المجلة العراقية للعلوم الاحصائية (19) 2011
ص ص [101 – 122]

استخدام طريقة البرمجة الخطية في تقدير معالم نموذج الانحدار البسيط بالاعتماد على الانحرافات المطلقة

احمد محمود السبعاوي *   احمد ادريس المشهداني**

الملخص

يتناول هذا البحث عرضاً موجزاً لموضوع البرمجة الخطية وعرض طريقة المربعات الصغرى في تقدير معلمات الانحدار ومن ثم استخدام طريقة أخرى لتقدير المعلمات لا تعتمد على تقليل التباين (أو تقليل مجموع مربعات الأخطاء) وإنما تعتمد على تقليل انحرافات القيم المطلقة عن المتوسط، اذ يهدف البحث إلى إيجاد طريقة سهلة مبسطة ورصينة لتقدير الانحراف المطلق لحاجتنا إليه عندما تظهر عيوب للتباين كمقياس لتشتت البيانات. كما تعتمد طريقة المربعات الصغرى على تقليل مجموع المربعات فقد جعلنا تقليل انحرافات القيم المطلقة عن المتوسط هدفاً في نموذج البرمجة الخطية وفق قيود يتم من خلالها تقدير معالم الانحدار التي تحقق دالة الهدف. ويخرج البحث باستنتاجات منها أنه من الممكن استخدام نموذج البرمجة الخطية لتقدير الانحراف المطلق بجعل المقياس يمثل دالة الهدف في النموذج وأن هذه الطريقة سهلة وواضحة. وقد تم تطبيق ذلك على أمثلة بسيطة تم عرضها في الجانب العملي.

---

* استاذ مساعد/ قسم بحوث العمليات والتقنيات الذكائية .
** مدرس/ قسم بحوث العمليات والتقنيات الذكائية .

تاريخ التسلم :4 / 2 /2009 ـــــــ تاريخ القبول : 6/28 /2009

Using linear programming method to estimate parameters of linear regression model by absolute deviation

Abstract

This work deals with descriptive summary of linear programming subject, and describes least square method for estimating regression parameters. As well as using other methods in order to estimate the parameters unconventional on reduced variance (or reduce sum of square error) , but depend on minimizing variation of absolute values from the median , the aim of this work is to find an easy and precise way to estimate the absolute deviation , which is important when the variance fail in precise estimate and lead to enlarge the variation of the data , while least square method depend on minimizing sum of the square , therefore , the goal is to reduce the deviation of the absolute values from the mean in model of linear programming in order to estimate absolute deviation which is a simple and precise method , and this manuscript includes examples of this application.

المقدمة

لا يخفى لكل مطلع على علم الإحصاء وجود صنفين من المقـاييس للبيانـات وهي مقاييس التوسط ومقاييس التشتت، ولفترة طويلة مـن الـزمن بقـي الوسـط الحسابي في مقدمة المقدرات الإحصائية للتوسط اذ أنه كان نتيجةً لمعظـم طرائـق التقدير لمقدرات معلمات المجتمع، وكان يلي الوسط الحسابي من مقاييس التـشتت،

التباين أو الانحراف المعياري اذ كان المقياس الأفضل من بين مقـاييس التـشتت. ولكن وكما نعلم أنه لم يتم التوصل الى حد الآن إلى مقدر يجمـع كـل الخـواص (خواص المقدرات) وخال من العيوب، فمثلاً الوسط الحسابي يتأثر بـالقيم الـشاذة والتباين هو مجموع المربعات فهو يضخم من تشتت البيانات (أو يقللها فـي حـال كون قياسات الانحرافات اقل من الواحد)، وحتى عندما عالج الاحـصائيون هـذه المشكلة بالانحراف المعياري فهو يمثل الجذر ألتربيعي لمجموع مربعات القـيم أي أن الجذر يؤخذ لمجموع المربعات لتلافي مشكلة الصفر (مجموع انحرافات القـيم عن وسطها الحسابي).

وكان هناك مقياسا توسط و تشتت منافسين نوعاً ما للمقدرين السابقين ودائمـاً بعدهما في سلم المقدرات وهما الوسيط والانحراف المطلق على التوالي، اذ تم بهما تجاوز عيوب الوسط الحسابي وأهمها تأثره بالقيم الشاذة وعيوب التبـاين وأهمهـا التضخيم في التباين بسبب استخدام التربيع للمجموع تلافياً للصفر اذ أن مجمـوع انحرافات القيم عن وسطها الحسابي لا تكون صفراً بأخذ القيم المطلقة.[11]

يهدف البحث إلى استخدام طريقة سهلة مبسطة ورصـينة لتقـدير الانحـراف المطلق لحاجتنا إليها عندما تظهر عيوب للتباين بوصفه مقياسا لتشتت البيانات.

