مجلة العلوم الاقتصادية والادارية   المجلد 13/ ع46/ لسنة 2007

تقدير النموذج اللوجستي باستخدام أوزان بيز المتسلسل

Estimating of Logistic Model by using Sequential Bayes Weights

م. طه حسين علي
كلية الإدارة والاقتصاد/ قسم الإحصاء
جامعة صلاح الدين / أربيل

م. تارا احمد حسن
كلية الإدارة والاقتصاد/ قسم الإحصاء
جامعة صلاح الدين / أربيل

الملخص

تناول البحث عملية تقدير معلمات النموذج اللوجستي (Logistic Model) عندما يكون المتغير المعتمد متغيرا نوعيا (متغير وهمي) باستخدام طريقة المربعات الصغرى الموزونة (WLS) بالاعتماد على الأوزان المقدرة بأسلوب بيز المتسلسل (Bayesian Sequential Analysis) بدلا من الأسلوب التقليدي المستخدم من قبل العديد من الباحثين، والذي سوف يوفر آلية استخدام المعلومات الأولية والتي نحصل عليها من الخبرة أو التجارب السابقة فضلا عن معلومات المشاهدات وبشكل متسلسل في تقدير الأوزان المستخدمة في طريقة المربعات الصغرى الموزونة لمعالجة مشكلة عدم تجانس تباين الخطأ العشوائي والحصول على أفضل تقدير ممكن لمعلمات النموذج اللوجستي، ومن ثم أجراء مقارنة بين أسلوب بيز المتسلسل والأسلوب التقليدي في تقدير هذه الأوزان من خلال تطبيق عملي يتناول دراسة العلاقة بين عدد المرضى المشافين من مرض معين والتراكيز المختلفة من دواء معين أعطي إليهم في مستشفى رزكاري/ أربيل .

Abstract

This research deals with the process of estimating Logistic Model Parameters when the depended variable is variable attribute (Dummy variable) using Weighted Least Square (WLS) depending on the estimated weights estimated by Bayes Sequential Analysis Approach instead of the Classical Approach used by many researchers, which will provide the mechanism of using primary information, which can be got from the experience or past experiences, In addition to observations information in a sequential way in estimating weights used in (WLS) Approach to remedy the problem of unlikeness of disparity random line and getting the best estimation of the Logistic Model Parameters, then making a comparison between Bayesian Sequential Approach and the classical approach in estimating these weights through practical application depending on the study of relation between number of patients recovered of certain disease and the various concentrations of certain medicines they are given in Razgar Hospital / Erbil .

1- المقدمة:

في كثير من التطبيقات العملية نحتاج إلى دراسة العلاقة بين المتغير المعتمد عندما يكون متغيرا نوعيا مع متغير مستقل كمي، وتكون دالة الانحدار أو الاستجابة (لوجستية) أي غير خطية، ويكون المتغير المعتمد النوعي على شكل نسبة تمثل معلمة توزيع ذي الحدين (د. محمد ود. أموري، 1990، ص149) وباستخدام التحويل اللوغارتمي يمكن تحويلها إلى دالة استجابة خطية ، في حين تعاني هذه الدالة المحولة من مشكلة عدم تجانس تباين الخطأ العشوائي والتي لا يمكن تقدير معلماتها باستخدام طريقة المربعات الصغرى الاعتيادية لذلك يمكن استخدام طريقة المربعات الصغرى الموزونة لمعالجة هذه المشكلة في تقدير معلمات النموذج اللوجستي.

من جانب آخر يعتبر أسلوب بيز المتسلسل الذي قدمه العالم والد عام (1947) من المواضيع المهمة التي توفر آلية تقدير المعلمات بالاعتماد على المعلومات الأولية والتي نحصل عليها من الخبرة أو التجارب السابقة فضلا عن المعلومات المتأتية من العينة المسحوبة من المجتمع وبشكل متسلسل، أي أن كل مشاهدة جديدة سوف تضيف معلومة جديدة في تقدير معلمة التوزيع النهائي، واستخدام تباين التوزيع النهائي لهذه المعلمة في تقدير الأوزان المستخدمة في طريقة المربعات الصغرى الموزونة لتقدير معلمات النموذج اللوجستي.

