AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

استعمال أسلوب الجاكنايف في أنموذج انحدار كوكس (Cox) الشبه معلمي لمرضى سرطان الرئة (تطبيق عملي)

م .عقيل حميد فرحان[1]
Aqeel_hameed@ymail.com
جامعة ديالى/ كلّية الإدارة والاقتصاد/ قسم الاحصاء

م . مرتضى منصور عبد الله[2]
ilhasme@yahoo.com
جامعة واسط / كلّية الإدارة والاقتصاد/ قسم الاحصاء

المستخلص :

طبق انموذج انحدار كوكس(Cox Regression) والجاكنايف على البيانات من أجل دراسة العلاقة بين العناية المقدّمة للمريض خلال فتره رقوده في المستشفى وبين نوع العلاج و الجنس والعمر والحالة الاجتماعية وتحديد أيّ من تلك العوامل الأكثر تأثيرا على فتره بقاء المريض في المستشفى وتوضيح تأثير استعمال أسلوب الجاكنايف على مقاييس الانموذج شبه المعلمي كوكس، وقد بيّنتْ النتائج ان استعمال أسلوب الجاكنايف يعطي نتائج أكثر دقّة للبيانات المستعملة في البحث التي تَمّ الحصول عليها من مستشفى مدينة الطب المتمثلة بمستشفى الاورام التعليمية وكذلك مستشفى بعقوبة التعليمي ضمنت (*100*) شخصا مصابا بسرطان الرئة. التي تصيب أيّ إنسان في أيّ وقت كما بيّنتْ النتائج ان قيم الأخطاء المعيارية لأنموذج انحدار Cox عند متغيّر نوع العلاج والجنس والحالة الاجتماعية يكون أكبر من أسلوب الجاكنايف وهذا يدّل على دقّة أسلوب الجاكنايف ،أمّا عند متغيّر العمر فإنّه يحث العكس، واستبعاد متغيّر الجنس يمكن القول بأن استعمال أسلوب الجاكنايف يعطي نتائج أكثر دقّة لتمثيل البيانات

الكلمات المفتاحية :(الجاكنايف ، تراكمية ، الانتعاش، الاستراتيجية ، التحوّلات ، النسيجي ،الانتقال ، التقليدي ، البوتستراب)

Abstract :

Cox regression model and Jacknife were applied to the data in order to study the relationship between the care provided to the patient during his hospitalization period and between the type of treatment, gender, age and marital status, and to determine which of these factors had the most influence on the patient's stay in the hospital and to clarify the effect of using the Jacknife method On the scales of the semi-parametric Cox model, the results showed that the use of the Jagnief method gives more accurate results for the data used in the research obtained from the Medical City Hospital represented by the Teaching

Oncology Hospital as well as the Baquba Teaching Hospital, which included (100) people with lung cancer. The results showed that the standard errors of the Cox regression model for the variable of type of treatment, sex and marital status are greater than the Jacknaive method, and this indicates the accuracy of the Jacknaive method, but for the age variable it induces the opposite, and excluding the sex variable can be said that Using the Jacknife method gives more accurate results for data representation

Keywords= (Jackknife, cumulative, recovery, strategy, transitions, texture, transition, classic, bootstrap)

1- المقدّمة

بعد تطوير وتوسيع الخدمات الصحية التي هي أحد الأركان الأساسية الاستراتيجية التنمية الاقتصادية والاجتماعية والتي تستهدف تطويق الأوبئة والأمراض والاهتمام بتطوير الموارد البشرية، كما أنها تأتي استجابة لمشكلات الاجتماعية والبيئية التي يواجها الفرد، ولاسيما في مرحلة التحوّلات والتغيّرات الجذرية الواسعة التي يشهدها المجتمع في المعالجة الطبية الفعالة لكثير من الأمراض من شأنها أنْ ترفد المجتمع بعناصر منتجه تسهم في تطور الناتج القومي والحركة التموينية في المجتمع، فضلا عن توفير افضل كانت الاساس في البحث لوجود حلول ممكنة تساعد على بقاء حياة الإنسان ومن هذه الطرائق أو الأساليب نماذج الانحدار وبأسلوب أكثر تطوراً بحيث تلائم حالة المتغيّر التابع الذي يكون ثنائي الاستجابة، وان احد هذه النماذج واكثرها استعمالاً هو نموذج انحدار(Cox) الذي اقترحه العالم الانجليزي(Dived Cox) عام 1972م، فهو يُعدّ من النماذج المناسبة للبيانات الثنائية والذي من خلاله يتَمّ دراسة زمن البقاء من المصاب والعوامل المؤثرة على زمن البقاء للمصاب، إذْ يهدف هذا النموذج إلى معرفة عوامل الخطورة التي تؤثّر بشكل معنوي على داله الخطورة خلال مده الزمن ويتضمن تحليل البقاء لنموذج انحدار(Cox) دراسة الوقت منذ دخول المصاب الي فتره الدراسة حتى ظهور الحدث (الوفاة) والمراقبة، إذْ وجد ثلاثة انواع من نماذج الانحدار التي تطورت مع مرور الزمن وهي النماذج المعلمية ونماذج اللامعلمية و النماذج شبه المعلمية كل منها طرائق خاصه في التطبيق .

2 -مشكلة البحـــــــــــث

تُعدّ الأورام من الأمراض التي تصيب حياة الإنسان في أيّ زمان ومكان ولم يجد حلول للكشف عن اسرار هذا المرض الخبيث الذي يكمن في خلايا الإنسان لذا تَمّ استعمال طرائق الانحدار لتحليل البيانات وتفسيرها لإيجاد العلاقة بين المتغيّرات التوضيحية والمتغيّرات والمتغيّر المعتمد الذي يمثّل الوقت و المدّة الزمنية لدراسة الظاهرة إذْ تبرز الحاجة إلى استعمال أساليب انحدار أخرى لتحليل هذا النوع من البيانات وتُعدّ مشكلة تعذر تطبيق نماذج الانحدار الخطي الاعتيادي عندما تكون المتغيّرات المعتمدّة تمثّل مدّة البقاء .

3 -هدف البحـــــــث

تطبق انموذج انحدار كوكس(Cox Regression) والجاكنايف على البيانات من أجل دراسة العلاقة بين العناية المقدّمة للمريض خلال مدّة رقوده في المستشفى وبين نوع العلاج و الجنس والعمر والحالة الاجتماعية وتحديد أيّ من تلك العوامل الأكثر تأثيرا على مدّة بقاء المريض في المستشفى وتوضيح تأثير استعمال أسلوب الجاكنايف على مقاييس الانموذج شبه المعلمي كوكس.

