الفصل الثاني: الانحدار الخطي البسيط

1.2 مقدمـة:

إن دراســة العلاقة بين المتغيرات الاقتصــادية يتطلب تحديد المتغيرات المؤثرة في تلك العلاقة ومن أبســط وأسـهل أنواع العلاقات في التقدير والتحليل الإحصـائي والاقتصـادي، العلاقة بين متغيرين أحدهما المتغير التابع Dependent Variable والثاني المتغير المستقل Independent Variable ، وإذا رمزنا للمتغيرين بـY وX وعلى التوالي، فإن العلاقة الدالية التي تجمعهما تكون كالآتي:

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

حيث يشـير الرمز F إلى كون المتغير التابع Y يعتمد على المتغير المسـتقل X. ولتحديد شـكل العلاقة هذه – ما إذا كان خطياً أم غير خطي – يمكن الاسـتعانة بالنظرية الاقتصـادية، كما يسـتعان بالاقتصـاد الرياضـي والإحصـــاء لصـــياغة العلاقة واختبار المتغيرات، كما لابد من تحديد شـــكل العلاقة هذه إذ تحكم العلاقة بين المتغيرات بعدد من الأشـكال (الصـيغ) أبسـطها وأكثرها شـيوعًا الصـيغة الخطية، وتسـمى العلاقة الخطبة بين متغيرين بالانحدار الخطي البســيط ، Simple Linear Regression، فالعلاقة الخطية بين X ولتكن دخل الأسرة و Y ولتكن الإنفاق على سلعة معينة يمكن أن تكتب بالصيغة الرياضية الآتية:

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

حيث B_1 وB_0 عبارة عن معلمات مجهولة القيم وثوابت يُشرحان من وجهة النظر الرياضية كالآتي:

B_0 : تمثل تقاطع خط الانحدار مع المحور العمودي وهي عبارة عن القيمة التي تتخذها Y عندما تكون قيمة X مساوية للصفر.

B_1: تمثل الميل.

ومن وجهة النظر الاقتصادية تمثل (B_0) حالة الكفاف و(B_1) الميل الحدي للاستهلاك، وقيمة الميل عبارة عن مقدار الزيادة المتحققة في قيمة المتغير التابع Y نتيجة زيادة المتغير المستقل بمقدار وحدة واحدة.

غير أن العلاقة أعلاه (2.2) لا يمكن أن تشــرح العلاقة بين المتغيرين بشــكل دقيق، فهناك أســباب مهمة تجعل هذه المعادلة غير معبرة عن العلاقة بين X و Y تعبيراً كاملاً فقد يكون هناك انحراف بين العلاقة الحقيقية والمعادلة الإحصـائية التي تمثلها نتيجة أخطاء في القياسـات أو في اختيار المتغير المستقل، مما يتطلب إضـافة متغير جديد يسمى بالحد العشوائي Random Variable ويرمز له عادة بالرمز (U) ودوره امتصـاص العوامل غير القابلة للقياس، وكذلك أخطاء القياس، عليه فإن العلاقة من الصــيغة (2.2) يجب أن تعدل لكي تضــم حد الخطأ العشوائي حيث يصبح:

$$\mathrm{Y}_i = \mathrm{B}_0 + \mathrm{B}_i \mathrm{X}_i + U_i \ldots\ldots (3.2)$$

2.2 الفرضيات الخاصة بالمتغير العشوائي:

1 –إن المتغير العشـــوائي (U_i) هو متغير تعتمد قيمته في أية فترة زمنية على عامل الصـــدفة (بخيت علي ,2009صـفحة30) ، فقد تكون أكبر أو أصـغر أو مسـاوية إلى الصـفر، إلا أنها في المتوسـط تسـاوي صفر، أي $E(U_i) = 0$، ويمكن توضيح ذلك على النحو الآتي:

$$Y_i = B_0 + B_1X_i + U_i \quad \dots\dots\dots(4.2)$$

$$U_i = Y_i - B_0 - B_1X_i \dots\dots\dots(5.2)$$

وبإدخال Σ على طرفي المعادلة 5.2:

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

$$\sum U_i = \sum Y_i - n\,B_0 - B_1\sum X_i \dots\dots(6.2)$$

$$\therefore B_0 = \bar{Y} - B_1\bar{X}$$

نعوض عن B_0 بما يساويها في المعادلة (6.2) :

$$\sum U_i = \sum Y_i - n(\bar{Y} - B_1\bar{X}) - B_1\sum X_i \dots\dots(7.2)$$

$$\therefore \ \bar{Y} = \frac{\sum Y_i}{n} \ , \ \ \bar{X} = \frac{\sum X_i}{n}$$

وبحاصل ضرب الطرفين في الوسطين نحصل:

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

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– إن المتغير العشـوائي (U_i) يتوزع توزيعاً طبيعياً ، Normally distributed حول القيمة المتوقعة أو حول الوسط الحسابي المساوي للصفر عند كل قيمة من قيم المتغير المستقل X أي بشكل جرس.

3–إن تباين Variance ، المتغير العشوائي (حد الخطأ)، حول الوسط الحسابي مقدار ثابت عند كل قيمة من قيم X أي:

$$var(U_i) = E[U_i - E(U_i)^2]$$

$$\therefore E(U_i) = 0$$

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

وإذا كـان تبـاين الخطـأ غير ثـابـت عنـدئـذ تظهر مشـــكلـة تســـمى مشـــكلـة عـدم تجـانس التبـاين Heteroscedasticity، والتي سنتناولها بشيء من التفصيل لاحقاً.

الفرضيات الثلاث السابقة يمكن جمعها بشكل مختصر وتمثيلها كالآتي:

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

أي بمعنى أن الخطأ العشوائي، U_i، يتوزع ~ توزيعاً طبيعياً، N، بوسط حسابي مساوي للصفر، 0، وتباين ثابت قيمته σ^2.

