن القيم التقديرية $\hat{Y}_i$ اقل ما يمكن، أي جعل مجموع مربعات الأخطاء العشوائية عند نهايتها الصغرى وبما ان طريقة OLS تشترط تصغير القيمة $(\sum e_i^2)$ إلى الحد الأدنى فأنها عبارة عن مشكلة النهايات الصغرى أي:

$$\min \rightarrow \sum_{i=1}^{n} e_i^2$$

حيث ان:

$$e_i = Y_i - \hat{Y}_i$$

إذن:

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

بما ان معادلة الخط المستقيم الحقيقية غير المعروفة هي:

$$Y_i = B_0 + B_1 X_i$$

فان معادلة الخط المستقيم التقديرية تكون:

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i$$

بالتعويض عن $\hat{Y}_i$ بما يساويها نحصل:

$$\sum e_i^2 = \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i)^2$$

وكشـــرط رياضـــي لتصـــغير $\sum e_i^2$ تؤخـــذ المشتقات الجزئية لكل من $\hat{B}_0$ و $\hat{B}_1$ ومساواة كل منها بالصفر.

وان الشـــرط الجوهري للتصغير هو اخذ التفاضل الجزئي لمجموع المربعات بالنســـبة لمعـــاملات، Coefficient، الـــنموذج $\hat{B}_0$ و $\hat{B}_1$ ومساواة المشتقة الأولى بالصفر.

أي بتطبيق الشرط الضروري:

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_0} = 2\sum(Y_i - \hat{B}_0 - \hat{B}_1 X_i)(-1) = 0$$

$$-2\sum(Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0$$

بالقسمة على (2−) وفك القوس وترتيب المعادلة نحصل على:

$$\sum(Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0$$

$$\sum Y_i - n\hat{B}_0 - \hat{B}_1 \sum X_i = 0$$

$$\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_i \qquad \ldots(14.2)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_1} = 2\sum(Y_i - \hat{B}_0 - \hat{B}_1 X_i)(-X_i) = 0$$

$$-2\sum X_i (Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0$$

بالقسمة على (2−) نحصل على:

$$\sum X_i (Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0$$

$$\sum X_i Y_i - \hat{B}_0 \sum X_i - \hat{B}_1 \sum X_i^2 = 0$$

$$\sum X_i Y_i = \hat{B}_0 \sum X_i + \hat{B}_1 \sum X_i^2 \qquad \ldots(15.2)$$

تســـمى المعادلتين (14.2) و(15.2) بالمعادلتين الطبيعيتين (الآنيتين) حيث (n) عدد المشـــاهدات، X_i وY_i هي معلومة دائماً باعتبارهما قيم المشاهدات الحقيقية وبمجرد

تعويضـــهمـــا في المعادلتين (14.2) و(15.2) وبحلهما آنياً نحصل على قيم $\hat{B}_0$ و $\hat{B}_1$ اللتان تمثلان مقدرتان للمعلمتين الحقيقيتين B_0 و B_1.

1.3.2 طرق تقدير معاملات النموذج:

ولتقدير معاملات النموذج B_0 و B_1 نستعين بعدة طرق منها:

1- طريقة الحذف والتعويض.

2- طريقة المحددات.

3- طريقة التقدير حول نقطة المتوسط.

4- طريقة المصفوفات.

وسنبين هذه الطرق من خلال المثال الآتي، من دون تكرارها هنا وهناك.

مـــثال 1.2: الجدول الآتي يمثل عدد سنوات الخدمة (X_i) ومعدل الأجر السنوي (Y_i) بآلاف الدنانير لعينة تمثل (8) موظفين في أحد الدوائر.

X_i	Y_i	$X_i Y_i$	X_i^2
4	25.6	102.4	16
8	32.7	261.6	64
12	45.4	544.8	144
16	53.9	862.4	256
20	59.0	1180	400
24	62.6	1502.4	576
28	65.0	1820	784
32	65.5	2105.6	1024
$\sum X_i = 144$ $\overline{X} = 18$	$\sum Y_i = 410$ $\overline{Y} = 51.25$	$\sum X_i Y_i = 8379.2$	$4\sum X_i^2 = 326$

المطلوب: تقدير خط الانحدار بواسطة المعادلتين الطبيعيتين أعلاه.

