Use a bootstrap method in parametric regression models (*Logistic Model*)

استخدام اسلوب البوتستراب في نماذج الانحدار المعلمية (النموذج اللوجستي)

ا.م.د. شروق عبد الرضا السباح نور عباس عمران
جامعة كربلاء /كلية الادارة والاقتصاد

(بحث مستل من رسالة ماجستير)

المستخلص :

الصحة هي نعمة من نعم الله تعالى الكثيرة علينا ، والتي تمكن الانسان ان يحيا حياة طبيعية وتمكنه من التمتع بها ، فمنذ الازل دأب الانسان على تطوير العلاج والبحث عن الشفاء ، ومن بين الامور التي تعمل كل الحكومات وكل العاملين بالصحة والتنمية والمجتمع الدولي على تحسينها وتطويرها هي الرعاية الصحية والعناية بالمرضى .

لذا كان هذا البحث دراسة للعلاقة بين العناية المقدمة الى المريض خلال مدة الرقود في المستشفى وبين التدخين ، الجنس ، العمر والوزن عن طريق سحب عينة عشوائية بسيطة حجمها (182) شخصا مصابا بالجلطة الدماغية من مستشفى الحسيني في محافظة كربلاء المقدسة خلال الفترة(2014-2015) باستخدام انموذج الانحدار اللوجستي التقليدي والبوتسترابي ، وتحديد اي من تلك العوامل (التدخين ، الجنس ، العمر ، الوزن) الاكثر تأثيرا على فترة رقود المريض في المستشفى ، و توضيح اثر استخدام الاسلوب البوتسترابي على مقاييس الانموذج المعلمي اللوجستي.

ومن استنتاجات البحث ان استخدام اسلوب البوتستراب يعطي نتائج اكثر دقة .

Abstract :

Health is a blessing from God many we have , which enables man to live a normal life and being able to enjoy his life , so since time immemorial has repeatedly rights to develop and search for healing treatment , and among the things that are working on all governments and all workers to health and development, and the international community to improve The development is the health care and patient care.

So according to this research study of the relationship between the care provided to the patient during the period of lying in the hospital and between smoking, sex, age and weight by simple random sample size pull (182) infected with stroke people from the al-Husseini hospital in the holy city of Karbala during the period (2014- 2015) using a specimen traditional logistic regression and bootstrap , identifying any of those factors (smoking , sex , age , weight) the most impact on the patient's stay in hospital , and clarify the impact of the use of style Bootstrap on the form Teachers logistical measures and then compared with real data results with non-real (generated) in order to determine which one best use real data or non-real (generated).

And research findings that the use of style bootstrap gives more accurate results , and that the specimen logistic regression Bootstrap real data is the best specimen in the representation of the data.

المقدمة : *Introduction*

من المعلوم ان دراسة اي ظاهرة (اقتصادية ، اجتماعية ، علمية ،.....) يتطلب تحليل تلك الظاهرة وايجاد العوامل التي تؤثر فيها وترتبط بها ، وذلك من خلال ايجاد صيغة (او انموذج) تعبر عنها وتكون مدخل رئيسي لفهمها ، وهذا يعد من اهم اهداف نماذج الانحدار حيث يعرف تحليل الانحدار على انه اسلوب احصائي يستخدم لاستكشاف العلاقة بين متغير يعرف بالمتغير التابع ومتغير واحد او عدة متغيرات تعرف بالمتغيرات التوضيحية او التفسيرية ، وتوجد ثلاثة انواع من نماذج الانحدار التي تطورت مع مرور الزمن وهي النماذج المعلمية والنماذج اللامعلمية والنماذج شبه المعلمية ولكل منها طرق خاصة في التطبيق .

يعد انموذج الانحدار اللوجستي احد النماذج المعلمية النوعية التي يكون فيها المتغير التابع متغير ثنائي يأخذ قيمتين (واحد عند وقوع حدث معين و صفر عند عدم وقوع ذلك الحدث)، ويعتبر انموذج الانحدار اللوجستي الانموذج الاكثر شيوعا في معالجة البيانات الثنائية او النوعية ، اذ تعد الدالة اللوجستية من اكثر الدوال ملائمة لوصف البيانات التي يكون فيها متغير الاستجابة متغير ثنائي .
اما اسلوب البوتستراب الذي تم تطبيقه مع انموذج الانحدار المعلمي اقترحه العالم (*Efron*) عام (1979) وهو احد اساليب اعادة المعاينة بالإرجاع (اي اعادة وحدة المعاينة فقد المسحوبة قبل سحب الوحدة التالية لها) وهو نوع من طريقة مونتكالو حيث ان كلتا الطريقتين استخدمت بكثرة في الاحصاء التطبيقي .

مشكلة البحث : Research Problem

ان معيار تقدم الامم لا يبنى في ضوء قوتها الاقتصادية والعسكرية ولا المعمارية فقط ، بل في ضوء رفاهية شعوبها وما تقدمه من الخدمات الصحية والتعليمية وخدمات التغذية المقدمة الى مواطنيها ، ولقد عمل الانسان منذ الازل على تطوير وتوسيع الخدمات الصحية التي تعتبر احد الاركان الاساسية المهمة في تطوير الامم وتقدمها ، ولذا فأن الاهتمام بالخدمات والرعاية الصحية من الامور التي تعمل كل الحكومات والعاملين بالصحة والمجتمع الدولي على تحسينها وتطويرها .

هدف البحث : Aim Of The Research

تتحدد اهداف البحث بالتالي :
1 – تطبيق انموذج الانحدار اللوجستي التقليدي والبوتستراربي على البيانات من اجل دراسة العلاقة بين العناية المقدمة للمريض خلال فترة رقوده في المستشفى وبين التدخين ، الجنس ، العمر والوزن ، وتحديد اياً من تلك العوامل الاكثر تأثيرا على فترة بقاء المريض في المستشفى .
2 – توضيح اثر استخدام الاسلوب البوتستراربي على مقاييس الانموذج المعلمي اللوجستي .

منهجية البحث : Research Methodology

تم العمل على المنهج الاستقرائي و الذي يبدأ بملاحظة المشكلة ثم وضع الفروض لها وبعد ذلك اختبارها . وقد تم استخدام الاسلوب الاحصائي وفق هذا المنهج .

الجلطة الدماغية :Stroke

تعريف مختصر بالجلطة الدماغية [1] : *Manual Definition Of Stroke*

يتكون الدماغ من مجموعة كبيرة من الخلايا العصبية التي تقوم بالتحكم في جميع اجهزة الجسم ، وان هذه الخلايا غير قابلة للتجدد او التكاثر بشكل طبيعي ، حيث يولد الانسان وهو يمتلك عددا محددا من هذه الخلايا ، وهي مغذاة بالأكسجين والطاقة عن طريق شبكة من الاوعية الدموية ، وان هذه الخلايا تموت عند توقف وصول الدم اليها ، يؤدي فقدان عدد من هذه الخلايا الى فقدان الشخص المصاب درجة معينة من الوظائف العصبية ، وبالتالي فان الوظائف التي كانت تقوم بها هذه المنطقة من الخلايا تنعدم او تتأثر ، وهذا ما يسمى بالجلطة الدماغية .
وتسمى ايضا بالسكتة ومن الناحية الطبية تعرف بالحادث الوعائي الدماغي ، وهي "حدوث اضطراب مفاجئ في التروية الدموية لجزء من الدماغ يؤدي الى عجز في وظيفة عصبية او اكثر حركية او حسية او حاسية او استعرافية ".