1-1 استعراض المراجع:

يستخدم أسلوب البرمجة الخطية عموما لتعظيم أو لتقليل دالة خطية لعدد مـن المتغيرات تدعى دالة الهدف مقيدة بعدد من المعادلات أو المتباينات الخطية تـدعى القيود. وتعتبر العلاقة خطية عندما يؤدي أي تغير في قيمة المتغير التوضيحي إلى عدد ثابت من التغير في قيمة متغير الاستجابة. كما أن الخطية تعني أن أس المتغير هو واحد أي أن العلاقة تكون غير خطية عندما يكون أحد المتغيرات ذا أس غيـر الواحد الصحيح. عليه يمكن القول بأن مسألة البرمجة الخطية تنشأ عنـدما يكـون هناك عدد من الأنشطة يمكن إنجازها بطرائق متعددة وموارد أو إمكانيات محـددة ويكون المطلوب إنجاز هذه الأنشطة بطريقة فعـالة.[2]

درس العديد من الباحثين ترشيح انحرافات القيم المطلقة عن المتوسط بوصـفه مقياسا رئيسيا لحساب التشتت لأنه يعبر عن الانحرافات الحقيقية للقيم بعيـدا عـن التحريفات التي تجرى بعد اخذ التربيع لها كمـا فـي التبـاين ومـنهم Adler & Berenguer (1981). في بحـثهم الموسـوم " Random linear programs, Operations Research Center Report, U.C. "Technical Report Berkeley والتي ناقشو فيها مشكلة اشتقاق دالة القيمة المطلقة لانحرافات القيم عن المتوسـط، كمـا تحـدث كـل مـن Dodge, Y., ed. (1987), فـي كتابـه الموسـوم" Statistical Data Analysis Based on The L_1-Norm and Related Methods" North-Holland, Amsterdam و Gonin, R. & Money, A. (1989) فـي كتـابهم الموسـوم "Nonlinear Lp-Norm Estimation" Marcel Dekker, New York-Basel عـن التقـدير للدالـة الخطيـة (L_1) (الانحراف المطلق عن المتوسط) والدالة التربيعية (L_2) المتمثلة بطريقة المربعات الصغرى.[8]

ويعتبر أسلوب البرمجة الخطية أداة لتخطيط الأنشطة الاقتـصادية مـن أجـل البحث عن الحل الأمثل الذي قد يستلزم تخفيض الكلف أو زيادة الأرباح أو العوائد أو تعظيم كمية الإنتاج. فعندما نفترض علاقة مدخلات ومخرجات خطية مع دالـة هدف خطية فهذا يعني مسألة برمجة خطية.

2-1 تعريف البرمجة الخطية :

يمكن تعريف البرمجة الخطية بأنها عبارة عـن أسـلوب رياضـي يـستخدم للمساعدة في التخطيط واتخاذ القرارات المتعلقة بالتوزيع الأمثل للمـوارد المتاحـة وذلك بهدف زيادة الأرباح وتخفيض التكاليف.

وأن كلمة البرمجة تعني وضع المشكلة بصيغة رياضية أو نمـوذج رياضـي وحلها. وبناء على ذلك فأن البرمجة الخطية تتضمن تخطيط الأنشطة للحصول على

نتائج أمثل. وبمعنى أوسع فأن هذا المصطلح يعني أيضا التنفيذ المـنظم والأفـضل للأعمال.[4]

3-1 شروط البرمجة الخطية:

أوضحت التعاريف السابقة الشروط الواجب توافرها في مشكلة البرمجة الخطية وهي:

دالة الهدف - والموارد النادرة - والقيود - وبدائل متعددة للانجاز.

1-3-1 دالة الهدف :

وتبين العلاقة بين المتغيرات والهدف الذي تسعى الإدارة الى تحقيقه كأن يكون الهدف تعظيم الأرباح أو الدخل أو كمية الإنتاج أو تقليل التكاليف أو أي هدف أخر تحدده الإدارة. كما يجب أن تكون الدالة المعبرة عن هذا الهدف خطية كما وتعتبـر دالة الهدف مقياسا لكفاءة نموذج البرمجة الخطية.

2-3-1 الموارد النادرة :

وهي المدخلات التي لا يمكن من دونها إتمام العملية أو النشاط، وهـي نـادرة لأنها متوافرة بكميات محدودة لا يمكن زيادتها في المدى القصير مثل طاقة المعدات وأرصدة الموارد ومواصفات المنتوج ومتطلبات الكميـة ومحـددات التكنولوجيـا وغيرها.

3-3-1 القيود:

القيد عبارة عن علاقة رياضية بين المتغيرات والموارد النادرة ويكـون علـى شكل معادلة خطية. فإذا كان القيد معادلة فهي ذات جانبين تفصل بينهمـا إشـارة المساواة. أما إذا كان القيد متباينة فهي ذات جانبين تفصل بينهما إشارة أقل مـن أو يساوي أو إشارة أكبر من أو يساوي. ومن أمثلة القيود كالأتي :

$$X_1+2X_2 = 100$$

$$2X_1+3X_2 \leq 50$$

$$3X_1+ X_2 \geq 70$$

4-3-1 بدائل مختلفة للانجاز :

وهذا الشرط يعني وجود عدة بدائل لانجاز العملية أو النشاط وأن واحدا أو أكثر من هذه البدائل هو البديل الأفضل الذي يحقق الهدف المطلوب. وأن استخدام أسلوب البرمجة الخطية هو لإيجاد ذلك البديل أو البدائل التي تحقق الهدف المطلوب.[10]

1-4 افتراضات البرمجة الخطية:

ويقصد بالافتراضات هنا الشروط العلمية الأساسية الواجب توافرها في المشكلة حتى نستطيع حلها بواسطة البرمجة الخطية. ويمكن القول بأن هناك خمسة افتراضات يمكن تلخيصها كما ياتي :

1-4-1 التأكد:

ويعني ذلك أن القيم الموجودة في دالة الهدف (مساهمات العوامل) والمحددات أو القيود (احتياجات العوامل والمصادر المتوافرة) معروفة وثابتة وغير قابلة للتغير أثناء فترة معالجة المشكلة موضوع البحث.

