Use A Bootstrap Method In Parametric Regression Models (*Cox Model*)

استعمال اسلوب البوتستراب في نماذج الانحدار شبه المعلمية (انموذج *Cox*)

أ.م. د. شروق عبد الرضا السباح نور عباس عمران
جامعة كربلاء / كلية الادارة و الاقتصاد

(بحث مستل من رسالة ماجستير)

المستخلص :-

البيانات المستعملة في البحث تم الحصول عليها من مستشفى الحسيني في محافظة كربلاء المقدسة تضمنت (182) شخصا مصابا بالجلطة الدماغية خلال الفترة (2014 – 2015) .

وتم تحليل البيانات باستعمال انموذج انحدار Cox التقليدي والبوتسترابي وقد بينت النتائج ان استعمال اسلوب البوتستراب يعطي نتائج اكثر دقة .

Abstract :-

The data used in the research were obtained from al-Husseini hospital in the holy city of Karbala included (182) people infected with stroke during the period (2015-2014).

The data was using a specimen decline of traditional Cox and Bootstrap analysis of the results shows that the use of style bootstrap gives a more accurate and .

المقدمة :- *Introduction*

لنماذج الانحدار اهمية كبيرة في تحليل الظواهر المختلفة وقد تطورت مع الزمن بتطور علم الاحصاء اذ تعتبر النماذج المعلمية بشقيها الخطي و اللاخطي اول انواع نماذج الانحدار و استمر تطور نماذج الانحدار حتى ظهرت النماذج اللامعلمية ومن ثم النماذج شبه المعلمية و التي تكون حل وسط بين النماذج المعلمية و اللامعلمية اذ ان افتراضاتها اقوى من افتراضات النماذج اللامعلمية واقل تعقيدا من تللك الافتراضات التي تضعها النماذج المعلمية ، ومنها انموذج انحدار (*Cox*) الذي يهتم بدراسة العلاقة بين المتغيرات التوضيحية واوقات البقاء لها .

مشكلة البحث :- *Research Problem*

يحقق تحليل الانحدار اغلب اهداف البحث العلمي ، وتعد اساليبه الجزء المهم والاساسي لأي تحليل للبيانات يكون الهدف منه هو دراسة ووصف وتفسير العلاقة بين المتغير المعتمد والمتغيرات التوضيحية ، وعلى الرغم من ذلك فهو يعجز عن دراسة وتفسير العلاقة بين المتغيرات التفسيرية او التوضيحية والمتغير المعتمد الذي يمثل الوقت او المدة الزمنية ، مع الاخذ بنظر الاعتبار شيوع مثل هذا النوع من المتغيرات المعتمدة في دراسة الظواهر المختلفة ، ومن هنا تبرز الحاجة الى اساليب انحدار اخرى تكون عونا كبيرا لمثل هذا النوع من الدراسات ، وتمتلك قوة الانحدار الخطي الاعتيادي في ايجاد افضل المعادلات ، وتحل في الوقت ذاته مشكلة تعذر تطبيق نماذج الانحدار الخطي الاعتيادي عندما تكون المتغيرات المعتمدة تمثل مدة البقاء .

هدف البحث :- *Aim Of The Research*

تتحدد اهداف البحث بما يلي :-

1 – دراسة العلاقة بين العناية المقدمة للمريض خلال مدة الرقود في المستشفى وبين التدخين ، الجنس ، العمر والوزن من خلال تطبيق انموذج انحدار *Cox* التقليدي والبوتسترابي على البيانات ، وتحديد اي من تلك العوامل (التدخين ، الجنس ، العمر ، الوزن) الاكثر تأثيرا على مدة بقاء المريض في المستشفى .

2 – توضيح اثر استعمال الاسلوب البوتسترابي على مقاييس الانموذج شبه المعلمي *Cox* .

منهجية البحث :- *Research Methodology*

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

الجلطـة الـدمـاغيـة :-

تعريف مختصر بـالجلطـة الـدمـاغية [1] :-*Manual Definition Of Stroke*

يولد الانسان وهو يمتلك عدد محدد من الخلايا العصبية التي تتحكم في جميع وظائف الجسم وان هذه الخلايا غير قابله للتمدد او التكاثر ومن مجموع هذه الخلايا يتكون دماغ الانسان ، وان هذه الخلايا تموت عند توقف وصول الدم اليها ويؤدي فقدان عدد منها الى فقدان الشخص عدد من الوظائف العصبية هذا ما تسببه الجلطة الدماغية حيث تسبب بحرمان الخلايا الدماغية من الطاقة والاوكسجين اللذان يصلان عن طريق الدم عبر الاوعية الدموية وذلك اما لتمزق هذه الاوعية او لانسداد مجرى الدم .

وتعرف من الناحية الطبية 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 ".

علاجات الجلطة الدماغية [2]:- *Stroke Treatments*

عند الاصابة بالجلطة الدماغية يجب ان يتلقى المصاب العلاج الفوري و العاجل للحد من اضرارها قدر الامكان، ويختلف نوع العلاج باختلاف نوع السكتة الدماغية.

