The Administration & Economic College Journal
For Economics & Administration & Financial Studies
Vol.13,N2, P P. 249-268
ISSN PRINT 2312-7813
ISSN ONLINE 2313-1012
مجلة كلية الإدارة والاقتصاد للدراسات الاقتصادية
والإدارية والمالية
المجلد ١٣، العدد٢ ٢٠٢١ ص ص٢٤٩-٢٦٨

The use of logistic regression to determine the most important causes of the spread of colon cancer for the purpose of diagnosis and prevention with practical application.

استخدام الانحدار اللوجستي لتحديد أهم أسباب انتشار مرض سرطان القولون لغرض التشخيص والوقاية مع تطبيق عملي .

*م.م أمير كامل حمزة قرغولي

Abstract:-

The latest development of the concept of logistic regression and its uses in medical research, or what is known as the logit regression model, is a qualitative leap in the process of data analysis, access to the desired results of research, and knowledge of the causes of the spread of diseases and how to treat them in advance before reaching severe cases of disease that require intervention with chemotherapy or even intervention And since Cox's coefficient is one of the very important treatments in the field of medical tests, the idea came to study the logistic regression and with the help of the Cox regression coefficient, where a sample taken from the Medical City in Baghdad was studied with a size of (130) views, with (72 males) and (58 females) diagnosed. 100) Watching people with colon cancer and (30) watching people who are not infected with the disease and have similar symptoms to the symptoms of the disease, where (10) independent variables that have a direct effect on the studied variable adopted (colon cancer) were studied, and logistic regression was used to reach the analytical results

* الجامعه التقنية الوسطى / المعهد الطبي التقني – المنصور

and it was applied with the help of the ready-made statistical application (SPSS), and also was used Cox-regression test to explain the phenomenon, where its value appeared after conducting the test, which is equivalent to (0.424), which is a good percentage as an explanation for the phenomenon studied. Then the factor (x10 (glandular appendages)) comes in the second place with a strong effect and an average error value of (1.014) and the results continue according to what is shown in Table No. (5) from the applied side of this research.

المستخلص :-

احدث تطور مفهوم الانحدار اللوجستي واستخداماته في البحوث الطبية او مايعرف بنموذج انحدار لوجيت (logit) نقلة نوعية في عملية تحليل البيانات والوصول للنتائج المرجوة من البحوث ومعرفة اسباب انتشار الامراض وكيفية علاجها بشكل مسبق قبل الوصول للحالات المستعصية من المرض التي تتطلب التدخل بالعلاج الكيمياوي او حتا التدخل الجراحي وبما انه معامل كوكس من المعاملات المهمة جدا في مجال الاختبارات الطبية جاءت الفكرة لدراسة الانحدار اللوجسني وبمساعدة معامل انحدار كوكس حيث تم دراسة عينة مأخوذة من مدينة الطب في بغداد بحجم (١٣٠) مشاهدة بواقع (٧٢ ذكور) و (٥٨ اناث) وتم تشخيص (١٠٠) مشاهدة مصابين بمرض سرطان القولون و (٣٠) مشاهدة غير مصابين بالمرض ولديهم اعراض مشابهة لاعراض المرض حيث تم دراسة (١٠ متغيرات) مستقله ذات تأثير مباشر على المتغير المدروس المعتمد وهو (سرطان القولون (colon canser)) وتم استخدام الانحدار اللوجستي للوصول الى النتائج التحليلية وتم التطبيق بمساعدة التطبيق الاحصائي الجاهز (spss) ، وايضا تم استخدام اختبار انحدار كوكس (cox-regression) لتفسير الظاهرة حيث ظهرت قيمته بعد اجراء الاختبار بما يعادل (٠,٤٢٤)وهي نسبه جيده كتفسير للظاهرة المدروسة وايضا اتضح ان العامل (x3 (السمنة المفرطه)) اكثر العوامل المدروسة تأثيراً وبقيمة متوسط خطأ تعادل (1.034) ومن ثم يأتي العامل (x10 (الزوائد الغدية)) بالمرتبة الثانية بتأثير قوي وبقيمة متوسط خطأ (1.014) وتستمر النتائج وفق ما موضح في الجدول رقم (٥) من الجانب التطبيقي من هذا البحث .

المقدمة .