اعراض وعلامات الجلطة الدماغية [2] : *Symptoms And Signs Of Stroke*

ان اهم العلامات المبكرة التي ينبغي الانتباه اليها هي :
1 – صعوبات في المشي : فان الشخص المصاب قد يشعر بدوخة ويتعثر وفي اغلب الاحيان يفقد توازنه او يفقد قدرة التنسيق بين الحواس او الحركة والكلام .
2 – صعوبات في النطق : حيث انه الاصابة بالجلطة الدماغية قد تؤدي الى حصول صعوبات في كلام الشخص المصاب او قد يفقد القدرة على ايجاد الكلمات المناسبة لما يحصل معه .
3 – شلل او اخدرار: في جانب واحد من الجسم اذ ان الشخص المصاب قد يفقد الاحساس او يشعر بشلل نصفي .
4 – صعوبات في الرؤية : اذا اصيب الشخص بالجلطة الدماغية فقد يعاني وبشكل مفاجئ من تشوش الرؤية وقد يفقد الرؤية للحظات قليلة او قد يعاني من الشفع (الرؤية المزدوجة) .
5 – الصداع : قد تسبب الاصابة بالجلطة الدماغية صداع شديد احيانا يكون مصحوبا بتشنج بتشنج في الرقبة ، والام في الوجه والآلام بين العينين ، بالإضافة الى ذلك فقد يكون مصحوبا بالتقيؤ المفاجئ او تغيرات بالحالة الادراكية .
تعتبر العلامة الاولى عند معظم الناس و التي تشير الى اصابة محتملة بالسكتة الدماغية هي النوبة الاقفارية العابرة وهي توقف مؤقت لوصول الدم الى جزء واحد من الدماغ ، وتمتلك النوبة الاقفارية العابرة نفس الاعراض والعلامات التي تمتلكها الجلطة الدماغية ، الا انها تستمر لفترة زمنية قصيرة من بضع دقائق الى اربع وعشرون ساعة بعدها تتلاشى وتزول بدون ان تترك اضرار مستديمة ، وقد يكون الشخص معرض للإصابة بأكثر من نوبة اقفارية عابرة واحدة ، وتختلف الاعراض والعلامات المصاحبة لها فقط تختلف او تكون متماثلة ، وان الشخص الذي

يصاب بها يكون معرض اكثر لخطر الاصابة بسكتة دماغية قوية ، وهو معرض اكثر للإصابة بها ممن لم يتعرض لها من قبل .

انموذج الانحدار :Regression Model

نماذج الانحدار المعلمية [3] : Parametric Regression Model

وهي من اقدم نماذج الانحدار واكثرها شيوعا ويمكن تعريفها على انها واحدة من الاساليب الاحصائية التي تستخدم لوصف وتقدير العلاقة بين متغير عشوائي معتمد ومتغيرات عشوائية توضيحية ولهذه النماذج انواع منها النماذج الكمية (*Quantitative Models*) بأنواعها الخطية وغير الخطية والنماذج النوعية (*Qualtifiction Models*)التي يكون فيها المتغير التابع متغير ثنائي القيم ، وتتميز هذه النماذج بان جميع الحالات التي يتم دراستها تفترض ان توزيعها توزيعا طبيعيا وقابلة للقياس ، وان دالة الانحدار تتحدد بالمعلمات التي لا يمكن تغيرها الا بتغير عدد المتغيرات الداخلة في الدراسة ، وان معامل التحديد (R^2) الذي يعتبر مقياس لدرجة تمثيل الانموذج للبيانات قيد الدراسة لا يمكن تحسينه الا من خلال تغيير عدد المتغيرات الداخلة بالدراسة او بإضافة بيانات جديدة .

الانحدار اللوجستي [4][5][6][7] : Logistic Regression

هو طريقة احصائية لتحليل مجموعة من البيانات تحتوي على متغير واحد او اكثر من المتغيرات المستقلة المراد معرفة تأثيرها على المتغير التابع ، ويتميز المتغير التابع بأنه يحتوي على قيم ثنائية (نتيجتان محتملتان وحيدتان) .

بمعنى اخر فأن المتغير المعتمد يحتوي على قيم اسمية مثل (فشل ، نجاح) او (موت حياة) الى اخره .وهذا ما يسمى بالانحدار اللوجستي الثنائي(binary Logistic Regression).

وتوجد انواع عديدة للانحدار اللوجستي الا ان الانحدار اللوجستي الثنائي يعد اكثرها شيوعا ، حيث يوجد الانحدار اللوجستي المتعدد (Multinomial Logistic Regression) وهذا يستخدم عندما يكون المتغير التابع متغير اسمي متعدد الاستجابة ، والنوع الثالث يسمى الانحدار اللوجستي الرتبوي (الترتيبي) (Ordinal Logistic Regression) وهذا يستخدم عندما يكون المتغير التابع رتبوي .

ويتميز الانحدار اللوجستي بان العلاقة بين المتغير التابع والمتغيرات المستقلة غير خطية ويمكن تحويلها بسهولة الى الشكل الخطي باستخدام ما يعرف بتحويلة اللوجت (Logit Transformation) والتي تكون بالشكل التالي :-

$$Logit\ Transformation = ln\frac{P(x)}{1-P(x)} = \beta_0 + \beta X \qquad \dots\dots\dots (1)$$

اذ ان :

$$P(x) = \frac{e^{\beta_0+\beta X}}{1+e^{\beta_0+\beta X}} \qquad \dots\dots\dots (2)$$

$$-\infty < X < \infty$$

β_0 : تمثل الحد الثابت لأنموذج الانحدار اللوجستي .

β : تمثل معلمات انموذج الانحدار اللوجستي المراد تقديرها .

X : يمثل المتغيرات التوضيحية او التفسيرية للنموذج .

ويوجد فرقين مهمين بين الانحدار الخطي والانحدار اللوجستي ، الاول يتعلق بطبيعة العلاقة بين المتغير التابع والمتغيرات المستقلة ، اذ ان اي معادلة انحدار تكون بالشكل التالي :

$$\mathrm{E}(Y/x) = \beta_0 + \beta \mathrm{X} \qquad \dots\dots\dots (3)$$

اذ ان :

$E(Y/x)$: يمثل المتوسط الشرطي .

Y : ترمز الى المتغير التابع .

X : تمثل المتغيرات المستقلة .

β وβ_0 : تمثل معاملات الانحدار .

عندما يأخذ المتغير التابع قيم مستمرة فان { $E(Y/x)$ } يأخذ اي قيم لـ X ويكون ضمن المدى

$$-\infty < E(Y/x) < \infty$$

اما عندما يأخذ المتغير التابع قيم اسمية فأن { $E(Y/x)$ } تكون ضمن المدى

$$0 \le E(Y/x) \le 1$$

في هذه الحالة التغير في { $E(Y/x)$ } لكل وحدة تتغير فيها (X) يصبح اصغر بتقدم تدريجي بينما يقترب { $E(Y/x)$ } بشكل تدريجي من (1 او 0) ، ففي هذه الحالة يستخدم انموذج الانحدار اللوجستي الذي يكون فيه { $E(Y/x)$ } مساوي لـ { $P(x)$ } و تكون صيغتها الرياضية كما عبرنا عنها في المعادلة رقم (2).