4- أن قيم U_i غير مرتبطة بأي من المتغيرات المستقلة، أي انعدام التباين المشترك Covariance بين U_i و X_i أي:

$$Cov(U_i X_i) = \mathrm{E}(U_i X_i)$$
$$Cov(U_i X_i) = X_i E(U_i)$$
$$\therefore\ E(U_i) = 0$$
$$\therefore E(U_i X_i) = 0$$

ويمكن توضيح ذلك على النحو الآتي:

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

وبضرب طرفي المعادلة بـ $\sum X_i$:

$$\sum X_i\, U_i = \sum X_i\, Y_i - B_0 \sum X_i - B_1 \sum X_i^2$$
$$\therefore\ B_0 = \bar{Y} - B_1 \bar{X}$$

وعند تعويض ذلك :

$$\sum X_i U_i = \sum X_i Y_i - \sum X_i\,(\bar{Y} - B_1 \bar{X}) - B_1 \sum X_i^2 \ldots\ldots\ldots (10.2)$$
$$\therefore \bar{Y} = \frac{\sum Y_i}{n} \quad , \quad \bar{X} = \frac{\sum X_i}{n}$$

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– القيم المختلفة للمتغير العشوائي (U_i) تكون مستقلة عن بعضها البعض، بعبارة أخرى التباين المشترك لـــ U_i مع U_j مساوٍ للصفر، وعليه فإن قيمة العنصر العشوائي في أي فترة لا تعتمد على قيمته في فترة أخرى أي:

$$cov\ (U_iU_j) = E(U_iU_j) = 0 \qquad (i, j\ =\ 1,2,3, \ldots, n, i \neq j)$$

وإذا حدث وجود ارتباط بينها تظهر مشكلة تسمى مشكلة الارتباط الذاتي Autocorrelation، وسيتم شرحها لاحقاً.

6 –انعدام العلاقة بين المتغيرات المستقلة وفي حالة وجود علاقة قوية بينها تظهر مشكلة تسمى مشكلة الارتباط الخطي المتعدد Multicollinearity، والتي سيتم تناولها فيما بعد (بخيت علي ,2009صفحة39) .

3.2 طريقة المربعات الصغري The Ordinary Least Squares (OLS):

بالرجوع إلى العلاقة الخطية بين دخل الأسرة X وإنفاقها على سلعة معينة Y:

$$Y_i = B_0 + B_1X_i + U_i \ldots \ldots \ldots \ldots (11.2)$$

يتبين لنا بأن تأثير الدخل في الإنفاق على السلعة موضوع البحث يتحدد من خلال العلاقة المنتظمة $(B_0 + B_1X_1)$، أما تأثير العوامل الأخرى فإنه متجسد في (U_i).

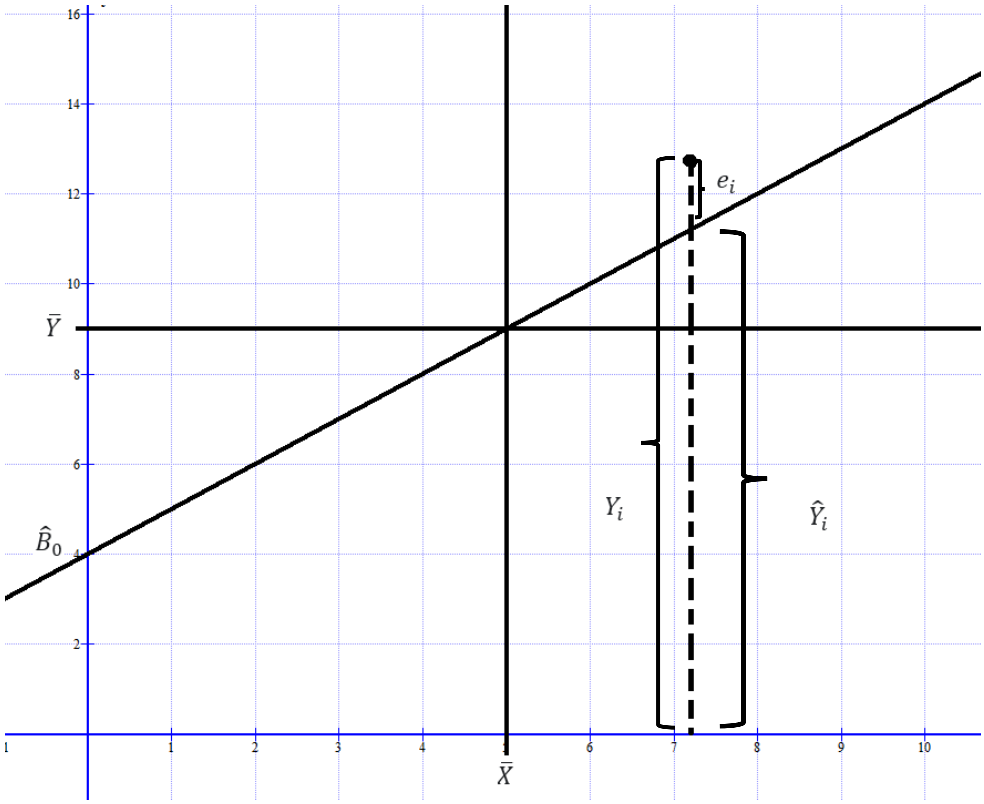
وعليه فإنه لمعرفة العلاقة الحقيقية بين دخل الأسرة وإنفاقها على السلعة في القطر يتطلب احتساب B_0 وB_1 ، إلا أن احتساب المعالم المذكورة لا يمكن أن يتم إلا في حالة الحصول على دخل وإنفاق جميع الأسر في ذلك القطر وهذا أمر غير ممكن بسبب صعوبة العملية الإحصائية اللازمة ولتسهيل العمل تسحب عينة من أسر القطر، ومن ثم تقدر قيم المعالم ويتم التقدير بواسطة المعادلة:

$$Y_i = \hat{B}_0 + \hat{B}_1X_i + e_i \ldots \ldots \ldots (12.2)$$

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

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1X_i \ldots \ldots \ldots \ldots (13.2)$$

تسمى المعادلة (13.2) بمعادلة خط الانحدار، وتشير العلامة (^) إلى كون القيم تقديرية وليست حقيقية وكل نقطة من نقاطه (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: الشكل (1.2)



