مقدرات الإمكان الأعظم الموزونة الحصينة لأنموذج اللوجستك مع تطبيق عملي

ا.م.د نزار مصطفى جواد
جامعة بغداد/ كلية الإدارة والاقتصاد

شيرين علي حسين
ماجستير إحصاء

المقدمة وهدف البحث

1- المقدمة

يعد استخدام النماذج الخطية من أكثر الطرائق الإحصائية استعمالاً في مختلف العلوم ، لأنه يصف العلاقة بين المتغيرات على هيئة معادلة خطية ، ويستعمل تحليل النماذج الخطية لعدة أغراض أهمها التنبؤ (Prediction) ، حيث يمكن تقدير الاستجابة والتنبؤ بما يفيد كثيراً في التخطيط و اتخاذ القرارات، وان من متطلبات تحليل الظواهر تحديد الأنموذج الملائم للظاهرة ، لأن اختيار الأنموذج المناسب يؤدي إلى اتخاذ القرار الصحيح بشأن المشكلة المراد دراستها، ومن ضمن نماذج الانحدار هو أنموذج انحدار اللوجستك الذي يؤدي دوراً خاصاً في التحليلات الإحصائية .

تبرز استخدامات أنموذج انحدار اللوجستك بشكل خاص في الدراسات المتعلقة بعلوم الحياة والعلوم الزراعية والطبية وخاصة في الدراسات ذات الطابع التجريبي، لكونه من النماذج الملائمة للبيانات الثنائية (Binary Data) والهدف الرئيس لتحليلات انحدار اللوجستك هو لتوضيح ما يسمى بالاستقلالية أو المتغيرات المستقلة وتأثيره على المتغيرات المعتمدة [3] .

ولأن أنموذج انحدار اللوجستك له الدور الأساس في دراسة وتحليل اي ظاهرة اعتماداً على البيانات الإحصائية المتوفرة عن هذه الظاهرة من جهة ، وعلى نقاوة هذه البيانات أو خلوها من الشواذ من جهة أخرى ، وعليه يجب اختيار الطرائق المناسبة لتقدير معلماته لغرض المقارنة ، وإيجاد أفضل طريقة للتقدير بحيث يمكن الاعتماد عليها للوصول إلى نتائج أكثر دقة ، وهناك أكثر من طريقة لتقدير المعالم منها الطرائق التقليدية (الكلاسيكية) ولكن مقدرات هذه الطرائق قد تخفق في حالة احتواء البيانات على قيم شاذة ، لذلك انصب اهتمام الباحثين بكيفية التعامل مع البيانات التي تحتوي على

قيم شاذة ، وذلك عن طريق استخدام طرائق التقدير الحصينة ، إذ يتم الحصول على مقدرات حصينة أكثر كفاءة من الطرائق الاعتيادية في حالة وجود الشواذ.

وعلى الرغم من اختلاف صيغ هذه التقديرات إلا أن هدفها واحد ، وهو استخدام اسلوب الموازنة بين المشاهدات وذلك من خلال أقران المشاهدات التي يعتقد أنها شواذ بأوزان أقل من تلك التي تقرن مع بقية المشاهدات للتقليل من تأثير الشواذ ، وكذلك استخدامها أسلوب التكرار في الحساب.

2- هدف البحث

يهدف هذا البحث إلى معرفة أهم العوامل المؤثرة على مرض (الذبحة الصدرية) من خلال طريقة مقدرات الإمكان الأعظم الموزونة (WMLE) على المرضى الراقدين في مستشفى ابن النفيس باستخدام أنموذج انحدار اللوجستك .

3- الاستعراض المرجعي The Review of Literature

هناك الكثير من الباحثين الذين قدموا دراسات في موضوع أنموذج انحدار اللوجستك والمقدرات الحصينة، وفيما يلي عرض لبعض هذه الدراسات:

في عام (2004) قدم كل من (Wang ,VanEeden and Zidekُ) [13] بحثاً تناولوا فيه استخدام طريقة الإمكان الأعظم الموزونة (WMLE) وطريقة المربعات الصغرى الموزونة (WLS) وقاموا بتطبيقها على بيانات واقعية عن الإمراض التنفسية وبعينة حجمها(733) ، إذ عمل مفاضلة بين هذه الطرائق باستخدام متوسط مربعات الخطأ (MSE) والتحيز (Biase) .