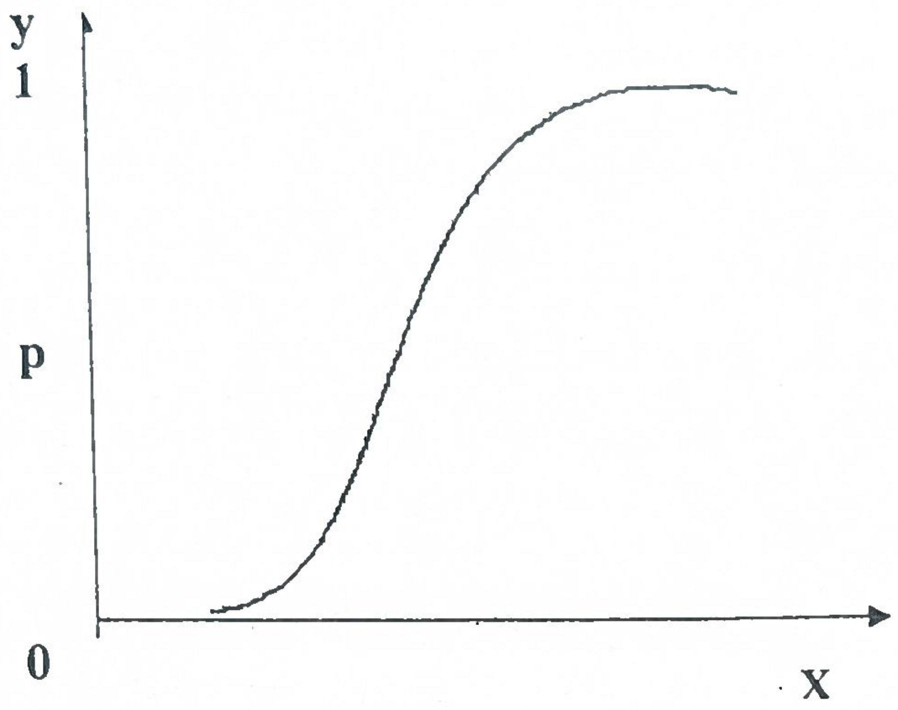
لذلك تناول البحث أعطاء فكرة عن الأسلوب التقليدي في تقدير الأوزان ومن ثم أسلوب بيز المتسلسل في تقدير هذه الأوزان وكيفية نمذجتها لاستخدامها في طريقة المربعات الصغرى الموزونة لتقدير معلمات النموذج اللوجستي، ومن ثم أجراء تطبيق عملي لدراسة العلاقة بين عدد المرضى المشافين من مرض معين والتراكيز المختلفة من دواء معين أعطي إليهم في مستشفى رزكاري/ أربيل، يهدف هذا التطبيق إلى إجراء مقارنة بين الأسلوبين أعلاه في التقدير من خلال إيجاد متوسط مجموع المربعات (MSE) لقياس كفاءة النموذجين المقدرين (وليس إلى معالجة مشكلة معينة في المجتمع) وذلك باستخدام الحاسبة الالكترونية وبالتحديد البرنامج الإحصائي الجاهز (SPSS).

2 : تقدير النموذج اللوجستي باستخدام الأوزان التقليدية :

في بعض حالات استخدام تحليل الانحدار نحتاج إلى متغيراً وهمياً للمتغير المعتمد ونلاحظ أيضا انه ربما تكون هذه العلاقة بين هذا المتغير والمتغير المستقل عبارة عن علاقة غير خطية (Nonlinear Model) وغالبا ما تأخذ دالة الانحدار شكل الحرف (S) ، أي أن هذه الدالة تسمى الدالة اللوجستية (Logistic Function) ونموذجها العام (د.خاشع، 1987، ص442) هو كما يأتي:

$$E\left(y/x \right) = \frac{\mathrm{Exp}(\ \beta_0 + \beta_1\, x)}{1 + \mathrm{Exp}(\ \beta_0 + \beta_1\, x)} \qquad \dots (2.1)$$

تكون الدالة اللوجستية محددة بين الصفر والواحد وتكون دالة رتيبة (Monotonic function)، وأن الاعتبارات الطبيعية المتعددة على مفاهيم قيم البداية تعطي تبرير لاستخدام الصيغة (2.1) للتعبير عن العلاقة غير الخطية بين المتغير المعتمد الوهمي والمتغير المستقل (د. أموري، 2002، ص155) كما نلاحظ من خلال الشكل الآتي (د.محمد وأموري، 1990، ص148):



الشكل رقم (2.1)

الشكل يوضح طبيعة الدالة اللوجستية، دالة (S)

إن ميزة هذه الدالة هي سهولة جعلها خطية، (خاشع، 1987، ص443) وان الوسط الحسابي لهذه الدالة (حيث أن المتغير المعتمد y متغيرا نوعيا يمكن تمثيله بمتغير وهمي يأخذ القيم صفر وواحد وله توزيع برنولي) هو:

$E(y/x) = p$

وبما أن p هو قيمة احتمالية لذا فيمكن جعل الدالة اللوجستية خطية باستخدام التحويل الآتي :

$$\hat{p}^* = Ln\left[\frac{E(y/x)}{(1-E(y/x))}\right]$$

وبالتعويض بقيمة الوسط الحسابي نحصل على الصيغة الآتية :

$$\hat{p}^* = Ln\left[\frac{p}{1-p}\right] \quad \text{... (2.2)}$$

أن هذا التحويل يسمى التحويل اللوجستي للاحتمال p لذا فأن الصيغة (2.1) سوف تختزل إلى الشكل الآتي:

$$\hat{p}_i^* = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad \text{... (2.3)}$$