معظم حالات سرطان القولون والمستقيم تحدث بسبب عوامل ذات علاقة بنمط الحياة وتقدم السن مع وجود عدد قليل من الحالات بسبب عوامل وراثية غير معروفة. تشمل عوامل الخطورة المسببة لهذا المرض أيضا النظام

الغذائي والسمنة والتدخين وقله النشاط البدني العوامل الغذائية التي تزيد خطر الإصابة بهذا المرض تشمل تناول اللحوم الحمراء واللحوم المعالجة (مثل اللحوم المملحة والمجففة والمخمرة والمدخنة واللحوم المعلبة والصلصات التي تحتوى على اللحوم) وكذلك أمراض المشروبات الكحولية. من العوامل الخطيرة أيضاً التهابات الأمعاء وتشمل مرض كرون والتهاب القولون التقرحي ، بعض الاضطرابات الجينية الموروثة التي قد تسبب الإصابة بسرطان القولون والمستقيم تشمل مرض السلائل الورمي الغدي العائلي ومتلازمة لينش إلا أنها لا تمثل سوى أقل من ٥ ٪من الحالات ، عادة ما تبدأ الإصابة على شكل ورم حميد (غالباً على شكل زائدة لحمية) وتتحول مع مرور الوقت إلى سرطانة.

يتم تشخيص سرطان الأمعاء عن طريق الحصول على عينة من القولون (تسمى خزعة) من خلال إجراء التنظير السيني أو منظار القولون. يعقب ذلك إجراء التصوير الطبي لتحديد مدى انتشار المرض. يعد فحص السرطان إجراء فعالاً لمنع وتقليل الوفيات الناجمة عن سرطان القولون والمستقيم، وينصح بإجراء الفحوصات الطبية ابتداء من سن الخمسين وحتى الخامسة والسبعين. أثناء فحص القولون بالمنظار يمكن إزالة الزوائد اللحمية الصغيرة إن وجدت، أما إذا تم اكتشاف زوائد كبيرة أو ورم فيتم استئصال خزعة لإخضاعها للفحص لمعرفة إن كانت سرطانية ، تساعد بعض العقاقير الطبية مثل الأسبرين وغيره من مضادات الالتهاب اللاستيرويدية على تقليل مخاطر الإصابة بهذا المرض، غير أنه لا ينصح باستخدامها بشكل منتظم لهذا الغرض بسبب آثارها الجانبية ، وبالعودة تأريخياً يعتبر الطبيب والإحصائي جوزيف بيركسون أول من عرف مفهوم الانحدار اللوجستي وطبقه في مجال الإحصاء الحيوي سنة ١٩٤٤ ونشرت أول ورقة بحثية تعنى بدراسة الاحصاء الحيوي يذكر فيها الانحدار اللوجستي باسمه في مجلة الجمعية الإحصائية الأمريكية كماً ان الانحدار الوجستي استخدم سابقا في مجال النمو السكاني حيث كانت تسمى بدالة النمو (growth function) حيث تم استخدامها اول مره في عام ١٩٢٠ ومن ثم تحول اسمها الى (logistic function) وتم تعميمها على باقي المجالات العلمية ومنها الطبية وكما هو الحال في بحثنا هذا ومن الجدير بالذكر انه يتم علاج سرطان القولون والمستقيم عن طريق الجمع ما بين عدة وسائل علاجية وهي الجراحة والعلاج الإشعاعي والعلاج الكيميائي والعلاج الموجه ، تكون السرطانات المحصورة داخل جدار القولون أكثر قابلية للعلاج عن طريق الجراحة بينما يصبح العلاج غير ممكن في حالة انتشار الإصابة وانتقالها إلى أعضاء صعبة أخرى في الجسم

اولا - مشكلة البحث

تشير العديد من البحوث الطبية إلى أن هناك علاقة وثيقة بين الجوانب التشخيصية للامراض و بين بعض الدارسات والتطبيقات الاحصائية المستخدمة في هذا المجال وبما يساعد على سرعة التشخيص المبكر للامراض الصعبة .