وفي عام (2005) قام الباحث (Simeckova) [9] بدراسة أنموذج انحدار اللوجستك اعتماداً على مقدرات (WMLE) ، حيث استعرض أسلوب (WML – Estimate) وقدر المعالم باستخدام هذه الطريقة وبوجود نسبة من التلوث في البيانات وقام بعملية المقارنة باستخدام (MSE) وبأسلوب المحاكاة .

وفي عام (2006) قدم الباحثان (Vieira and Skinner) [12] دراسة حول طريقة مقدرات الإمكان الأعظم الموزونة(Weighted Maximum Likelihood Estimators) للنماذج الخطية ، حيث تضمنت هذه الدراسة استخدام أسلوب المحاكاة والمقارنة بين طريقة مقدرات(WMLE) وطرائق مقدرات (Asymptotically Distribution Free Methods) ومختصره(ADF) ، وقد أثبتت كفاءة التقدير باستخدام طريقة(WMLE)على الطرائق الموزونة (ADF) من خلال المؤشرات الإحصائية التباين والتحيز.

الجانب النظري

4- أنموذج انحدار اللوجستك Logistic Regression Model

يعد إنموذج إنحدار اللوجستك من النماذج القابلة للتحويل الى نماذج خطية ، وان هناك علاقة بين متغير الاستجابة (Y_i) وعدد من المتغيرات التوضيحية نطلق عليها ($X_1, X_2, \ldots\ldots, X_k$) التي يعتقد بان لها تاثيراً كبيراً على احتمال الاستجابة (P).

لكون الاستجابة ثنائية، فأن المتغير المعتمد (متغير الاستجابة Y_i) له قيمتان ، أما الواحد وباحتمال نجاح (P_i) أو الصفر و وباحتمال فشل (q_i) ، (إما حدوث استجابة أو عدم حدوث استجابة) [13]

وبتفسير آخر :

P_i: احتمال الاستجابة عندما Y=1

q_i: احتمال عدم الاستجابة عندما Y=0

يتبع متغير الاستجابة (Y) توزيع برنولي B ($1, P_i$)

P_i : دالة الاستجابة اللوجستية

$$P_i = \frac{e^{\beta_0 + \sum_{j=1}^{k} \beta_j X_{ij}}}{1 + e^{\beta_0 + \sum_{j=1}^{k} \beta_j X_{ij}}} \quad \ldots\ldots\ldots\ldots(1)$$

$$q_i = 1 - P_i = \frac{1}{1 + e^{\beta_0 + \sum_{j=1}^{k} \beta_j X_{ij}}} \quad \ldots\ldots\ldots\ldots(2)$$

$\beta_0, \beta_1, \ldots\ldots, \beta_k$: معالم مجهولة يتم تقديرها .

X_{ij} : متغيرات مستقلة (متغيرات توضيحية) .

$i = 1,..... , m$ ، **(**m**)** عدد المشاهدات .

$j = 1,....., k$ ، **(** k **)** عدد المتغيرات المستقلة .

يعرف أنموذج انحدار اللوجستك أيضا بنموذج Logit [5] ويمكن التعبير عنه رياضياً كالآتي :

$$\frac{P(Y=1)}{1-P(Y=1)} = \frac{P_i}{1-P_i} = \exp(\beta_0 + \sum_{j=1}^{k} \beta_j X_{ij}) \qquad \text{..... (3)}$$

ونظراً لصعوبة تطبيق اساليب تقدير النماذج الخطية لتقدير معلمات هذا النموذج بصورة مباشرة ، لذلك اقترح الباحث Berkson (1944) [1] بتحويل إنموذج اللوجستك الى دالة خطية وذلك بأخذ اللوغاريتم الطبيعي لكلا الطرفين نحصل على

$$Log[\frac{P_i}{1-P_i}] = \beta_0 + \sum_{j=1}^{k} \beta_j X_{ij} = Z \text{ .....(4)}$$