اما الفرق الثاني فهو يتعلق بالتوزيع الشرطي للمتغير المعتمد ، ففي الانحدار الخطي نفترض بان المشاهدات للمتغير المعتمد قد تكون بالشكل التالي :-

$$Y = E(Y/x) + \varepsilon \qquad \dots\dots\dots (4)$$

حيث انه من المعلوم ان (ε) تمثل الخطأ وهي تظهر انحراف المشاهدات عن المتوسط الشرطي ، وهذه الكمية يجب ان تتوزع توزيعا طبيعيا بمتوسط مساوي للصفر وتباين ثابت لكل مستويات المتغيرات المستقلة ، وعليه يكون التوزيع الشرطي للمتغير المعتمد بوجود المتغيرات المستقلة هو التوزيع الطبيعي بمتوسط { $E(Y/x)$ } وتباين ثابت .

وفي الحالة التي يكون فيها المتغير المعتمد ذو قيم ثنائية فأننا نفترض بان المشاهدات للمتغير المعتمد قد تكون بالشكل التالي :-

$$Y = P(x) + \varepsilon \quad \dots\dots\dots\dots (5)$$

وعليه فان (ε) يكون لها قيمتين احتماليتين ، فاذا كانت ($Y=1$) فان $\varepsilon = 1 - P(x)$ مع احتمال $P(x)$ ، واذا كانت ($Y = 0$) فان { $\varepsilon = -P(x)$ } مع احتمال ($1 - \mathrm{P}(x)$) ، اي ان (ε) تتوزع بمتوسط مساوي للصفر وتباين ($P(x)[1 - P(x)]$) وهذا يعني بان التوزيع الشرطي للمتغير المعتمد هو توزيع (*Binomial*) مع احتمال شرطي $P(x)$.

طرائق تقدير معلمات انموذج الانحدار اللوجستي :

Methods For Estimating The Logistic Regression Model Parameters

توجد طرائق عدة لتقدير معلمات الانحدار اللوجستي ومنها :

طريقة الامكان الاعظم [4][6][8]: ***Maximum Likelihood Method***

وهي الطريقة الاكثر شيوعا وهي "طريقة تكرارية (*Iterative*) تعتمد على تكرار العمليات الحسابية عدة مرات حتى يتم الوصول الى افضل تقديرات للمعلمات "، وهي تعتبر من اكثر الطرق ملائمة لكافة النماذج الخطية وغير الخطية .

و بما ان (y_i) يتبع توزيع ثنائي الحدين فان الامكان الاعظم تصبح بالشكل التالي :

$$L(\beta, y) = \prod_{i=1}^{N} \frac{n!}{y!\,(n! - y!)}\, P(x)^{y_i}\, (1 - P(x))^{n_i - y_i} \quad \dots\dots\dots\dots (6)$$

و بتبسيط المعادلة اعلاه تصبح :

$$L(\beta, y) = \prod_{i=1}^{N} \left(\frac{P(x)}{1 - P(x)}\right)^{y_i}\, (1 - P(x))^{n_i} \quad \dots\dots\dots\dots (7)$$

و بما ان :

$$\log\left(\frac{P(x)}{1 - P(x)}\right) = \beta_0 + \beta x \quad \dots\dots\dots\dots (8)$$

و بأخذ (e) للمعادلة (8) تصبح :

$$\frac{P(x)}{1 - P(x)} = e^{\beta_0 + \beta x} \quad \dots\dots\dots\dots (9)$$

وبالتعويض في المعادلة (7) تصبح :

$$L(\beta, y) = \prod_{i=1}^{N} \left(e^{\beta_0 + \beta x}\right)^{y_i} \left(1 - \frac{e^{\beta_0 + \beta x}}{1 + e^{\beta_0 + \beta x}}\right)^{n_i} \quad \dots\dots\dots\dots (10)$$

وبأخذ (Log) للمعادلة اعلاه تصبح :

$$L(\beta) = \sum_{i=1}^{N} y_i(\beta_0 + \beta x) - n_i \log\left(1 + e^{\beta_0 + \beta x}\right) \quad \dots\dots\dots (11)$$

ومن اجل الحصول على قيم للمعلمات (β) نقوم باشتقاق المعادلة اعلاه نسبتا للمعلمة المراد تقديرها ومن ثم نساويها للصفر لنحصل على مجموعة من المعادلات التي لا نستطيع حلها الا عن طريق خوارزمية تكرارية تعرف بخوارزمية المربعات الصغرى التكرارية الموزونة .

طريقة تصغير مربع كاي[9] :- ***Minimum Chi – Square Method***

واحدة من طرق تقدير معلمات انموذج الانحدار اللوجستي التي تعتمد على تصغير مجموع (x^2) وجعله اقل ما يمكن اذ ان :-

$$x^2 = C(\beta) = \sum_{i=1}^{n} \frac{(y_i - E_i)^2}{S_i} \quad \dots\dots\dots\dots (12)$$

اذ ان :

y_i : تمثل القيمة المشاهدة للمتغير y عند المستوى i

E_i : تمثل القيمة المتوقعة للمتغير y عند المستوى i

S_i : يمثل التباين للمتغير y عند المستوى i

وبما ان المتغير y يتبع توزيع ثنائي الحدين (*Binomial*) وعليه فأن

$$E_i = np(x) \quad \dots\dots\dots\dots (13)$$

$$S_i = np(x)(1 - p(x)) \quad \dots\dots\dots\dots (14)$$

اذ ان :

n : تمثل حجم العينة .

$p(x)$: كما ذكرناها سابقا تمثل احتمال النجاح وصيغتها حسب المعادلة رقم (*2*) .

وبالتعويض عن قيم E_i و S_i في المعادلة (*12*) تصبح بالشكل التالي :

$$C(\beta) = \sum_{i=1}^{n} \frac{(y_i - np(x))^2}{np(x)(1 - p(x))} \quad \dots\dots\dots\dots (15)$$

وبالتبسيط تصبح :-

$$C(\beta) = \sum_{i=1}^{n} \frac{1}{np(x)(1 - p(x))} \{y_i^2 - 2nyp(x) + n^2 p(x)^2\} \quad \dots\dots (16)$$

وبالتعويض عن قيمة ($p(x)$) في المعادلة اعلاه ومن ثم تطبيق طريقة نيوتن – رافسون نحصل على التقديرات المطلوبة لمعلمات انموذج الانحدار اللوجستي .