4 - تحليل البقاء (Survival Analysis)[8][6]

هو حدث معين ومن أمثلتها وتطبيقاتها في هذا المجال الزمن الذي يُعدّ عنصراً أساساً في التحليل لغرض الدراسة بين الزمن الذي يسبق حدوث الحدث من المتغيّرات أو أكثر ويمكن أنْ تكون بين المتغيّرات الكمية أو الوصفية أو المختلطة وإنّ مِنْ أهمّ مصادرها هو انموذج كوكس الذي يستعمل في العديد من المجالات .

فتحليل البقاء يتعامل مع حدوث حالة الموت للمصاب بمرض معين في التجارب الطبية ويتعامل مع حدوث الفشل في الأنظمة الميكانيكية، فتحليل البقاء له تسميات مختلفة ففي الحقول الطبية والعلوم الصحية يطلق عليه لتحليل دوال البقاء لسبب ان الحدث الحرج هو الموت و في الدراسات الهندسية يطلق عليه مصطلح المعولية أو الموثوقية، أيّ اننا في تحليل البقاء نستعمل مصطلح (الفشل) عند حدوث أو وقوع الحدث المعني (على الرغم مِنْ أن هذا الحدث قد يكون في الواقع هو حدوث النجاح فمثلا الانتعاش أو الشفاء جراء العلاج).

ويمكن تطبيق طرائق تحليل بقاء في العديد من المجالات التطبيقية. مثل(الطب، الصحة العامة، علم الأحياء، علم الأوبئة، الهندسة، الاقتصاد، المالية، الأبحاث الاجتماعية، علم النفس، و الديموغرافيا) و تحليل البيانات أوقات الفشل عاده ما يعني معالجة واحدة من المشاكل الآتية : تقدير دوال البقاء، مقارنة معالجات دوال البقاء والتقييم العوامل المتغيرة أو الاعتماد على وقت الفشل للمتغيرات التفسيرية ويُعدّ الأسلوب الوحيد الذي يتعامل مع البيانات المراقبة وغير المراقبة.

5 - مفاهيم أساس في تحليل البقاء[1]

1-أصل الوقت

هو قياس الوقت الذي يدعى بوقت الصفر وعلى سبيل المثال التجارب الطبية إذْ ان بداية العلاج تعرف ببداية الوقت ، وانا لاحظت أنّه أصلا الوقت لا يشترط ان يكون متطابقة لمفردات الدراسة جميعها .

2-مدّة الوقت

هي مدّة الوقت الحقيقي وغالبا ما يقاس ب (Clock Time) أو (Real Time).

3- الحدث

إنّ اختلاف الدراسة الزمن أيّ (الوقت) فالحدث في الدراسات الطبية يعني الوفاة وان تعددت اسبابها .

4- وقت الدراسة ووقت المريض

إنّ المرضى لا يدخلون الدراسة في الوقت نفسه بالضبط، ولكن على مدى شهور أو الأحيان سنوات، بعد ذلك تتَمّ متابعة المرضى حتى الوفاة أو حتى نهاية الدراسة عند تحليل البيانات على الرغم مِنْ أنّ أوقات البقاء تلاحظ لبعض المرضى و أنّ البعض الأخر قد يفقد المتابعة والبعض الاخر ما زال على قيد الحياة حتى نهاية الدراسة المدّة الزمنية للمفردة تعرف بوقت الدراسة أيّ إنّ احتمال البقاء هو على الأقل (١)عند الوقت صفر ويكون صفر عند ما لانهاية.

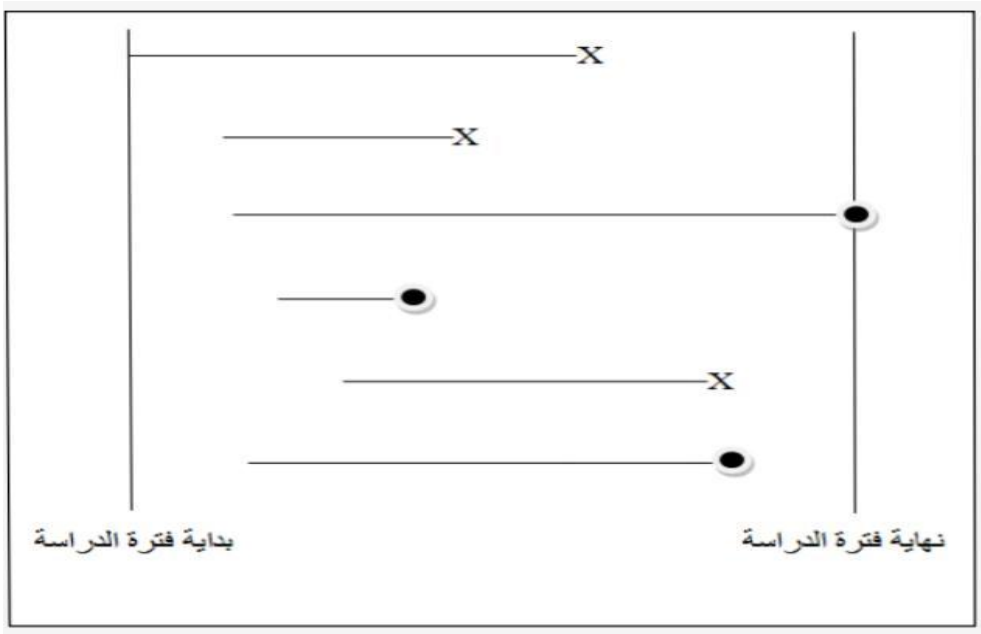
6- تعريف تحليل البقاء (Definition of Analysis Survival)

هو تحليل إحصائي يستعمل في حالة كون المتغيّر المعتمد هو الوقت حتى حدوث الحدث الآخر هذا الوقت قد يكون يوم ، شهر ، سنه منذ بداية المفردة وحتى حدوث الحدث أيّ ان أيّ تحليل يهتَمّ بمجموعة من المفردات[4]

ونلاحظ وجود العديد من المسميات لها مثل الإحصاء الطبي ويعرف بتحليل البقاء أو تحليل الأخطار Analysis Hazard ، وفي دراسات الهندسة يعرف بتحليل وقت الفشل Analysis Time Failure أو تحليل المصداقية Analysis Reliability وفي دراسات أخرى مثل علم النفس يعرف بتحليل تأريخ الحدث History Event Analysis، وتحليل المدّة Analysis Duration ، وفي علم الاقتصاد بتحليل الانتقال Transition Analysis [7].