Z : تمثل العلاقة الخطية الناتجة من اخذ اللوغاريتم للنسبة $\left(\frac{P_i}{1-P_i} \right)$

5- مقدرات الإمكان الأعظم الموزونة (WMLE) [8]

Weighted Maximum Likelihood Estimates

تعد طريقة **(** WMLE **)** واحدة من أهم الطرائق الحصينة ، وإن الفكرة الأساسية لهذه الطريقة نشأت من خلال تحوير الاجراءات المتبعة في طريقة الامكان الاعظم ، حيث اقترح الباحثان **(**1993**)** **(**Carroll and Pederson**)** [2] طريقة بسيطة لتحويل اسلوب **(**MLE**)** الى مقدر حصين وذلك من خلال اقران المشاهدات باوزان للتقليل من تأثير الشواذ .

تعتمد هذه الطريقة على دالة الكثافة الاحتمالية للمتغيرات العشوائية **(** $Y_1, Y_2, ..., Y_m$ **)** والذي يتوزع برنو لي

$$P(Y_i = y_i) = P_i^{y_i} (1 - P_i)^{1-y_i} \text{ .....(5)} \qquad i = 1, ...m$$

وبأخذ اللوغاريتم للطرفين نحصل على

$$LogP(Y_i = y_i) = y_i LogP_i + (1-y_i) Log(1-P_i)$$

.....(6)

ونظراً لوجود علاقة بين احتمال الاستجابة (P_i) والمتغيرات المستقلة في انموذج انحدار اللوجستك

وبتعويض المعادلتين (1) و (2) في المعادلة (6) نحصل على

$$LogP(Y_i = y_i) = y_i Log[\frac{\exp(\beta_0 + \sum_{g=1}^{k} \beta_g X_{ij})}{1+\exp(\beta_0 - \sum_{j=1}^{k} \beta_j X_{ij})}] + (1-y_i) Log[\frac{1}{1+\exp(\beta_0 + \sum_{j=1}^{k} \beta_j X_{ij})}] \quad \text{....(7)}$$

وهذا التقدير(7) يتطلب تصغير المقدار المطلوب [11]

$$Min \sum_{i=1}^{m} W_i L_i(\beta) \text{.....(8)}$$

$L_i(\beta)$: تمثل لوغاريتم الدالة

$$L_i(\beta) = Log\ P(Y_i = y_i)$$

$i=1,\ldots..,m$

W_i: تمثل دالة الوزن .

ولغرض الحصول على مقدرات (WMLE) التي تصغر المقدار في المعادلة (8) يتم استخدام إحدى الطرائق العددية او باستخدام طريقة المربعات الصغرى الموزونة وكالآتي

$$\hat{\beta} = (X'WX)^{-1} X'WZ \text{ .....(9)}$$

X: مصفوفة المتغيرات المستقلة(التوضيحية) وبدرجة ($m * (k+1)$) .

W: تمثل مصفوفة قطرية عناصر قطرها الرئيس دالة الوزن W_i.

وبما أن حصانة المقدر الناتج عن هذه الطريقة تعتمد على دالة الوزن W_i فأن هناك صيغاً عديدة لهذه الدالة ، لذلك تم استخدام الصيغة المقترحة من قبل الباحثان (2003) (Meuller and Neykov) [7] والتي وضعت باستخدام دالة الوزن

$$W_1(t) = (at + b) \text{ .....(10)}$$

مع الأخذ بنظر الاعتبار قيم الثوابت $a=0.8$, $b=0.2$

t **تمثل دالة** (function) **ويمكن إيجادها من الصيغة التالية**

$$t = [(X - M^{\wedge})' \ \Sigma^{\wedge -1} (X - M^{\wedge})' \quad \text{.....(11)}$$

$$X = [\, X_{i_1}, X_{i_2}, ..., X_{i_k} \,]'$$ حيث

وبتكرار الصيغة (11) لـ m من المرات نحصل على t_i

حيث $i=1,.....,m$

اذ أن استخدام هذا الاسلوب يحسن الاسلوب الكلاسيكي ، ويتم ذلك بأتباع الخطوات الآتية:

1. تحسب التقديرات الاعتيادية لموجه الوسط الحسابي ومصفوفة التباين المشتركة $\hat{M}$ و $\hat{\sum}$ على التوالي

$$M^{\wedge} = \frac{1}{m}\sum_{i=1}^{m} X_i \qquad .....(12)$$

$$\Sigma^{\wedge} = \frac{1}{m}\sum_{i=1}^{m} ((X_i - M^{\wedge})(X_i - M^{\wedge})') \qquad .....(13)$$

2. تستخدم هذه التقديرات في حساب الصيغة (11) لكل مجموعة.