من الشكل يتبين أن : $e_i = Y_i - \hat{Y}_i$

حيث يمكن للبواقي e_i أن تكون سالبة أو موجبة حسب موضع نقطة المشاهدة من الخط المقدر . ولإيجاد أفضل خط مستقيم لعينة مشاهدات Y ، X من بين خطوط لا نهائية العدد تصف المعادلة الخطية تستخدم طريقة المربعات الصغرى (OLS)، ويتضمن ذلك في محاولة جعل مجموع مربع انحرافات القيم الحقيقية Y_i عن القيم التقديرية $\hat{Y}_i$ أقل ما يمكن، أي جعل مجموع مربعات الأخطاء العشوائية عند نهايتها الصغرى وبما أن طريقة OLS تشترط تصغير القيمة $(\sum e_i^2)$ إلى الحد الأدنى فإنها عبارة عن مشكلة النهايات الصغرى أي:

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

حيث أن: $$e_i = Y_i - \hat{Y}_i$$

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

بما أن معادلة الخط المستقيم الحقيقية غير المعروفة هي: $Y_i = B_0 + B_1 X_i$

فإن معادلة الخط المستقيم التقديرية تكون: $\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i$

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

$$\sum e_i^2 = \sum \left(\hat{Y}_i - \hat{B}_0 - \hat{B}_1 X_i\right)^2$$

وكشرط رياضي لتصغير $\sum e_i^2$ تؤخذ المشتقات الجزئية لكل من $\hat{B}_0$ و $\hat{B}_1$ ومساواة كل منها بالصفر.

وإن الشرط الجوهري للتصغير هو أخذ التفاضل الجزئي لمجموع المربعات بالنسبة لمعاملات Coefficient ، النموذج $\hat{B}_0$و$\hat{B}_1$ ومساواة المشتقة الأولى بالصفر.

أي بتطبيق الشرط الضروري:

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

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

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

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

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

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

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

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

بالقسمة على (2−) نحصل على :

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

$$\sum X_i Y_i - \hat{B}_0 \sum X_i - \hat{B}_1 \sum X_1^2 = 0$$

$$\sum X_i Y_i = \hat{B}_0 \sum X_i - \hat{B}_1 \sum X_1^2 \ldots\ldots\ldots (15.2)$$

تسمى المعادلتين (14.2) و (15.2) بالمعادلتين الطبيعيتين (الآنيتين) حيث (n) عدد المشاهدات X_i و Y_i هي معلومة دائمًا باعتبارهما قيم المشـــاهدات الحقيقية وبمجرد تعويضـــهما في المعادلتين (14.2) و (15.2) وبحلهما آنيًا 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.

<u>مثال 1.2</u>: 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_iY_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
$\sum X_i = 144$ $\bar{X} = 18$	$\sum Y_i = 410$ $\bar{Y} = 51.25$	$\sum X_iY_i = 8379.2$	$4\sum X_i^2 = 326$

المطلوب : تقدير خط الانحدار بواسطة المعادلتين الطبيعيتين أعلاه.

1.1.3.2 طريقة الحذف والتعويض Substitution Method:

$$\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_i \dots\dots\dots..(16.2)$$
$$\sum X_iY_i = \hat{B}_0 \sum X_i + \hat{B}_1 \sum X_i^2 \dots\dots(17.2)$$

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

$$410 = 8\hat{B}_0 + 144\hat{B}_1 \dots\dots\dots\dots\dots.(18.2)$$
$$8379.2 = 144\hat{B}_0 + 3264\hat{B}_1 \dots\dots.(19.2)$$

وبضرب المعادلة (18.2) في 18 نحصل:

$$7380 = 144\hat{B}_0 + 2592\hat{B}_1 \dots\dots\dots\dots\dots.(20.2)$$
$$8379.2 = 144\hat{B}_0 + 3264\hat{B}_1 \dots\dots\dots\dots\dots.(21.2)$$

وبطرح المعادلة (20.2) من (21.2) نحصل:

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

للحصول 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 فإن 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 أجره السنوي بمقدار 1486 دينار.

2.1.3.2 طريقة المحددات Determinates Method:

ويمكن الحصول على قيم $\hat{B}_0$ و $\hat{B}_i$ 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:

$$|D| = \begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix} = (n)\left(\sum X_i^2\right) - \left(\sum X_i\right)\left(\sum X_i\right)$$
$$|A_0| = \begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix} = \left(\sum Y_i\right)\left(\sum X_i^2\right) - \left(\sum X_i Y_i\right)\left(\sum X_i\right)$$
$$|A_1| = \begin{vmatrix} \mathrm{n} & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix} = (\mathrm{n})\left(\sum X_i Y_i\right) - \left(\sum X_i\right)\left(\sum Y_i\right)$$

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

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

$$or: \quad \hat{B}_0 = \bar{Y} - \hat{B}_1 \bar{X}$$

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

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

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

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

$$|D| = 26112 - 20736$$

$$|D| = 5376$$

$$|A_0| = \begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix} = \begin{vmatrix} 410 & 144 \\ 8379.2 & 3254 \end{vmatrix}$$

$$|A_0| = (410)(3264) - (144)(9379.2)$$

$$|A_0| = 1338240 - 1206604.8$$

$$|A_0| = 131635.2$$

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

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

$$|A_1| = 67033.6 - 59040$$

$$|A_1| = 7993.6$$

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

$$\hat{B}_1 = \frac{|A_1|}{|D|} = \frac{7993.6}{5376} = 1.486901762$$

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 $\bar{Y}$ و $\bar{X}$ باستخدام فكرة البواقي e_i فبقسمة المعادلة الطبيعية رقم (1) على n نحصل على:

$$\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_i \dots\dots (22.2)$$
$$\frac{\sum Y_i}{n} = \hat{B}_0 \frac{n}{n} + \hat{B}_1 \frac{\sum X_i}{n}$$
$$\bar{Y} = \hat{B}_0 + \hat{B}_1 \bar{X} \dots\dots\dots (23.2)$$
$$\therefore \hat{B}_0 = \bar{Y} - \hat{B}_1 \bar{X} \dots\dots (24.2)$$