1.1.3.2 طريقة الحذف والتعويض، Substitution Method:

$$\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_i \qquad \ldots(16.2)$$

$$\sum X_i Y_i = \hat{B}_0 \sum X_i + \hat{B}_1 \sum X_i^2 \qquad \ldots(17.2)$$

وبتعويض القيم من الجدول في المعادلتين 16.2، 17.2 نحصل على:

$$410 = 8\hat{B}_0 + 144\hat{B}_1 \qquad \ldots(18.2)$$

$$8379.2 = 144\hat{B}_0 + 3264\hat{B}_1 \qquad \ldots(19.2)$$

وبضرب المعادلة (18.2) في 18 نحصل:

$$7380 = 144\hat{B}_0 + 2592\hat{B}_1 \qquad \ldots(20.2)$$

$$8379.2 = 144\hat{B}_0 + 3264\hat{B}_1 \qquad \ldots(21.2)$$

وبطرح معادلة (20.2) من (21.2) نحصل:

$$999.2 = 672\hat{B}_1$$

$$\hat{B}_1 = \frac{999.2}{672} = 1.486904762$$

للحصـــول على قيمة $\hat{B}_0$ نعوض عن قيمة $\hat{B}_1$ في أحد المعادلتين الرئيستين ولتكن معادلة (18.2).

$$410 = 8\hat{B}_0 + 144\hat{B}_1$$

$$410 = 8\hat{B}_0 + 144(1.486904762)$$
$$410 = 8\hat{B}_0 + 214.1142857$$
$$410 - 214.1142857 = 8\hat{B}_0$$
$$195.8857143 = 8\hat{B}_0$$

$$\hat{B}_0 = \frac{195.8857143}{8} = 24.48571429$$

وعليه فان المعادلة المقدرة ، Estimated، للعلاقة بين عدد سنوات الخدمة X_i ومعدل الأجر السنوي Y_i للعينة المعنية تكون:

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i$$
$$\hat{Y}_i = 24.48571429 + 1.486904762 X_i$$

تشير المعادلة التقديرية إلى وجود علاقة طردية بين المتغير التابع Y_i الذي يمثل معدل الأجر السنوي للموظف والمتغير المستقل X_i الذي يمثل عدد سنوات الخدمة فبزيادة خدمته الوظيفية بمقدار سنة واحدة يزداد معدل اجره السنوي بمقدار 1486 دينار.

2.1.3.2 طريقة المحددات، Determinates Method:

ويمكن الحصول على قيم $\hat{B}_0$ و $\hat{B}_1$ باعتماد المحددات (قاعدة كرايمر) وذلك بإعادة كتابة المعادلتين الطبيعيتين (16.2) و(17.2) في صيغة مصفوفة وعلى النحو الآتي:

$$\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_i$$

$$\sum X_i Y_i = \hat{B}_0 \sum X_i + \hat{B}_1 \sum X_i^2$$

$$\begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \end{bmatrix}$$

ولتقدير $\hat{B}_0$ و $\hat{B}_1$ ينبغي تكوين المحددات الآتية:

$$|D| = \begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix} = (n)(\sum X_i^2) - (\sum X_i)(\sum X_i)$$

$$|A_0| = \begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix} = (\sum Y_i)(\sum X_i^2) - (\sum X_i Y_i)(\sum X_i)$$

$$|A_1| = \begin{vmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix} = (n)(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)$$

$$\hat{B}_1 = \frac{|A_1|}{|D|} = \frac{n(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{n(\sum X_i^2) - (\sum X_i)^2}$$

$$\hat{B}_0 = \frac{|A_0|}{|D|} = \frac{(\sum Y_i)(\sum X_i^2) - (\sum X_i Y_i)(\sum X_i)}{n(\sum X_i^2) - (\sum X_i)^2}$$

or

$$\hat{B}_0 = \overline{Y} - \hat{B}_1 \overline{X}$$

بالـــرجوع إلى بيانات المثال (1.2) وباعتماد المحددات نحصل على قيم $\hat{B}_0$ و $\hat{B}_1$ وكما يلي:

$$\begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \end{bmatrix}$$

$$|D| = \begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix} = \begin{vmatrix} 8 & 144 \\ 144 & 3264 \end{vmatrix}$$

$$|D| = (8)(3264) - (144)(144)$$

$$|D| = 26112 - 20736$$

$$|D| = 5376$$

$$|A_o| = \begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix} = \begin{vmatrix} 410 & 144 \\ 8379.2 & 3264 \end{vmatrix}$$