لكن على الرغم من ذلك نجد ان اغلب الدراسات والتطبيقات الاحصائية على المستوى العربي عموماً والمستوى المحلي خصوصا التي تناولت موضوع إستخدام التطبيق الاحصائي لنموذج الانحدار اللوجستي في المجال الطبي قد ركزت على دراسة العوامل المؤثرة في الاصابة بالامراض المختلفة (اي دراسة عوامل الخطورة ومسببات الامراض المختلفة) ولم تتناول الجانب التشخيصي للامراض

(وهذا ما يمثل مشكله البحث حيث تم استخدام الانحدار اللوجستي للكشف عن التشخيص والمسببات على حد سواء)

لذا جاء هذا البحث بهدف توضيح العلاقة بين العوامل المساعدة في تشخيص حالات الاصابة بمرض(سرطان القولون) والحالات المرضية المشكوك بإصابتها بهذا المرض ، وتحديد أولوية تلك العوامل بإستخدام أنموذج الانحدار اللوجستي واستخدام اختبار كوكس .

ثانياً - اهمية البحث

لقد تم الاعتماد على عدد من العوامل ودراستها (السمنة المفرطة ،الوراثة ، الكحول ، علاج شعاعي لسرطان سابق ، تدخين ، امراض مزمنة ، امراض معوية التهابية ، زوائد غدية) حيث ان هذه العوامل يعتقد انها من المسببات الاساسية للاصابة بمرض سرطان القولون لذا جاءت اهمية هذا البحث لغرض الاستفادة من الاعراض وتحديد العوامل المؤثره للاصابة وتشخيص المرض قبل الوصول الى حالة المرض والانتقال الى مرحلة العلاج

ثالثاً - هدف البحث

يهدف البحث الى استخدام إنموذج الانحدار اللوجستي وبالاعتماد على اختبار كوكس لتوضيح العلاقة بين العوامل المؤثرة في تشخيص حالات الاصابة بسرطان القولون مثل ، عمر المريض ، جنس المريض الاعراض الرئيسة التي يشكو منها المريض(والحالات المرضية المشكوك بإصابتها بهذا المرض . وتحديد أولوية تلك العوامل بمعنى تحديد العوامل الأكثر تأثيرا في إمكانية تشخيص حالات الإصابة بهذا المرض قبل اجراء التحليلات المختبرية المتعددة وبنفس الوقت الوقاية من الاصابة بالمرض من خلال تشخيص الاعراض الممكنه للانتقال الى حالة الاصابة من خلال تحديد العوامل المؤثرة واتخاذ الحيطة والحذر من هذه العوامل ومحاولة الابتعاد عنها قبل الانتقال الى مرحلة الاصابة .

رابعاً - فرضية البحث

H_0: فرضية العدم / بمعنى عدم وجود علاقة بين المتغير التابع والمتغير المستقل

H_1 : الفرضية البديلة / عدم قبول فرضية العدم بمعنى وجود علاقة بين المتغير التابع والمتغير المستقل

خامساً - عينة البحث

تمثلت عينه البحث بأخذ ١٣٠ مشاهدة لأشخاص راقدين في مستشفى مدينة الطب مشكوك بإصابتهم بالمرض بواقع (٧٢ ذكور ، ٥٨ اناث) تم الحصول عليها بصورة عشوائية ولسلسلة زمنية من سنه (٢٠١٧- ٢٠٢٠) في مدينة الطب (مستشفى الأورام التعليمي) حيث تم جمع البيانات الخاصة للعينة المذكوره اعلاه وتبين ان ١٠٠ مشاهدة مصابين بمرض سرطان القولون و ٣٠ مشاهدة غير مصابين بالمرض .

سادسا – أداة جمع البيانات وتحليلها

تم جمع البيانات بشكل عينة عشوائية مأخوذه من مجتمع مدينة الطب في بغداد / باب المعظم وتم التحليل بإستخدام التطبيق الإحصائي الجاهز (spss) .

الجانب النظري

أولا- مرض سرطان القولون

ثانيا – الانحدار اللوجستي

ثالثا – اختبار كوكس

اولا - سرطان القولون (colon cancer)[٢،٤،٥]

- ويعرف بسرطان الأمعاء- هو أحد أنواع مرض السرطان وينشأ من القولون أو المستقيم (أجزاء من الأمعاء الغليظة) نتيجة لحدوث نمو غير طبيعي للخلايا التي لديها القدرة على المهاجمة والانتشار إلى الأعضاء الأخرى في الجسم ، ومن علامات وأعراض هذا المرض وجود دم في البراز وتغير في حركة الأمعاء (إما إمساك أو إسهال) وفقدان الوزن والشعور بالتعب والإرهاق طوال الوقت .