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

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

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

$$\hat{Y}_i - \bar{Y} = \hat{B}_0 + \hat{B}_1 X_i - \hat{B}_0 - \hat{B}_1 \bar{X} \dots\dots\dots (26.2)$$
$$\hat{Y}_i - \bar{Y} = \hat{B}_1 X_i - \hat{B}_1 \bar{X}$$
$$\hat{Y}_i - \bar{Y} = \hat{B}_1 (X_i - \bar{X})$$
$$\hat{y}_i = \hat{B}_1 X_i \dots\dots\dots\dots (27.2)$$
$$\because e_i = y_i - \hat{y}_i$$
$$e_i = y_i - \hat{B}_1 x_i \dots\dots\dots (28.2)$$

وبإدخال Σ وتربيع الطرفين:

$$\sum e_i^2 = \sum \left(y_i - \hat{B}_1 x_i\right)^2 \dots\dots\dots\dots (29.2)$$

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

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

وبالقسمة على (2−):

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

حيث أن $\sum x_i y_i$ تمثل مجموع حاصل ضرب انحرافات القيم عن الوسط الحسابي للمتغيرين X_i و Y_i وأن $\sum x_i^2$ تمثل مجموع مربع الانحرافات لقيم المتغير X_i عن وســـطه الحســـابي. وعليه فإن المعادلتين (24.2) و(31.2) هي المعادلات الأساسية التي تستخدم في إيجاد قيم $\hat{B}_0$ و$\hat{B}_1$ بموجب طريقة الانحرافات.

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

X_i	Y_i	x_i	Y_i	x_iy_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
$\sum X_i = 144$ $\bar{X} = 18$	$\sum Y_i = 410$ $\bar{Y} = 51.25$	$\sum X_i = 0$ $x_i = X - \bar{X}$	$\sum Y_i = 0$ $y_i = Y - \bar{Y}$	$\sum x_iy_i = 999.2$	$\sum x_i^2 = 672$

$$\hat{B}_1 = \frac{\sum x_iy_i}{\sum x_i^2} = \frac{999.2}{672} = 1.486904762$$
$$\hat{B}_1 = \bar{Y} - \hat{B}_1\bar{X}$$
$$\hat{B}_0 = 51.25 - (1.486904762)(18)$$
$$\hat{B}_0 = 51.25 - 26.76428572$$
$$\hat{B}_0 = 24.48571428$$
$$\therefore \hat{Y}_i = 24.485 + 1.486\, X_i$$

4.1.3.2 طريقة المصفوفات Matrices Method:

إضـافة إلى الطرق السـابقة فإنه يمكن تقدير قيم $\hat{B}_0$ و $\hat{B}_1$ باعتماد صـيغة المصـفوفات حيث يمكن كتابة المعـادلات الطبيعيـة على شـــكـل مصـــفوفـات Matrices ومتجهـات Vectors، وكمـا يلي (بخيـت علي ,2009صفحة50) :

نفترض أن هناك علاقة تحتوي على i من المشـــاهدات من $1 \leftarrow n$ وأن هناك عدد من المتغيرات من $1 \leftarrow K$.

$$Y_i = B_0 + B_1X_i + U_i \dots\dots\dots\dots\dots.(32.2)$$
$$Y_1 = B_0 + B_1X_{11} + B_2X_{12} + \cdots ..+B_kX_{1K} + U_1$$
$$Y_1 = B_0 + B_1X_{11} + B_2X_{12} + \cdots ..+B_kX_{1K} + U_1$$
$$Y_1 = B_0 + B_1X_{21} + B_2X_{22} + \cdots ..+B_kX_{2K} + U_2$$
$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$
$$Y_n = B_0 + B_1X_{n1} + B_2X_{n2} + \cdots ..+B_kX_{nK} + U_n$$

يمكن تمثيل هذه المعادلات بصيغة رمز المصفوفات:

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

$$Y = XB + U \dots\dots.. (33.2)$$

حيث أن:

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

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

B: منحه عمودي أبعاده $(k + 1 \times 1)$ تحتوي المعالم المجهولة.

U: سبحه عمودي أبعاده (nx1) يحتوي الخطأ العشوائي.

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

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

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

$$\therefore \mathrm{e} = \mathrm{Y} - \mathrm{X}\hat{\mathrm{B}} \dots\dots\dots\dots. (34.2)$$

وباستخدام المبدلة لـ e .

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

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

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

بما أن الحد الثاني والثالث يمثلان قيمة واحدة وأن كل حد يمثل مبدلة للآخر فإن:

$$\left[\hat{B}'X'Y = \left(\hat{B}'X'Y\right)' = \hat{B}XY'\right]$$

$$\therefore e'e = Y'Y - 2\hat{B}'X'Y + \hat{B}'X'X\hat{B} \dots\dots\dots.. (36.2)$$

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

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

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

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 أن حاصل ضرب المعكوس $(X'X)^{-1}$ في المصفوفة $X'X$ يساوي مصفوفة الوحدة 1 ، إذًا المعادلة أعلاه تصبح: $\therefore \hat{B} = (X'X)^{-1}X'Y \ldots\ldots\ldots (37.2)$

حيث أن:

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

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

$X'Y$: تمثل المتجه.

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

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

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

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

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

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

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

$$|X'X| = 5376$$

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

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

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

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

$$X'Y = \begin{bmatrix} 410 \\ 8379.2 \end{bmatrix}$$
$$\hat{B} = (X'X)^{-1}X'Y$$
$$\hat{B} = \begin{bmatrix} 0.607142857 & -0.026785714 \\ -0.026785714 & 0.001488095 \end{bmatrix} \begin{bmatrix} 410 \\ 8379.2 \end{bmatrix}$$
$$\begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \end{bmatrix} = \begin{bmatrix} 248.9285714 & +(-224.4428547) \\ -10.98214274 & +12.46904562 \end{bmatrix}$$
$$\begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \end{bmatrix} = \begin{bmatrix} 24.48571 \\ 1.48690 \end{bmatrix}$$
$$\therefore \ \hat{Y}_i = 24.485 + 1.486\, X_i$$

يتضح مما سبق بأن النتائج التي تم التوصل إليها في المعادلة التقديرية هي نفسها تمامًا بالطرق الأربعة الآنفة الذكر.