1-4-2 التناسبية:

ويعني ذلك أن كل نشاط قد يعتبر مستقلا عن الأخر، ذلك أن معيار الإنجاز هو حاصل جمع المساهمات المختلفة، كذلك أن الكميات التي يتم استخدامها من المواد المختلفة تتناسب مع احتياجات العوامل المختلفة من كل من هذه الموارد. فمثلا إذا كنا نحتاج إلى وحدتين من المواد الأولية لإنتاج وحدة واحدة تامة الصنع من منتج معين فأننا نحتاج إلى أربعين وحدة من العوامل الأولية لإنتاج عشرين وحدة من هذا المنتج وهذا الافتراض هو أساس افتراض الإضافية.

1-4-3 الإضافية:

ويعني هذا الافتراض أنه لا يوجد تداخل بين الفعاليات أو الأنشطة المختلفة.

1-4-4 قابلية القسمة أو الكسرية :

والمقصود هنا أن الحل لمشكلة البرمجة الخطية ليس بالضرورة أن يكون أعدادا صحيحة، وهذا يعني قبول كسور كقيم لعوامل القرار. وإذا كان من الصعب إنتاج أجزاء من المنتج فعند ذلك نلجأ إلى استخدام برمجة الاعداد الصحيحة.

1-4-5 اللاسلبية :

وهذا يعني أن قيم عوامل أو متغيرات القرار يجب أن تكون موجبة(غير سالبة) فالقيم السالبة للكميات المادية مستحيلة. فعلى سبيل المثال لا نستطيع إنتـاج عـدد سالب من الكراسي أو القمصان أو المصابيح أو إطارات السيارات وغيرها.[3]

5-1 الصياغة الرياضية لنموذج البرمجة الخطية :

تتركز الصياغة الرياضية لأسلوب البرمجة الخطية حـول نـاحيتين مهمتـين هما:-

الأولى:- صياغة دالة الهدف.

الثانية:- صياغة القيود المفروضة على المشكلة.

والصياغة الرياضية لدالة الهدف تأخذ الشكل الأتي :-

$$\text{Max(or Min) } Z=C_1X_1+C_2X_2+\ldots\ldots\ldots\ldots\ldots+C_nX_n$$

طبقاً للقيود الآتية :-

$$a_{11}x_1 + a_{12}x_2 +\ldots\ldots\ldots\ldots\ldots\ldots+ a_{1n}x_n\ (\leq,=,\geq)\ b_1$$

$$a_{21}x_1 + a_{22}x_2 +\ldots\ldots\ldots\ldots\ldots\ldots+ a_{2n}x_n\ (\leq,=,\geq)\ b_2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_{m1}x_1 + a_{m2}x_2 +\ldots\ldots\ldots\ldots\ldots+ a_{mn}x_n\ (\leq,=,\geq)\ b_m$$

فضلا عن قيود عدم السالبية التي تكون بالشكل الأتي :-

$$X_j \geq 0$$

لجميع قيم (j)

حيث أن:

c_j : معامل المتغير الذي تسلسله (j) في دالة الهدف.

X_j : متغيرات النموذج (المشكلة).

A_{ij} : معامل المتغير الذي تسلسله (j) في القيد الذي تسلسله (i).

$i=1,2,\ldots\ldots\ldots,m$ ، $j=1,2,\ldots\ldots\ldots,n$

n : تمثل عدد المتغيرات، m : تمثل عدد القيود.[2],[1]

6-1 استخدام البرمجة الخطية في إيجاد معلمات الانحدار

في هذا الجزء سوف نقوم بدراسة تطبيق أسلوب البرمجة الخطية على فرع من فروع علم الإحصاء وهو تحليل الانحدار. كمثال على ذلك سنأخذ حجما معينا وعدد تكرار معين لبيانات جمعت لمشاكل البرمجة الخطية لغرض اشتقاق معلمات معادلة خط الانحدار لعدد التكرارات المطلوبة لحل معظم مشاكل الحجم.

7–1 مقدرا الوسط والوسيط:

الوسط الحسابي لا يمثل مقياس التوسط تمثيلا جيدا لأنه يتأثر بـالقيم الـشاذة والوسط الحسابي ينجر للقيم الدنيا أو العليا وهذه إحدى عيوبه. لـذا فـان الوسـيط مقياسا بديلا اذ يمثل القيمة الوسطى من البيانات.