7- بيانات البقاء (Data Survival)

هو وقت البقاء الذي يُعدّ متغيّراً حقيقياً موجب للقيم دائماً وتستعمل لوصف البيانات في حدوث الحدث ولها ثلاث شروط أساس يتَمّ تحديدها بدقّة عند دراستها وهي اساس أو أصل الوقت Origin Time ، للمدّة الزمنية Scale Time والحدثEvent [3][11].



الشكل (1) يمثل بيانات البقاء

8 -أنواع بيانات البقاء وخواصها [6]

توجد أنواع عدّة من البيانات غير الكاملة (Censored Data) تظهر عند تحليل بيانات البقاء وتتمثّل في الأنواع الآتية :

1) المشاهدة الفاشلة التامة (Complete or Failed Observation):

المشاهدة المخفقة أو الفاشلة هي تلك المشاهدة أو الوحدة والتَّي تكون وقت وقوع الحدث تَمّ قياسه بصورة مؤكدة قبل انتهاء مدّة الملاحظة ووقوع حدثها ، مثال ذلك موت فأر التجارب تحت الدراسة للتأثر مرض معّين أو توقف الماكنة وعطلها قبل انتهاء مدّة الملاحظة.

2) بيانات غُير الكاملةً يمينا(Right censored Data)

إنّ البيانات غير الكاملة يمينا تزودنا بالحد الادنى لوقت الفشل (الاخفاق) الحقيقي، ويعرف بأنّه وقت الفشل الواقع(أو الذي سيقع) عند نقطه زمنيه وبعد تحديد فترة زمنيه للملاحظة، و البيانات غير الكاملة يمينا تحدث عندما تنتهي المدّة الزمنية المحددة للملاحظة قبل أنْ تتحقق عناصر المشاهدات وتفشل، كذلك تحدث هذه البيانات غير الكاملة بعيدا عندما يكون الاخفاق أو الفشل بسبب حدث غير الحدث تحت الدراسة (مثل عطل ماكنة نتيجة انقطاع التيار الكهربائي).

3) البيانات غير الكاملة يسارا (left censored data)

المشاهدات غير الكاملة يسارا تزودنا بالحد الأدنى بوقت الفشل الحقيقي، وإذْ إنْ وقت الفشل يحدث عند نقطه زمنيه قبل المدّة الزمنية المحددة للفشل عند ذلك فإنّ فشل البيانات يسارا يحدث عندما يكون فحص الاخفاق للوحدات لا يتَمّ الا بعد مرور فترة زمنية من بداية المتابعة و عندما تفشل الوحدة فإنّنا نعرف بالضبط متى سوف تفشل.

4) البيانات غير الكاملةObservation Censored Internal) [7]

مدّة البيانات غير الكاملة هي تلك المدّة والتي فيها نعلم أن الاخفاق (الفشل)سٌيحدث بين وقتين ولكن لا نعلم بالضبط وقت حدوث الاخفاق داخل المدّة المعنٌية هذا النوع من البٌيانات يٌدعى أٌيضا (Readout Data) وأنها تؤشر في الحالات $ta \leq ft \leq tb$ الحالات التي عندها تكون الوحدات تفحص بصورة دورية .والجدول رقم (1)مثل أنواع البٌيانات لدراسة البقاء لتجربة حدث بداية مدّة الملاحظة لها عند الزمن Tb=15 ولعدد(20) مفردة وأن نهايٌة مدّة الملاحظة عن المدّة الزمنٌية . Te=152 الا أن حدوث (4)حالات فشل قبل بدأٌية مدّة الملاحظة وهذا ما اطلق علٌيه (Data Censored Left)و (5) حالات فشل في الفترات (144.8,120.4,100,33.7,15.4) وبذلك فهي تدعى Complete or Failed) Observation) وبالانتهاء مدّة الملاحظة Te=150 نجد أنّ الوحدة لم تفشل ضمن مدّة الملاحظة وهي.

	Time	Censored	Count	
بداية الفترة $T=0$ ←	T < 15	Left	4	← Left Censored Data
وقت بداية ملاحظة العطلات بعد مرور 15 ساعة على التجربة ←	15.4	1		
	33.7	1		
	100	1		
	120.4	1		
	144.8	1		
وقت انتهاء ملاحظة العطلات بعد مرور 150 ساعة على التجربة ←	T > 150	Right	11	← Right Censored Data

الشكل رقم (2) تمثيل بيانات البقاء[2]

9 - دالة المخاطرة (Hazard Function)[6]

إنّ دالة المخاطرة هي داله الاحتمال الشرطي لحصول الحدث (الموت) خلال مدّة (tΔ+t, t) بشرط أنْ يبقى الفرد حتى الزمن(t) أو أكثر ويمكن تسميته بداله الفشل أو معدّل الفشل ويرمز له بالرمز h (t) و يمكن تعبير عنه رياضيا النحو الآتي:

$$h_{(t)} = \lim_{\Delta t \to 0} p \frac{t \le T \le t + \Delta t / T \ge t}{\Delta t} \quad , t \ge 0 \qquad \dots\dots 1$$

والتي يمكن أنْ تعطى من خلال دالة التوزيع التراكمي F(t) ودالة كثافة الاحتمال f(t) كالاتي

$$h_{(t)} = \frac{f_{(t)}}{1 - f_{(t)}} \dots\dots\dots\dots\dots\dots\dots 2$$

ونلاحظ من الدالة الخصائص الآتية:

1-دائما غير سالبة (موجبة).

2-ليس لها حد أعلى بالنسبة للدالتين (دالة البقاء، دالة المخاطرة) و حالة البقاء هي الأكثر طبيعة عند تحليل البيانات لأنّها تصف مباشرة تطور البقاء لمجموعة من البيانات. و مزايا دالة الخطر هي :

1- تقيس الامكانية الاتية (الفورية) إذْ إنّ دالة البقاء مقياس تراكمي مع الوقت.

2 -تستعمل للتعرف على شكل النموذج.

3- دالة الخطر هي الوسيلة التي تساعدنا في إيجاد النموذج الرياضي لبيانات البقاء فإن نموذج البقاء يكتب في شكل حدود لدالة الخطر.