معايير حسن المطابقة : ***Criteria Of Goodness Of Fit***

تقييم القوة التفسيرية للأنموذج : **Evaluate The Explanatory Power Of The Model**

من المعلوم في الانحدار الخطي معامل التحديد (R^2) هو مقياس يوضح نسبة التغير في المتغير التابع (Y) الذي سببها التغير في المتغير المستقل (X) اما في الانحدار اللوجستي فان هناك العديد من المقاييس لحساب معامل التحديد (R^2) اقترحت لقياس حسن المطابقة وتسمى احيانا شبه معامل التحديد ومنها :-

1 – مؤشر الامكان (R_L^2)[10]: ***The Likelihood Index (R_L^2)***

وهذا المؤشر يستخدم مع بيانات البقاء ، وهو يحسب بسهولة من الانحرافات للنموذجين ، وحسب الصيغة التالية :-

$$R_L^2 = \frac{D_0 - D_p}{D_0} \quad \dots\dots\dots (17)$$

اذ ان :

D_0 : تمثل انحراف الانموذج عندما يحتوي الثابت فقط .

D_p : تمثل انحرف الانموذج عندما يحتوي جميع المتغيرات .

وان قيمته محصورة بين الصفر والواحد ، فهو يساوي صفر اذا كانت المتغيرات لا تؤدي الى تحسن الانموذج عن الانموذج الاولي ، ويساوي واحد عندما يكون الانموذج ملائم جدا .

2 – مؤشر *Mcfadden's Pseudo* [11]: ***Mcfadden's Pseudo R Square***

وهذا المؤشر اقترحه العالم (*Mcfadden*) في عام (*1973*) للمقارنة بين نموذجين احدهما لا يحتوي على المتغيرات التوضيحية والاخر يحتوي عليها ، وصيغته :-

$$R^2_{Mcfadden} = 1 - \frac{\log(L_{null})}{\log(L_c)} \quad \dots\dots\dots (18)$$

اذ ان :

L_c : تمثل قيمة الامكان الاعظم عندما الانموذج يحتوي على جميع المتغيرات .

L_{null} : تمثل قيمة الامكان الاعظم عندما الانموذج يحتوي على الثابت فقط .

واذا كانت قيم المعلمات تميل نحو الصفر فان قيمة ($R^2_{Mcfadden}$) هي ايضا تميل نحو الصفر، وفي عام (*1983*) طور هذا المؤشر من قبل العالم (*Maddala*) وهذا المؤشر يمكن تطبيقه على أي انموذج تقدر معلماته بطريقة الامكان الاعظم ، وصيغته :

$$R^2_{Maddala} = 1 - \left(\frac{L_{null}}{L_c}\right)^{2/n} \quad \dots\dots\dots\dots (19)$$

اذ ان :

L_c : تمثل قيمة الامكان الاعظم عندما الانموذج يحتوي على جميع المتغيرات .

L_{null} : تمثل قيمة الامكان الاعظم عندما الانموذج يحتوي على الثابت فقط .

n : تمثل حجم العينة .

واذا عبرنا عن الصيغة اعلاه باستخدام نسبة الامكان الاعظم فإنها تساوي :-

$$R^2_{Maddala} = 1 - \ e^{-\theta/n} \quad \dots\dots\dots\dots (20)$$

اذ ان :

$$\theta = -2\,log\left(L_{null}/L_c\right) \quad \dots\dots\dots\dots (21)$$

اذ ان :

θ : تمثل نسبة الامكان الاعظم .

3 – مؤشر *Cox and Snell* [10]:

وصيغته :

$$R^2_{Cox\ \&\ snell} = 1 - \left[\frac{L_{null}}{L_c}\right]^{2/n} \quad \dots\dots\dots\dots (22)$$

اذ ان :

L_c : تمثل قيمة الامكان الاعظم عندما يحتوي الانموذج على كل المتغيرات .

L_{null} : تمثل قيمة الامكان الاعظم عندما يحتوي الانموذج على الثابت فقط .

n : تمثل حجم العينة .

وهو مؤشر صعب لان أكبر قيمة له هي الواحد ، او بالأحرى أعلى قيمة له (0.75) عندما تكون احتمالات النجاح للعينة (0.50) .

4 – مؤشر *Nagelkerke* [10]:

ويعتبر هذا المؤشر مصحح او معدل لـ ($R^2_{Cox\ \&\ Snell}$) وهو يعتبر مقياس لحسن المطابقة العامة وصيغته :

$$R^2_{Nag} = R^2_{COx\ \&\ Snell}/R^2_{Max} \quad \dots\dots\dots\dots (23)$$

اذ ان :

$$R^2_{Max} = 1 - (L_{null})^{2/N} \quad \dots\dots\dots\dots (24)$$

وهذا المؤشر يكون دائما اكبر من ($R^2_{Cox\ \&\ Snell}$) ، لان في الحقيقة مصحح له وان اعلى قيمة لهما هي الواحد ، وبشكل عام (R^2) في الانحدار اللوجستي تعتبر مقياس للنماذج الجيدة ، وتكون قيمها اقل من (R^2) في النماذج الخطية المقدرة بطريقة المربعات الصغرى الاعتيادية وذلك لأنها تعتمد على المدى وتوزيع المتغيرات المستقلة وايضا لان المتغيرات الثنائية تميل الى ان يكون (R^2) منخفض حتى في علاقة الانحدار المثالية . ولمعرفة نسبة التغير في المتغير التابع التي سببها التغير في المتغير المستقل استخدمنا المؤشرين($R^2_{Cox\ \&\ Snell}$ و $R^2_{Nagelkerke}$).

اختبار معنوية المعالم المقدرة[12] :

Test The Significantly Of Estimated Parameters

في الانحدار الخطي يستخدم اختبار (*t*) لتقييم معنوية تأثير المتغيرات المستقلة في المتغير التابع ، اما في انموذج الانحدار اللوجستي فتستخدم احصاءة والد (*Wald Statistic*) لتقيم معنوية المعالم وتحسب حسب المعادلة التالية :

$$W^2 = \left[\frac{\widehat{\beta}}{S.E_{\widehat{\beta}}}\right]^2 \quad \dots\dots\dots (25)$$

اذ ان :

$S.E_{\widehat{\beta}}$: يمثل الانحراف المعياري لمعامل الانحدار اللوجستي للمتغير X.

وان احصاءة والد تتبع توزيع (x^2) بدرجة حرية مساوية للواحد ($df = 1$) واذا تم احتساب جذر هذه القيمة فان المعادلة (*25*) تصبح بالشكل التالي :

$$w = \left[\frac{\hat{\beta}}{S.E_{\hat{\beta}}}\right] \quad \ldots\ldots\ldots\ldots (26)$$

وان هذه الاحصاءة تتبع التوزيع الطبيعي القياسي N(0,1) .

اختبار هوسمر – ليميشو [6][13]: *The Hosmer- Lemeshow Test*

وهذا الاختبار مشابه لاختبار(x^2) لحسن المطابقة حيث يستخدم لتقيم حسن المطابقة للأنموذج وهو يسمح بأي عدد من المتغيرات التوضيحية سواء كانت مستمرة او متقطعة ، وان هذه الاحصاءة تتوزع توزيع (x^2) بدرجة حرية (*10-2*) واذا كانت قيمتها اكبر من مستوى المعنوية فان هذا يؤكد جودة التوفيق للأنموذج بالكامل ، وهذا الاختبار يعتمد على تجميع الحالات على اساس قيم الاحتمالات المتوقعة ويتم التجميع بطريقتين :-

1 – تجميع الحالات اعتمادا على المئينات للاحتمالات المتوقعة .