العلامات والاعراض

علامات وأعراض سرطان القولون والمستقيم تعتمد على مكان وجود الورم في الأمعاء وعلى مدى انتشاره في الجسم و هناك علامات تحذيرية تظهر عند الإصابة بهذا المرض وتشمل تكرر حالات الإمساك الشديد ، ووجود دم في البراز، وصغر في قطر البراز (السماكة) ، وفقدان الشهية ، وفقدان الوزن ، لدى الأشخاص الذين تتجاوز أعمارهم الخمسين عاما أو حدوث نزيف وقد يصاحبه الغثيان أو التقيؤ خصوصاً لسرطان المستقيم أو الإصابة بفقر الدم يعدان من العلامات الخطيرة لدى من هم فوق سن الخمسين، إلا أن الأعراض الشائعة الأخرى مثل

فقدان الوزن وحدوث تغير في نشاط الأمعاء لا تدعو للقلق إلا إذا كانت مرتبطة بحدوث نزيف ، إن كلاً من البوليبات غير السرطانية والسرطانية لا يسبب عادة أية أعراض ، حتى إذا تحولت البوليبات إلى سرطان، فإن الأعراض تكون نادرة إلى أن يصبح الورم ضخما عندما يسد الأمعاء الغليظة أو ينزف إلى البراز وعندها حين يحدث هذا قد يكون السرطان قد اخترق جدار الامعاء وانتشر إلى الغدد اللمفية في البطن أو إلى أعضاء أخرى و الأعراض التي قد تظهر هي الامساك ، إخراج دم من المؤخرة ، وأوجاع البطن

إن سرطان القولون يؤدي إلى الإمساك كما يؤدي أحياناً إلى ألمٍ في القولون إلا أن معظم حالات الإمساك وآلام البطن ليس سببها سرطان القولون، ولكن هناكٌ اعراض أخرى تترافق مع سرطان القولون

- وجود الدم في البراز أو عليه سواء كان لون الدم قاتماً او شديد الفتوحة
- يكون البراز أرفع بكثير من المعتاد ، وأنت تقول بأن هذه المشكلة مزمنة وليست حديثة
- .آلام غازية متكررة
- .الشعور بأن الأمعاء غير فارغة تماماً
- انخفاض الوزن بدون أسباب.
- شعورٌ دائم بالتعب

المسببات

أكثر من ٧٥-٩٥ %من حالات الإصابة بسرطان القولون تحدث لدى الأشخاص الذين لديهم القليل أو ليس لديهم أي عوامل وراثية ، عوامل الإصابة تزيد لدى الذكور وكبار السن كما من مسببات الإصابة تناول كميات كبيرة من الدهون والكحول واللحوم الحمراء واللحوم المعالجة والسمنة والتدخين وعدم ممارسة الرياضة البدنية ، ترتبط حوالي ١٠ % من الحالات بعدم ممارسة نشاط كافي كما تزيد نسبة الإصابة لدى من يكثرون من تعاطي الكحول و وجد أن شرب ٥ أكواب من الماء يوميا يقلل من خطورة الإصابة بسرطان القولون والمستقيم والأورام الغدية

إمراض التهابات الأمعاء

الأشخاص الذين يعانون من مرض التهاب الأمعاء (التهاب القولون التقرحي ومرض كرون) معرضون بدرجة كبيرة للإصابة بسرطان القولون، وتزيد خطورة الإصابة كلما طالت فترة الإصابة بالتهاب الأمعاء وكلما زادت شدة الالتهابات. في هذه الفئات عالية الخطورة ينصح بالوقاية عن طريق استخدام الأسبرين وكذلك إجراء تنظير للقولون بصورة منتظمة و الأشخاص المصابون بمرض التهاب الأمعاء يمثلون نسبة أقل من ٢%من حالات الإصابة بسرطان القولون سنوياً حيث وجد أن المصابين بمرض كرون يصاب ٢ % منهم بسرطان القولون

والمستقيم بعد ١٠ سنوات، و٨ %بعد ٢٠ سنة، و١٨ %بعد ٣٠ سنة، كما أن المصابين بالتهاب القولون التقرحي يتطور المرض لدى ١٦ % منهم إلى خلل في الأنسجة أو إلى سرطان في القولون على مدى ٣٠ عاما