10 -انحدار كوكس (Cox Regression)

في القسم الثاني من الجانب النظري سوف نناقش نموذج تحليل بيانات البقاء الأكثر استعمالاً ، نموذج انحدار كوكب للمخاطر النسبية للعالم كوكس cox عام 1972م، إذْ إنّ بيانات البقاء تُعدّ بيانات مهمّة في المجال الطبي والحيوي وكما انها تتأثّر بعوامل مختلفة، الأمر الذي استدعى من الباحثين والمختصين في هذا المجال الجهود من أجل معرفة مده بقاء الإنسان المريض عند إصابته بمرض معين على قيد الحياة وكذلك زادت الحاجه لمعرفه العوامل الأساسية المؤثرة في بقاء المريض على قيد الحياه خلال مدّة الدراسة، وذلك من خلال استعمال وتطوير الأساليب والوسائل الاحصائية بهدف زياده الدقّة والشمولية في المعرفة، ومن هنا وصلنا إلى أسلوب الانحدار الذي شملناه في هذا الفصل بمعرفه اركانه واساسياته والخوض في تفصيل جانبه الرياضي، والجدير بالذكر في البداية ان أسلوب انحدار كوكس يستخدم بشكل واسع في المجال الطبي والحيوي، ولكن لم يقتصر فقط على مثل هذه المجالات التي تعنّى بالدرجة الاولى الامراض والانتكاسات ومرضى العلاجات البيولوجية، فتَمّ استعمال هذا الأسلوب بشكل جيد و اظهر نتائج مقبولة في جميع المجالات المختلفة[11].

نموذج محضر كوكس المخاطر النسبية يعني بالدرجة الاولى الخطر والذي تتميز بأنها غير محددة وكما أنّها نسبية ومستقلة عن الزمن لمختلف الأفراد المشاركين في الدراسة.

ويتميز نموذج انحدار كوكس بالعديد من المميزات فهو بداية نموذج شبه معلمي لديه مجموعة من الفروض الاساسية، ويُعدّ نموذجاً قوياً وجيد الاستعمال في العديد من التطبيقات وكما أنّه يتعامل جيدا مع مشكله بيانات البقاء الأساس، الا وهي بيانات المراقبة، وتَمّ التطرق للجانب الرياضي لنموذج انحدار كوكس من إذْ تقدير معالم الاختبارات المتعلقة به.

وقد اقترح هذا الانموذج العالم الانكليزي الإحصائي (David Cox) عام (1972) وهو الانموذج الأكثر شيوعا لدراسة وقت البقاء للمتغيرات التوضيحية ، لمعرفة العلاقة بين نسبة الخطر ومجموعة المتغيّرات وكما مبين بالشكل الآتي[9] :

$$h\,(t/x) = h_0\,(t) * \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \ldots\ldots 3$$

$$t \geq 0$$

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

إذْ إنّ :

$h_0(t)$: تمثّل الدالة الأساس للمخاطرة لقيم المتغيّرات التوضيحية عندما تساوي صفر (x=0) وهي دالة غير سالبة وغير محدودة .

$\beta_1, \beta_2, \dots, \beta_n$: تمثّل المعلمات المقدرة لأنموذج الانحدار عندما تزداد بمقدار وحدة واحدة

$x_1, x_2, \dots, x_n$: تمثّل المشاهدات الأساس التوضيحية .

علماً أنّ الانموذج لا يحتوي على الحد الثابت أيّ إنّه جزء من $h_0(t)$.

10-1- - شروط استعمال نموذج كوكس[10]

1-متغيّر تابع من جزأين (متغيّر ثنائي القيمة وصفي + متغيّر زمن حدوث الحدث المسبق).

2 –متغيرات مستقلة وإنْ كانت (كمّية أو وصفية أو مختلطة) ويكون لها تأثير على الظاهرة المعنية.

10 -2 - خصائص نموذج كوكس [9]

1-لا يتطلب أن تختار بعض نماذج الاحتمال المعينة لتمثّل عدد مرات البقاء .

2-نموذج شبه معلمي (semi parametric)

3 -من الممكن أنْ يستوعب كُلٍّ من القياسات المنفصلة والمستمرة من مرات الحدث .

4- من السهل دمج المتغيّرات التي تعتمد على الوقت، وهي المتغيّرات التي تتغير قيمتها على

مدار مدّة المراقبة.

5 -يعني نموذج كوكس بتأثير المتغيّرات على معدّل الخطر و لكنه يترك معدّل الخطر الأساس غير محدد.

٦ -لا تعني بمعرفة المخاطر المطلقة.

7 -يقدّر المخاطر النسبية أكثر منها المطلقة .

10 -3 -فروض نموذج كوكس[2]

- مفردات قيم المتغيّرات المفسّرة لها نفس مفردات دالة الخطر.
- الفرضية التناسلية هي اختلاف قيم متغيّر معين من مفردتين يترتب أن يكون للمفردتين دالتا خطر متوازيتان.

11- طريقة الإمكان الاعظم لتقدير معلمات انموذج انحدار Cox

وهي أول طريقة تَمّ استعمالها لتقدير معلمات انموذج *Cox* ، التي اقترحها العالم (*Cox*) عام (*1972*) ، وبسبب كثرة التعديلات التي تحتاجها لم تعتمد هذه الطريقة كثيرا ، إذْ اقترح العالم (Cox) ان تقدير المعلمات (β) يعتمد على مجموعة الوقت (t_i) التي تحدث عندها الوفاة أو الفشل إذْ يشترط وجود هذا الوقت ضمن مجموعة الخطر ح($R(t_i)$(Risk Set)) فإذا كان احتمال الوفاة أو الفشل للشخص هو[4]:

$$\frac{h_0(t_i)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}{\sum_{I\epsilon R(t_i)} h_0(t_i)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)} \dots\dots\dots\dots 4$$

واذا كـانت عناصر دالة الإمكان الشرطي تتمثّل بكُل فشل فإنّه بعد اخذ اللوغاريتم لها تصبح بالشكل الآتي :

$$L(\beta) = \sum_{i=1}^{k} (\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) - \sum_{i=1}^{k} \log \left[\sum_{I \in R(t_i)} (\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \right] \dots\dots\dots 5$$

وللحصـول على تقديـرات للمعلمـات (β) نـقـوم بتعـظـيـم الـدالـة ثُـم تطبيـق طـريقة نيوتن رافسون.