1 – الجلطة الدماغية الاقفارية :-

اهم خطوة في علاج السكتة الدماغية الاقفارية هي استئناف تزويد الدماغ بالدم بأسرع وقت ممكن وذلك عن طريق علاج الطوارئ بواسطة الادوية حيث يجب اعطاء الشخص المصاب ادوية تزيد من تخثر الدم خلال ثلاث الساعات الاولى من الاصابة بها ، اما العلاج الجراحي فقد يوصي الطبيب المعالج بإجراء عملية جراحية لفتح الشريان المسدود كليا او جزئيا و منها عملية فتح الشريان (*CEA*) او عملية تثبيت دعامة شبكية مرنة (*Stent*) داخل التضييق (*CAS*).

2 – الجلطة الدماغية النزفية :-

من الضروري اجراء الجراحة في حالة الاصابة بالجلطة الدماغية النزفية من اجل معالجتها او لمنع الاصابة بها مرة اخرى ، ومن الاجراءات الجراحية الاكثر انتشارا في حالة الاصابة بها هي ازالة الاوعية الدموية المشوهة (*AMV*) .

نموذج الانحدار :- *Regression Model*

نماذج الانحدار شبه المعلمية [3][4] :- *Semi parametric Regression Model*

وهي اسلوب احصائي يجمع بين النماذج المعلمية والنماذج اللامعلمية ، حيث ان النماذج اللامعلمية تتمتع بمرونة عالية لكونها لا تعتمد على الفروض او القيود التي تعتمد عليها النماذج المعلمية ، و تعرف النماذج شبه المعلمية على انها انموذج انحدار يتضمن جزئين الاول معلمي بأبعاد محدودة واخر لامعلمي بأبعاد غير محدودة ، وان اهم ما يميز النماذج شبه المعلمية هو كونها تجمع بين مميزات نموذجي الانحدار المعلمي واللامعلمي ،حيث انها تتضمن القيود الصارمة التي يفترضها الانموذج المعلمي وايضا المرونة العالية التي توفرها النماذج اللامعلمية .

تحليل البقاء [5]:- *Survival Analysis*

يهتم تحليل البقاء بتحليل الوقت المنقضي ، حيث يمثل متغير الاستجابة او المتغير المعتمد الوقت بين اصل الوقت والنقطة الاخيرة ، وان اصل الوقت يشير الى بداية حدوث الحدث كحدث الولادة او بداية اخذ علاج معين او بداية متابعة ماكنة ما ،اما النقطة الاخيرة فهي تمثل حدث النهاية كالموت او الشفاء او فشل الماكنة ، ان هذه الاوقات المنقضية لها خاصيتان تبطلان تطبيق التقنيات الاحصائية القياسية كاختبار (t) او تحليل التباين او الانحدار المتعدد ، الخاصية الاولى هي ان قيم الوقت تنحرف في اغلب الاحيان نحو اليمين من التوزيع اي انها موجبة الالتواء ، وان الطرق القياسية تتطلب بان يكون توزيعها طبيعيا ، والخاصية الثانية فهي تتعلق بالبيانات حيث ان جزء من هذه البيانات يكون مراقب (*Censored Data*) وهي على عدة انواع وبسبب هاتان الخاصيتان لا يمكن تحليل بيانات البقاء باستعمال الطرق القياسية، ولأجل تحليل بيانات البقاء يجب التعرف على بيانات المراقبة ودالتان هما دالة البقاء ودالة المخاطرة .

البيانـات المـراقبـة :- *The Censored Data*

تحدث هذه البيانات عندما نمتلك بعض المعلومات عن وقت الفشل ، وهي على ثلاثة انواع :-

اولا :- البيانـات المـراقبة يمينـا :- *Right Censored Data*

وهي تظهر الحد الادنى لوقت الفشل وانواعها :-

1 – النـوع الاول I مـن المـراقبـة :-

و ينقسم :-

النـوع الاول I من المـراقبـة (الثابت) [6]:-

وهي تحدث عند تحديد فترة زمنية محددة لنهاية المراقبة وتنتهي تلك الفترة قبل ان تفشل جميع المشاهدات ، لنفترض ان الحدث المهم او حدث الدراسة هو الاصابة بمرض معين فان الباحث يحدد فترة زمنية للمراقبة بدلا من ان ينتظر لحين اصابة جميع الافراد .

النوع الاول I من المراقبة (العشوائي) [7]:-

وهي تحدث عند تحديد فترة زمنية لنهاية المراقبة ، لكن ليس كل الافراد يمتلكون نفس وقت المراقبة ، أي ان المراقبة تحدث بصورة عشوائية فمثلا في التطبيقات الطبية وفي التجارب السريرية يدخل المرضى الى الدراسة في اوقات مختلفة و بالتالي فهم لا يمتلكون نفس وقت المراقبة .