ولتوفيق الدالة اللوجستية نفرض بان هناك تكرارات لكل مستوى من مستويات المتغير المستقل x ، أي أن لها المستويات ($x_1, x_2, x_3, \ldots, x_m$) وأن هناك n_i من قيم المشاهدات y قد تكرر لكل مستوى من مستويات x (حيث أن y تأخذ القيم صفر أو الواحد ، n_i من المرات لكل مستوى من مستويات x) ، فإذا فرضنا أن y_i هي عدد ظهور الواحد لكل مستوى من مستويات x ، فأن نسبة ظهور الواحد في كل مستوى من مستويات x (خاشع ، 1987 ، ص443) هي :

$$\theta \quad 1 \qquad ; \quad i = 1, 2, \ldots, m \qquad \ldots (2.4)$$

وعلى هذا الأساس يمكن أن نوفق الدالة اللوجستية بطريقة المربعات الصغرى باستخدام التحويل على النسب المشاهدة $\hat{p}_i$ (Cox , 1970 , p.155) أي :

$$p_i^* = Ln\left(\frac{p_i}{1-p_i}\right) \qquad \ldots (2.5)$$

كمتغير معتمد (نوعي) وهمي.

وبما أن التباين للخطأ غير متجانس لذا فان طريقة المربعات الصغرى الموزونة (التقليدية) يمكن أن تستخدم لتقدير $\hat{\beta}_0$ و $\hat{\beta}_1$ ، وأن الوزن في هذه الحالة هو معكوس التباين (Draper & Smith , (1981) , pp.108-116) ، أي أن :

$$\hat{w}_i = \frac{1}{V(p_i^*)} \qquad \ldots (2.6)$$

إن النسبة في العينة $\hat{p}_i$ تمثل معلمة توزيع ذي الحدين (Binomial distribution) بمتوسط مقداره p_i وتباين $p_i(1-p_i)/n_i$ حيث أن p_i هو احتمال وجود الصفة المدروسة التي تأثرت بالمتغير المستقل ذات مستوى معين أي x_i ، وأن العلاقة بين p_i (التي تمثل المتغير المعتمد النوعي أو الوهمي) و x_i الذي يمثل المتغير المستقل في النموذج اللوجستي (Cox , 1970 , p.148) تعتمد على النظرية الآتية :

$$p_i = f(x_i) \qquad \ldots (2.7)$$

وأن المتغير المحول p_i^* له متوسط يساوي إلى $Ln(p_i/(1-p_i))$ والتباين هو مقلوب تباين توزيع ثنائي الحدين أي أن :-

$$V(p_i^*) = \frac{1}{n_i \, p_i(1-p_i)} \qquad \ldots (2.8)$$

لذا فان الأوزان التقليدية التي يمكن استخدامها في طريقة المربعات الصغرى الموزونة يمكن تقديرها بالشكل الآتي:-

$$\hat{w}_i = n_i\, p_i\left(1 - p_i\right) \quad ; \quad i = 1, 2, 3, \ldots, m \qquad \ldots (2.9)$$

وعلى هذا الأساس يمكن تقدير معلمات الانحدار $\hat{\beta}_0$ و $\hat{\beta}_1$ باستخدام طريقة المربعات الصغرى الموزونة التقليدية (د.محمد ود.أموري، 1990، ص150) وكما يأتي:-

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{m} w_i p_i^* x_i - (\sum_{i=1}^{m} w_i p_i^*)(\sum_{i=1}^{m} w_i x_i) / \sum_{i=1}^{m} w_i}{\sum_{i=1}^{m} w_i x_i^2 - (\sum_{i=1}^{m} w_i x_i)^2 / \sum_{i=1}^{m} w_i} \qquad \ldots (2.10)$$

كذلك لدينا:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad \ldots (2.11)$$

حيث أن:-

$$\bar{y} = p^* = \frac{\sum_{i=1}^{m} w_i p_i^*}{\sum_{i=1}^{m} w_i} \qquad \ldots (2.12)$$

كذلك لدينا:-

$$\overline{x} = \frac{\sum_{i=1}^{m} w_i x_i}{\sum_{i=1}^{m} w_i} \quad \ldots (2.13)$$

وعلى هذا الأساس نحصل على معادلة الانحدار الخطية المقدرة معلماتها بواسطة طريقة المربعات الصغرى الموزونة التقليدية للتعبير عن الدالة اللوجستية، أي أن:-

$$\hat{p}_i^* = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad \ldots (2.14)$$

ومن خلال هذه المعادلة يمكن إيجاد القيم التقديرية $\hat{p}_i^*$ بالتعويض عن مستويات قيم x_i ، ومن ثم تحول قيم $\hat{p}_i^*$ إلى قيمتها الأصلية من خلال الصيغة الآتية :-

$$\hat{p}_i = \frac{Exp(\hat{p}_i^*)}{1 + Exp(\hat{p}_i^*)} \quad \ldots (2.15)$$

نلاحظ أيضا بأنه من الممكن شمول أكثر من متغير واحد للدالة اللوجستية ، فمثلا لو كانت تحوي على متغيرين x_1 و x_2 (خاشع ، 1987 ، ص448) فأن الدالة اللوجستية المحولة ستكون :-

$$p^* = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 \quad \ldots (2.16)$$

3 : تقدير النموذج اللوجستي باستخدام أوزان بيز المتسلسل :

نلاحظ من خلال الصيغة (2.9) بأن أوزان طريقة المربعات الصغرى التقليدية هي عبارة عن تباين توزيع ذي الحدين ، وعلى هذا الأساس يمكن تقدير هذه الأوزان باستخدام أسلوب بيز المتسلسل من خلال إيجاد تباين التوزيع النهائي (Variance of Posterior distribution) لهذا التوزيع .