12- طريقة الجاكنايف Jackknife Method

إنّ الجاكنايف هو أسلوب لا معلمي وطريقة من طرائق إعادة المعاينة ، ويهدف إلى تقليل التحيز في تقدير المعلمات بنماذج الانحدار إلى اقل ما يمكن، وعليه فأن أسلوب الجاكنايف يعرف بأنّه تلك الطريقة التي تقوم أساساً على استبعاد جزء من البيانات في كل مرة عند إجراء عملية التقدير وتَمّ معالجة الجزء المستبعد بحالتين، أمّا إهمال الجزء المستبعد من البيانات بشكل كلي عند إجراء عملية التقدير ، أو أخذ الجزء المستبعد من البيانات بنظر العناية عند إجراء عملية التقدير إذ يتَمّ احتساب وزن هذا الجزء وإدخاله في عملية التقدير ، وبالإمكان تطبيق أسلوب الجاكنايف في تقدير معلمات نماذج الانحدار الخطية وغير الخطية وسنتطرق في هذا المبحث إلى طريقة الجاكنايف الاعتيادية للتقدير كما في المعادلة[11]:

12-1 - طريقة الجاكنايف الاعتيادية

The Ordinary Jackknife Method

بالرجوع إلى المعادلة (2-4) وعلى فرض أن هناك عينة عشوائية هي $x_1, x_2, \dots, x_n$ ولنفرض أن $\hat{\beta}$ هو التقدير الذي تَمّ الحصول عليه للمعلمة (β) من تلك العينة ، وعلى أساس هذه الطريقة يتَمّ تجزئة العينة ذات الحجم (n) من المفردات إلى(N) من(المجاميع) وبحجم (m) من المشاهدات بحيث إنّ (n=Nm) ويتَمّ استبعاد (m) من مشاهدات العينة الأصلية بشكل عشوائي ويتَمّ استخراج (β^*_{-i}) التي هي تقدير للمعلمة (β) عند استبعاد المجموعة رقم (i) من المجاميع الجزئية ، فتصبح معادلة في نموذج انحدار كوكس بالاعتماد على أسلوب الجاكنايف بالشكل الآتي[11] :

$$\hat{\beta}_i = N\hat{\beta} - (N-1)\beta^*_{-i} \dots\dots\dots 6 \qquad i = 1, \dots, N$$

والذي يدعى بالقيم الكاذبة(الوهمية)ً Pseudo – values لتقدير الجاكنايف والوسط الحسابي لهذه التقديرات يكون :

والتي تمثّل تقديرات الجاكنايف للمعلمة (β) و N تمثّل عدد المجاميع

$\hat{\beta}$: تقديرات معلمات أنموذج دالـة الانحدار بالاعتمـاد على البيانـات المتاحة جميعها لـ (n) من المشاهدات باستعمال طريقة المربعات الصغرى الموزونة، وبالإمكان إيجاد تباين تقديرات الجاكنايف وبحسب الصيغة الآتية[20] -:

$$VG = \frac{1}{N(N-1)} \sum_{i=1}^{N} [\hat{\beta}_i - J(\hat{\beta})][\hat{\beta}_i - J(\hat{\beta})]' \ldots\ldots\ldots\ldots 7$$

ثُمّ يتَمّ استخراج MSEعلى وفق الصيغة الآتية :

$$JMSE = \frac{\sum_{i=1}^{n} \hat{e}_i^2(J)}{n-k-1} \ldots\ldots\ldots\ldots 8$$

إذ إنّ:

$\hat{e}_i = Y_i - \hat{Y}_i$ وتمثّل البواقي.

13- اختبـار معنـوية المعلمات المقدرة[9]

تستعمل احصاءات والـد (Wald Statistic) لاختبار معنـوية المعالم المقـدرة وفق الصيغـة الآتية :-

$$Z_j = \left[\frac{\hat{\beta}_j}{S.E_{\hat{\beta}_j}}\right]^2 \quad \ldots\ldots\ldots\ldots 9$$

إذْ إنّ :

$\hat{\beta}_j$:- المَعلمة المقدّرة لأنموذج انحدار Cox

$S.E_{\hat{\beta}_j}$:- الخطـأ المعياري للمَعْلَمة المقدَرة لأنموذج انحدار Cox

وبدرجة حرية واحدة وجد أنّ البيانات تتبع توزيع (x^2) وإيجـاد مدّة الثقـة للمعلمة المقـدرة بـاستعمال الصيغـة الآتية :-

$$\hat{\beta}_j \mp Z_{\alpha/2} S.E_{\hat{\beta}_j} \quad \ldots\ldots\ldots\ldots 10$$

14- تقييم القـوة التفسيريـة للنمـوذج [3]

ويتَمّ ذلك من خـلال احتساب قيمة (R^2) التي تحسب حسب الصيغة الآتيـة :-

$$R^2 = 1 - \exp\left[\frac{2}{n}\left(\log L_0 - \log L_p\right)\right] \quad \dots\dots\dots\dots 11$$

إذْ إنّ .

L_0 :- تمثّل مقدار قيمة الإمكان الاعظم للأنموذج الذي يحتوي على القيمة الثابتة فقط .

L_p :- تمثّل مقدار قيمة الإمكان الاعظم للأنموذج الذي يحتوي على كل المتغيّرات التوضيحية .

n :- هي عدد المشاهدات اذ كانت مراقبة ام لا .

15- الجانب التطبيقي

15-1-- وصف البيانات

تَمّ تطبيق الموضوع على بيـانات واقعيـة أخذت من مستشفى الزهراء التعليمي في محافظة واسط ، من خـلال مراجعـة ردهـة الطوارئ ومراقبة حياة (100) شخص مصاب بسرطان الرئة ولما لها من اهمية لحياة الإنسان وهي احدى الأعضاء المهمة والمحمية في القفص الصدري ويوجد زوج منها وتُعدّ الجزء الاساس للجهاز التنفسي للإنسان وان سرطان الرئة هو الذي ينشأ في الخلايا داخل الرئة وعادتاً ما يصيب الاعمار فوق 45 ويكون على نوعين اساس هما سرطان الرئة ذو الخلايا الكبيرة و الخلايا الصغيرة .

إذْ تَمّ تحليل البيانات ودراستها باستعمال برنامج SPSS لغرض استخراج نتائجها الخاصة بالأنموذج وأهمّ الإحصاءات الوصفية ودونت النتائج علما ان المتغيّرات التالية تمثّل البيانات الحقيقة وكما يأتي .

γ: يمثّل مدّة الرقود في المشفى (العناية بالمريض) المصابين بسرطان الرئة، وهو يمثّل المتغيّر التابع لأنموذج انحدار Cox و الجاكنايف .