2-3-2 الخواص الأساسية لمقدرات المربعات الصغرى:

في كل تقدير يتم الحصول عليه، هناك خصائص عدة مرغوب فيها لذلك التقدير، ومن هذه الخصائص خاصية أفضل مقدر خطي غير متحيز، Best Linear Unbiased Estimator (BLUE) فكل مقدر $(\hat{B})$ يمكن التعبير عنه كدالة خطية بالنسبة لمشاهدات المتغير التابع (Y)، أي أن:

$$\hat{B} = K_1Y_1 + K_2Y_2 + \cdots . + K_nY_n$$

حيث أن: K_i ،$(i = 1,2,\ldots,n)$ عبارة عن أوزان أو قيم ثابتة. ومن بين جميع المقدرات الخملية نبحث عن المقدرات غير المتحيزة، ونقصد بعدم التحيز هو أن يكون الفرق بين القيمة المتوقعة للمقدر وقيمة المعلمة الحقيقية يساوي صفر، أي: $I:(B) = B - 0$

وأفضل مقدر هو ذلك المقدر الذي يكون تباينه حول الوسط الحسابي أقل ما يمكن فإذا كان B^*، $\hat{B}$ مقدرات خطية غير متحيزة فإن $\hat{B}$ أفضل مقدر إذا تحققت العلاقة الآتية:

$$var(\hat{B}) < var\,(B^*)$$

وسوف نتناول أدناه هذه الخصائص بشيء من التفصيل:

1.2.3.2 الخاصية الخطية Linearity Property:

مقدرات المربعات الصغرى خطية في المتغير التابع حيث نلاحظ أن تلك المقدرات يمكن وصفها في صورة دالة أو ترتيب خطي من قيم المتغير Y، أي:

$$\hat{B}_1 = K_1Y_1 + K_2Y_2 + \cdots + K_nY_n$$
$$\therefore \hat{B}_1 = \sum_{i=1}^{n} K_iY_i$$

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

وأن: للمشاهدة الواحدة: $y_i = Y_i - \bar{Y}$

$$\therefore \hat{B}_1 = \frac{\sum x_i (Y_i - \bar{Y})}{\sum x_i^2} = \frac{\sum x_i Y_i}{\sum x_i^2} - \frac{\bar{Y} \sum x_i}{\sum x_i^2}$$

$$\therefore \sum x_i = \sum (X_i - \bar{X}) = 0$$

$$\therefore \frac{\bar{Y} \sum x_i}{\sum x_i^2} = 0$$

لذلك تكون: $\hat{B}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \sum \left[\frac{x_i}{\sum x_i^2} \ Y_i \right]$

ولما كانت قيمX ثابتة نجد أن: $\frac{\sum x_i}{\sum x_i^2}$ هي مقادير ثابتة، ويمكن أن نرمز لها بالرمز K_i، وهي ثابتة في كل العينات، أي: $K_i = \frac{x_i}{\sum x_i^2}$

$\hat{B}_1$ مجموع مرجح لقيم المتغير التابع، Y_i

$$\therefore \hat{B}_1 = \sum K_i Y_i$$

$$\therefore \hat{B}_1 = F(Y_i)$$

إذًا $\hat{B}_1$ هو مقدر خطي.

خصائص الأوزان Weighted Properties :

وبما أن الأوزان تعتمد على قيم X الثابتة فقط فإنها تعتبر ثابتة أيضًا، وتخضع K_i للشروط الآتية:

1-مجموع الأوزان يساوي صفر، أي أن:

$$\sum_{i=1}^{n} K_i = 0$$

إذًا:

$$\therefore K_i = \frac{x_i}{\sum x_i^2}$$

$$\therefore x_i = \sum (X_i - \bar{X}) = 0$$

وبالتعويض، فإن:

$$\therefore \sum K_i = \frac{0}{\sum x_i^2} = 0$$

2- مجموع حاصل الضرب للأوزان (K) في قيم المتغير المستقل (X_i) أو في انحرافاتها عن متوسطها الحسابي (x_i) يساوي الواحد صحيح، أي:

$$\sum_{i=1}^{n} K_i X_i = \sum_{i=1}^{n} k_i x_i = 1$$

$$\sum K_i x_i = \sum K_i (X_i - \bar{X}) = \sum K_i X_i - \bar{X} \sum K_i$$

وقد بينا سابقاً أن:

$$\because \sum K_i X_i = 0$$

$$\therefore \sum K_i x_i = \sum K_i X_i$$

وبما أن:

$$\because K_i = \frac{x_i}{\sum x_i^2}$$

إذًا:

$$\sum K_i x_i = \sum \frac{x_i}{\sum x_i^2} x_i$$

لذلك تكون:

$$\sum K_i x_i = \frac{\sum x_i^2}{\sum x_i^2} = 1$$

$$\therefore \sum_{i=1}^{n} K_i x_i = \sum_{i=1}^{n} K_i X_i = 1$$

3- مجموع مربعات الأوزان $\sum K_i^2$، يساوي معكوس مجموع مربعات انحرافات المتغير المستقل $\frac{1}{\sum x_i^2}$، أي أن:

$$K_i = \frac{x_i}{\sum x_i^2} = \frac{(X_i - X)}{\sum (X_i - \bar{X})^2}$$

وعند تربيع الطرفين، ثم جمعهما، يكون:

$$\sum K_i^2 = \frac{\sum (X_i - \bar{X})^2}{[\sum (X_i - \bar{X})^2]^2}$$

وبعد الاختصار والترتيب:

$$\sum K_i^2 = \frac{1}{\sum (X_i - \bar{X})^2} = \frac{1}{\sum x_i^2}$$

وبطريقة مماثلة يمكن البرهنة بأن $\hat{B}_0$ دالة خطية من قيم المتغير Y.