نحن لدينا مشاهدات هي عبارة عن المتغير المعتمد الوهمي y_i الذي له توزيع ثنائي الحدين، فأن دالة الترجيح (Likelihood Function) للمشاهدة الأخيرة y_m وبالاعتماد على النموذج الحركي (Smith , 1979 , p.42) تكون كالآتي :

$$P(y_m/p) = C_{y_m}^{n_m} p^{y_m}(1-p)^{n_m - y_m} \qquad ...(3.1)$$

بحيث أن:- $y_i = 0 , 1 , ..., n_m$

وأن التوزيع الأولي (Prior Distribution) للمشاهدة قبل الأخيرة y_{m-1} هو عبارة عن توزيع بيتا (Beta Distribution) بالمعلمات a_{m-1} و b_{m-1} ، (Harrison & Stevens , (1971) , pp.341-362) أي أن :

$$P(p / y_1, y_2,...,y_{m-1}) \,\alpha\, p^{a_{m-1}-1}(1-p)^{b_{m-1}-1} \qquad ...(3.2)$$

وباستخدام نظرية بيز المتسلسل الآتية (Wald , 1947 , p.42) :-

$$P(p / y_1, y_2,...,y_m) \,\alpha\, P(p / y_1, y_2,...,y_{m-1}) P(y_m / p) \qquad ...(3.3)$$

يمكن الحصول على التوزيع النهائي (posterior distribution) للمعلمة p ، و كما يأتي:-

$$P(p / y_1, y_2,...,y_m) \,\alpha\, p^{a_{m-1}-1}(1-p)^{b_{m-1}-1} . p^{y_m}(1-p)^{n_m - y_m}$$

$$= p^{(a_{m-1}+y_m)-1}(1-p)^{(b_{m-1}+n_m-y_m)-1}$$

والتي هي نواة توزيع بيتا (The Kernel of Beta Distribution) ، (Box & Tiao , 1973 , p.82) لذلك فأن التوزيع النهائي الكامل (The Complete Posterior Distribution) يكون بالشكل الآتي :-

$$P(p / y_1, y_2,...,y_m) = \frac{\Gamma(a_m + b_m)}{\Gamma(a_m)\Gamma(b_m)} p^{a_m - 1}(1-p)^{b_m - 1} \qquad ...(3.4)$$

حيث أن :-

$$a_m = a_{m-1} + y_m \quad \dots (3.5)$$

$$b_m = b_{m-1} + n_m - y_m \quad \dots (3.6)$$

يمكن تحليل المعادلتين أعلاه (طه ، 2005 ، ص248-249) وكما يأتي :-

عندما $m = 1$ فان :

$$a_1 = a_0 + y_1$$
$$b_1 = b_0 + n_1 - y_1$$

وعندما $m = 2$ فان :

$$a_2 = a_1 + y_2$$
$$b_2 = b_1 + n_2 - y_2$$

وبالتعويض بقيم a_1 و b_1 نحصل على :-

$$a_2 = a_0 + y_1 + y_2 = a_0 + \sum_{i=1}^{2} y_i$$

$$b_2 = b_0 + n_1 - y_1 + n_2 - y_2 = b_0 + \sum_{i=1}^{2} n_i - \sum_{i=1}^{2} y_i$$

وهكذا إلى أن نحصل على المعادلتين الآتيتين :-

$$a_m = a_0 + \sum_{i=1}^{m} y_i \quad \dots (3.7)$$

$$b_m = b_0 + \sum_{i=1}^{m} n_i - \sum_{i=1}^{m} y_i \quad \dots (3.8)$$

مع العلم أن قيم المعلمات a_0 و b_0 هي قيم أولية يمكن أن نحصل عليها من خبرة أو التجارب السابقة حول قيم المشاهدات التي تمثل المتغير المعتمد الوهمي المستخدم فيها هذا التوزيع .
بناءً على ذلك يمكن الحصول على المعدل النهائي لتوزيع بيتا (Dr. Dhafir & Dr. Abdul Majid , 1988 , p.48) وكما يأتي :-

$$\mu = \frac{a_m}{a_m + b_m} = \frac{a_0 + \sum_{i=1}^{m} y_i}{a_0 + b_0 + \sum_{i=1}^{m} n_i} \qquad \ldots(3.9)$$

أي أن النسب باستخدام أسلوب بيز المتسلسل هي :-

$$p_{B_m} = \frac{a_0 + \sum_{i=1}^{m} y_i}{a_0 + b_0 + \sum_{i=1}^{m} n_i} \qquad \ldots(3.10)$$