لنفرض انه لدينا المشاهدات $b_1,b_2,\ldots,b_n$ حيث أن n فردية لكي يكون تحديـد الصيغ بسيطا جدا. هناك ارتباط وثيق بين هذه المفاهيم الإحصائية وتحقيق الأمثلية، فعلى سبيل المثال الوسط الحسابي ($\bar{x}$) (والذي يمثل مجموع القيم b_i على عـددها) يطرح من كل الأعداد الحقيقية أي هو القيمـة التـي تجعـل مجمـوع مربعـات الانحرافات بين نقاط البيانات والأعداد الحقيقية أقل ما يمكن أي أن [7]:

$$\bar{x} = \arg\min_{x\in R} \sum_{i=1}^{k} (x-b_i)^2 \qquad \ldots(1)$$

ولتحقيق هذا الادعاء نفرض أن لدينا الدالة:-

$$f(x) = \sum_{i=1}^{k} (x-b_i)^2 \qquad \ldots(2)$$

نشتق هذه الدالة بالنسبة الى (x) ونساوي المشتقة بالصفر نحصل على :-

$$f'(x) = \sum_{i=1}^{k} 2(x-b_i) = 0 \qquad \ldots(3)$$

وبحل هذه المعادلة بالنسبة الى النقطة الحرجة (x) نحصل على:-

$$x = \frac{1}{k}\sum_{i=1}^{k} b_i = \hat{x} \qquad \ldots(4)$$

وتكون هذه النقطة الحرجة نقطة نهاية صغرى إذا كانت :-

$f''(x) > 0$

وتكون نقطة نهاية عظمى إذا كانت :-

$f''(x) < 0$

وبما أن $f''(x) > 0$ لكل ($x \in r$) عليه تكون هذه النقطة الحرجة هي نقطـة نهاية صغرى (النقطة الصامدة) أي أن الوسط الحسابي هو المقدر لمتوسط المجتمع (μ).

كما ان الوسيط يتمتع بارتباط وثيق بتحقيق الأمثلية لأنه يطرح من كل الأعداد أيضا أي هو القيمة التي تجعل مجموع اختلافات القيم المطلقة بين نقـاط البيانـات والأعداد الحقيقية (X) أقل ما يمكن أي أن:-

$$M = min\ x \in r \sum_{i=1}^{k} |x - b_i| \qquad ...(5)$$

هي دالة خطية مستمرة ومحدبة. كذلك فأن الاشتقاق بالنسبة الـى (X) يكـون بالشكل الأتي:-

$$f'(x) = \sum_{i=1}^{k} \text{sgn}(x - b_i) \qquad ...(6)$$

حيث أن:-

$$\text{sgn}(x) = \begin{cases} 1 & if\ x > 0 \\ 0 & if\ x = 0 \\ -1 & if\ x < 0 \end{cases} \qquad ...(7)$$

وهذا الاشتقاق يكون أعدادا فقط على يسار (X) سالبة وعلى يمينها موجبة. أي أن نتيجة الاشتقاق تقفز من الصفر إلى الواحد أو ناقص واحد حيث يشير الـصفر إلى اشتقاق الدالة عند نقطة الوسيط.[11]

8-1 دالة الانحدار الخطي:

التحليل السابق يمكن إعادة صياغته على النحو الأتي :-

نعطي مشاهدة عشوائية ولتكن (b) نفترض بأنها تشمل جزأين هما:

جزء ثابت لكنه مجهول ونرمز له (x).

وجزء يمثل متغيرا عشوائيا حول هذا الجزء الثابت وهو (ε).

$b = x + \varepsilon$

إذا أخذنا عدة مشاهدات ونضعها كمتغير $i=1,2,......,k$. فأن المشاهدات ($b_1,b_2,...,b_k$) والثوابت (ε) سوف تختلف لكن (x) قيمة ثابتة في كل المشاهدات. أي أن:-

$$b_i = x + \in_i \quad , \quad i = 1,2,........,k$$

نلاحظ بأن الوسط الحسابي هو قيمة (x) المطروحة من مجموع الانحرافات عن (ε_i,s)، وبنفس الطريقة نلاحظ قيمة الوسيط هي قيمة (x) التي تطرح من مجموع انحرافات القيمة المطلقة عن (ε_i,s)، وكمثال على ما ذكر أعلاه نفرض أن باحثا طبيا قام بجمع بيانات عن ضغط الدم على ألاف من المرضى وهدفه الوحيد هو معرفة كيف أن ضغط الدم يعتمد على العمر، السمنة (زيادة في الوزن)، الجنس،...الخ. ارتبطت كل هذه المشاهدة (b) بضغط الدم وقيم هذه المتغيرات تسيطر على ضغط الدم، نقوم بإعطاء كل متغير رمزا معينا ويفضل أن تكون الرموز بشكل متسلسل أي أننا نرمز للعمر بـ (a_1) وللسمنة بـ (a_2) وللجنس (a_3) وهكذا لبقية المتغيرات وليكن (n) عدد المتغيرات المسيطرة والتي تعتبر مهمة لدى الباحث لأن قيمة المشاهدة (b) تعتمد بشكل خطي على عدد المتغيرات المسيطرة، وبما أننا نفترض بأن هناك مجاهيل (أعداد)n، $j=1,2,.........,$ ، x_j

أي أن:-

$$b=\sum_{j=1}^{n} a_j x_j + \in$$

$$b = a_1 x_1 + a_2 x_2 + a_3 x_3 + ... + a_n x_n + \in$$

حيث أن:-

b: متغير الاستجابة (المتغير المعتمد).

a: المتغير المستقل (التوضيحي).

x: ميل المتغير a_j.