$$\hat{B}_0 = W_1Y_1 + W_2Y_2 \ldots \ldots . + W_nY_n$$

$$\therefore \hat{B}_0 = \sum_{i=1}^{n} W_iY_i$$

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

وأن:

$$\bar{Y} = \frac{\sum Y_i}{n}, \hat{B}_1 = \sum K_iY_i$$

$$\therefore \hat{B}_0 = \frac{\sum Y_i}{n} - \bar{X}\sum K_iY_i$$

ومن خلال إعادة الترتيب، نحصل:

$$\hat{B}_0 = \sum \left[\frac{1}{n} - \bar{X}K_i\right] Y_i$$

ولما كانت $K_i, \bar{X}$ مقادير ثابتة، نرمز لها بالرمز W_i، أي أن:

$$\mathrm{W}_i = \left[\frac{1}{n} - \bar{X}K_i\right]$$

وبتعويض ذلك في المعادلة أعلاه، نجد أن $\hat{B}_0$ تعتمد على قيم Y_i فقط، بعبارة أخرى:

$$\hat{B}_0 = \sum W_oY_i$$

$$\therefore \hat{B}_0 = \mathrm{F}(Y_i)$$

إذاً $\hat{B}_0$ هو مقدر خطي.

2.2.3.2 Unbiasedness Property: (نظير مدكور, 2006-2007صفحة22)

تقصـــد بعدم التحيز هو أن يكون الفرق بين القيمة المتوقفة للمقدر $\mathrm{E}(\widehat{\mathrm{B}})$ وقيمة المعلمة الحقيقية للمجتمع الإحصائي (B)، يساوي صفر، أي: $\mathrm{E}(\widehat{\mathrm{B}}) - B = 0$

بعبارة أخرى يعتبر المقدر غير متحيز إذا كان وسطها يساوي القيمة الحقيقية للمعلمة: $\mathrm{E}(\widehat{\mathrm{B}}) = B$

ولإثبات خاصية عدم التحيز بالنسبة $\widehat{\mathrm{B}}_1$ نتبع ما يلي:

من خاصية الخطية: $\widehat{\mathrm{B}}_1 = \sum K_iY_i$

وباستحضار المعادلة: $Y_i = B_0 + B_1X_i + U_i$

وبتعويض ذلك في أعلاه: $\widehat{\mathrm{B}}_1 = \sum K_i\,(B_0 + B_1X_i + U_i)$

وبفتح القوس نحصل:

$$\widehat{\mathrm{B}}_1 = B_0\sum K_i + B_1\sum K_iX_i + \sum K_1U_i \ldots \ldots \ldots (38.2)$$

وباستخدام شروط الأوزان فإن:

$$\sum K_i = 0$$
$$B_0 \sum K_i = 0$$
$$\sum K_i X_i = 1$$
وبتعويض ذلك في المعادلة (38.2)، نحصل :
$$\therefore \hat{B}_1 = B_1 + \sum K_1 U_i \dots \dots \dots (39.2)$$
وبأخذ توقع طرفي المعادلة:
$$E(\hat{B}_1) = B_1 + \sum K_i E(U_i)$$
من فرضيات الخطأ العشوائي:
$$\because E(U_i) = 0$$
$$\because EK_i E(U_i) = 0$$
$$\therefore E(\hat{B}_1) = B_1 \dots \dots \dots \dots \dots \dots \dots (40.2)$$
ويعني ذلك أن $\hat{B}_1$ هو تقدير غير متحيز للقيمة الأصلية، B.

إثبات خاصية عدم التحيز لـ $\hat{B}_0$:

وبنفس المنهجية يمكن البرهنة بأن $\hat{B}_0$ تعتبر مقدرة غير متحيزة للمعلمة الحقيقية B_0.

من الخاصية الخطية يتبين لنا :
$$\hat{B}_0 = \sum W_i Y_i$$
$$W_i = \left[\frac{1}{n} - \bar{X} K_i\right]$$
وأن:
$$Y_i = B_0 + B_1 X_i + U_i$$
وعند التعويض نحصل:
$$\therefore \hat{B}_0 = \sum \left[\frac{1}{n} - \bar{X} K_i\right] [B_0 + B_1 X_i + U_i]$$
وعند فك الأقواس والتعويض، نحصل:
$$\hat{B}_0 = B_0 - B_0 \bar{X} \sum K_i + B_1 \bar{X} - \bar{X} B_1 \sum K_i X_i + \sum \left(\frac{1}{n} - \bar{X} K_i\right) U_i \dots \dots (41.2)$$
وباستخدام شروط الأوزان:
$$\because \sum K_i = 0$$
$$\therefore B_0 \bar{X} \sum K_i = 0$$

$$\because \sum K_i X_i = 1$$

$$\therefore \hat{B}_0 = B_0 + \sum \left(\frac{1}{n} - \bar{X} K_i\right) U_i \ldots \ldots . (42.2)$$

وبأخذ توقع طرفي المعادلة:

$$E(\hat{B}_0) = B_0 + \sum \left(\frac{1}{n} - \bar{X} K_i\right) E(U_i)$$

$$\because E(U_i) = 0$$

$$\therefore \sum \left(\frac{1}{n} - \bar{X} K_i\right) E(U_i) = 0$$

$$\therefore E(\hat{B}_0) = B_0 \ldots \ldots \ldots . . (43.2)$$

ويعني ذلك أن $\hat{B}_0$ هو تقدير غير متحيز للقيمة الأصلية، B_0

3.2.3.2 خاصية أفضل مقدر أقل تباين Best Minimum Variance:

إن مفهوم تباين المعالم يُحدد بواسطة الانحراف بين المعالم المقدرة لـ B وقيمتها المتوقعة .$E(\hat{B})$

فبالنسبة لـ $\hat{B}_1$ فإن (نظير مذكور, 2006−2007صفحة28) :

$$var(\hat{B}_1) = E[\hat{B}_1 - E(\hat{B}_1)]^2 \ldots \ldots \ldots \ldots (44.2)$$