أيّ إنّ دالة البقاء تمثلت بشفاء المريض أو استقرار الحالة الصحية وخروج المصاب من المشفى واعطيت كل حالة منها ما يلي (1) الشخص على قيد الحياة (0) الشخص متوفي

المتغيّرات المستقلّة الآتية تمثّل المتغيّرات المـراد قيـاس تـأثيرها على المتغيّر التابع

X_1 : حالة الوفاة و المراقبة وصنف إلى (1) وفاة و (2) فقدان المراقبة .

X_2 : نوع العلاج وصنف إلى (1) معالج بالإشعاع و (2) معالج كيميائي .

X_3 : عمر الشخص .

X_4 : جنس الشخص .

X5: الحالة الاجتماعية للمريض وصنف إلى (1) متزوج (2) غير متزوج.

وتَمّ تحليل البيانات باستعمال برنامج (SPSS) ودونت النتائج الخاصة بالنموذج والإحصاءات الوصفية في الجداول التالية

2-15- الإحصاءات الوصفية

الجدول (1) يمثّل الإحصاءات الوصفية للنسب المئويـة والمتوسّطات والانحـراف المعيـاري للمتغير(Y).

Y	Frequency	Percent %	Mean	SD
1-10	25	25%	2.4300	1.06605
11-20	26	26%		
21-30	30	30%		
31-40	19	19%		
Total	100	100%		

تبين من النتائج اعلاه ان اقل نسبة للأشخاص الراقدين في المشفى تتراوح بين (31-40) هي 19 % في حين سجلت اعلى نسبة للراقدين بين (21-30) بنسبة 30% وهي تشكل تقريباً أكثر من ثلاث ارباع عينة البحث وهي نسبة عالية جداً

الجدول (2) يمثّل الإحصاءات الوصفية والنسب المئويـة والمتوسّطات والانحـراف المعيـاري للمتغير(X1)

Died(X1)	Frequency	Percent %	Mean	SD
censored	46	46%	0.54	0.501
died	54	54%		
Total	100	100%		

يلاحظ من الجدول اعلاه ان نسبة مراقبة المرضى هي 46% في حين بلغت نسبة الوفيات 54 % تمثّل نسبة كبيرة من نسبة المراقبة للمرضى وذلك بسبب أمّا فقدان السيطرة عليهم أو خروجهم من المستشفى

الجدول (3) يمثّل الإحصاءات الوصفية والنسب المئوية والمتوسّطات والانحراف المعياري لمتغير(X2)

Drug(X2)	Frequency	Percent %	Mean	SD
control	20	20%	0.80	0.402
treatment	80	80%		
Total	100	100%		

تبين من الجدول اعلاه ان نسبة الذين تَمّ علاجهم كيميائياً هي 80% وهي نسبة كبيرة مقارنَة بالمرضى الذين تَمّ علاجهم في الاشعاع بنسبة 20% وتَمّ شفاؤهم وهذا يدّل على ان العلاج الكيميائي افضل بكثير

الجدول (4) يمثّل الإحصاءات الوصفية والنسب المئوية والمتوسّطات والانحراف المعياري للمتغير(X3)

Age (x6)	Frequency	Percent %	Mean	SD
41-50	29	29%	3.84	0.63118
51-60	58	58%		
61--70	13	13%		
Total	100	100%		

تبين من الجدول اعلاه أنّ نسبة اعداد المصابين أو المرضى 58 للأعمار (41-50) وهي اعلى نسبة بالمقارنة بالنسب الأخرى أو الأعمار الأخرى

الجدول (5) يمثّل الإحصاءات الوصفية والنسب المئوية والمتوسّطات والانحراف المعياري للمتغير(X4)

Sex(X3)	Frequency	Percent %	Mean	SD
Male	63	63%	1.37	0.49889
Female	37	37%		
Total	100	100%		

تبين من الجدول اعلاه ان نسبة الاصابة في اعداد الذكور وهي 63% وهي نسبة كبيرة مقارنة بنسبة الإصابة عند الاناث 37 %

الجدول (6)يمثّل الإحصاءات الوصفية والنسب المئويـة والمتوسّطات والانحـراف المعيـاري للمتغير(X5)

X5	Frequency	Percent %	Mean	SD
Marred	56	56%	1.44	0.49889
unmarred	44	44%		
Total	100	100%		

وتبين من نتائج الجدول اعلاه أنّ نسبة الاصابة بأعداد المتزوجين 56%اعلى من نسبة الاصابة بأعداد الغير متزوجين 44 % وهي نسبة متقاربة نوعاً ما .

ومن أجل الحصول على نتـائج انموذج انحدار *Cox* التقليدي و الجاكنايف، تَمّ تحديد عينـة الجاكنايف العينـة البسيطـة ، وان عددها (100) ، ومن تطبيق أنموذج الانحدار حصلنا على النتائج الآتية :-

3-15 -نتائج تطبيق انموذج انحدار *Cox* التقليدي

يستعمل انموذج انحدار Cox لتحديد أو معرفة العلاقة بين المتغيّرات و معرفـة تـأثير المتغيّرات المستقلّة المحـددة مسبقاً على مدّة رقود المريض في المستشفى المصابين سرطان الرئة ، ولتطبيق انمـوذج انحـدار Cox يجب معرفـة هل ان الانمـوذج معنوي ام لا لذا يجب وضع الفرضيات الآتية :

فرضية العدم :- المتغيّرات المستقلّة لا تؤثّر على المتغيّر التابع .

الفرضية البديلة :-. المتغيّرات المستقلّة تؤثّر على المتغيّر التابع

$$H_0: \beta = 0$$
$$H_1 : \beta \neq 0$$

إذْ تبين نتائج اختبار (Chi-Square) أنّ قيمة p.value تساوي (.000) وهي اقل من مستوى المعنوية ($\propto = 0.05$) وهـذا يـؤكد معنوية الانموذج أيّ اننا نرفض فرضية العدم أيّ ان المتغيّرات المستقلّة لا تؤثّر على المتغيّر التابع وتقبل الفرضية البديلة أيّ ان المتغيّرات المستقلّة تؤثّر وبشكل مباشر على المتغيّر التابع .

وان مـلاءمة الانمـوذج للبيانات يمكـن معرفتها من حساب قيمة (R^2) لأنمـوذج انحـدار Cox باستعمال المعادلة (-15 2) إذْ وجدت قيمتها التي تساوي .