حيث أن معدل النسبة لبيز المتسلسل للمشاهدة الأولى y_1 هو كما يأتي :-

$$p_{B_1} = \frac{a_0 + y_1}{a_0 + b_0 + n_1}$$

كذلك يمكن أن نحصل على معدل النسبة لبيز المتسلسل للمشاهدة الثانية y_2 كما يأتي :-

$$p_{B_2} = \frac{a_0 + \sum_{i=1}^{2} y_i}{a_0 + b_0 + \sum_{i=1}^{m} n_i}$$

وهكذا إلى أن نحصل على معدل النسبة لبيز المتسلسل للمشاهدة الأخيرة y_m كما هو موضح في الصيغة (3.10)، والتي تناولت المشاهدة الواحدة تلوا الأخرى فضلاً عن المعلومات الأولية a_0 و b_0 في تقدير معدلات النسب النهائية .

وعلى هذا الأساس يمكن أن توفق الدالة اللوجستية بطريقة المربعات الصغرى الموزونة باستخدام التحويل على معدلات النسب المقدرة باستخدام أسلوب بيز المتسلسل كمتغير معتمد وهمي وكما يأتي:-

$$p^*_{B_i} = Ln\left(\frac{p_{B_i}}{1 - p_{B_i}}\right) \qquad \ldots(3.11)$$

وبما أن التباين للخطأ هنا أيضا غير متجانس لذا فان طريقة المربعات الصغرى الموزونة يجب أن تستخدم لتقدير $\hat{\beta}_{B_0}$ و $\hat{\beta}_{B_1}$ ، لذلك فان أوزان بيز المتسلسل يمكن تقديرها بالاعتماد على التباين للتوزيع النهائي لتوزيع بيتا **(Dr. Dhafir & Dr.Abdul Majid , 1988 , p.48)** وكما يأتي :-

$$\sigma^2_{B_m} = \frac{a_m b_m}{(a_m + b_m)^2 (a_m + b_m + 1)}$$

أي أن :

$$\sigma^2_{B_m} = \frac{(a_0 + \sum_{i=1}^{m} y_i)(b_0 + \sum_{i=1}^{m} n_i - \sum_{i=1}^{m} y_i)}{(a_0 + b_0 + \sum_{i=1}^{m} n_i)^2 (a_0 + b_0 + \sum_{i=1}^{m} n_i + 1)} \qquad ...(3.12)$$

ولإيجاد تباين النسبة للمتغير المحول $p_{B_i}^*$ باستخدام أسلوب بيز المتسلسل يمكن أخذ مقلوب الصيغة أعلاه (كما نلاحظ في الصيغة (2.8)) و كما يأتي :-

$$V(p^*_{B_m}) = \frac{1}{\sigma^2_{B_m}} = \frac{(a_0 + b_0 + \sum_{i=1}^{m} n_i)^2 (a_0 + b_0 + \sum_{i=1}^{m} n_i + 1)}{(a_0 + \sum_{i=1}^{m} y_i)(b_0 + \sum_{i=1}^{m} n_i - \sum_{i=1}^{m} y_i)} \qquad ...(3.13)$$

لذلك فان أوزان بيز المتسلسل التي يمكن استخدامها في طريقة المربعات الصغرى الموزونة بأوزان بيز المتسلسل يمكن تقديرها كما يأتي :-

$$\hat{w}_{B_m} = \frac{1}{V(p^*_{B_m})} = \frac{(a_0 + \sum_{i=1}^{m} y_i)(b_0 + \sum_{i=1}^{m} n_i - \sum_{i=1}^{m} y_i)}{(a_0 + b_0 + \sum_{i=1}^{m} n_i)^2 (a_0 + b_0 + \sum_{i=1}^{m} n_i + 1)} \qquad ...(3.14)$$

حيث أن وزن المشاهدة الأولى يمكن تقديره كما يأتي:-

$$\hat{w}_{B_1} = \frac{(a_0 + y_1)(b_0 + n_1 - y_1)}{(a_0 + b_0 + n_1)^2 (a_0 + b_0 + n_1 + 1)}$$

كذلك يمكن تقدير وزن المشاهدة الثانية كما يأتي :-

$$\hat{w}_{B_2} = \frac{(a_0 + \sum_{i=1}^{2} y_i)(b_0 + \sum_{i=1}^{2} n_i - \sum_{i=1}^{2} y_i)}{(a_0 + b_0 + \sum_{i=1}^{2} n_i)^2 (a_0 + b_0 + \sum_{i=1}^{2} n_i + 1)}$$

وهكذا إلى أن نحصل على وزن المشاهدة الأخيرة y_m كما هو موضح في الصيغة (3.14) .