3. يعاد حساب المتوسط ومصفوفة التباين المشتركة باستخدام اوزان مقدرة تعتمد على قيم t_i وكالاتي[10]

$$\hat{M}^{*} = \frac{\sum_{i=1}^{m} W_{0i} \; X_i}{\sum_{i=1}^{m} W_{0i}} \qquad ..... (14)$$

$$\hat{\sum}^{*} = \frac{\sum_{i=1}^{m} W_{0i}^{\;2} [(X_i - \hat{M}^{*})(X_i - \hat{M}^{*})']}{\sum_{i=1}^{m} W_{0i}^{\;2} - 1} \qquad ..... (15)$$

حيث

$\hat{M}^{*}$: متجه موقع حصين (robust location vector) ، وبدرجة (k*1).

$\hat{\sum}^{*}$: مصفوفة قياس حصين (robust dispersion matrix) ، وبدرجة (k*k).

إما) W_{0i} (فتمثل دالة وزن ، وهناك عدة صيغ لهذه الدالة ، لذلك تم استخدام إحدى هذه الصيغ وهي دالة Huber [6] ، حيث يكون المقدر الناتج عنها غير شديد الحساسية والتأثر بالشواذ ، والصيغة كما يأتي :

$$W_{0i} = \min\{1, \frac{g}{|t_i|} \quad \text{.....(16)}$$

حيث $g = 1.37$

تمثل هذه الخطوات التكرار الاول ، ثم يعاد حساب الخطوة (2) بصورة تكرارية بالاعتماد على نتائج التكرار السابق وتتوقف عندما يكون الفرق بين نتائج تكرارين متعاقبين غير ملحوظ او قليل .

الجانب التطبيقي

هناك الكثير من الدراسات التي تناولت الإصابة بإمراض القلب في مختلف إنحاء العالم ، وقد استخدمت طرقا متعددة ، منها المسوحات والتقديرات بالاعتماد على مقاييس مختلفة ، كما تناولت هذه الدراسات تأثير عوامل متعددة ومختلفة منها العمر والجنس والكولسترول وغير ذلك ، وان الدراسات والبحوث الإحصائية حول المرض قد تساهم في مجال تقليل الإصابة به . كما إن المعلومات الأفضل عن إمراض القلب تفتح تدريجياً أفاقا مشرقة تبشر بتخفيف وطأة هذه الإمراض الشديدة الفتك في يومنا هذا .

في الحقيقة إن وفاة أو شفاء الشخص المصاب بهذا المرض تتأثر بشكل كبير ومباشر بعوامل كثيرة أهمها ضغط الدم ، ونسبة السكر في الدم ، وعمر المصاب ، والتدخين ، وتناول الكحول ، وطبيعة العمل ، وعامل الوراثة ، ونسبة الكولسترول في الدم، وغيرها من العوامل .

وتكمن خطورة هذا المرض في كون نسبة كبيرة من الذين يصابون به يتعرضون إلى الوفاة بعد أيام قليلة من إصابتهم به . وفي أحيان أخرى كثيرة ونتيجة للعناية المركزة التي يتلقاها المريض في المستشفى والبيت قد تؤدي إلى الشفاء . يهدف هذا الجانب من البحث إلى تقدير مدى تأثير هذه العوامل على متغير الاستجابة (شفاء أو وفاة المريض) وذلك باستخدام أنموذج انحدار اللوجستك . جمعت البيانات الخاصة بالبحث من مستشفى ابن النفيس باعتبارها من المستشفيات الكبيرة المتخصصة بإمراض القلب والأوعية الدموية في العراق. وقد تمت الاستعانة بمجموعة من الأطباء المتخصصين بإمراض القلب لتصنيف أهم العوامل المؤثرة على المرض ، اذ تم جمع البيانات الخاصة بالسنوات 2007 ، 2008 وبعينة حجمها (200) طبلة مريض ، ومن خلال البطاقة

الخاصة لكل مريض أخذت متغيرات الآتية وهي استجابة المريض ، عمر المريض ، وجنس المريض ، والسكر في الدم ، وضغط الدم ، والتدخين.