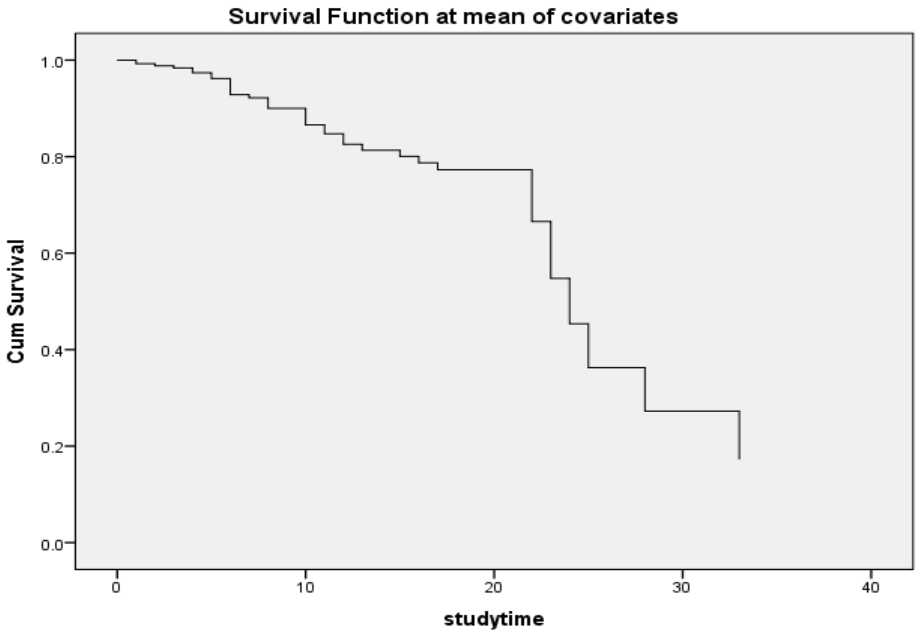
$$R^2 = 1 - \exp\{\frac{2}{182}((432.927) - (377.284))\}$$
$$R^2 = 0.66212$$

وهـذه القيمة (66.212%) تشير إلى التغير الحاصل للمتغيّر التابع ويتَمّ تفسيره من خلال انموذج انحدار Cox .

الجدول (7) يمثّل نتائج أنموذج انحدار Cox التقليدي

المتغيرات	معاملات انحدار Cox β	الخطأ المعياري S.E ي	احصاءة Wald	درجات الحرية Df	القيمة الاحتمالية Sig.	القيمة المتوقعة Exp(β)
نوع العلاج	-2.604	0.372	48.934	1	0.000	.074
العمر	0.114	0.030	14.237	1	0.000	1.121
الجنس	-0.617	0.305	4.098	1	0.043	0.540
الحالة الاجتماعية	0.048	0.286	0.029	1	0.865	1.050

من ملاحظة الجدول اعلاه يتبين ان نوع العلاج للقيمة الاحتمالية لإحصاءات Wald للمتغيّر اقل من مستوى المعنوية وهذا يدّل على ان نوع العلاج له تأثير معنوي على المتغيّر التابع أمّا بخصوص القيمة الاحتمالية لمتغيّر العمر فأنها اقل من مستوى المعنوية أيّ انه لها تأثير معنوي في المتغيّر التابع وتبين ايضا ان متغيّر الجنس ان له تأثير معنوي في المتغيّر التابع أمّا بخصوص الحالة الاجتماعية ان القيمة الاحتمالية لها أكبر من مستوى المعنوية وهذا يدّل على ان الحالة الاجتماعية المتغيّر الوحيد الذي ليس له تأثير معنوي على المتغيّر التابع .



الشكل (3) يوضح دالة البقاء للمرضى المصابين

4-15 - نتائج تطبيق أسلوب الجاكنايف

عند تطبيق هذا الأسلوب بالاعتماد على طريقة الإمكان الاعظم نحصل على نتائج مقاربة للنتائج باستعمال انموذج انحدار Cox التقليدي وبقيت النتائج دونت في جدول رقم (8)

الجدول (8) يمثّل نتائج أنموذج انحدار الجاكنايف

المتغيرات	تقدير القيم الكاذبة		تقدير المعلمات		اختبار	القيمة الاحتمالية
	B	Std. Error	Beta	Std. Error	T	Sig
نوع العلاج	10.780	1.668	0.582	0.090	6.465	0.000
العمر	-0.414	0.148	-0.206	0.073	-2.799	0.006
الجنس	2.490	1.722	0.104	0.072	1.446	0.151
الحالة الاجتماعية	-0.428	1.575	-0.020	0.072	-0.272	0.787

تبين من خلال النتائج وحساب قيمة اختبار (t test) تبين ان القيمة الاحتمالية لمتغيّر (نوع العلاج) اقل من مستوى المعنوية ، وبين ان نوع العلاج ذو تأثير معنوي على المتغيّر التابع ، أمّا القيمة الاحتمالية لمتغيّر (العمر) اقل من مستوى المعنوية ، وهذا يدّل على تأثيرها المعنوي ايضاً على المتغيّر التابع ، ومتغيّر (الجنس) يبين ان القيمة الاحتمالية لها أكبر من مستوى المعنوية أيّ ان ليس لها تأثير معنوي على المتغيّر

التـابـع ، أمّا لمتغيـر(الحالة الاجتماعية) تبين ان القيمة الاحتمالية لها أكبر مـن مسـتـوى المـعـنـويـة ، اذن الحالة الاجتماعية والجنس هما الوحيدان ليس لـهم تـأثـيـر معنـوي على المتغيّر التابع .

ومن خلال القيم في الجدول رقم (3) يمكـن كتابـة انمـوذج انحـدار الجاكنايف بالشكـل الآتي :-

$$\hat{h}(t,x) = \widehat{h_0}(t)\text{Exp}\{(0.582X_1) + (-0.206\,X_2) + (0.104X_3) + (-0.020X_4)$$

ويتبين من المعادلة اعلاه زيادة المتغيرين x_1 وx_2 بمقدار وحـدة واحدة ما يؤدي إلى تناقص دالـة المخاطـرة ، وان زيـادة المتغيرين x_3 و x_4 بمقدار وحـدة واحـدة تـؤدي إلى زيادة دالـة المخاطرة .