من خاصية عدم التحيز : $E(\hat{B}_1) = B_1$

$$var(\hat{B}_1) = E[\hat{B}_1 - B_1]^2 \ldots \ldots \ldots \ldots (45.2)$$

بالرجوع إلى المعادلة (39.2) من خاصية عدم التحيز والخاصة بـ $\hat{B}_1$ نجد أن:

$$\hat{B}_1 = B_1 + \sum K_i U_i$$

$$\hat{B}_1 - B_1 = \sum K_i U_i$$

بتربيع الطرفين:

$$(\hat{B}_1 - B_1)^2 = \left(\sum K_i U_i\right)^2$$

وبأخذ توقع طرفي المعادلة:

$$E(\hat{B}_1 - B_1)^2 = E\left(\sum K_i U_i\right)^2$$

$$var(\hat{B}_1) = E\left[\sum K_i U_i\right]^2 \ldots \ldots \ldots \ldots \ldots (46.2)$$

وبفك القوس، نحصل:

$$var(\hat{B}_1) = \sum K_i^2 (U_i)^2 + 2 \sum_{i<j} \sum K_i U_j \, E(U_i U_j)$$

ومن فرضيات الخطأ العشوائي أن:

$$E(U_i)^2 = \sigma^2$$

وأن:

$$E\left(U_i U_j\right) = 0 \qquad For\ all\ i \neq j$$

$$\therefore var\left(\hat{B}_1\right) = \sum K_i^2\, \sigma^2 \ldots \ldots \ldots (47.2)$$

$$\because K_i = \frac{x_i}{\sum x_i^2}$$

$$\sum K_i^2 = \frac{\sum x_i^2}{(\sum x_i^2)^2} = \frac{1}{\sum x_i^2}$$

$$\therefore var\left(\hat{B}_1\right) = \frac{1}{\sum x_i^2} . \sigma^2 = \frac{\sigma^2}{\sum x_i^2}$$

وبالأسلوب نفسه يمكن تحديد التباين لـ $\hat{B}_0$ على النحو الآتي:

$$\therefore var\left(\hat{B}_0\right) = E\left[\hat{B}_0 - E\left(\hat{B}_0\right)\right]^2 \ldots \ldots \ldots \ldots \ldots (48.2)$$

من خاصية عدم التحيز :

$$E\left(\hat{B}_0\right) = B_0$$

$$var\left(\hat{B}_0\right) = E\left[\hat{B}_0 - B_0\right]^2$$

وبالرجوع إلى المعادلة (42.2) من خاصية عدم التحيز والخاصة بـ $\hat{B}_0$ نجد أن:

$$\hat{B}_0 = B_0 + \sum \left(\frac{1}{n} - \bar{X} K_i\right) U_i$$

$$\hat{B}_0 - B_0 = \sum \left(\frac{1}{n} - \bar{X} K_i\right) U_i$$

$$var\left(\hat{B}_0\right) = E\left[\sum \left(\frac{1}{n} - \bar{X} K_i\right) U_i\right]^2$$

$$var\left(\hat{B}_0\right) = \sum \left(\frac{1}{n} - \bar{X} K_i\right)^2 E(U_i)^2$$

$$E(U_i)^2 = \sigma^2$$

$$\therefore var\left(\hat{B}_0\right) = \sigma^2 \left[\sum \left(\frac{1}{n} - \bar{X} K_i\right)^2\right]$$

وبفك القوس وباستخدام شروط الأوزان نحصل على:

$$\sum K_i = 0$$

$$\therefore \frac{2\bar{X}^2}{n} \sum K_i = 0$$

وأن:

$$\sum K_i^2 = \frac{1}{\sum X_i^2}$$

$$\therefore var(\hat{B}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum X_i^2}\right)$$

ولإثبات أن مقدرات (OLS) هي أفضـل مقدرات خطية غير متحيزة أي أنها تتسـم بالكفاءة، ونتيجة لنسـبية معيار الأفضـلية فإننا نقوم بتعريف مقدرة أخرى ولتكن B_1^* تختلف عن مقدرة OLS ، $\hat{B}_1$ حيث سيتضـح لاحقاً أن تباين تلك المقدرة B_1^* لابد أن يفوق تباين مقدرة OLS ، $\hat{B}_1$ وبالتالي نفضـــل المقدرة الأخيرة $\hat{B}_1$صـــاحبة التباين الأقل.

$$B_1^* = \sum C_i Y_i \ldots (49.2)$$
$$C_i = K_i + d_i$$

وأن $d_i \neq 0$ ، حيث d_i كميات ثابتة:

$$\therefore Y_i = B_0 + B_1 X_i + U_i$$
$$B_1^* = \sum C_i \,(B_0 + B_1 X_i + U_i) \ldots \ldots \ldots (50.2)$$
$$B_1^* = B_0 \sum C_i + B_1 \sum C_i X_i + \sum C_i X U_i$$

ولكي تكون B_1^* غير متحيزة أي $E(B_1^*) = B_1$ يجب أن تتصف C_i بالصفات الآتية:

أ) $\sum C_i = 0$

أو: $\sum K_i + \sum d_i = 0$

معروفة مسبقًا : $\therefore \sum K_i = 0$

– $\sum d_i = 0$

أي إن: $\therefore \sum C_i = 0$

ب) $\sum C_i X_i = 1$

أو: $\because \sum K_i X_i + \sum d_i X_i = 1$

معروفة مسبقًا.