∈: الحد الثابت (نقطة تقاطع مستوى الانحدار مع المحور *y*).

المعادلة السابقة تسمى (نموذج الانحدار). وبما أن الباحث قد جمع البيانات من ألاف المرضى عليه فأن هذه البيانات تشمل (b_i)، (a_j) ويجب أن تفهرس حسب هؤلاء المرضى أي أن:-

$$b_i=\sum_{j=1}^{n} a_{ij} x_j + \in_i \quad ، \; i=1,2,........,k \;، \; j=1,2,........,n \quad ..(8)$$

إذا فرضنا بأن (*b*) يمثل متجه المشاهدات، وأن (∈) يمثل متجه المتغير العشوائي، وأن (*A*) مصفوفة تمثل صفوفها قيم المتغيرات المسيطرة لـ (*i*) من المرضى، عليه يمكن التعبير عن نموذج الانحدار بالمصفوفات بالشكل الأتي :-

$$b = Ax + \in$$

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \in_1 \\ \in_2 \\ \vdots \\ \in_k \end{bmatrix} \qquad ....(9)$$

إن الهدف من تحليل الانحدار هو إيجاد المتجه (*x*) الذي يمثل أفضل المشاهدات بالنسبة لـ (*b*)، لذلك نحن نرغب في جعل قيم متجه الثوابت (∈) الذي

يمثل الوسط الحسابي أو الوسيط في نهايته الصغرى، ويمكن إيجاد هذا التقليل إمـا بواسطة مجموع مربعات الأخطاء ($\in_i$'s) أو مجموع القيم المطلقـة لـ ـ ($\in_i$'s) .[11]

9-1 طريقة المربعات الصغرى :

لدينا الصيغة العامة:-

$$\|y\|_p = (\sum_{i=1}^{k} y_i^{\,p})^{1/p} \quad ...(10)$$

وعندما (*p=1*) نحصل على:-

$$\bar{x} = \bar{x} = \arg\min \sum_{\substack{i=1 \\ x\in R}}^{k} \|b - Ax\| \quad ...(11)$$

لإيجاد قيمة ($\bar{x}$) وباستخدام طريقة المربعات الصغرى وعندما (*p=2*) :-

$$f(x) = \|b - Ax\|_2^2 = \sum_{i=1}^{k} (b_i - \sum_{j=1}^{n} a_{ij} x_j)^2 \quad ...(12)$$

وبما أن النقطة الحرجة تعرف على أنها النقطة التي اشتقاقها يتعلق بكل متغير مجهول. فإذا رمزنا للنقطة الحرجة بالرمز ($\bar{x}$) عليه يجـب أن تحقـق المعادلـة الآتية[9] :-

$$\frac{\partial f}{\partial x_r}(\bar{x}) = \sum_{i=1}^{k} 2(b_i - \sum_{j=1}^{n} a_{ij}\bar{x}_j)(-a_{ir}) = 0 \quad , \quad r = 1,2,.........,n$$

$$\frac{\partial f}{\partial x_r}(\bar{x}) = \sum_{i=1}^{k} b_i a_{ir} - \sum_{i=1}^{k}\sum_{j=1}^{n} a_{ij}\bar{x}_j a_{ir} = 0 \quad , \quad r = 1,2,.........,n$$

$$\sum_{i=1}^{k} b_i a_{ir} = \sum_{i=1}^{k}\sum_{j=1}^{n} a_{ij}\bar{x}_j a_{ir}$$

$$A^T b = A^T A\bar{x}$$

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}$$

$$A^T b = A^T A\bar{x} \;\Rightarrow \bar{x} = (A^T A)^{-1} A^T b \qquad \ldots(13)$$

نلاحظ في عملية الاشتقاق تم الضرب بـ ($-a_{ir}$) وذلك لوجود (n) من المعادلات أو المشاهدات حسب العلاقة:-

$$b_i = \sum_{j=1}^{n} a_{ij} x_j + \in_i \qquad ، \; i = 1,2,\ldots\ldots,k \; ، \; j = 1,2,\ldots\ldots,n$$

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \in_1 \\ \in_2 \\ \vdots \\ \in_k \end{bmatrix}$$

عليه يصبح لدينا (k) من المعادلات وعند عملية الاشتقاق نقوم بأخذ معاملات المتغير (x_i) لكل معادلة. كما تدعى هذه الصيغة (بانحدار المربعات الصغرى).

مثال:- إذا كانت لدينا البيانات الآتية الناتجة من معادلة خط الانحدار:-

$$\begin{bmatrix} 1 \\ 2.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} \in_1 \\ \in_2 \\ \in_3 \end{bmatrix}$$

سوف نقوم بحل هذا المثال بالاعتماد على الصيغة :-

$$\bar{x} = (A^T A)^{-1} A^T b$$

$$A^T A = \begin{bmatrix} 0 & 2 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 6 \\ 6 & 3 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 0 & 2 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 17 \\ 6.5 \end{bmatrix}$$

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \frac{1}{24} \begin{bmatrix} 3 & -6 \\ -6 & 20 \end{bmatrix} \begin{bmatrix} 17 \\ 6.5 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 7/6 \end{bmatrix}$$

10-1 طريقة القيمة المطلقة :