2 – تجميع الحالات اعتمادا على القيم الثابتة للاحتمالات المتوقعة .

وحسب الطريقة الاولى يتم حساب احصاءة الاختبار بعد ترتيب الحالات (m) تصاعديا حسب القيم المتوقعة للاحتمالات ، وان القيم المشاهدة والمتوقعة للحالات يتم تجميعها وفقا لقيم المتغير التابع Y (0,1) ، وبعد ذلك يتم توزيعها الى مجاميع (K) وان عدد الحالات في كل مجموعة يكون $(m/_K)$ ويجب ان تحتوي المجموعة الاولى على الحالات التي تمتلك اقل قيمة للاحتمالات المتوقعة $(m_1 = m/K)$ ، واخر مجموعة يجب ان تحتوي على الحالات التي تمتلك اعلى القيم للاحتمالات المتوقعة $(m_k = m/K)$ ،وهذه الاحصاءة يرمز لها بالرمز C ويتم احتسابها وفقا للصيغة التالية :-

$$C = \sum_{i=1}^{k} \frac{(O_i - m_i'\overline{\pi_i})^2}{m_i'\overline{\pi_i}(1-\overline{\pi_i})} \quad \ldots\ldots\ldots\ldots\ldots (27)$$

اذ ان :

$$O_i = \sum_{j=1}^{m_i'} Y_j \quad \ldots\ldots\ldots\ldots\ldots (28)$$

اي انها تمثل عدد الاستجابات عندما (*Y=1*) و(m_i') تمثل العدد الكلي للحالات في المجموعة (i) وان :

$$\overline{\pi_i} = \sum_{j=1}^{m_i'} \frac{n_j\widehat{\pi_j}}{m_i'} \quad \ldots\ldots\ldots\ldots\ldots (29)$$

وهو يمثل المعدل للاحتمالات المتوقعة للمجموعة(i) .

البوتستراب [14][15]: *Bootstrap*

طور العالم (*Efron*) في عام (*1979*) طريقة جدية لإعادة المعاينة بالإرجاع سميت بطريقة البوتستراب (*Bootstrap*) و هي اعادة عينة تشتمل على (n) من العناصر المسحوبة بالإرجاع بشكل عشوائي من (N) من البيانات الاصلية .

و هي نوع من طريقة المونتكالو ، و المونتكالو هي احدى طرق الاحصاءات الاستنتاجية التي تتضمن تقدير معلمات المجتمع او حساب فترات الثقة او اختبار الفرضيات او غيرها ، حيث ان الاحصاءات الاستنتاجية تستعمل للاستدلال حول المجتمع ولقياس ثقة هذه الاستنتاجات تستخدم معلومات مكتسبة من العينة العشوائية لذلك المجتمع .

وان طريقة البوتستراب تزودنا بتقديرات للخطأ المعياري والتحيز عن طريق (B) من العينات البوتسترابية المأخوذة بالإرجاع من العينة الاصلية.

والبوتستراب يشبه اسلوب الجاكنايف (*Jackknife*) الذي هو ايضا احد طرق اعادة المعاينة بالإرجاع الذي يستخدم لتقدير التحيز والاخطاء المعيارية عند فشل الطرق القياسية، وان طريقة الجاكنايف مشابهة لطريقة (*Cross-Validation*) في حذف واحدة من المشاهدات حيث انه من العينة ($X = (x_1, x_2, \ldots\ldots, x_n)$) نحذف واحدة من المشاهدات لنحصل على عينة الجاكنايف وبتكرار هذه العملية يكون لدينا ($x_1, x_{i-1}, x_{i+1}, \ldots\ldots, x_n$) التي يتبين من خلالها ان (i-*th*) من عينات الجاكنايف التي هي العينة الاصلية التي لها (i-*th*) من نقاط البيانات التي حذفت ، ثم نقوم بحساب الاحصاءة المطلوبة لكل عينة من عينات الجاكنايف ، ويمكن اعتبار الجاكنايف الطريقة الاكثر توافقا مع طريقة البوتستراب .

والخطوات المتبعة في استخدام طريقة البوتستراب هي :

1 – من العينة الحقيقية ($X = (x_1, x_2, \ldots\ldots\ldots, x_n)$) نسحب عينات عشوائية جزئية بالإرجاع وبحجم معين نرمز لها بـ ($X^* = (x_1^*, x_2^*, \ldots\ldots\ldots, x_m^*)$) عددها يساوي ($m$) والذي يكون على الاقل (*1000*) مرة.

2 – ايجاد تقدير لمعالمات انموذج الانحدار اللوجستي بطريقة الامكان الاعظم ولكل عينة من عينات البوتستراب .

التطبيق العملي : Practical Application

وصف البيانات : Description Of Data

تم الحصول على بيانات البحث من مستشفى الحسيني في محافظة كربلاء المقدسة تمثلت بـ (182) شخصا مصابا بالجلطة الدماغية .

متغيرات البحث : Research Variables

y_i : يمثل العناية المقدمة للمريض متمثلة مدة الرقود في المستشفى للمرضى المصابين بالجلطة الدماغية ، وهو يمثل المتغير التابع لأنموذج الانحدار اللوجستي التقليدي و البوتسترابي ويتمثل بالقيم التالية :-

0 : عندما تكون مدة الرقود في المستشفى للشخص المصاب بالجلطة الدماغية اقل من ستة ايام .

1 : عندما تكون فترة الرقود في المستشفى للشخص المصاب بالجلطة الدماغية ستة ايام فأكثر.

المتغيرات التالية تمثل المتغيرات التوضيحية او المستقلة المراد قياس تأثيرها على المتغير التابع :

X_1 : يمثل التدخين وصنف الى (1)مدخن و (2) غير مدخن .

X_2 : يمثل جنس الشخص وصنف الى (1) ذكر و (2) انثى .

X_3 : يمثل عمر الشخص وصنف الى الفئات العمرية التالية :-

الفئة العمرية (30-40) يرمز لها بالرمز (1).

الفئة العمرية (40-50) يرمز لها بالرمز (2).

الفئة العمرية (50-60) يرمز لها بالرمز (3).

الفئة العمرية (60-70) يرمز لها بالرمز (4).

الفئة العمرية (70-80) يرمز لها بالرمز (5).

X_4 : يمثل وزن الشخص وتم تصنيفه الى الفئات التالية :-

اذا كان وزن الشخص (50-60) كيلو غرام يرمز له بالرمز (1).

اذا كان وزن الشخص (60-70) كيلو غرام يرمز له بالرمز (2).

اذا كان وزن الشخص (70-80) كيلو غرام يرمز له بالرمز (3).

اذا كان وزن الشخص (80-90) كيلو غرام يرمز له بالرمز (4).

اذا كان وزن الشخص (90-100) كيلو غرام يرمز له بالرمز (5).

وبتطبيق برنامج (SPSS) على البيانات حصلنا على النتائج الخاصة بنموذج الانحدار اللوجستي التقليدي والبوتسترابي :

نتائج تطبيق انموذج الانحدار اللوجستي التقليدي :

The Results Of The Application Of The Logistic Regression Model

يستخدم الانحدار اللوجستي لإظهار العلاقة بين المتغير التابع الثنائي وبين واحد او اكثر من المتغيرات التوضيحية التي يمكن ان تكون منفصلة او مستمرة ، وفي هذا البحث استخدمنا انموذج الانحدار اللوجستي لإظهار العلاقة بين المتغير التابع الذي يمثل مدة الرقود في المستشفى للمصابين والمتغيرات المستقلة التي تم تحديدها سابقا .