وبغية تسهيل مهمة تحليل هذه البيانات فقد تم وضعها على شكل مجاميع وبواقع (86) مجموعة على ضوء التوافيق الممكنة لجميع فئات المتغيرات المستقلة .

إما المتغيرات المستقلة وفئاتها كالآتي :

١. **متغير العمر (X_1) كان على أربع فئات هي :**

الفئة الأولى ($X_1=1$) لعمر اقل من (45) .

الفئة الثانية ($X_1=2$) لعمر من (45-60) .

الفئة الثالثة ($X_1=3$) لعمر من (60-75) .

الفئة الرابعة ($X_1=4$) لعمر من (75) فأكثر.

٢. **متغير الجنس (X_2) وضم فئتين هما :**

الفئة الأولى ($X_2=1$) ذكور .

الفئة الثانية ($X_2=2$) إناث .

٣. **متغير نسبة السكر في الدم (X_3) وقد ضم أربع فئات هي :**

الفئة الأولى ($X_3=1$) نسبة السكر اقل من (165) .

الفئة الثانية ($X_3=2$) نسبة السكر بين (165-230) .

الفئة الثالثة ($X_3=3$) نسبة السكر بين (230-295) .

الفئة الرابعة ($X_3=4$) نسبة السكر من (295) فأكثر.

٤. **متغير ضغط الدم (X_4) وقد ضم أربع فئات هي :**

الفئة الأولى ($X_4=1$) الضغط اقل من (120) .

الفئة الثانية ($X_4=2$) الضغط بين (120-160) .

الفئة الثالثة ($X_4=3$) الضغط بين (160-200) .

الفئة الرابعة ($X_4=4$) الضغط من (200) فأكثر.

٥. **متغير التدخين (X_5) وضم فئتين هما :**

الفئة الأولى ($X_5=1$) يدخن .

الفئة الثانية ($X_5=2$) لا يدخن.

إما متغير الاستجابة Y يمثل شفاء المريض (Y=0) ووفاة المريض (Y=1)
ولغرض معرفة ملائمة البيانات للأنموذج فقد تم اختبار الفرضية الآتية :

H_0: البيانات تتوزع توزيع ثنائي الحدين.

H_1: البيانات لا تتوزع توزيع ثنائي الحدين.

كانت قيمة χ^2 المستخرجة هي (48.248) وإما قيمة χ^2 الجد ولية بمستوى معنوية (0.05) وبدرجة حرية (85) فهي (57.5) و عليه تعد البيانات ذات توزيع ثنائي الحدين وبذلك يمكن تطبيق أنموذج انحدار اللوجستك على هذه البيانات .

ويبين الجدول(4-1) مجاميع مرضى القلب (n_i) والمتوفين منهم (Y_i) بحسب فئات العمر (X_1) ، والجنس (X_2) ، ونسبة السكر في الدم (X_3) ، وضغط الدم (X_4) ، والتدخين (X_5) .

جدول(1)

مجاميع مرضى القلب (n_i) **والمتوفين منهم** (Y_i) **بحسب فئات العمر** (X_1) ، **والجنس** (X_2) ، **ونسبة السكر في الدم** (X_3) ، **وضغط الدم** (X_4) ، **والتدخين** (X_5)