بعد أن عرفنا بأن الوسيط يعطي معلمة أكثر دقة للقيمة المعدلة لمجموعة من الأعداد من الوسط الحسابي (لعدم تأثره بالقيم الشاذة) وأن مشكلة الانحدار هي إيجاد المتجه (x) الذي يعطي أفضل المشاهدات للمتجه (b) وهنا يمكن صياغة مشكلة الانحدار بأسلوب البرمجة الخطية وفي الحقيقة من السهولة رؤية ذلك في مشكلة الانحدار الآتية[11] :-

$$\min \sum_{i=1}^{k} \left| b_i - \sum_{j=1}^{n} a_{ij} x_j \right| \qquad \ldots(14)$$

يمكن إعادة صياغتها بأسلوب البرمجة الخطية كالأتي :-

$$\min \sum_{i=1}^{k} t_i \qquad \ldots(15)$$

$$sub.to \ -t_i \le (b_i - \sum_{j=1}^{n} a_{ij} x_j) \le t_i \qquad ، \ i = 1,2,3,.........,k$$

حيث أن t_i تمثل ($b_i - \Sigma a_{ij} x_j$)

ولحل مشكلة الانحدار نكتفي فقط بحل مشكلة البرمجة الخطية، بالرغم من أن حساب التفاضل والتكامل لا يمكن أن يستعمل للحصول على صيغة واضحة لحل مشكلة الانحدار(مشكلة البرمجة الخطية) إلا أنه يمكن أن يستعمل للحصول على صيغة رياضية، وعندما يستخدم استخداما صحيحا يؤدي إلى حل مشكلة الانحدار(مشكلة البرمجة الخطية).

إن الصيغة الرياضية الناتجة منافسة للصيغة الناتجة من اقل المربعات، ولمناقشة هذه الطريقة نبدأ باعتبار الدالة الموضوعية للانحدار:-

$$f(x)=\|b-Ax\| \quad \ldots(16)$$

$$f(x)=\sum_{i=1}^{k}\left|b_i-\sum_{j=1}^{n}a_{ij}x_j\right| \quad ، \; i=1,2,\ldots..,k \quad ، \; j=1,2,\ldots..,n$$

تمييز الدالة الموضوعية أو اشتقاقها هو مشكلة إلا أنه يمكن أن نستعمل الصيغة المبسطة للاشتقاق أي أن:-

$$g(z)=|z|$$

$$\bar{g}(z)=\frac{z}{|z|} \quad \ldots(17)$$

وباستعمال هذه الصيغة في اشتقاق الدالة وبمساواة المشتقة بالصفر نحصل على المعادلات الآتية للنقاط الحرجة :-

$$\frac{\partial f}{\partial x_r}=\sum_{i=1}^{k}\frac{b_i-\sum_{j=1}^{n}a_{ij}x_j}{\left|b_i-\sum_{j=1}^{n}a_{ij}x_j\right|}(-a_{ir})=0 \quad ، \; r=1,2,\ldots\ldots..,n$$

وإذا افترضنا بأن:-

$$e_i(x)=\left|b_i-\sum_{j=1}^{n}a_{ij}x_j\right| \quad \ldots(18)$$

عليه يمكن أن نعيد كتابة معادلة الاشتقاق كالأتي:-

$$\sum_{i=1}^{k}\frac{a_{ir}b_i}{e_i(x)}=\sum_{i=1}^{k}\sum_{j=1}^{n}\frac{a_{ir}a_{ij}x_j}{e_i(x)} \quad ، \; i=1,2,\ldots..,k \; ، \; j=1,2,\ldots..,n \; ، \; r=1,2,\ldots..,n$$

وعلى فرض أن (E_x) تمثل المصفوفة القطرية بحيث تحتـوي علـى المتجـه [$e_i(x)$] على القطر، عليه يمكن كتابة المعادلات الـسابقة باسـتخدام المـصفوفات كالأتي:-

$$A^TE_x^{-1}b=A^TE_x^{-1}Ax \quad \ldots(19)$$

وهذه المعادلة يمكن حلها بالنسبة لـ (X) وذلك بضرب طرفي المعادلة بـ ـ $^{-1}$ ($A^T E_x^{-1} A$) فتكون النتيجة كالأتي:-

$$x = (A^T E_x^{-1} A)^{-1} A^T E_x^{-1} b \qquad \ldots(20)$$

تقترح هذه الصيغة مخطط تكراري والذي يصل الى حل بعد أن نبدأ بتحريـك (X_0) وبعد ذلك سنستعمل الصيغة السابقة على التوالي ونحسب تقريبات جديدة، فإذا افترضنا أن (X_r) يدل على التقريب (r) من التكرارات، ثم صيغة التحديث يمكن أن تبدو كالأتي :-

$$x^{r+1} = (A^T E_{x^k}^{-1} A)^{-1} A^T E_{x^k}^{-1} b \qquad \ldots(21)$$

وبافتراض أن المصفوفة المعكوسة يتم إيجادها في كل تكـرار، عليـه يمكـن ملاحظة أن هذا يؤدي لمخطط تكراري لحل مشكلة الانحـدار(مـشكلة البرمجـة الخطية) [11].

الجانب العملي

تم اخذ البيانات من ملحق المصدر [7](Applied Regression Analysis) والمتضمنة 24 زوجا من المشاهدات (X,Y) يمثل X المتغير التوضيحي (المستقل) و Y متغير الاستجابة.