2 – النوع الثاني II من المراقبة [8]:-

وفي هذا النوع من المراقبة يتم تحديد نهاية المراقبة بعد الحصول على عدد محدد سلفا من حالات الفشل ، فمثلا لوكان هدف الدراسة هو الحصول على (*80*) شخص مصاب بمرض معين فتنتهي المراقبة عند الحصول على هذا العدد.

ثانيا :- البيانات المراقبة يسارا[9] :- *Left Censored Data*

وهذه البيانات تحدث عندما يكون لدينا علم بوقت حدوث حدث معين (كتشخيص المرض) لكننا لا نعلم الوقت الحقيقي للحدث (أي متى اصيب الشخص بالمرض).

ثالثا :- فترة البيانات المراقبة [6]:- *Interval Censored Data*

وهي تحدث عندما يكون لدينا علم بأن الحدث (الفشل) يحدث خلال فترة زمنية معينة ، وكمثال على ذلك لو كان الافراد يستهلكون منتج ما خلال فترة زمنية لكن لا نعلم بالضبط متى يستهلكون ذلك المنتج باعتبار الاستهلاك هو حدث الفشل.

دالة البقاء [10]:- *Survivor Function*

وهي احتمال البقاء لفترة تتجاوز قيمة محددة وهي مكملة لدالة التوزيع التجميعية (*Cumulative Distribution Function* *F(t)*) ويعبر عنها رياضيا :-

$$S(t) = 1 - F(t) = \Pr[T \geq t] = \int_t^{\infty} f(x)dx \quad \dots\dots\dots\dots (1)$$

اذ ان :

S(t) : تمثل دالة البقاء .

F(t) : تمثل دالة التوزيع التجميعية (*c.d.f*) للمتغير العشوائي (T).

f(x) : تمثل دالة الكثافة الاحتمالية (*p.d.f*) .

وبما ان دالة البقاء احتمال فأنها موجبة واعلى قيمة لها هي الواحد واقل قيمة هي الصفر ، واذا كان الوقت (t) مساويا للصفر فان :

$$S(0) = 1 \quad \dots\dots\dots\dots (2-31)$$

وكلما اقترب الوقت (t) من (∞) فان دالة البقاء تقترب من الصفر.

دالة المخاطرة [11][12]:- *Hazard Function*

وهي احتمال الفشل خلال فترة صغيرة جدا من الوقت(t) وتعرف ايضا بدالة الفشل الشرطية وصيغتها الرياضية :-

$$h(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} pr[t \leq T < t + \Delta t \backslash T \geq t] \quad \dots\dots\dots\dots (2)$$

اذ ان البسط يمثل الاحتمال الشرطي للحدث الذي سيحدث في الفترة $[t, t+\Delta t)$ والمقام يمثل التغير في الوقت، وتسمى احيانا نسبة الفشل الانية او قوة الفناء او نسبة فشل العمر المحددة .

وان العلاقة بين الدالة الاحتمالية (*f(t)*) ودالة البقاء (*S(t)*) ودالة المخاطرة (*h(t)*)تكون بالشكل التالي :-

$$h(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} pr[t \leq T < t + \Delta t \backslash T \geq t]$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{pr[(t \leq T < t+\Delta t) \cap (T \geq t)]}{pr(T \geq t)}$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{pr[t \leq T < t+\Delta t]}{pr(T \geq t)}$$

$$h(t) = \frac{f(t)}{s(t)} \quad \dots\dots\dots\dots (3)$$

انموذج انحدار Cox [13]:- *Cox Regression Model*

اقترح هذا الانموذج من قبل العالم الانكليزي الاحصائي (*David Cox*) عام (1972) وهو الانموذج الاكثر شيوعا في دراسة تبعية اوقات البقاء على المتغيرات التوضيحية ، وقد اظهر العالم Cox ان العلاقة بين نسبة الخطر ومجموعة المتغيرات التوضيحية بالشكل التالي :

$$h(t/x) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \quad \ldots\ldots\ldots (4)$$
$$t \geq 0$$
$$-\infty < X < \infty$$

اذ ان :

$h_0(t)$: تمثل دالة المخاطرة الاساسية عندما تكون قيم المتغيرات التوضيحية مساوية للصفر ($x=0$) وهي دالة غير محددة وغير سالبة .

$\beta_1, \beta_2, \ldots, \beta_n$: تمثل معلمات انموذج الانحدار المراد تقديرها وهي تفسر التغير في نسبة الخطر عندما تزداد المتغيرات التوضيحية بمقدار وحدة واحدة.

$x_1, x_2, \ldots, x_n$: تمثل المتغيرات التوضيحية .

ويلاحظ ان هذا الانموذج لا يحتوي على الحد الثابت حيث انه يصبح جزء من $h_0(t)$.