ان تأثير المتغيرات التوضيحية على المتغير التابع يقاس من خلال معاملات الانحدار β، ولمعرفة فيما اذا كان للمتغيرات التوضيحية لها تأثير على المتغير التابع نفترض الفرضيات التالية :

فرضية العدم : لا يوجد تأثير للمتغيرات التوضيحية علي المتغير التابع $H_0: \beta = 0$

الفرضية البديلة : يوجد تأثير للمتغيرات التوضيحية على المتغير التابع $H_1: \beta \neq 0$

ولاختبار الفرضيات اعلاه تستخدم احصاءة (*Chi-Square*) اذ بينت النتائج ان قيمتها الاحتمالية عند مستوى معنوية ($\alpha = 0.05$) ودرجة حرية (*4*) مساوية لعدد المتغيرات التوضيحية تساوي (*P-Value = 0.00*) وهي اقل من مستوى المعنوية ، لذا نرفض فرضية العدم و لا نرفض الفرضية البديلة ، مما يؤكد كفاءة ومعنوية الانموذج ككل .

ومن اجل معرفة مدى ملائمة انموذج الانحدار اللوجستي للبيانات قيد الدراسة نستخدم معامل التحديد اذ اظهرت النتائج بان قيمة ($R^2_{Cox\ \&\ Snell}$) تساوي (*0.503*) وهي تشير الى ان (*50.3%*) من التغير الحاصل في المتغير التابع يتم تفسيره من خلال انموذج الانحدار اللوجستي ، وكذلك قيمة ($R^2_{Nagelkerke}$) تساوي (*0.671*) وهي تشير ايضا الى ان (*67.1%*) من التغير الحاصل في المتغير التابع يتم تفسيره من خلال انموذج الانحدار اللوجستي ، وايضا بينت النتائج بان قيمة (*-2 Log Likelihood*) كانت تساوي (*125.014*) وهي تعني ان (*125.014*) من قرارات الانموذج نتوقع بأنها اخذت بشكل خاطئ .

ولاختبار فيما اذا كان الانموذج يمثل البيانات بشكل جيد ام لا نفترض الفرضيتان التاليتان :

فرضية العدم : لا يوجد فرق معنوي بين القيمة المشاهدة والقيمة الحقيقية ، او ان الانموذج يمثل البيانات بشكل جيد .

الفرضية البديلة : يوجد فرق معنوي بين القيمة المشاهدة والحقيقية ، او ان الانموذج لا يمثل البيانات بشكل جيد .

ولاختبار الفرضيات اعلاه يستخدم اختبار(*Hosmer And Lemeshow*) حيث كانت القيمة الاحتمالية لاحصاءة (*Chi-Square*) تساوي (*0.398*) وهي اكبر من مستوى المعنوية ($\alpha = 0.05$) ، لذلك نرفض الفرضية البديلة ولا نرفض فرضية العدم اي ان الانموذج يمثل البيانات بشكل جيد ، ونتائج الجدول(1) تؤكد ذلك .

جدول (1) القيم المشاهدة و المتوقعة لفترة رقود المريض (المصاب)

الخطـــــــوات	زمن الرقود =اقل من 6ايام		زمن الرقود= 6ايام فاكثر		المجمـــــــوع
	المشاهدة	المتوقعة	المشاهدة	المتوقعة	
1	20	20.161	1	0.839	21
2	16	17.008	2	0.992	18
3	16	15.805	1	1.195	17
4	17	15.673	0	1.327	17
5	15	16.196	3	1.804	18
6	16	15.156	1	1.844	17
7	17	16.430	2	2.570	19
8	15	16.993	5	3.007	20
9	18	15.253	1	3.747	19
10	11	11.877	5	4.123	16

جدول (2) نتائج انموذج الانحدار اللوجستي التقليدي

المتغيرات	معاملات الانحدار اللوجستي β	الاخطـاء المعيارية *S.E*	احصاءة *Wald*	درجات الحرية *Df*	القيمـــة الاحتماليـة *Sig.*	القيمة المتوقعة *Exp(B)*	فـتـــــرات الـثـقــة	
							الدنيا *Lower*	العليا *Upper*
التدخين	-0.108	0.420	0.066	1	0.798	0.898	0.394	2.045
الجنس	0.389	0.403	0.930	1	0.335	1.475	0.669	3.251
العمر	-0.370	0.226	2.678	1	0.102	0.690	0.443	1.076
الوزن	-0.434	0.209	4.306	1	0.038	0.648	0.430	0.976

يظهر في الجدول اعلاه المتغيرات التوضيحية وتأثيرها المعنوي على المتغير التابع ويتبين لنا التالي :

من خلال ملاحظة الجدول (2) نجد ان متغير الوزن X_4 له تأثير معنوي على المتغير التابع ، اذ كانت القيمة الاحتمالية لاحصاءة (*Wald*) تساوي (*P-Value =0.038*) وهي اصغر من مستوى المعنوية ، و لا تأثير لبقية المتغيرات .

ويظهر ايضا في الجدول (2) قيم ($\exp(\beta)$) التي تمثل نسبة الارجحية (*Add Ratio*) وهي تشير الى مقدار التغير الحاصل في نسبة ارجحية وقوع الحدث (فترة بقاء المريض المصاب بالجلطة الدماغية على قيد الحياة في المستشفى اكثر من ستة ايام) عند حدوث تغير في قيمة المتغير المستقل (x_i) المرتبط بالمعلمة (β) .

وتزداد نسبة ارجحية وقوع الحدث (مدة رقود المريض في المستشفى اكثر من ستة ايام) اذا كانت قيمة ($\exp(\beta)$) اكبر من الواحد ، بينما تنخفض هذه النسبة مع كل زيادة في قيمة المتغير المستقل (x_i) عندما تكون قيمة ($\exp(\beta)$) اقل من الواحد .

من ذلك نجد ان نسبة ارجحية بقاء المريض في المستشفى اكثر من ستة أيام متحققة لكل المتغيرات (التدخين ، الجنس ، العمر ، الوزن) فبالنسبة :

1 – لمتغير التدخين تشير الى فترة رقود المريض اعلى بمقدار (2.4547) مما هو عليه عند غير المدخنين .

2 – و لدى الرجال اعلى من النساء بمقدار (4.371) .

3 – وان الزيادة في عمر المريض بمقدار وحدة واحدة تؤدي الى زيادة بقاءه في المستشفى بمقدار (1.9937) مرة .

4 – اما بالنسبة للعمر فان زيادة وزن المريض ودة واحدة تؤدي الى زيادة فترة رقوده في المستشفى فترة أطول بمقدار (1.9117) مرة .