وبعد إجراء التحليل كانت النتائج كالآتي:

الجدول (1) تحليل الانحدار بالاعتماد على تصغير مربعات الخطأ

Y	X	Ŷ	e_i	$\|e_i\|$	e_i^2
2.3	1.3	1.8754	-0.4246	0.4246	0.180285
1.8	1.3	1.8754	0.0754	0.0754	0.005685
2.8	2	2.112	-0.688	0.688	0.473344
1.5	2	2.112	0.612	0.612	0.374544
2.2	2.7	2.3486	0.1486	0.1486	0.022082
3.8	3.3	2.5514	-1.2486	1.2486	1.559002
1.8	3.3	2.5514	0.7514	0.7514	0.564602
3.7	3.7	2.6866	-1.0134	1.0134	1.02698
1.7	3.7	2.6866	0.9866	0.9866	0.97338
2.8	4	2.788	-0.012	0.012	0.000144
2.8	4	2.788	-0.012	0.012	0.000144
2.2	4	2.788	0.588	0.588	0.345744
5.4	4.7	3.0246	-2.3754	2.3754	5.642525
3.2	4.7	3.0246	-0.1754	0.1754	0.030765
1.9	4.7	3.0246	1.1246	1.1246	1.264725
1.8	5	3.126	1.326	1.326	1.758276
3.5	5.3	3.2274	-0.2726	0.2726	0.074311
2.8	5.3	3.2274	0.4274	0.4274	0.182671
2.1	5.3	3.2274	1.1274	1.1274	1.271031
3.4	5.7	3.3626	-0.0374	0.0374	0.001399
3.2	6	3.464	0.264	0.264	0.069696
3	6	3.464	0.464	0.464	0.215296
3	6.3	3.5654	0.5654	0.5654	0.319677
5.9	6.7	3.7006	-2.1994	2.1994	4.83736
Σ				**16.9196**	**21.19367**

في الجدول (1) تم الاعتماد على نظرية تقليل مجموع مربعات الخطأ في تقدير معلمات نموذج الانحدار الخطي البسيط وكانت معادلة الانحدار التقديرية هي:

$$\hat{Y}_i = 1.436 + 0.338 X_i$$

وكان مجموع انحرافات القيم المطلقة للقيم الحقيقية لـ Y_i عن القيم التقديرية $\hat{Y}_i$ هو (16.9196) وكان مجموع انحرافات مربعات القيم Y_i عن القيم $\hat{Y}_i$ هي (21.19367)

بعد استخراج قيم معالم نموذج الانحدار بالاعتماد على فرضية تصغير مربعات انحرافات القيم عن المتوسط تم وضع نموذج برمجة خطية كان الهدف منه تصغير القيمة المطلقة لانحرافات القيم عن المتوسط كانت القيود وفق المعادلة (16) وأظهرت نتائج التحليل الجاهز وفق برنامج winQSB قيمة الحد الثابت 1.4128 وقيمة معامل الانحدار الخطي هو 0.2979 والنتائج معروضة في الجدول (2).

في الجدول (2) تم الاعتماد على مبدأ تقليل انحرافات القيم المطلقة في تقدير معلمات نموذج الانحدار الخطي البسيط وكانت معادلة الانحدار التقديرية هي:

$$\hat{Y}_i = 1.4128 + 0.2979 X_i$$

وكان مجموع انحرافات القيم المطلقة للقيم الحقيقية للمتغير التوضيحي Y_i عن القيم التقديرية $\hat{Y}_i$ هو (16.46418) وكان مجموع انحرافات مربعات القيم Y_i عن القيم $\hat{Y}_i$ هي (22.16579)

الجدول (2) تحليل الانحدار بالاعتماد على تصغير انحرافات القيم المطلقة عن المتوسط

Y	X	$\hat{Y}$	e_i	$\|e_i\|$	e_i^2
2.3	1.3	1.80007	-0.49993	0.49993	0.24993
1.8	1.3	1.80007	7E-05	7E-05	4.9E-09
2.8	2	2.0086	-0.7914	0.7914	0.626314
1.5	2	2.0086	0.5086	0.5086	0.258674
2.2	2.7	2.21713	0.01713	0.01713	0.000293
3.8	3.3	2.39587	-1.40413	1.40413	1.971581
1.8	3.3	2.39587	0.59587	0.59587	0.355061
3.7	3.7	2.51503	-1.18497	1.18497	1.404154
1.7	3.7	2.51503	0.81503	0.81503	0.664274
2.8	4	2.6044	-0.1956	0.1956	0.038259
2.8	4	2.6044	-0.1956	0.1956	0.038259
2.2	4	2.6044	0.4044	0.4044	0.163539
5.4	4.7	2.81293	-2.58707	2.58707	6.692931
3.2	4.7	2.81293	-0.38707	0.38707	0.149823
1.9	4.7	2.81293	0.91293	0.91293	0.833441
1.8	5	2.9023	1.1023	1.1023	1.215065
3.5	5.3	2.99167	-0.50833	0.50833	0.258399
2.8	5.3	2.99167	0.19167	0.19167	0.036737
2.1	5.3	2.99167	0.89167	0.89167	0.795075
3.4	5.7	3.11083	-0.28917	0.28917	0.083619
3.2	6	3.2002	0.0002	0.0002	4E-08
3	6	3.2002	0.2002	0.2002	0.04008
3	6.3	3.28957	0.28957	0.28957	0.083851
5.9	6.7	3.40873	-2.49127	2.49127	6.206426
Σ				**16.46418**	**22.16579**