وعلى هذا الأساس يمكن تقدير معلمات الانحدار $\hat{\beta}_{B_0}$ و $\hat{\beta}_{B_1}$ باستخدام طريقة المربعات الصغرى الموزونة بأوزان بيز المتسلسل وكما يأتي :-

$$\hat{\beta}_{B_1} = \frac{\sum_{i=1}^{m} \hat{w}_{B_i} p^*_{B_i} x_i - (\sum_{i=1}^{m} \hat{w}_{B_i} p^*_{B_i})(\sum_{i=1}^{m} \hat{w}_{B_i} x_i) / \sum_{i=1}^{m} \hat{w}_{B_i}}{\sum_{i=1}^{m} \hat{w}_{B_i} x_i^2 - (\sum_{i=1}^{m} \hat{w}_{B_i} x_i)^2 / \sum_{i=1}^{m} \hat{w}_{B_i}} \quad ...(3.15)$$

كذلك لدينا :-

$$\hat{\beta}_{B_0} = \bar{y}_B - \hat{\beta}_1 \bar{x}_B \quad ... (3.16)$$

حيث أن :-

$$\bar{y}_B = p^*_B = \frac{\sum_{i=1}^{m} \hat{w}_{B_i} p^*_{B_i}}{\sum_{i=1}^{m} \hat{w}_{B_i}} \quad ...(3.17)$$

كذلك لدينا :-

$$\bar{x}_B = \frac{\sum_{i=1}^{m} \hat{w}_{B_i} x_i}{\sum_{i=1}^{m} \hat{w}_{B_i}} \quad ... (3.18)$$

نات أعلاه يمكن أن نحصل على معادلة الانحدار الخطية المقدرة معلماتها بواسطة طريقة ، الصغرى الموزونة بأوزان بيز المتسلسل للتعبير عن المعادلة اللوجستية ، أي أن :

$$\hat{p}^{*}_{B_i} = \hat{\beta}_{B_0} + \hat{\beta}_{B_1} x_i \qquad \dots (3.19)$$

و من خلال هذه المعادلة يمكن إيجاد القيم التقديرية $p^{*}_{B_i}$ بالتعويض بمستويات قيم x_i ، ومن ثم تحول قيم $\hat{p}^{*}_{B_i}$ إلى قيمها الأصلية (كما نلاحظ في الصيغة(2.14)) من خلال الصيغة الآتية:

$$\hat{p}_{B_i} = \frac{Exp\,(\hat{p}^{*}_{B_i})}{1 + Exp\,(\hat{p}^{*}_{B_i})} \qquad \dots (3.20)$$

و التي من خلالها نحصل على القيم الأصلية المقدرة التي تمثل المتغير المعتمد الوهمي .

4 : الجانب التطبيقي

لغرض توضيح كيفية استخدام أسلوب بيز المتسلسل في تقدير الأوزان المستخدمة في طريقة المربعات الصغرى لتقدير معلمات نموذج الانحدار المحول من النموذج اللوجستي وإجراء مقارنة بين الأسلوب التقليدي وأسلوب بيز في تقدير هذه الأوزان، قام الباحثان بدراسة العلاقة بين نسبة المرضى المشافين من مرض معين وتراكيز دواء معين أعطي أليهم في مستشفى رزكاري/ أربيل لثمانية عينات مختلفة وكانت البيانات كما يأتي :

الجدول رقم (1)

بيانات تمثل تراكيز الدواء والاستجابات المقابلة لها لعينات مختلفة

عدد المرضى المشافين y_i	الاستجابة (1= يشفى ، 0= لم يشفى)	حجم العينة n_i	تراكيز الدواء x_i
3	0,0,0,1,0,0,1,1,0,...	60	0.5
9	1,0,0,1,1,0,0,0,1,...	80	1
12	1,0,0,1,1,0,0,0,1,...	100	1.5
14	0,0,1,0,0,0,0,1,0,...	110	2
18	0,1,1,0,0,0,0,1,1,...	115	2.5
7	0,0,0,1,0,0,1,1,0,...	90	3
5	1,1,0,0,0,1,0,0,0,...	80	3.5
4	0,0,1,1,0,1,0,1,0,...	70	4

المصدر: تم جمع البيانات من قبل الباحثان في وحدة التسجيل لمستشفى رزكاري لعام 2006 .

4.1 : تقدير النموذج اللوجستي باستخدام الأوزان التقليدية :

نجد قيم النسب p_i لكل مستوى من مستويات الدواء بالاعتماد على الصيغة (2.4) ، ومن ثم نجد قيم النسب المحولة p_i^* بالاعتماد على الصيغة (2.5) ، ومن ثم نقدر الأوزان التقليدية $\hat{w}_i$ بالاعتماد على الصيغة (2.9) ، ويمكن تلخيص النتائج من خلال الجدول الآتي :

الجدول رقم (2)

قيم التحويل اللوغارتمي والأوزان التقليدية لبيانات الجدول أعلاه

تراكيز الدواء x_i	حجم العينة n_i	عدد المرضى المشافين y_i	p_i	p_i^*	$\hat{w}_i$
0.5	60	3	0.0500	-2.9444	2.8500
1	80	9	0.1125	-2.0655	7.9875
1.5	100	12	0.1200	-1.9924	10.5600
2	110	14	0.1273	-1.9250	12.2204
2.5	115	18	0.1565	-1.6845	15.1809
3	90	7	0.0778	-2.4726	6.4572
3.5	80	5	0.0625	-2.7081	4.6875
4	70	4	0.0571	-2.8042	3.7688

المصدر:- تم إعداد الجدول من قبل الباحثان .