i	X_1	X_2	X_3	X_4	X_5	Y_i	n_i	i	X_1	X_2	X_3	X_4	X_5	Y_i	n_i
1	1	1	1	1	2	2	4	24	2	1	3	2	2	2	5
2	1	1	1	2	1	0	5	25	2	1	3	4	١	1	2
3	1	1	1	2	2	1	2	26	2	1	4	1	1	0	2
4	1	1	1	3	1	1	2	27	2	1	4	1	١	1	1
5	1	1	1	4	1	0	3	28	2	1	4	2	1	1	1
6	1	1	1	4	2	1	1	29	2	1	4	4	٢	1	1
7	1	1	2	1	1	0	1	30	2	2	1	1	١	1	2
8	1	1	3	1	2	0	1	31	2	2	1	1	٢	0	2
9	1	1	3	4	1	0	2	32	2	2	1	2	١	0	2
10	1	1	4	2	1	0	1	33	2	2	1	2	٢	1	9
11	1	1	4	4	1	1	1	34	2	2	1	3	٢	2	2
12	1	2	1	1	2	0	4	35	2	2	1	4	٢	0	1
13	1	2	1	2	1	0	1	36	2	2	1	4	٢	0	1
14	1	2	1	2	2	0	3	37	2	2	2	2	٢	0	2
15	1	2	1	4	2	0	2	38	2	2	3	1	١	0	1
16	2	1	1	1	1	1	4	39	2	2	3	2	٢	1	1
17	2	1	1	1	2	1	5	40	2	2	3	3	٢	0	1
18	2	1	1	2	1	2	6	41	2	2	4	1	2	0	2
19	2	1	1	2	١	0	4	42	2	2	4	2	2	0	4
20	2	1	1	3	1	0	4	43	2	2	4	3	2	2	2
21	2	1	1	3	١	0	1	44	2	2	4	4	٢	0	1
22	2	1	1	4	1	3	5	45	3	1	1	1	١	2	6
23	2	1	1	4	2	2	2	46	3	1	1	1	٢	0	2

تابع لجدول(1)

i	X_1	X_2	X_3	X_4	X_5	Y_i	n_i	i	X_1	X_2	X_3	X_4	X_5	Y_i	n_i
47	3	1	1	2	1	1	5	70	3	2	4	3	2	2	3
48	3	1	1	2	2	0	4	71	4	1	1	1	1	0	2
49	3	1	1	4	١	2	4	72	4	1	1	1	2	1	1
50	3	1	2	1	1	0	2	73	4	1	1	2	1	0	1
51	3	1	2	2	2	0	2	74	4	1	1	3	1	2	3
52	3	1	3	2	1	1	3	75	4	1	1	3	2	1	1
53	3	1	3	3	١	0	1	76	4	1	1	4	2	0	1
54	3	1	3	4	1	1	1	77	4	1	2	1	2	0	1
55	3	1	4	1	1	0	1	78	4	1	3	3	1	1	1
56	3	1	4	1	2	0	1	79	4	1	3	3	2	1	1
57	3	1	4	2	1	2	2	80	4	1	4	1	2	0	1
58	3	1	4	2	١	1	4	81	4	1	4	4	2	1	1
59	3	1	4	3	2	1	1	82	4	2	1	1	2	1	1
60	3	2	1	1	1	0	1	83	4	2	1	2	2	2	2
61	3	2	1	1	2	2	6	84	4	2	1	4	2	1	1
62	3	2	1	2	1	0	5	85	4	2	4	1	2	0	1
63	3	2	1	2	2	1	5	86	4	2	4	3	2	2	2
64	3	2	1	3	2	3	4								
65	3	2	1	4	1	1	1								
66	3	2	2	2	١	0	1								
67	3	2	2	4	2	0	1								
68	3	2	3	3	2	2	2								
69	3	2	4	2	2	0	2								

لقد تم استخدام طريقة ($WMLE(w_1)$) في الجانب التطبيقي وبحسب الصيغ المبينة أعلاه ،إذ إن هذه الطريقة تحتاج إلى تقديرات أولية لذلك تم استخراج قيم هذه التقديرات باستخدام طريقة المربعات الصغرى الاعتيادية وكانت كما في الجدول الآتي :

جدول (2)

القيم التقديرية للمعلمات بطريقة المربعات الصغرى

Param. / Method	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
(OLS)	-9.76321	0.998941	-1.732466	-0.076485	1.216757	0.776051