طرائق تقدير معلمات انموذج انحدار Cox :-

Methods For Estimating The Cox Regression Model

طريقة الامكان الاعظم الشرطي[8] :- *Conditional Likelihood Method*

وهي اول طريقة من طرق تقدير معلمات انموذج انحدار *Cox* ، اقترحها العالم (*Cox*) عام (*1972*) ، وبسبب كثرة التعديلات التي تحتاجها لم تعتمد هذه الطريقة كثيرا ، حيث اقترح العالم (*Cox*) ان تقدير المعلمات (β) يعتمد على مجموعة الوقت (t_i) التي تحدث عندها الوفاة او الفشل حيث يشترط وجود هذا الوقت ضمن مجموعة الخطر ($R(t_i)$ ($Risk\ Set$)) فاذا كان احتمال الوفاة او الفشل للشخص هو:

$$\frac{h_0(t_i)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}{\sum_{I \epsilon R(t_i)} h_0(t_i)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)} \quad \ldots\ldots\ldots (5)$$

واذا كانت عناصر دالة الامكان الشرطي تتمثل بكل فشل فأنه بعد اخذ اللوغاريتم لها تصبح بالشكل التالي :-

$$L(\beta) = \sum_{i=1}^{k}(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) - \sum_{i=1}^{k}\log\left[\sum_{I \epsilon R(t_i)}(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)\right] \ldots\ldots\ldots (6)$$

وللحصول على تقديرات للمعلمات (β) نقوم بتعظيم الدالة ومن ثم تطبيق طريقة نيوتن رافسون .

طريقة الامكان الجزئي [14][15]: *Partial Likelihood Method*

وهي واحدة من الطرائق التي تستعمل لتقدير معلمات انموذج انحدار(*Cox*) وقد اقترحها العالم (*Cox*) عام (*1975*) وتعتبر الطريقة الاكثر شيوعا التي لا تعتمد على ($h_0(t)$) للحصول على (β)، وهي التقنية التي طورت للاستدلال حول معلمات انموذج انحدار (*Cox*) في وجود $h_0(t)$.

فاذا كانت (t_i) تمثل اوقات الفشل و($R(t_i)$) تمثل مجموعة الخطر وهي تحتوي على كل المفردات المعرضة لخطر الفشل وهي ($R(t_i) = \{j: t_j \geq t_i\}$)، وان احتمال الفشل للشخص كما عرف في المعادلة (*5*) وبافتراض ان (S) تمثل عدد حالات الفشل فان دالة الامكان الجزئية :-

$$\prod_{i=1}^{S}\frac{\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}{\sum_{I \epsilon R(t_i)}\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)} \quad \ldots\ldots\ldots (7)$$

ولتكن (v_i) تمثل حالات الفشل عند الوقت (t_i) وان ($v_i = 1$) عند حالات الفشل و ($v_i = 0$) عندما تكون المشاهدات مراقبة ، بالإمكان اعادة كتابة المعادلة اعلاه

$$L(\beta) = \prod_{i=1}^{k}\left[\frac{\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}{\sum_{j=1} Y_j(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}\right]^{v_i} \quad \ldots\ldots\ldots (8)$$

اذ ان $Y_j(t) = 1$ عندما $t \leq t_j$ ماعدا ذلك $Y_j(t) = 0$

وبأخذ اللوغاريثم الطبيعي للمعادلة (*32-2*) تصبح

$$\log L(\beta) = \sum_{i=1}^{k} v_i \left[\frac{\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}{\sum_{j=1} Y_j(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)} \right] \dots \dots (9)$$

وبأخذ المشتقة للمعادلة اعلاه ومساويتها للصفر نحصل على المعلمات المطلوبة .

طريقة الامكان الجزئي الموزون[8] :- *Weighted Partial Likelihood*

وتعتبر هذه الطريقة تطوير بسيط لطريقة الامكان الجزئي ، وقد اقترحت عام (*1993*) من قبل العالم (*Peter Sajisni*) ، فهي تفترض ان لدالة الامكان الجزئي اوزان هذه الاوزان تقترح حسب اوقات الفشل المراقبة ، وان دالة الامكان الجزئي الموزون تعرف بالشكل التالي :-

$$L(\beta) = \prod_{i=1}^{k} C_i \frac{h_i(t)}{\sum_{i=1}^{n} h_i(t)} \quad \dots \dots \dots \dots (10)$$

اذ ان :-

$$h_i(t) = y_i(t) h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \dots \dots \dots (11)$$

وان

$$y_i = \begin{cases} 1 & t \text{ تشير الى ان الشخص } i \text{ في حالة مراقبة عند الوقت} \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

وان

$$C_i = \begin{cases} 1 & \text{اذا كان وقت الفشل مراقب} \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

وبأخذ اللوغاريتم الطبيعي وتعظيم المعادلة (*10*) نحصل على التقديرات المطلوبة .