وعلى هذا الأساس يمكن تقدير معلمات الانحدار $\hat{\beta}_0$ و $\hat{\beta}_1$ باستخدام طريقة المربعات الصغرى الموزونة التقليدية وبالاعتماد على الصيغ (2.10) ، (2.11) ، (2.12) ، (2.13) والحصول على نموذج الانحدار (كما في الصيغة (2.14)) وكما يأتي :-

$$\hat{p}_i^* = -1.8518 - 0.1175 x_i$$

هذا ولإعادة التحويل إلى المشاهدات الأصلية يمكن من خلال ما يأتي :

فمثلا لتركيز الدواء $x_1 = 0.5$ فان القيمة المتوقعة $\hat{p}_1^*$ نحصل عليها من خلال ما يأتي :

$$\hat{p}_1^* = -1.8518 - 0.1175\,(0.5) = -1.9106$$

$$\hat{p}_1 = \frac{\mathrm{Exp}(\beta_0 + \beta_1 x_1)}{1 + \mathrm{Exp}(\beta_0 + \beta_1 x_1)} = \frac{\mathrm{Exp}\left(\hat{p}_1^*\right)}{1 + \mathrm{Exp}\left(\hat{p}_1^*\right)} = \frac{\mathrm{Exp}\left(-1.9106\right)}{1 + \mathrm{Exp}\left(-1.9106\right)} = 0.1289$$

وذلك عندما تكون $x_1 = 0.5$ وهكذا بقية المشاهدات .

4.2 : تقدير النموذج اللوجستي باستخدام أوزان بيز المتسلسل :

في البداية نقدر نسب بيز المتسلسل p_{B_m} حيث أن (m = 1 , 2 , ... , 8) بالاعتماد على الصيغة (3.10) وعلى فرض قيم أولية $a_0 = 9$ التي تمثل عدد حالات الشفاء و $b_0 = 80$ التي تمثل عدد حالات عدم الشفاء (أخذت القيم الأولية من أحد كبار الأطباء في مستشفى رزكاري) ، ومن ثم إيجاد نسب بيز المحولة $p^*_{B_m}$ من خلال الصيغة (3.11) ، ومن ثم نقدر أوزان بيز المتسلسل $\hat{w}_{B_m}$ من خلال الصيغة (3.14) ، ويمكن تلخيص النتائج من خلال الجدول الآتي :

الجدول رقم (3)

نسب بيز المتسلسل والقيم المحولة وأوزان بيز المتسلسل لبيانات الجدول رقم(1)

تراكيز الدواء x_i	حجم العينة n_i	عدد المرضى المشافين y_i	p_{B_m}	$p^*_{B_m}$	$\hat{w}_{B_m}$
0.5	60	3	0.0805	-2.4356	0.00049
1	80	9	0.0917	-2.2931	0.00036
1.5	100	12	0.1003	-2.1939	0.00027
2	110	14	0.1071	-2.1207	0.00022
2.5	115	18	0.1173	-2.0183	0.00019
3	90	7	0.1118	-2.0725	0.00015
3.5	80	5	0.1064	-2.1281	0.00013
4	70	4	0.1020	-2.1752	0.00012

المصدر:- تم إعداد الجدول من قبل الباحثان .

وعلى هذا الأساس يمكن تقدير معلمات الانحدار $\hat{\beta}_{B_0}$ و $\hat{\beta}_{B_1}$ باستخدام طريقة المربعات الصغرى الموزونة (المقدرة بأسلوب بيز المتسلسل) وبالاعتماد على الصيغ (3.15)، (3.16) ، (3.17) ، (3.18) والحصول على نموذج الانحدار (كما في الصيغة (3.19)) وكما يأتي:-

$$\hat{p}^* = -2.0462 - 0.1082\, x$$

فمثلا لتركيز الدواء $x_1 = 0.5$ فأن القيمة المتوقعة $\hat{p}_1^*$ نحصل عليها من خلال ما يأتي :

$$\hat{p}_1^* = -2.0462 - 0.1083(0.5) = -2.1004$$

والآن تحول $\hat{p}_1^*$ إلى قيمتها الأصلية كالآتي :

$$\hat{p}_1 = \frac{\operatorname{Exp}(\beta_0 + \beta_1 x_1)}{1 + \operatorname{Exp}(\beta_0 + \beta_1 x_1)} = \frac{\operatorname{Exp}(\hat{p}_1^*)}{1 + \operatorname{Exp}(\hat{p}_1^*)} = \frac{\operatorname{Exp}(-2.1004)}{1 + \operatorname{Exp}(-2.1004)} = 0.1091$$

وذلك عندما تكون $x_1 = 0.5$ وهكذا بقية المشاهدات .