كما أنه تم استخدام نظام البرنامج الجاهز (SPSS) ، وكذلك تم كتابة البرنامج بلغة (Visual Basic) وتطبيقها على بيانات التجربة إذ تم حساب مقدرات المعالم بطريقة ($WMLE(w_1)$) وحساب MSE للأنموذج وكما في الجدول الآتي :

جدول (3)

القيم التقديرية للمعلمات بطريقة $WMLE(w_1)$

Param. / Method	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	MSE(MOD)
$WMLE(w_1)$	-7.248097	0.27032	-0.394097	0.044704	0.380192	0.168247	0.85015454

وفي الجدول (4) تظهر قيم (P_i) (احتمال الاستجابة) التي تم حسابها باستخدام العلاقة (1) المذكورة أعلاه.

جدول (4)

قيم (P_i) (احتمال الاستجابة)

i	P_i	i	P_i
1	0.134437	25	0.59988
2	0.166122	26	0.170191
3	0.1965	27	0.201312
4	0.242778	28	0.24872
5	0.35468	29	0.627134
6	0.419396	30	0.100425
7	0.118831	31	0.118804
8	0.146992	32	0.14681
9	0.387724	33	0.173663
10	0.189919	34	0.25379
11	0.405378	35	0.313517
12	0.906891	36	0.370749
13	0.112075	37	0.181588
14	0.132583	38	0.109807
15	0.28318	39	0.189874
16	0.148862	40	0.27746
17	0.17609	41	0.135831
18	0.217571	42	0.198538
19	0.257334	43	0.290109
20	0.317893	44	0.423734
21	0.375921	45	0.194977
22	0.464257	46	0.23062
23	0.548855	47	0.28491
24	0.281334	48	0.336938

تابع جدول (4)

i	P_i	i	P_i
49	0.717988	73	0.373012
50	0.203873	74	0.544615
51	0.352288	75	0.643762
52	0.311474	76	0.938749
53	0.537798	77	0.315748
54	0.663918	78	0.595248
55	0.222898	79	0.703547
56	0.263632	80	0.345177
57	0.325667	81	0.01072
58	0.38511	82	0.203825
59	0.562247	83	0.297827
60	0.131554	84	0.634929
61	0.155621	85	0.233011
62	0.19229	86	0.497114
63	0.227443		
64	0.332301		
65	0.410428		
66	0.237816		
67	0.507339		
68	0.363267		
69	0.260001		
70	0.379812		
71	0.25534		
72	0.301985		

4- ٤ الاستنتاجات

إن الهدف من هذا التحليل هو إعطاء إدارة المستشفى صورة واضحة عن مدى تأثير هذه العوامل على متغير الاستجابة (شفاء أو وفاة المريض) ويعد عامل مساعد للتشخيص، بحيث تتمكن هذه الإدارة من التخطيط الجيد لتوفير المستلزمات الضرورية والأطباء والكوادر التمريضية لرعاية هؤلاء المرضى. ومن خلال دراسة النتائج للبيانات المدروسة توصلنا الى الآتي :

١. **كانت قيم المعلمات الخاصة بمتغيرات الجنس ، والسكر المستخدمة بطريقة المربعات الصغرى سالبة ، وهذه مخالفة للواقع إذ من المعلوم إن طبيعة عامل السكر تعد من العوامل المساعدة في وفاة مرضى الذبحة الصدرية ، والذي أثبتته طريقة ($WMLE(w_1)$) .**

٢. **نلاحظ من خلال الجدول (4) أن احتمال الاستجابة الذي يمثل احتمال وفاة المريض تكون عالية (اكبر من 0.5) للمجموعات التي تكون فيها متغيرات الضغط ، والسكر، والعمر ضمن الفئات العليا والمبينة في الجدول (1) .**

الـمـلـخص

إن معظم الدراسات في موضوع أنموذج انحدار اللوجستك تأخذ طابعاً أكثر تقدماً في عملية التحليل الإحصائي الدقيق الذي يهدف إلى الحصول على مقدرات ذات مستوى عالٍ من الكفاءة .