اختبار معنوية المعالمات المقدرة[16]:-

Test The Significantly Of Estimated Parameters

كما في انموذج الانحدار اللوجستي نستخدم احصاءة والد (*Wald Statistic*) لاختبار معنوية المعالم المقدرة وحسب الصيغة التالية :-

$$Z_j = \left[\frac{\widehat{\beta_j}}{S.E_{\widehat{\beta_j}}} \right]^2 \quad \dots \dots \dots \dots (12)$$

اذ ان :-

$\widehat{\beta_j}$:- تمثل المعلمات المقدرة لأنموذج انحدار *Cox*

$S.E_{\widehat{\beta_j}}$:- يمثل الخطأ المعياري لمعلمات انموذج انحدار *Cox*

وان هذه الاحصاءة تتبع توزيع (x^2) بدرجة حرية واحدة .

ويمكن ايجاد فترات الثقة للمعلمات المقدرة باستعمال الصيغة التالية :-

$$\widehat{\beta_j} \mp Z_{\alpha/2} S.E_{\widehat{\beta_j}} \quad \dots \dots \dots \dots \dots (13)$$

تقييم القوة التفسيرية للنموذج[5]:-

ويتم ذلك من خلال احتساب قيمة (R^2) التي تحسب حسب الصيغة التالية :-

$$R^2 = 1 - \exp\left[\frac{2}{n}\left(\log L_0 - \log L_p\right)\right] \quad \dots \dots \dots \dots (14)$$

اذ ان :-

L_0 :- تمثل قيمة الامكان الاعظم عندما الانموذج يحتوي على الثابت فقط .

L_p :- تمثل قيمة الامكان الاعظم عندما يحتوي الانموذج على كل المتغيرات التوضيحية .

n :- تمثل عدد المشاهدات سواء كانت مراقبة ام لا .

البوتستراب[17]:- *Bootstrap*

هي احدى طرق اعادة المعاينة بالإرجاع " اي اعادة وحدة المعاينة المسحوبة قبل سحب الوحدة التي تليها "، وهي تعتبر اكثر اعتمادا وانتشارا من طريقة الجاكنايف التي هي ايضا واحدة من طرق اعادة المعاينة بالإرجاع التي تكون اكثر توافقا مع طريقة البوتستراب الا ان الجاكنايف اسهل من طريقة البوتستراب عندما يكون تكرار البوتستراب المستعمل لحساب التقديرات المطلوبة اكبر من حجم العينة .

والخطوات المتبعة في استعمال طريقة البوتستراب هي :
1 – من العينة الحقيقية $(X = (x_1, x_2, \dots\dots\dots, x_n))$ نسحب عينات عشوائية جزئية بالإرجاع وبحجم معين نرمز لها بـ $(X^* = (x_1^*, x_2^*, \dots\dots\dots, x_m^*))$ عددها يساوي (m) والذي يكون على الاقل (200-20) مرة.
2 – ايجاد تقديرات لمعالم انموذج انحدار (Cox) بطريقة الامكان الجزئي ولكل عينة من عينات البوتستراب .

التطبيق العملي :-

وصف البيانات :-

تم تطبيق الموضوع على بيانات واقعية اخذت من مستشفى الحسيني في محافظة كربلاء المقدسة ، من خلال مراجعة ردهة الطوارئ ومراقبة حياة (182) شخص مصاب بالجلطة الدماغية .

متغيرات البحث :-

T_i: يمثل مدة الرقود (العناية بالمريض) في المستشفى للمرضى المصابين بالجلطة الدماغية ، وهو يمثل المتغير التابع لأنموذج انحدار Cox البوتسترابي .

دالة البقاء تمثلت بحالة خروج المصاب من المستشفى وقد تم ترميزها بالشكل التالي :-

1 :- اذا كان الشخص على قيد الحياة .

0 :- اذا كان الشخص متوفي .

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 .

X_4 : يمثل وزن الشخص .

تم تحليل البيانات باستعمال برنامج $(SPSS)$ و حصلنا على النتائج الخاصة بالنموذج وكذلك الاحصاءات الوصفية التالية :-

الاحصاءات الوصفية :-

الجدول التالي يبين الاحصاءات الوصفية كالنسب المئوية والمتوسطات والانحراف المعياري لمتغيرات البحث كافة .

جدول (1) الاحصاءات الوصفية لمتغيرات البحث كافة

المتغيرات	عدد الحالات	النسبة المئوية	المتوسط Mean	الانحراف المعياري S.d
العناية (فترة الرقود) Y	182	100 %	1.12	0.320
اقل من 6 ايام	161	88.5 %		
6 ايام فأكثر	21	11.5 %		
التدخين X_1	182	100 %	1.45	0.498
مدخن	101	55.5%		
غير مدخن	81	44.5%		
الجنس X_2	182	100 %	1.43	0.497
ذكر	103	56.6%		
انثى	79	43.4%		
العمر X_3	182	100 %	3.46	0.944
30-40	1	0.5%		
40-50	32	17.6%		
50-60	55	30.2%		
60-70	71	39%		
70-80	23	12.6%		
الوزن X_4	182	100 %	2.96	0.957
50-60	9	4.9%		
60-70	51	28 %		
70-80	70	38.5%		
80-90	43	23.6 %		
90-100	9	4.9 %		

من الجدول (1) نلاحظ التالي :-

1 – نسبة الاشخاص اللذين مدة رقودهم في المستشفى هي (6) ايام فأكثر تساوي (11.5%) وهي نسبة قليلة جدا بالمقارنة مع نسبة الاشخاص اللذين يقل زمن رقودهم عن (6) ايام التي تساوي (88.5%) وهي تشكل تقريبا اكثر من ثلاثة ارباع عينة البحث .