$\sum K_i X_i = 1$

أي أن: $\sum d_i X_i = 0$

$$\therefore \sum C\, X_i = 1 + 0 + 1$$
$$B_1^* = B_1 + \sum C_i U_i \ldots \ldots .. (51.2)$$
$$\because E(U_i) = 0$$
$$\therefore \sum C_i \, E(U_i) = 0$$
$$E(B_1^*) = B_1 \ldots \ldots \ldots \ldots (52.2)$$

هنا المقدر B_1^* غير متحيز ولتحديد تباين هذه المقدرة الخطية غير المتحيزة فإننا نعوض في قانون var.

$$varB_1^* = E[B_1^* - E(B_1^*)]^2$$

من المعادلة رقم (52.2): $\because\ E(B_1^*) = B_1$

$$varB_1^* = E[B_1^* - B_1]^2$$

وباستخدام المعادلة رقم (51.2):

$$B_1^* = B_1 + \sum C_i U_i$$

$$B_1^* - B_1 = \sum C_i U_i$$

$$.\,var(B_1^*) = E\left[\sum C_i U_i\right]^2 ..(53.2)$$

وبفك القوس، نحصل على:

$$var(B_1^*) = \sum C_i^2\, E(U_i)^2 + 2\sum_{i>j} C_i C_j\, E\left(C_i C_j\right)$$

ومن فرضيات الخطأ العشوائي: $E(U_i)^2 = \sigma^2$

وأن: $E\left(U_i U_j\right) = 0$

$$var(B_1^*) = \sum C_i^2\, \sigma^2 \ldots\ldots.(54.2)$$

ولما كانت: $C_i = K_i + d_i$

$$var(B_1^*) = \sum \left(K_i + \sum d_i\right)^2 \sigma^2$$

$$var(B_1^*) = \sigma^2 \left[\sum K_i^2 + \sum d_i^2 + \sum K_i d_i\right]$$

$$var(B_1^*) = \sigma^2 \sum K_i^2 + \sigma^2 \sum d_i^2 + 2\sigma^2 \sum K_i d_i$$

$$\because \sum K_i = 0$$

$$\because \sum K_i d_i = 0$$

$$var(B_1^*) = \sigma^2 \sum K_i^2 + \sigma^2 \sum d_i^2 \ldots\ldots\ldots.(55.2)$$

ولما كانت: $\sum K_i^2 = \frac{1}{\sum x_i^2}$

$$var(B_1^*) = \sigma^2 \frac{1}{\sum x_i^2} + \sigma^2 \sum d_i^2 \ldots\ldots\ldots\ldots(56.2)$$

$$\sigma^2 \frac{1}{\sum x_i^2} = var(B_1^*)$$

$$\therefore\ var(B_1) = var\left(\hat{B}_1\right) + \sigma^2 \sum d_i^2 \ldots\ldots\ldots.(57.2)$$

وحيث أن: $\sum d_i^2 > 0$

عليه فإن: $var(B_1^*) > var(\hat{B}_1)$

ومن ثم فإننا نرى ان مقدرة OLS الخاصة بالمعلمة $\hat{B}_1$ تتميز بأصغر تباين من بين المقدرات الخطية غير المتحيزة للمعلمة B_1^* .

4.2 تقدير تباين حد الخطأ العشوائي:

يستعمل تباين $\hat{B}_0$ و $\hat{B}_1$ في إجراء الاختبارات المعنوية الخاصة بتلك المقدرات، غير أن تباين المقدرات يحتوي على معلمة مجهولة هي (σ^2) تباين حد الخطأ العشوائي الذي يرمز له بالرمز (S_e^2) و يمكن اشتقاقه كالآتي:

بما أن المعادلة التقديرية تأخذ الصيغة الآتية:

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i \ldots\ldots..(58.2)$$

وبإدخال Σ على طرفي المعادلة :

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

وبالقسمة على n:

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

$$\overline{\hat{Y}} = \hat{B}_0 + \hat{B}_1 \bar{X} \ldots\ldots\ldots\ldots\ldots (59.2)$$

وبطرح المعادلة (59.2) من المعادلة (58.2) نحصل:

$$\hat{Y}_i - \overline{\hat{Y}} = \hat{B}_0 + \hat{B}_1 X_i - \hat{B}_0 - \hat{B}_1 \bar{X}$$

وبعد الاختصار، نحصل:

$$\hat{Y}_i = \hat{B}_1 X_i - \hat{B}_1 \bar{X}$$

إذًا:

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

$$\hat{Y}_i = \hat{B}_1 x_i \ldots\ldots\ldots (60.2)$$

$$\because e_i = y_i - \hat{y}_i$$

وبالتعويض من المعادلة (60.2):

$$\therefore e_i = y_i - \hat{B}_1 x_i$$

وبتربيع طرفي هذه المعادلة:

$$e_i^2 = \left(y_i - \hat{B}_1 x_i\right)^2$$

وعند فك الأقواس، نحصل:

$$e_i^2 = y_i^2 + \hat{B}_1^2 x_i^2 - 2\hat{B}_1 x_i y_i$$

وبإدخال Σ على طرفي المعادلة:

$$\sum e_i^2 = \sum y_i^2 + \hat{B}_1^2 \sum x_i^2 - 2\hat{B}_1 \sum x_i y_i \ldots\ldots..(61.2)$$

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

وإن حاصل ضرب الطرفين في الوسطين لذلك سيكون:

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

وبالتعويض في المعادلة (61.2) بما يساويها:

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

وبعد الاختصار نحصل:

$$\sum e_i^2 = \sum y_i^2 + \hat{B}_1 \sum x_i y_i \ldots\ldots\ldots\ldots\ldots (62.2)$$

وقد جرى تعريف σ^2 على أنها: $\sigma^2 = \frac{\sum e_i^2}{n-2}$

وبالتعويض عن البسط من المعادلة 62.2 ، نحصل:

$$\sigma^2 = \frac{\sum y_i^2 - \hat{B}_1 \sum x_i y_i}{n-2} \ldots\ldots..(63.2)$$

وهذا يعني أن تباين أخطاء العينة تمثل النسبة بين مجموع مربعات انحرافات القيم المشاهدة عن خط انحدار العينة إلى درجة الحرية للخطأ.

وإذا رمزنا إلى التقدير الخطي غير المتحيز لتباين الخطأ بالرمز (S_e^2)، فأن: $S_e^2 = \frac{\sum e_i^2}{n-2}$

ذلك يعني أن S_e^2 أفضل مقدر غير متحيز لتباين المتغير العشوائي أوحد الخطأ، U_i.