$$|A_o| = (410)(3264) - (144)(9379.2)$$

$$|A_o| = 1338240 - 1206604.8$$

$$|A_o| = 131635.2$$

$$|A_1| = \begin{vmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix} = \begin{vmatrix} 8 & 410 \\ 144 & 8379.2 \end{vmatrix}$$

$$|A_1| = (8)(8379.2) - (410)(144)$$

$$|A_1| = 67033.6 - 59040$$

$$|A_1| = 7993.6$$

$$\hat{B}_0 = \frac{|A_0|}{|D|} = \frac{131635.2}{5376} = 24.48571429$$

$$\hat{B}_1 = \frac{|A_1|}{|D|} = \frac{7993.6}{5376} = 1.486904762$$

وبذلك تكون المعادلة التقديرية:

$$\therefore \hat{Y}_i = 24.4857 + 1.486 X_i$$

3.1.3.2 طريقة التقدير حول نقطة المتوسط:

كما يمكن تقدير $\hat{B}_0$ و $\hat{B}_1$ بواسطة انحرافات المتغيرين Y_i و X_i عن وسطهما الحسابي $\overline{Y}$ و $\overline{X}$ باستخدام فكرة البواقي e_i فبقسمة المعادلة الطبيعية رقم (1) على n نحصل:

$$\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_i \qquad ...(22.2)$$

$$\frac{\sum Y_i}{n} = \hat{B}_0 \frac{n}{n} + \hat{B}_1 \frac{\sum X_i}{n}$$

$$\overline{Y} = \hat{B}_0 + \hat{B}_1 \overline{X} \qquad ...(23.2)$$

$$\therefore \hat{B}_0 = \overline{Y} - \hat{B}_1 \overline{X} \qquad ...(24.2)$$

ولإيجاد $\hat{B}_1$ نعود إلى معادلة الخط المستقيم التقديرية حيث:

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i \qquad ...(25.2)$$

وبطرح المعادلة رقم (23.2) من (25.2) نحصل على:

$$\hat{Y}_i - \overline{Y} = \hat{B}_0 + \hat{B}_1 X_i - \hat{B}_0 - \hat{B}_1 \overline{X} \qquad \ldots(26.2)$$

$$\hat{Y}_i - \overline{Y} = \hat{B}_1 X_i - \hat{B}_1 \overline{X}$$
$$\hat{Y}_i - \overline{Y} = \hat{B}_1 (X_i - \overline{X})$$

$$\hat{y}_i = \hat{B}_1 x_i \qquad \ldots(27.2)$$

$$\because e_i = y_i - \hat{y}_i$$
$$e_i = y_i - \hat{B}_1 x_i \qquad \ldots(28.2)$$

وبإدخال $\sum$ وتربيع الطرفين:

$$\sum e_i^2 = \sum (y_i - \hat{B}_1 x_i)^2 \qquad \ldots(29.2)$$

بإيجاد المشتقة الجزئية لـــ $\hat{B}_1$ ومساواتها بالصفر:

$$\frac{\partial \sum e_i^2}{\partial \hat{B}_1} = 2\sum (y_i - \hat{B}_1 x_i)(-x_i) = 0 \qquad \ldots(30.2)$$
$$-2\sum x_i (y_i - \hat{B}_1 x_i) = 0$$

وبالقسمة على (2-):

$$\sum x_i (y_i - \hat{B}_1 x_i) = 0$$
$$\sum x_i y_i - \hat{B}_1 \sum x_i^2 = 0$$

$$\sum x_i y_i = \hat{B}_1 \sum x_i^2$$

$$\hat{B}_1 = \frac{\sum x_i y_i}{\sum x_i^2} \qquad ...(31.2)$$

حيث ان:

$\sum x_i y_i$ تمثل مجموع حاصل ضرب انحرافات القيم عن الوسط الحسابي للمتغيرين X_i و Y_i، وان $\sum x_i^2$ تمثل مجموع مربع الانحرافات لقيم المتغير X_i عن وسطه الحسابي. وعليه فان المعادلتين (24.2) و(31.2) هي المعادلات الأساسية التي تستخدم في أيجاد قيم $\hat{B}_0$ و $\hat{B}_1$ بموجب طريقة الانحرافات.