2 – نسبة الاشخاص المدخنين المصابين بالجلطة الدماغية (55.5%) وهي مقاربة لنسبة الاشخاص غير المدخنين .

3 – نسبة الذكور المصابين بالجلطة الدماغية (56.6%) وهي ايضا مقاربة لنسبة الاناث المصابات بالجلطة الدماغية (43.4%) .

4 – اكبر نسبة للأشخاص المصابين بالجلطة الدماغية هم الأشخاص اللذين تتراوح اعمارهم بين (60-70) سنة ،بينما اقل نسبة كانت للأشخاص اللذين تتراوح اعمارهم بين (30-40) سنة وهذا يعني انه كلما زاد عمر الشخص كلما زاد خطر الاصابة بالجلطة الدماغية .

5 – نسبة الاشخاص المصابين بالجلطة الدماغية اللذين تتراوح اوزانهم بين (50-60) كيلو غرام وهي مساوية لنسبة الاشخاص اللذين تتراوح اوزانهم بين (90-100) كيلو غرام بينما كانت اكبر نسبة هي للأشخاص اللذين تتراوح اوزانهم بين (70-80) كيلو غرام .

ومن اجل الحصول على نتائج انموذج انحدار *Cox* التقليدي و البوتستراب ، قمنا بتحديد عينة البوتستراب وهي العينة البسيطة ، وعدد عينات البوتسراب وهي (100) ، ومن تطبيق أنموذج الانحدار حصلنا على النتائج التالية :-

نتائج تطبيق انموذج انحدار *Cox* التقليدي :-

Results of applying Of The Cox regression model

انموذج انحدار *Cox* يستعمل لتحديد العلاقة بين المتغيرات التفسيرية المتوافرة واوقات البقاء لها، وفي هذا البحث نحاول معرفة تأثير المتغيرات المستقلة المحددة سابقا على مدة الرقود في المستشفى للمرضى المصابين بالجلطة الدماغية ، ومن اجل تطبيق انموذج انحدار *Cox* يجب اولا معرفة فيما اذا كان الانموذج معنوي ام لا ويتم ذلك من خلال اختبار الفرضيات التالية :-

فرضية العدم :- لا يوجد تأثير للمتغيرات المستقلة على المتغير التابع . $H_0: \beta = 0$

الفرضية البديلة :- يوجد تأثير للمتغيرات المستقلة على المتغير التابع . $H_1 : \beta \neq 0$

ولقد اوضحت النتائج بان القيمة الاحتمالية لاحصاءة (*Chi-Squaer*) تساوي (0.044) وهي اقل من مستوى المعنوية ($\propto = 0.05$) وهذا يؤكد معنوية الانموذج أي اننا نرفض فرضية العدم و لا نرفض الفرضية البديلة.

ومدى ملائمة الانموذج للبيانات يمكن معرفتها من خلال احتساب قيمة (R^2) لأنموذج انحدار *Cox* باستعمال المعادلة (14) حيث كانت قيمته تساوي

$$R^2 = 1 - exp\{\frac{2}{182}((-424.7655) - (-419.879))\}$$

$$R^2 = 0.05228145$$

وهذه القيمة تشير الى (5.228 %) من التغير الحاصل في المتغير التابع يتم تفسيره من خلال انموذج انحدار *Cox* .

جدول (2) نتائج أنموذج انحدار Cox التقليدي

المتغيرات	معاملات انحدار Cox β	الخطأ المعياري S.E	احصاءة Wald	درجات الحرية Df	القيمة الاحتمالية Sig.	القيمة المتوقعة Exp(β)	فترات الثقة 95% الدنيا Lower	فترات الثقة 95% العليا Upper
التدخين	-0.143	0.212	0.457	1	0.499	0.867	0.572	1.312
الجنس	-0.075	0.258	0.085	1	0.770	0.927	0.559	1.538
العمر	0.010	0.012	0.678	1	0.410	1.010	0.986	1.034
الوزن	0.028	0.013	4.829	1	0.028	1.028	1.003	1.054

من ملاحظة الجدول اعلاه يتبين لنا التالي :-

القيمة الاحتمالية لاحصاءة *Wald* للمتغير X_1 (التدخين) اكبر من مستوى المعنوية ، وهذا يعني ان التدخين ليس له تأثير معنوي على المتغير التابع ، والقيمة الاحتمالية لاحصاءة *Wald* للمتغير X_2 (الجنس) اكبر من مستوى المعنوية ، اي انه ليس للجنس تأثير معنوي على المتغير التابع ، وايضا لمتغير X_3 (العمر) ليس له تأثير معنوي على المتغير التابع ، اذ كانت القيمة الاحتمالية لاحصاءة *Wald* له ايضا هي اكبر من مستوى المعنوية ، اما القيمة الاحتمالية لاحصاءة *Wald* للمتغير X_4 (الوزن) اصغر من مستوى المعنوية ، اذن الوزن هو المتغير الوحيد الذي له تأثير معنوي على المتغير التابع .