الاستنتاجات

نلاحظ من النتائج المذكورة انفا أن انحرافات القيم المطلقة بالاعتماد على تصغير القيم المطلقة للخطأ (باستخدام البرمجة الخطية) كانت اقل من انحرافات القيم المطلقة بالاعتماد على طريقة تقليل مربعات الخطأ (باستخدام طريقة المربعات الصغرى) وهذا يدعم ما اشرنا إليه سابقا من أن الاعتماد على تقليل القيم المطلقة يعطي نتائج أدق لمقدرات معالم نموذج الانحدار الخطي البسيط.

أما بالنسبة لموضع التباين وكونه اكبر عندما تم الاعتماد على طريقة القيمة المطلقة فان هذا الفرق لا يقلل من أهمية الطريقة، والسبب في ذلك هو أن التباين بالحقيقة مقياس غير دقيق لقياس التشتت مقارنة بانحرافات القيم المطلقة، حيث انه لو كان لدينا على سبيل المثال مجموعة مشاهدات تم حساب المتوسط لها ومن ثم قررنا قياس مقدار تشتت المشاهدات عن المتوسط بطرح المتوسط من كل مشاهدة كانت إحد هذه الفروق هو القيمة 2، أي أن:

$$x_i - \bar{x} = 2$$

ففي أثناء حساب التباين سيتم تربيع هذه القيمة أي تصبح 4 مما يجعلها لا تمثل التشتت الحقيقي للبيانات وإنما اكبر من التشتت الحقيقي. وعلى العكس لو كانت قيمة الفرق بين المشاهدة والمتوسط اقل من الواحد، مثلا:

$$x_i - \bar{x} = 0.5$$

فإنها تصبح 0.25 بعد التربيع، صحيح أن التباين هنا قد قل عن القيمة الحقيقة ولكن المشكلة أن هذا النقصان أو الزيادة في الحالة الأولى قد جعل التباين مقياس غير دقيق في حساب مدى تشتت البيانات. على العكس من الانحراف المطلق الذي لا يغير من قيمة الفرق إلا الإشارة فقط والتي لا تؤثر في القيمة الرقمية الحقيقية للتشتت.

ومع ذلك بقى التباين متربعا على عرش أفضل مقياس للتشتت وتم تصليحه بأخذ الجذر ألتربيعي للمجموع وحساب ما يسمى الانحراف المعياري، والسبب في هذا البقاء هو تجاوب مجموع مربعات الخطأ مع العمليات الرياضية المستخدمة في تقدير المعالم. ولذلك كان أسلوب البرمجة الخطية قد حل مشكلة صعوبة التعامل مع

القيمة المطلقة للخطأ في الوقت نفسه حافظ على القيم الحقيقية للتشتت دون الحاجة إلى تربيع مجموع انحرافات القيم. وبذلك يمكن القول بأن:

1. أسلوب البرمجة الخطية مناسب جداً لتقدير المعلمات باستخدام انحرافات القيم المطلقة.

2. وجود طريقة سهلة للتقدير تمكننا من الاستعاضة عن الوسط الحسابي أو التباين كمقدرات في حالة وجود عيوب ظاهرة لهما (مثل التأثر بالقيم الشاذة أو التباين الضخم بسبب تربيع الانحرافات).

المصادر

المصادر العربية :

1- الشمرتي، حامد سعد نور و الزبيدي، علي خليل، "مدخل إلى بحوث العمليات "، دار مجدلاوي للنشر والتوزيع، عمان-الأردن، (2007).

2- جزاع، عبد ذياب، " مدخل إلى بحوث العمليات "، الطبعة الثانية، وزارة التعليم العالي والبحث العلمي، جامعة بغداد، (1988).

3- محمد، صادق ماجد، " بحوث العمليات "، الطبعة الأولى، مطبعة دار الحكمة، جامعة البصرة، (1991).

المصادر الأجنبية :

4- A. Ravi Ravindran (2009) "Operations Research Applications". CRC Press Taylor & Francis Group.

5- A. Billionnet and F. Calmels (1996). "Linear programming for the 0-1 quadratic knapsack Problem". European Journal of Operational Research , 92:310-325.

6- A. Ross, K. M. Bretthauer, and B. Shetty(1999). "Nonlinear integer programming for optimal allocation in stratified sampling". European Journal of Operational Research, 116:667-680.

7- David G. Luenberger (2008) Linear and Nonlinear Programming" Addison-Wesley publishing company (3ed).

8- G. Bernd, M. Jiri (2007). "Understanding and Using Linear Programming". springer.

9- Norman, Draper (1966). " Applied Regression Analysis" John Wiley & Sons (2nd).

10- Taha, Hamdy A. (2007) "Operations research: an introduction" Pearson Prentice Hall (8ed)

11- William L. Steiger (1983) "least Absolute Deviations Theory, Application, and Algorithms. Birkhijuser Boston, Inc.