و ان معادلة الانحدار اللوجستي التقليدي تكون بالشكل التالي :

$$\hat{Y} = \hat{g}(X) = (-0.108)X_1 + (0.389)X_2 + (-0.370)X_3 + (-0.434)X_4 \dots\dots\dots\dots\dots\dots (30)$$

نتائج تطبيق انموذج الانحدار اللوجستي البوتستراپي :

Results Of Applying Bootstrap Logistic Regression Model

عند استخدام الاسلوب البوتستراپي على بيانات البحث حصلنا على نفس النتائج التي حصلنا عليها عند استخدام الاسلوب التقليدي لـ (احصاءة (Chi-Square)) لاختبار معنوية الانموذج وقيم ($R^2_{Cox\ \&\ snell}$ و $R^2_{Nagelkerke}$) وايضا قيم اختبار (Hosmer And Lemeshow)) وباختلاف النتائج التي تخص المعاملات وكما في الجدول التالي :

جدول (3) نتائج انموذج الانحدار اللوجستي البوتستراپي

المتغيـــرات	معاملات الانحدار اللوجستي البوتستراپي β	Bootstrap[a]			
		الخطأ المعياري Std.Error	القيمة الاحتمالية Sig.(2-tailed)	فترات الثقة 95 %	
				الدنيا Lower	العليا Upper
التدخين	-0.108	0.400	0.703	-1.089	0.668
الجنس	0.389	0.374	0.257	-0.536	1.111
العمر	-0.370	0.194	0.040	-0.744	-0.042
الوزن	-0.434	0.221	0.029	-0.865	0.078

من ملاحظة الجدول (3) نجد ان :

كل من المتغيرين (العمر و الوزن) له تأثير على المتغير التابع لان القيمة الاحتمالية لكلاهما اصغر من ($\alpha = 0.05$) (مستوى المعنوية) ، اذ ان القيمة الاحتمالية لمتغير العمر تساوي (P-Value = 0.040) ، وان القيمة الاحتمالية لمتغير الوزن تساوي (P-Value =0.029) ، وبالتالي فان معادلة الانحدار اللوجستي البوتستراپي تكون بالشكل التالي :

$$\hat{Y} = \hat{g}(X) = (-0.108)X_1 + (0.389)X_2 + (-0.370)X_3 + (-0.434)X_4 \dots\dots\dots\dots\dots\dots (31)$$

المقارنة بين الاسلوب البوتستراپي و التقليدي :

The Comparison Between The Traditional Style And Bootstrap

اذا كانت المعنوية او القيمة الاحتمالية (*P-Value*) تمثل عدد او رقم يستخدم لتقييم المقاييس الاحصائية ، واذا كان الخطأ المعياري هو مقياس للخطأ المتوقع في درجة الاختبار وكلما قلت قيمته كلما زادت الدقة في درجات الاختبار ، وعلية سنقوم بمقارنة قيم الاخطاء المعيارية والمعنوية (*P-Value*) لنموذج الانحدار اللوجستي التقليدي مع قيمها لنموذج الانحدار اللوجستي البوتستراپي من اجل تحقيق الهدف الثاني للدراسة الذي يتمثل بتوضيح اثر الاسلوب البوتستراپي على مقاييس النموذج والنتائج يبينها الجدول التالي :

جدول (4) القيم المعنوية والاخطاء المعيارية

المتغيـرات	انموذج الانحدار اللوجستي التقليدي P-Value	انموذج الانحدار اللوجستي البوتستراپي P-Value	الخطأ المعياري اللوجستي التقليدي Std.Error	الخطأ المعياري اللوجستي البوتستراپي Std.Error
التدخين	0.798	0.703	0.420	0.400
الجنس	0.335	0.257	0.403	0.374
العمر	0.102	0.040	0.226	0.194
الوزن	0.038	0.029	0.209	0.221

من الجدول اعلاه يتبين :
1 – في انموذج الانحدار اللوجستي البوتسترابي كانت معظم قيم (*P-Value*) اقل من قيمها لنموذج الانحدار اللوجستي التقليدي .
2 – في أنموذج الانحدار اللوجستي البوتسترابي كانت معظم قيم الاخطاء المعيارية اقل من قيمها لأنموذج الانحدار اللوجستي التقليدي .
من النتائج اعلاه يتبين أن استخدام اسلوب البوتستراب يعطي نتائج اكثر دقة .

الاستنتاجات :

1 – بينت نتائج انموذج الانحدار اللوجستي التقليدي ان متغير الوزن هو المتغير الوحيد الذي له تأثير معنوي على فترة بقاء المريض بالمستشفى (العناية المقدمة) ، لان اصحاب الوزن العالي يكونون معرضين للإصابة بالجلطة الدماغية اكثر من غيرهم .
2 – اوضحت نتائج انموذج الانحدار اللوجستي البوتسترابي ان متغير العمر ومتغير الوزن لهما تأثير معنوي على فترة بقاء المريض في المستشفى ، اي ان كبار السن واصحاب الوزن العالي اكثر حاجة للعناية .
3 – تبين من النتائج ان غالبية قيم المعنوية (*P- value*) والاخطاء المعيارية لمتغيرات البحث لانموذج الانحدار اللوجستي البوتسترابي اقل منها لنموذج الانحدار اللوجستي التقليدي وهذا يعني ان استخدام الاسلوب البوتسترابي يعطي نتائج اكثر دقة .

التوصيات :

1– التوصية باستخدام (الانموذج المعلمي) انموذج الانحدار اللوجستي البوتسترابي في دراسة البيانات الثنائية .
2– التوصية بدراسة انموذج الانحدار اللوجستي المتعدد (البوتسترابي).
3– مقارنة انموذج الانحدار اللوجستي البوتسترابي مع اساليب احصائية اخرى مثل تحليل النماذج اللوغارتمية الخطية او تحليل الدوال التميزية .
4 – ضرورة عدم التلكؤ في الذهاب الى الطبيب المختص في حالة الشعور بأعراض ارتفاع ضغط الدم اذ انه يعتبر عامل الخطر الرئيسي للإصابة بالجلطة الدماغية .