نتائج تطبيق انموذج انحدار *Cox* البوتسترابي :- ***Results of applying bootstrap Cox regression model***

عند تطبيق انموذج انحدار *Cox* البوتسترابي على البيانات ايضا حصلنا على نتائج مماثلة للنتائج التي حصلنا عليها باستعمال انموذج انحدار *Cox* التقليدي وبقية النتائج مبينة بالجدول التالي :-

جدول (3) نتائج انموذج انحدار *Cox* البوتسترابي

المتغيرات	معاملات انحدار *Cox* البوتسترابي β	$Bootstrap^a$		
		الخطأ المعياري *Std.Errar*	القيمة الاحتمالية *Sig.(2-tailed)*	فترات الثقة % 95
				الدنيا *Lower* / العليا *Upper*
التدخين	*-0.143*	*0.203*	*0.495*	*-0.629* / *0.230*
الجنس	*-0.075*	*0.221*	*0.673*	*-0.527* / *0.382*
العمر	*0.010*	*0.012*	*0.455*	*-0.011* / *0.031*
الوزن	*0.028*	*0.010*	*0.020*	*0.007* / *0.050*

من خلال ملاحظة القيم الاحتمالية في الجدول رقم (3) يتبين لنا التالي :-

القيمة الاحتمالية للمتغير X_1 (التدخين) اكبر من مستوى المعنوية ، وهذا يعني ان التدخين ليس له تأثير على المتغير التابع ، وايضا المتغيران X_2 (الجنس) وX_3 (العمر) ليس لهما تأثير على المتغير التابع ، لان القيمة الاحتمالية لهما اكبر من مستوى المعنوية ، اما المتغير X_4 (الوزن) فان القيمة الاحتمالية له اقل من مستوى المعنوية ، اي ان الوزن هو المتغير الوحيد الذي له تأثير على المتغير التابع .

ومن خلال القيم في الجدول رقم (3) يمكن كتابة انموذج انحدار Cox البوتسترابي بالشكل التالي :-

$$\hat{h}(t,x) = \widehat{h_0}(t)\text{Exp}\{(-0.143\ X_1) + (-0.075\ X_2) + (0.010\ X_3) + (0.028\ X_4)$$

من ملاحظة المعادلة اعلاه يتبين ان زيادة المتغيرين x_1 وx_2 بمقدار وحدة واحدة يؤدي الى تناقص دالة المخاطرة ، وان زيادة المتغيرين x_3 و x_4 بمقدار وحدة واحدة تؤدي زيادة دالة المخاطرة .

المقارنة بين الاسلوب البوتسترابي والتقليدي :-

The Comparison Between The Traditional Style And Bootstrap

تحقيقا للهدف الثاني من الدراسة ومن اجل توضيح اثر استعمال الاسلوب البوتسترابي على مقاييس النموذج نقوم بمقارنة قيم المعنوية (P-Value) والاخطاء المعيارية لنموذج انحدار Cox التقليدي مع قيمها لنموذج انحدار Cox البوتسترابي ، علما اننا لم نستخدم قيم (R^2) في المقارنة لأننا نحصل على نفس القيم عند استعمال الاسلوبين التقليدي والبوتسترابي .

جدول (4) قيم المعنوية والاخطاء المعيارية

المتغيرات	انموذج انحدار *Cox* التقليدي *P-Value*	انموذج انحدار *Cox* البوتسترابي *P-Value*	الخطأ المعياري لنموذج انحدار *Cox* التقليدي *Std.Errer*	الخطأ المعياري لنموذج انحدار *Cox* البوتسترابي *Std.Error*
التدخين	*0.499*	*0.495*	*0.212*	*0.203*
الجنس	*0.770*	*0.673*	*0.258*	*0.221*
العمر	*0.410*	*0.455*	*0.012*	*0.12*
الوزن	*0.028*	*0.020*	*0.013*	*0.10*

من الجدول (4) يتبين :-

1 - ان قيم (*P-Value*) لأنموذج انحدار *Cox* البوتسترابي اقل منها لأنموذج انحدار *Cox* التقليدي .

2 – ان قيم الاخطاء المعيارية لأنموذج انحدار *Cox* البوتسترابي اقل منها لأنموذج انحدار *Cox* التقليدي.

اذن مما سبق يمكن القول بأن استعمال اسلوب البوتستراب يعطي نتائج اكثر دقة .

الاستنتاجات :-

1 – اوضحت نتائج انموذج انحدار (*Cox*) التقليدي والبوتسترابي ان متغير الوزن هو المتغير الوحيد الذي له تأثير معنوي على فترة بقاء المريض في المستشفى .