وتأتي أهمية الحاجة إلى تطبيق الطرائق الحصينة عندما تكون بيانات الظاهرة المدروسة ملوثة ، اي وجود بضع مشاهدات تنحرف وبشكل ملحوظ عن المشاهدات الأخرى تدعى بالشواذ، والتي ينتج بسببها مقدرات غير جديرة بالثقة ومن ثمَّ فأن القرار سيكون خاطئاً بشأن المشكلة المراد دراستها .

ومن هنا جاء هدف هذا البحث في الوصول إلى أهم العوامل المؤثرة على مرضى الذبحة الصدرية من خلال استخدام طريقة ($WMLE(w_1)$) على بيانات حقيقية لمرضى الذبحة الصدرية تم أخذها من مستشفى ابن النفيس لغرض تقدير معلمات أنموذج انحدار اللوجستك ثنائي الاستجابة .

Abstract

Robust weighted Maximum Likelihood Estimates and their comparison with other methods of logistic model with practical application

Assis. Prof. Dr. Nazar M. Jawad — **University of Baghdad**

Sherien Ali — **University of Baghdad**

In most of studies about the logistic regression model take in it's nature more progness in the procedure of acurate statistical analysis which aim getting estimators of a high level of efficiency , and importance of it's necessity to applicate of the robust methods, appears when the data of the studied phenomenon are contaminated , it means some of the observations vareit clearly from other observations called outliers which may produce estimators untrustworthy and so the decision will be false about the problem that we are trying to study it .

From this point was the goal of this research in reaching robust estimators to estimate parameters of binary response logistic regression model through study estimators (method (WMLE) , OLS)Also the researcher specified a chapter for applying and using method

WMLE (w_1) on real data of angina pectoris patients which taking from Ibn-al- Nafees hospital for the purpose of estimate parameters of binary logistic regression model.

المصادر

1. Berkson. J. (1944) "Application Of The Logistic Function To Bioassay " JASA Vol .39 , PP . 357 -365.

2. Carroll, R.J. Pederson, S. (1993) . On Robustness In The Logistic Regression Model. J.R. Statist. Soc .B 55,693 – 706.

3. Friedman, J.H., Hastie, T., & Tibshirani, R. (2000). Additive Logistic Regression: A statistical View Of Boosting. Annals Of Statistics, 28, 337-374.

4. Gramer, J.S. (1991). The LOGIT Model: An Introduction For Economists, Edward Arnold, London.

5. Hosmer, D.W., Jr. And Lemeshow, S. Applied Logistic Regression, Wiley, (2000), ISBN: 0 – 471 – 35632 - 8.

6. Huber , P.J. (1984), Finite Sample Breakdown Of M- And P-Estimators , The Annals Of Statistics, 12 , 119 – 126.

7. Mueller, Ch.H. Neykov, N. (2003). Breakdown Points Of Trimmed Likelihood Estimators And Related Estimators In Generalized Linear Models J. Statist. Plann .Inference 116, 503 – 519.

8. Ricard, A. M., R. Douglas, M. And Victor, J.Y. (2007). "Robust Statistics", Theory And Methods.

9. Simeckova. M. (2005). "Maximum Weighted Likelihood Estimator In Logistic Regression ". Charles University, Faculty Of Mathematics And Physics, Part I, 144 – 148 .

10. Stahel, W.A. (1981), Breakdown Of Covariance Estimators, Research Report 31, Fachgruppe .Für Statistic. ETH .Zurich.

www.sfsu .edu /~efc/ classes/ logistic. htm

11. Vandev, D. Neykov, N. (1998). About Regression Estimators with High Breakdown Point Statistics 32, 111 -129.

12. Vieira, Marcel D.T. And Skinner, Chris J.(2006), "Estimating Models For Panel Survey Data Under Complex Sampling", Southampton ,UK , University Of Southampton , Southampton Statistical Sciences Research Institute , 30PP .(S3R) ,Methodology. Working Papers, M06/17.

13. Wang, X., Van Eeden, C., Zidek, J.V. (2004). Asymptotic Properties Of Maximum Weighted Likelihood Estimators. 199, 37-54. University Of British Columbia.

www.Elsevier Mathematics.com

14. www.arabs2day.ws/forums/index.8433