المصادر :-

1 – ابو حامد ، سمير ، " الجلطة الدماغية فالج ... عالج ..! " ، الطبعة الاولى ، دمشق ،خطوط النشر والتوزيع ، 2009 .
2 – https://www.webteb.com/neurology/السكتة الدماغية
3 – طراد ، علاء جابر ، " نماذج الانحدار المعلمي وشبه المعلمي (دراسة مقارنة) " ، رسالة ماجستير مقدمة الى قسم الاحصاء ، كلية الادارة والاقتصاد ، الجامعة المستنصرية ، 2013 .
4 – سعيد ، رشا عادل ، " استخدام نموذج الانحدار اللوجستي في دراسة العوامل المساعدة على تشخيص حالات الاصابة بسرطان المثانة " ، بحث منشور في مجلة العلوم الاقتصادية والادارية ، كلية الادارة والاقتصاد ، جامعة بغداد ، 2015 .
5 – Liu , Y ."On Goodness Of Fit Of Logistic Regression ", Doctoral thesis submitted to the Department of Statistics, University of Kansas , 2007.
6 – عباس ، علي خضير ، " استخدام نموذج الانحدار اللوجستي في التنبؤ بالدوال ذات المتغيرات الاقتصادية التابعة النوعية " ، بحث منشور في مجلة جامعة كركوك للعلوم الادارية والاقتصادية ، كلية الادارة والاقتصاد ، جامعة تكريت ، 2012 .
7 – Hosmer , D. , Lemeshow , Stanley ," Applied Logistic Regression " ,Second Edition ,New York ,AWiley - Interscience Publication , 2000.
8 – Czepiel, S .," Maximum Likelihood Estimation Of Logistic Regression Model : Theory And Implementation ", 2002 .
9 – بيثون ، نغم نافع متي ، " خواص قوة الاختبار وحدود الثقة لمعلمات نموذج اللوجستك الخطي (دراسة مقارنة) " ، رسالة ماجستير مقدمة الى قسم الاحصاء ، كلية الادارة والاقتصاد ، جامعة بغداد ، 1992 .
10 – Jassim , H . , And , Low , H . , And , AL-Karkhi , A . " An Overview Of Evaluation Criteria In Logistic Regression Model " , International Conference on Mathematical Sciences, University Teknologi , Malaysia,2007.
11 – Hu , B. , And , Palta , M . " Pseudo - R^2 In Logistic Regression Model " , University Of Wisconsin , Madison , 2006 .
12 – Li , Z . " Power And Sample Size Culation For Logistic Regression Tests For Differential Item Functioning " , Boston College , 2014 .
13 – Bewick , V . , And , Cheek , L . , And , Ball , J . " Statistics Review14: Logistic Regression " , University Of Brighton , Australia , 2005 .
14 – Algamal, Z ., And , Rasheed, K.," Re – Sampling In Linear Regression Model Using Jackknife And Bootstrap" , Research Published In The Iraqi Journal Statistical Science, 2010 .
15 – Martinez , W. , And , Martinez , A. " Computational Statistics Handbook with MATLAB", United States of America,2011.

بيانات البحث :

ت	التدخين	الجنس	العمر	الوزن	مدة الرقود
1	1	1	3	3	0
2	1	1	2	3	0
3	2	2	3	1	1
4	2	1	2	2	0
5	1	2	4	3	0
6	1	2	2	2	0
7	1	1	5	2	0
8	2	1	2	3	0
9	2	1	4	4	0
10	1	2	3	2	0
11	2	1	4	1	0
12	1	2	2	2	1
13	1	1	3	3	0
14	1	1	4	2	0
15	1	2	5	2	0
16	1	2	4	1	0
17	2	1	2	2	0
18	1	2	4	1	0
19	2	1	3	2	0
20	1	1	4	2	0
21	2	1	4	3	1
22	1	2	3	2	0
23	2	1	3	3	0
24	1	2	4	2	0
25	2	1	4	2	0
26	1	1	3	4	0
27	2	1	4	4	0
28	2	2	3	2	0
29	2	1	4	3	0
30	2	2	4	2	0
31	1	1	4	3	0
32	1	1	4	2	0
33	2	2	3	2	0
34	2	2	2	2	0
35	1	1	2	3	0
36	1	2	4	2	0
37	2	1	5	4	0
38	2	1	4	2	0
39	1	2	3	2	0
40	1	1	2	3	1
41	1	1	5	4	0
42	1	2	2	2	0
43	2	1	4	4	0
44	1	1	4	2	0
45	1	2	4	4	0
46	1	2	3	3	0
47	1	1	4	2	0
48	2	2	3	3	0
49	2	1	3	4	0

50	1	1	5	4	0
51	2	1	4	3	0
52	1	1	4	4	0
53	2	1	3	4	0
54	1	1	2	5	0
55	1	2	4	3	0
56	1	2	4	4	0
57	2	1	3	4	0
58	1	1	4	5	0
59	1	1	4	5	0
60	1	1	4	3	0
61	1	2	4	4	0
62	2	1	1	3	0
63	2	2	4	3	0
64	1	1	3	4	0
65	2	1	3	3	0
66	1	2	2	3	0
67	2	1	4	2	0
68	1	1	3	3	0
69	1	2	4	2	0
70	1	1	2	4	0
71	1	2	2	4	0
72	1	1	5	3	0
73	2	2	3	1	1
74	1	1	2	4	0
75	2	2	4	4	0
76	1	2	2	3	0
77	2	1	4	5	0
78	1	1	4	4	0
79	1	1	3	5	0
80	2	2	3	3	0
81	2	1	4	4	0
82	2	2	3	3	0
83	2	1	2	4	0
84	1	2	5	3	0
85	1	2	3	3	0
86	2	1	4	3	0
87	1	1	4	4	1
88	2	1	3	5	0
89	2	2	4	3	0
90	1	2	3	4	0
91	2	2	3	2	0
92	1	2	3	3	1
93	2	2	4	1	0
94	2	1	5	4	0
95	1	1	5	3	0
96	1	2	4	2	1
97	1	2	4	3	0
98	2	1	2	3	0
99	1	1	3	3	0
100	1	2	2	3	0
101	2	1	3	3	1

102	2	4	2	2	0
103	1	2	2	2	0
104	1	2	3	4	0
105	2	1	5	3	0
106	2	2	3	2	0
107	1	1	2	3	0
108	1	1	2	3	0
109	2	1	4	4	0
110	1	2	5	2	0
111	1	1	4	3	0
112	2	2	3	1	1
113	1	2	2	3	0
114	2	2	4	2	1
115	2	1	5	2	0
116	1	1	2	3	1
117	2	2	4	2	0
118	1	2	4	3	0
119	1	2	3	1	0
120	1	1	3	3	0
121	2	1	4	4	0
122	2	2	3	2	1
123	1	1	4	3	0
124	1	1	4	3	0
125	1	2	4	2	0
126	2	2	2	1	0
127	1	1	3	4	0
128	2	1	4	3	0
129	1	1	2	3	0
130	1	2	3	2	1
131	2	2	5	2	0
132	2	1	4	3	0
133	2	1	5	4	0
134	1	2	5	3	0
135	2	1	4	3	0
136	2	2	5	2	0
137	1	1	2	4	1
138	1	1	4	3	0
139	2	1	3	2	0
140	1	1	4	3	0
141	2	2	3	2	0
142	2	1	4	4	0
143	2	2	3	2	0
144	1	1	5	4	0
145	1	1	4	4	0
146	1	2	3	3	0
147	2	1	3	2	0
148	1	2	4	3	0
149	2	2	5	2	1
150	2	2	2	3	0
151	1	1	3	4	0
152	1	1	2	2	0
153	2	1	4	4	0

154	1	1	4	3	0
155	2	2	3	3	1
156	2	2	3	2	0
157	2	1	3	3	0
158	1	2	3	3	0
159	2	2	2	2	0
160	2	2	4	2	0
161	1	1	2	3	0
162	1	1	3	2	1
163	1	2	5	2	0
164	1	1	3	3	0
165	1	2	5	3	0
166	2	1	4	4	0
167	2	2	5	3	0
168	1	2	3	4	0
169	2	1	2	5	0
170	1	1	4	4	0
171	2	2	5	3	0
172	1	1	4	4	0
173	2	1	4	4	0
174	1	1	3	5	0
175	2	2	3	4	0
176	1	1	4	3	0
177	1	2	5	4	1
178	1	1	4	3	0
179	2	2	3	3	0
180	1	1	4	3	0
181	2	1	3	5	1
182	1	2	4	3	1