عوامل الخطر

١. كبار السن - عمر أكبر من ٥٠ سنة
٢. الأجناس من أصل إفريقي
٣. لديك تاريخ شخصي لسرطان القولون والمستقيم أو الزوائد الغددية
٤. الأمراض المعوية الالتهابية و الأمراض الالتهابية المزمنة في القولون، مثل التهاب القولون التقرحي (Ulcerative (colitis)) وداء كرون ، يمكن أن يزيدان من خطر الإصابة بسرطان القولون
٥. إصابة أحد الأقارب بسرطان القولون سابقا
٦. نظامك الغذائي منخفض الألياف ، ومرتفع الدهون
٧. نمط الحياة الخامل ، إذا كنت غير نشط فأنت أكثر عرضة للإصابة بسرطان القولون فممارسة النشاط البدني بانتظام يقلل من خطر الاصابة بسرطان القولون
٨. السمنة ، الأشخاص الذين يعانون من السمنة المفرطة تزيد لديهم مخاطر الإصابة بسرطان القولون وزيادة خطر الوفاة بسبب سرطان القولون بالمقارنة مع الأشخاص أصحاب الوزن الطبيعي
٩. التدخين ايضأ من العوامل المهمة للاصابة بمرض سرطان القولون حيث تزيد نسبة المدخنين المصابين عن غير المدخنين .
١٠. شرب الكحول بكميات كبيرة قد يزيد من خطر الاصابة بسرطان القولون هناك دراسة وجدت أن "الأشخاص الذين يستهلكون أكثر من ٣٠ غ من الكحول يومياً (خاصةً الذين يستهلكون أكثر من ٤٥ غ يومياً) لديهم زيادة خفيفة في خطر الإصابة بسرطان القولون
١١. العلاج الإشعاعي للسرطان ، العلاج الإشعاعي الموجه على البطن لعلاج سرطان سابق قد يزيد من مخاطر الإصابة بسرطان القولون مستقبليا

خيارات العلاج

اذا كنت تعاني نزيفاً من المستقيم، فقد يجري لك الطبيب واحد أو أكثر من الاختبارات التشخيصية ، إن الفحص بالمنظار السيجمي المرن يكون غالبا هو أول وسيلة عندما يكون نزول دم احمر زاه هو العرض الغالب ، إذ يكون الارجح ان الدم آت من النهاية البعيدة من القولون (أي الاقرب إلى المستقيم والشرج). اذا كان الدم يظهر عند اجراء اختبار الدم المختفي في البراز، فقد يجرى اختبار حقنة الباريوم الشرجية، أو الفحص بمنظار القولون ، كل منهما يمكن ان يكشف عن سرطان القولون، بالإضافة إلى الأسباب الأخرى للنزيف وتشمل البواسير،

التهاب المستقيم، والبوليبات غير السرطانية أو البوليبات السرطانية التي تتحول فيما بعد إلى سرطان قولوني عندما تنمو إلى داخل تجويف الامعاء الغليظة وتنتشر خلال جدارها .كما وان الفحص بمنظار القولون يكون مفضلاً بصفة عامة لأنه يكون أفضل قليلا في الكشف عن السرطان ولأنه يمكن أثناء الفحص بمنظار القولون أخذ عينات من ورم – يحتمل ان يكون سرطانيا – لفحصه أو استئصال بوليبات نازفة و قد يقوم الطبيب بإزالة قطعة صغيرة من أي نسيج يبدو غير طبيعي وفحصها للكشف عن السرطان و ينصح بإجراء فحوص بمنظار القولون بانتظام للاشخاص المعرضيين بدرجة عالية للاصابة بسرطان القولون .

عوامل الخطر / خيارات العلاج يتم تصنيف سرطان القولون إلى درجات أو مراحل، ويعتمد العلاج على الدرجة، ينصح بإجراء جراحة لاستئصال جزء من المعي – أو المعي كله – لكل درجة منها، وهذا يشمل شق البطن واستئصال الجزء المصاب بالسرطان من المعي (استئصال القولون).

يمكن إجراء الجراحة أيضا باستخدام منظار داخلي وذلك بعد أن يصنع الجراح قطوعاً معدودة وصغيرة كثقب المفتاح في البطن، وهذ الاجراء يكتنفه الجدل .