15-5- المقارنة بين الأسلوب الجاكنايف و coxالتقليدي

يجب توضيح تأثير استعمال الأسلوب الجاكنايف على مقاييس الانموذج لذا نقوم بمقارنة الأخطاء المعيارية لا نموذج انحدار Cox التقليدي مع قيمها لنموذج انحدار الجاكنايف وقيم المعنوية (P-Value)، علما اننا لم نستعمل قيم (R^2) في المقارنة لأننا نحصل على نفس القيم عند استعمال الاسلوبين

الجدول (9) يمثّل قيم المعنوية والأخطاء المعيارية

المتغيـرات	انموذج انحدار Cox التقليدي P-Value	انموذج انحدار Cox الجاكنايف P-Value	الخطأ المعياري لنموذج انحدار Cox التقليدي Std.Errer	الخطأ المعياري لنموذج انحدار Cox الجاكنايف Std.Error
نوع العلاج	0.000	0.000	0.372	0.090
العمر	0.000	0.006	0.030	0.073
الجنس	0.043	0.151	0.305	0.072
الحالة الاجتماعية	0.865	0.787	0.286	0.072

ان قيـم (P-Value) لأنمـوذج انحـدار Cox والجاكنايف متساويان عند المتغيرين (نوع العلاج والعمر) أمّا عند متغيّر الجنس فانحدار cox اقل من الجاكنايف ،أمّا عند متغيّر الحالة الاجتماعية فهو في أسلوب الجاكنايف اقل من انحدار cox وقيم الأخطاء المعيارية لأنموذج انحدار Cox عند متغيّر نوع العلاج والجنس والحالة الاجتماعية يكون أكبر من أسلوب الجاكنايف وهذا يدّل على دقّة أسلوب الجاكنايف ،أمّا عند متغيّر العمر فيحدث العكس، واستبعاد متغيّر الجنس يمكن القول بأن استعمال أسلوب الجاكنايف يعطي نتائج أكثر دقّة .

16- الاستنتاجات :-

1 – اوضحت النتائج ان نسبة الراقدين والاعلى هي للأعمار (21 - 30) وبنسبة 30% وتُعدّ كبيرة جداً بالنسبة للعينة التي تَمّ اخذها في حين كانت اقل نسبة هي 19% للأعمار (31- 40) فضلاً عن نسبة عدد الوفيات 54% ونسبة الشفاء للمصابين باستعمال العلاج الكيميائي 80% افضل من العلاج باستعمال الاشعاع وبنسبة الإصابة عند الذكور اعلى من الإناث ونسبة الإصابة عند الازواج ايضا اعلى منها عند غير المتزوجين .

2- وبيّنتْ نتائج انموذج انحدار (Cox) التقليدي و الجاكنايف أنّ المتغيّر للحالة الاجتماعية المتغيّر الوحيد ليس له تاثير معنوي على مدّة بقاء المريض في المستشفى أمّا متغيّر الجنس فكان فقط في أسلوب الجاكنايف غير معنوي.

3- ان قيم الأخطاء المعيارية لأنموذج انحدار Cox عند متغيّر نوع العلاج والجنس والحالة الاجتماعية يكون أكبر من أسلوب الجاكنايف وهذا يدّل على دقّة أسلوب الجاكنايف ،أمّا عند متغيّر العمر فإنّه يحث العكس، واستبعاد متغيّر الجنس يمكن القول بأن استعمال أسلوب الجاكنايف يعطي نتائج أكثر دقّة .

4- تبين من نتائج المقارنة بين الطرائق ان افضل طريقة لتمثيل البيانات الحقيقية هو انموذج انحدار الجاكنايف.

17 -التوصيات :-

1 – الاخذ بنتائج البحث وتطبيق الطرائق على بيانات أخرى أكثر ومتغيرات متعددة

2- المراجعة المنتظمة للطبيب واجراء التحاليل اللازمة ولا سيما للأعمار التي بيّنتْ في البحث .

2 – ايجاد متغيرات جديدة مع زيادة حجم العينة و قد تكون لها تأثير معنوي على مدّة بقاء المريض بأيّ مرض سرطاني أو مرض يهدد بصحة الإنسان في المستشفى له علاقة وثيقة بالمرض.

3 – دراسة أسلوب الجاكنايف (Jackknife) نماذج البو تستراب ويفضل ان يكون مع في نماذج الانحدار اللوجستي.

المصادر

1- ابراهيم، سجى ميلاد،(٢٠١٨)م " انحدار كوكس مع تطبيق عملي باستعمال spss"، بحث تخرج مقدم لاستكمال متطلبات الحصول على درجه البكالوريوس، كليه الإدارة والاقتصاد، جامعة سبها.

2-السباح ، شروق عبد الرضا سعيد ،(2009)م " بناء انموذج انحدار لوجستي معدّل لحياة الاطفال الخدج في محافظة كربلاء " ، اطروحة دكتوراه مقدمة إلى قسم الإحصاء ، كلّية الادارة والاقتصاد ، جامعة بغداد .

3-السباح، شروق عبد الرضا سعيد" استعمال أسلوب البوتستراب في نماذج شبه المعلمية (نموذج كوكس)، رسالة ماجستير مقدمه إلى قسم الاحصاء، كليه الإدارة والاقتصاد، جامعه كربلاء،

4- الاعظمي، يسرى طارق،(1999)م " دراسة وتحليل نموذج كوكس Cox للانحدار مع تطبيق عملي"، رسالة ماجستير مقدمه إلى قسم الاحصاء، كليه الإدارة والاقتصاد، جامعه بغداد .

5-احمد، نمران سلطان علي،(2013)،" استعمال نموذج الانحدار اللوجستي وانحدار كوكس(Cox) العوامل المؤثرة في الإصابة بالسرطان الدم في المؤسسة الوطنية لمكافحة السرطان، مجله العلوم الادارية والاقتصادية، كليه الاقتصاد والعلوم الادارية، جامعه عدن.

6-كمال، اسماء جميل سالم،(٢٠١٧)م " استعمال الانحدار اللوجستي وانحدر كوكس في تحديد العوامل المؤثرة على السمنة لدى السيدات المتزوجات في قطاع غزه"، رساله ماجستير مقدمه إلى قسم الاحصاء، كليه الاقتصاد والعلوم الإدارية، جامعه الازهر- غزه.

ثانياً : المـــــصادر ألاجــنبية (Foreign References)

7- Abdulla, fayyadh ail, efficient regression estimator for ordinary least squares and Jackknife based on Jackknife algorithm by deleting one case, statistics department, college of administration and economics, university of wasit .

8-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. .

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