4.3 : المقارنة بين النموذجين :

من خلال النموذجان أعلاه يمكن المقارنة بين الأسلوب التقليدي وأسلوب بيز المتسلسل لاختيار أفضل طريقة لتقدير معلمات النموذجين والملائم للبيانات أعلاه، ولهذا الغرض تم إيجاد متوسط مجموع مربعات الخطأ (MSE) للنموذجان أعلاه، وكانت كما يوضحه الجدول الآتي:

الجدول رقم (4)

متوسط مجموع مربعات الخطاء للطرق المستخدمة

الطرق المستخدمة	MSE
طريقة المربعات الصغرى الموزونة باستخدام الأسلوب التقليدي في تقدير الأوزان	1.4918
طريقة المربعات الصغرى الموزونة باستخدام طريقة أسلوب بيز المتسلسل في تقدير الأوزان	0.0000022

المصدر:– تم إعداد الجدول من قبل الباحثان .

من خلال الجدول أعلاه نلاحظ أن أسلوب بيز المتسلسل في تقدير الأوزان أدى إلى حصولنا على متوسط مجموع مربعات للخطأ أقل وبشكل واضح جدا، مما يؤكد على أفضلية أسلوب تقدير المعلمات باستخدام أسلوب بيز المتسلسل وهي نتيجة طبيعية لما يتميز به هذا الأسلوب من أمكانية التقدير باستخدام معلومات المشاهدات فضلا عما توفره المعلومات الأولية المتأتية من الخبرة أو التجارب السابقة .

الاستنتاجات والتوصيات

1- أمكانية استخدام أسلوب بيز المتسلسل في تقدير الأوزان المستخدمة في طريقة المربعات الصغرى الموزونة لتقدير معلمات النموذج اللوجستي عندما يكون المتغير المعتمد متغيرا نوعيا ذا فئتين.

2- أسلوب بيز المتسلسل أدى إلى حصولنا على تقدير معلمات النموذج اللوجستي بشكل أفضل من الأسلوب التقليدي بالنسبة لدراسة العلاقة بين نسبة المرضى المشافين من مرض معين وتراكيز دواء معين أعطي أليهم في مستشفى رزكاري/ أربيل.

3- يوصي الباحثان باستخدام أسلوب بيز المتسلسل في تقدير الأوزان المستخدمة في طريقة المربعات الصغرى الموزونة لتقدير معلمات النموذج اللوجستي عندما يكون المتغير المعتمد متغيرا نوعيا ذا فئتين وعند توفر معلومات أولية يمكن أن نحصل عليها من الخبرة أو التجارب السابقة حول قيم المعلمات المقدرة.

4- يوصي الباحثان بأجراء دراسات مماثلة وذلك بأخذ قيم أولية مختلفة لاستخدامها كمعلومات أولية في تقدير معلمات النموذج اللوجستي للعلاقة أعلاه للحصول على أفضل تقدير ممكن لمعلمات هذا النموذج.

5- يوصي الباحثان بتناول دراسة مماثلة ولكن في تقدير معلمات النموذج اللوجستي المتعدد كما في الصيغة (2.16) التي تحتوي على متغيرين مستقلين أو أكثر.

6- يوصي الباحثان بأجراء دراسة مماثلة باستخدام أسلوب بيز ذات المعلومات الغنية أو أسلوب بيز التجريبي في تقدير معلمات النموذج اللوجستي.

المراجع

أولا- المراجع باللغة العربية :

1- د. أموري هادي كاظم، (2002)، "الاقتصاد القياسي"، مطبعة جامعة الموصل، ص155.
2- د. أموري هادي كاظم ود. سعيد علي هادي، (1990)، "القياس الاقتصادي التطبيقي"، جامعة بغداد، ص 149.
3- د. خاشع محمود الراوي، (1987)، "المدخل إلى تحليل الانحدار"، جامعة الموصل ص 442-448.
4- طه حسين علي، "نمذجة سلاسل ماركوف لعمليات برنولي"، بحث منشور في مجلة زانكو جامعة صلاح الدين/ أربيل، العدد (26)، كانون الأول (2005)، ص248-249.
5- د. محمد مناجد الدليمي ود.أموري هادي كاظم، (1990) ، " تحليل الانحدار بالأمثلة " ، جامعة بغداد ، ص148 .

ثانيا- المراجع باللغة الأجنبية :

1- Box & Tiao , (1973) , " Bayesian Inference in Statistical Analysis " Addison–Wesley publishing company , California , London , p.12 .
2- Cox, D.R., (1970), "The Analysis of Binary Data", Mehtuen, London , p.155 .
3- Dr. Dhafir & Dr. Abdul Majid , "Statistical Inference" , (1988) , p.48
4- Draper N.R. & Smith H. , (1981) , " Applied Regression Analysis " , John Wiley & Sons , Inc. Published simultaneously in Canada , pp.108-116 .
5- Harrison P. J. & Stevens C. F. , (1971) , " A Bayesian approach to short – term forecasting" , Operation Research ,Vol. 22 , No.2 , pp. 341-362 .
6- Smith , J.Q. , (1979), "A generalization of the Bayesian Steady Forecasting Model " , JRS-Society series B, Vol.41 .
7- Wald A. , (1947), " Sequential Analysis ", Wiley & Sons , Inc. , Newyork .