أحيانا ما تتبع الجراحة بالعلاج الاشعاعي أو العلاج الكيماوي أو الاثنين معا، معظم مرضى سرطان القولون لا يحتاجون إلى تقميم القولون، والاجراءات الجراحية التي تحافظ على العاصرة الشرجية تسمح لغالبية المرضى بالاحتفاظ بقدرتهم على التحكم في امعائهم والتخلص من فضلاتها بالطريق الطبيعي .

الوقاية من سرطان القولون

إن تغير نمط الحياة غير الملائم يمكن أن يمنع سرطان القولون

إن المحافظة على الوزن المثالي ، اللياقة البدنية والتغذية الصحيّة تقلّل من مخاطر الأمراض السرطانية بشكل عام ، وفقا لذلك فإن التغيير في نظام العيش يمكن أن يخفّف من مخاطر الإصابة بسرطان القولون بنسبة-80% 60

الاجراءات التالية يمكن ان تقلل من الاصابة

الأغذية أو النظم الغذائية منخفضة الدهون، عالية الألياف مع حصص يومية متعددة من الفواكه والخضراوات والحبوب النشوية يحتمل أن تقلل قابلية الاصابة بالمرض، إن تناول غذاء غني بالكالسيوم والفولات (مادة غذائية يحتاجها الجسم بكميات قليلة وتوجد في الخضراوات الصفراء والخضراء الورقية) قد يقلل ذلك الخطر أيضا. • العقاقير اللاستيرويدية المضادة للالتهاب : NSAIDs رغم أن الباحثين لا يعلمون السبب إلا أن تناول تلك الادوية لسنوات عديدة يبدو أن يخفض معدلات الاصابة بالسرطان المعوي ، مع ذلك فإن التناول المنتظم

للاسبرين، وهو نوع من تلك العقاقير، لا يقلل خطر سرطان القولون و الاقلاع عن التدخين إذ إن تدخين السجائر يزيد خطر سرطان القولون و العلاج التعويضي الهرموني حيث إن النساء اللاتي يتناولن هرمون الاستروجين بعد سن انقطاع الطمث ينخفض لديهن خطر الاصابة بسرطان القولون بنسبة تتروح بين ٣٠ %و٤٠ % واستئصال البوليبات حيث إن استئصال البوليبات الورمية الغدية يعمل على التخلص من مصدر محتمل للأورام الخبيثة.

١- عليك أن تأكل كل يوم أقل من خمسة مرات وتأخذ حصة من الفواكه والخضروات
٢- عليك أن تأكل الخبز ذات الألياف العالية.
٣- عليك أن تبتعد عن التدخين وشرب الكحول
٤- عليك أن تبتعد عن اللحوم المصنعة مثل المرتديلا ، الهوت دوق ، السجق ، وغيرها
٥- تناول الأسبرين أو البرفين يقي من سرطان القول

ثانيا - الانحدار اللوجستي[١،٢،٤،٦،٧]

يعرف الانحدار اللوجستي بانه طريقة احصائية لتحميل مجموعة البيانات التي يتحدد فيها متغير الاستجابة من خلال متغير مستقل واحد او اكثر.variable independent more or oneحيث يكون متغير الاستجابة (المتغير التابع) من النوع الثنائي(binary) وان الانحدار اللوجستي - هو أنموذج إحصائي ينتمي لنماذج الانحدار الخطي يمكن من نمذجة متغير ثنائي الحد بدلالة مجموعة من المتغيرات العشوائية المتوقعة ، رقمية كانت أو فئوية و يستخدم الانحدار اللوجستي للتنبؤ باحتمالية وقوع حدث ما بمعرفة إضافية لقيم متغيرات يمكن أن تكون مفسرة والتي يمكن أن تكون رقمية و يشتهر الانحدار اللوجستي أيضا بتسميات نموذج لوجيت) Logit)أو المصنف العام للأنتروبية. تستعمل هذه النماذج بشكل واسع في العديد من التطبيقات العلمية والتجارية وهي من طرق النماذج الأكثر تطبيقا في مجال التعلم الآلي ، حيث تصنف ضمن طرق التعلم الآلي المراقب .

وبشكل مختصر الانحدار اللوجستي هو حالة خاصة لمجموعة النماذج الخطية العامة ، رغم أنه تاريخيا، تعتبر الأخيرة تعميما لتقنية الانحدار اللوجستي.