2 - اوضحت النتائج ان قيم المعنوية (*P- value*) والاخطاء المعيارية لمتغيرات البحث كافة لنموذج انحدار (*cox*) البوتسترابي اقل منها لنموذج انحدار (*Cox*) التقليدي وهذا يعني ان استعمال الاسلوب البوتسترابي يعطي نتائج اكثر دقة .

3 - اكدت النتائج ان انموذج انحدار *Cox* البوتسترابي للبيانات الحقيقية هو الانموذج الافضل في تمثيله للبيانات .

التوصيات :-

1 – زيادة الوعي الصحي من خلال المراجعة المنتظمة للطبيب واجراء التحاليل اللازمة .

2 – دراسة إمكانية إدخال متغيرات جديدة قد تكون لها تأثير معنوي على مـدة بقاء المريض المصاب بالجلطة الدماغية في المستشفى مثل ضغط الدم لأنه من المتغيرات المؤثرة التي تزيد من احتمال حصول الجلطة الدماغية .

3 – دراسـة امكانيـة استعمال اسلوب الجاكنايف (*Jackknife*) مع نماذج الانحدار شبه المعلمية (*Cox*) .

المصادر :-

1 – ابو حامد ، سمير ، " الجلطة الدماغية فالج ... عالج ..! " ، الطبعة الاولى ، دمشق ،خطوط النشر والتوزيع ، *2009* .

2 – *https://ar.wikipedia.org/wiki%D8%B3%D9%83%D8%AA%D8%A9*

3 – Ichimura , Hidehiko, "Semiparmetric Least Squares (SLS) And Weighted SLS Estimation", Department Of Economics , University of Minnesota ,Minneapolis , 1991.

4 – طراد ، علاء جابر ، " نماذج الانحدار المعلمي وشبه المعلمي (دراسة مقارنة) " ، رسالة ماجستير مقدمة الى قسم الاحصاء ، كلية الادارة والاقتصاد ، الجامعة المستنصرية ، *2013* .

5 – https: // ncss.Wpengine.netdna– cdu.com / Wpcontent /theme /ncss / pdf/procdures/Ncss/Cox-Regression

6 –Ihwah , Azimmatul," The Use of Cox Regression Model to Analyze the Factors that Influence Consumer Purchase Decision on a Product", Department of Agro-industrial Technology, Brawijaya University, Indonesia,2015.

7- Biosi ," Introduction To Survival Analysis " , 2004 .

8 – الاعظمي ، يسرى طارق اسماعيل ، " دراسة وتحليل نموذج *Cox* للانحدار مع تطبيق عملي " ، رسالة ماجستير مقدمة الى قسم الاحصاء ، كلية الادارة والاقتصاد ، جامعـة بغداد ، *1999* .

9 – Samartzis , Lefteris , "Survival and censored data", first edition ,2006 .

10 – Tyler , Smith ,And Other , " Survival Analysis Using Cox Proportional Hazards Modeling For Single And Multiple Event Time Data " , Naval Health Research Center, San Diego.

11 – Klein , Johnp , And , Moeschberger , Melvin L. ," Survival Analysis Techniques For Censored And Truncated Data " , Second Edition , New York , 2003 .

12–https : // www.amstat.org / chapters/ northeasterni // inois /pastere-nts/presntations/summer05-Ibrahim-J.pdf

13 – Hu , Ping, And, Tsiatis, Anastasios A , And ,Davidian ,Marie," Estimating the Parameters in the Cox Model When Covariate Variables Are Measured with Error", North Carolina State University , North Carolina ,1998.

14 – Qi , Jiezhi," Comparison of Proportional Hazards and Accelerated Failure Time Models",Message Majsur Introduction to Mathematics and Statistics Department ،the Faculty of Graduate Studies and Research in Fulfillment , University of Saskatchewan, Saskatoon,2009.

15 – Liao , Hsi-Wen," A Simulation Study Of Estimators In Stratified Proportional Hazards Model ", Pharmaceutical Research Associates, Inc. , Shrewsbury, NJ, U.S.A.

16 – الجبلاوي ، عوض الكريم عبد الرحيم بشير ،و مدني ، مدينـة سر الختم خليل ، وعلي ، يسن علي عبد الله ، " تقدير دالـة المخـاطرة بـاستخدام نموذج انحدار كوكس لمرضى سرطان الثدي بمستشفى العلاج بالأشعة (الذرة) " ، بحث تكميلي لنيل درجة بكالوريوس الشرف في الاحصاء التطبيقي ، كلية العلوم ، جامعة السودان للعلوم و التكنلوجيا ، *2015* .

17– السباح ، شروق عبد الرضا سعيد ، " بناء انموذج انحدار لوجستي معدل لحياة الاطفال الخدج في محافظة كربلاء " ، اطروحة دكتوراه مقدمة الى قسم الاحصاء ، كلية الادارة والاقتصاد ، جامعة بغداد ، *2009* .