وبالرجوع إلى بيانات الجدول (1.2) والتعبير عنها بصيغة انحرافات يمكن الحصول على قيم $\hat{B}_0$ و $\hat{B}_1$ كما يأتي:

X_i	Y_i	x_i	Y_i	$x_i y_i$	x_i^2
4	25.6	-14	-25.65	359.1	196
8	32.7	-10	-18.55	185.5	100
12	45.4	-6	-5.85	35.1	36
16	53.9	-2	2.65	-5.3	4
20	59.0	2	7.75	15.5	4
24	62.6	6	11.35	68.1	36
28	65.0	10	13.75	137.5	100
32	68.5	14	14.55	203.7	196
$\sum X_i = 144$ $\overline{X} = 18$	$\sum Y_i = 410$ $\overline{Y} = 51.25$	$\sum x_i = 0$ $x_i = X - \overline{X}$	$\sum y_i = 0$ $y_i = Y - \overline{Y}$	$\sum x_i y_i =$ 999.2	$\sum x_i^2 = 672$

$$\hat{B}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{999.2}{672} = 1.486904762$$

$$\hat{B}_0 = \overline{Y} - \hat{B}_1 \overline{X}$$
$$\hat{B}_0 = 51.25 - (1.486904762)(18)$$
$$\hat{B}_0 = 51.25 - 26.76428572$$
$$\hat{B}_0 = 24.48571428$$
$$\therefore \hat{Y}_i = 24.485 + 1.486 X_i$$

4.1.3.2 طريقة المصفوفات، Matrices Method:

إضافة إلى الطرق السابقة فأنه يمكن تقدير قيم $\hat{B}_0$ و $\hat{B}_1$ باعتماد صيغة المصفوفات حيث يمكن كتابة المعادلات الطبيعية على شكل مصفوفات، Matrices ومتجهات، Vectors، وكما يلي:

نفترض ان هناك علاقة تحتوي على i من المشاهدات من $1 \leftarrow n$ وان هناك عدد من المتغيرات من $1 \leftarrow K$.

$$Y_i = B_0 + B_1 X_i + U_i \qquad \ldots(32.2)$$

$$\begin{array}{l}
Y_1 = B_0 + B_1 X_{11} + B_2 X_{12} + \cdots\cdots + B_k X_{1k} + U_1 \\
Y_2 = B_0 + B_1 X_{21} + B_2 X_{22} + \cdots\cdots + B_k X_{2k} + U_2 \\
\vdots \quad\;\; \vdots \quad\;\; \vdots \qquad\quad \vdots \qquad\qquad\qquad\quad \vdots \qquad\quad \vdots \\
\vdots \quad\;\; \vdots \quad\;\; \vdots \qquad\quad \vdots \qquad\qquad\qquad\quad \vdots \qquad\quad \vdots \\
Y_n = B_0 + B_1 X_{n1} + B_2 X_{n2} + \cdots\cdots + B_k X_{nk} + U_n
\end{array}$$

يمكن تمثيل هذه المعادلات بصيغة رمز المصفوفات:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots X_{1k} \\ 1 & X_{21} & X_{22} & \cdots X_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots X_{nk} \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_k \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix}$$

$$Y = XB + U \qquad \ldots(33.2)$$

حيث ان:

Y: متجه عمودي أبعاده (nx1) يحتوي على n مشاهدة للمتغير التابع Y.

X: مصفوفة أبعادها (n x k+1) تحتوي مشاهدات المتغيرات المستقلة X. وعمودها الأول يحتوي على قيم الواحد صحيح لأخذ الثابت بنظر الاعتبار.

B: متجه عمودي أبعاده (k+1x1) تحتوي المعالم المجهولة.

U: متجه عمودي أبعاده (n x 1) يحتوي الخطأ العشوائي.