خصائص الانحدار اللوجستي

أ- الانحدار اللوجستي لا يفترض وجود علاقة خطية بين المتغير التابع والمتغيرات المستقلة.
ب- المتغير التابع يجب ان يكون ثنائي التفرع (dichotomous) بحيث يحتوي على فئتين مثال ذلك نقي او غير نقي مصابين وغير مصابين وهكذا .
ج- الانحدار اللوجستي لا يشترط ان تكون المتغيرات المستقلة من النوع المستمر ولا ان تتبع التوزيع

الطبيعي) distributed normally not) ولا ان تكون العلاقة بين المتغير التابع و المتغيرات المستقلة

خطية ولا يفترض تساوي التباين ضمن كل فئة وهذا يجعل نموذج الانحدار اللوجستي أكثر مرونة من بقية نماذج التنبؤ والتصنيف

د- يجب ان تكون الفئات محددة وشاملة بحيث ان كل مفردة تنتمي الى فئة واحدة فقط.

ـ- يجب ان يكون حجم العينـة المستخدم في الانحدار اللوجستي اكبر من حجم العينـة المستخدم في

الانحدار الخطي لان معاملات نموذج الانحدار اللوجستي يتم تقديرها باستخدام طريقـة دالة الإمكان الأعظم

(method likelihood maximum)وهي طريقة تحتاج الى عينة كبيرة الحجم نسبياً .

مجالات التطبيق

في ما يلي عينة من مجالات تطبيق الانحدار اللوجستي مع الأمثلة:

- في مجالات الطب والإحصاء الحيوي : مثلا احتمال حدوث نوبة قلبية عند شخص ما خلال فترة زمنية معينة حسب المعرفة القبلية ببعض المعلومات الديمغرافية (عمره أو جنسه مثلا) أو الطبية (مؤشراته البدنية أو الصحية أو الغذائية) أو الوبائية (سلوكياته كالتدخين مثلا.)
- الصيدلة: في تقدير رد الفعل والمقارنة بين نجاح الأدوية من عدمه .
- التأمينات: لفرز وتقسيم مجموعات العملاء حسب المخاطر ومدى قابلية جذبهم لمنتجات تأمين معينة .
- المجال البنكي: خصوصا في تنقيط العملاء أثناء دراسة ملفات القروض).Scoring Credit)
- التسويق: حساب توقعات ميل المستهلك إلى شراء منتج ما أو امتناعه عن الشراء .
- في سير الآراء والعلوم السياسية: مثلا للتنبؤ بقرار التصويت في الانتخابات اعتمادا على تنميط قبلي للمصوتين (مستواهم الاجتماعي، توجهاتهم السياسية، مستواهم التعليمي.)
- مجالات الاحصاء السكاني (الديموكرافي) .

النموذج الرياضي

- نعتبر عينة Ω حجمها η، يشار لكل فرد (مشاهدة إحصائية) في العينة بالرمزw
- Yهو المتغير النوعي المراد تفسيره ويحتمل قيمتين(١،٠) قيمة Yبالنسبة للفرد w يرمز لها بـ(w).Y

($X=(X_1, X_2, ..., X_P)$هي المتغيرات المفسرة ويمكن أن تكون رقمية أو نوعية ، قيمة المتغيرات المفسرة بالنسبة للفرد w يرمز لها ب $X(w)=(x_{1(w)} x_{2(w),...,} x_{p(w)})$

- الاحتمالان القبليان لتحقق القيمتين ٠ أو ١ يشار إليهما ب $P(Y=1)$و$P(Y=0)$. وهما يساويان تباعاً النسب الملاحظة للقيمتين ٠ و ١ في العينة و الاحتمال هنا قبلي بمعنى أنه لا يستلزم معرفة قبلية بقيم . X
- الاحتمالان الشرطيان لتحقق قيمة معينة لY بمعرفة إضافية لقيم $X=(X_1,X_{2...,}X_P)$ يشار إليهما تباعاً ب $P(Y=1/X)$و$p(y=0/x)$. يشار للاحتمال الشرطي اللازم تقديره بالنسبة للفردw ب$\pi(w)$.