وللحصـــول على تقديرات المربعات الصغرى العادية لمتجه المعلمات يمكن كتابة المعادلة المقدرة التي يراد الحصول عليها وبصيغة المصفوفات كما يلي:

$$Y = \hat{Y} + e$$

$$Y = X\hat{B} + e$$

$$\therefore e = Y - X\hat{B} \qquad \ldots(34.2)$$

وباستخدام المبدلة لـــ e .

$$e'e = (Y - X\hat{B})'(Y - X\hat{B})$$

$$e'e = (Y' - \hat{B}'X')(Y - X\hat{B})$$
$$e'e = Y'Y - \hat{B}'X'Y - Y'X\hat{B} + \hat{B}'X'X\hat{B} \qquad \ldots(35.2)$$

بما ان الحد الثاني والثالث يمثلان قيمة واحدة، كما وان كل حد يمثل مبدلة للآخر فأن:

$$[\hat{B}'X'Y = (\hat{B}'X'Y)' = \hat{B}XY']$$
$$\therefore e'e = Y'Y - 2\hat{B}'X'Y + \hat{B}'X'X\hat{B} \qquad \ldots(36.2)$$

ولما كانت $\hat{B}' = (\hat{B}')' = \hat{B}$

وبأخذ المشتقة الجزئية لـــ $\hat{B}'$ ومساواتها بالصفر:

$$\frac{\partial e'e}{\partial \hat{B}'} = -2X'Y + 2X'X\hat{B} = 0$$

$$2X'X\hat{B} = 2X'Y$$

وبقسمة طرفي المعادلة على (2) نحصل:

$$X'X\hat{B} = X'Y$$

وبضرب طرفي المعادلة بالمعكوس $(X'X)^{-1}$ نحصل:

$$(X'X)^{-1}X'X\hat{B} = (X'X)^{-1}X'Y$$

وبما ان حاصل ضرب المعكوس $(X'X)^{-1}$ في المصفوفة $X'X$ يساوي مصفوفة الوحدة، I ، إذن المعادلة أعلاه تصبح:

$$\therefore \hat{B} = (X'X)^{-1}X'Y \qquad \ldots(37.2)$$

حيث ان:

$\hat{B}$: تمثل معاملات الانحدار المطلوب تقديرها.

$(X'X)^{-1}$: تمثل معكوس المصفوفة $(X'X)$.

$X'Y$: تمثل المتجه.

وبالـــرجوع إلى البيانات الواردة في الجدول (1.2) وباعتماد صيغة المصفوفات يمكن الحصول على قيم $\hat{B}_0$ و $\hat{B}_1$ وكالآتي:

$$\hat{B} = (X'X)^{-1} X'Y$$

$$(X'X) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 8 & 12 & 16 & 20 & 24 & 28 & 32 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 8 \\ 1 & 12 \\ 1 & 16 \\ 1 & 20 \\ 1 & 24 \\ 1 & 28 \\ 1 & 32 \end{bmatrix}$$

$$(X'X) = \begin{bmatrix} 8 & 144 \\ 144 & 3264 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{|X'X|} \operatorname{adj} X'X$$

$$|X'X| = \begin{bmatrix} 8 & 144 \\ 144 & 3264 \end{bmatrix} = (8)(3264) - (144)(144)$$

$$|X'X| = 26112 - 20736$$

$$|X'X| = 5376$$

$$\text{adj}X'X = \begin{bmatrix} 3264 & -144 \\ -144 & 8 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{5376}\begin{bmatrix} 3264 & -144 \\ -144 & 8 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 0.607142857 & -0.026785714 \\ -0.026785714 & 0.001488095 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 8 & 12 & 16 & 20 & 24 & 28 & 32 \end{bmatrix} \begin{bmatrix} 25.6 \\ 32.7 \\ 45.4 \\ 53.9 \\ 59.0 \\ 62.6 \\ 65.0 \\ 65.8 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 410 \\ 8379.2 \end{bmatrix}$$

$$\hat{B} = (X'X)^{-1}X'Y$$

$$\hat{B} = \begin{bmatrix} 0.607142857 & -0.026785714 \\ -0.026785714 & 0.001488095 \end{bmatrix} \begin{bmatrix} 410 \\ 8379.2 \end{bmatrix}$$

$$\begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \end{bmatrix} = \begin{bmatrix} 248.9285714 \quad + (-224.4428547) \\ -10.98214274 + 12.46904562 \end{bmatrix}$$

$$\begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \end{bmatrix} = \begin{bmatrix} 24.48571 \\ 1.48690 \end{bmatrix}$$

$$\therefore \hat{Y}_i = 24.485 + 1.486 X_i$$

يتضـــح ممـــا سبق بأن النتائج التي تم التوصل إليها في المعادلة التقديرية هي نفسها تماماً بالطرق الأربعة الآنفة الذكر.