نموذج لوجيت (Logit)

- نموذج الانحدار اللوجستي يحاول تفسير دالة لوجيت المتغير المراد تفسيره بدلالة توليف خطي للمتغيرات المفسرة
- $\ln\frac{p(Y=1/X)}{1-P(Y=\frac{1}{X})}$ ، بحيث ai (i من 0 الى p) هي معاملات النموذج التي يجب تقديرها .
- وهو ما يمكن من تقدير الاحتمال الشرطي حسب الصيغة التالية

$$P(Y=1/X)=\frac{e\ a0+a1x1+\cdots+apxp}{1+e\ a0+a1x1+\cdots+apxp}$$

تقدير النموذج

تقديرمعاملات النموذج يتم حسب طريقة الإمكان الاعظم (القصوى) (Maximum likelihood (estimation

- المتغيرYموزع حسب توزيع ثنائي الحدين : $Y\sim\beta(1,\pi)$،مما يمكن من وضع دالة تقدير الاحتمال الأرجح

للعينة على الشكل التالي$L=L\pi^{v}(1-\pi)^{(1-v)}\epsilon[0,1]$. الإمكانية القصوى تشير إلى احتمال تحقق

العينة Ω عبرسحب عشوائي

- تهدف الطريقة إلى إيجاد متجهة المعاملات $A=(a_0,a_1,\ldots,a_p)$ التي تحقق القيمة القصوى لL .
- لتسهيل التعامل مع هذا المؤشر يتم استعمال لوغاريتم تقدير الاحتمال :

$$LL=\sum y\ln(\pi)+(1-y)\ln(1-\pi)\epsilon[-\infty,0]$$

المقدر $\hat{A}$ الذي يحقق القيمة القصوى لLوLL أن يمتلك الخصائص التالية:

- أن يكون تقاربي وبدون تحيز.
- بأقل تباين ممكن.
- أن يكون تقاربي اذا توزيع طبيعي (غاوسي).

تقدير النموذج اللوجستي في البرامج الاحصائية

يمكن تطبيق تقنية الانحدار اللوجستي في (spss) يدويا (أي بدون استعمال وظائف معرفة قبليا) فقط عبر توليف الأدوات الحسابية المتوفرة واستعمال البرنامج المساعد Solver لإيجاد المقدرات المثلى للنموذج، وذلك عبر اتباع الخطوات التالية :

- تنظيم جدول البيانات بالتمييز بين المتغير النوعي المراد تفسيره Yوالمتغيرات المفسرة X(w) . $(X_1(w),X_2(w),\ldots,X_p(w))$

استخدام المسار التالي والذي سوف يتم شرحه تفصيلياً في الجانب التطبيقي

Analyze→Regression→Binary

تقييم نموذج الانحدار اللوجستي

على غرار طرق النماذج المنحدرة من الانحدار الخطي، يستلزم نموذج انحدار لوجيت القيام بمجموعة من عمليات التقييم باستعمال تقنيات الاستدلال الإحصائي لضمان نجاعة النماذج ولضمان مصداقيتها وضبط قوتها التوقعية ، علاوة على المقاربة الإحصائية ، يفرض استعمال الانحدار اللوجستي في التعلم الآلي المراقب القيام بطرق تقييم إضافية ، تهدف بالأساس إلى ضمان استمرارية النموذج وقابليته للتطور والتغير حسب تطور المعرفة بالمتغيرات المفسرة .

على العموم ، تنقسم طرق التقييم إلى صنفين :

- طرق التقييم الذاتي أو الداخلي) Evaluation Internal) والمرتبطة بطريقة النموذج في حد ذاتها (الانحدار اللوجستي) .
- طرق التقييم الخارجي (Evaluation External) المميزة لطرق التعلم الآلي المراقب والتي لا علاقة لها بماهية الطريقة المستعملة .

التصديق على النموذج يجب أن يستند على تحكيم عقلاني بين مخرجات طرق التقييم

ثالثا - نموذج انحدار كوكس[٨،٥]

اعتبر نموذج انحدار كوكس أحد أهم النماذج وأكثرها شيوعا في نماذج تحليل البقاء ، يستخدم هذا النموذج في الحالات التي يكون متغير الزمن الذي يسبق حدوث حدث معين ، له أهمية في تحليل الظاهرة قيد الدراسة ولكي يتم استخدام هذا النموذج يجب أن تتوفر الشروط التالية :

- المتغير التابع ويتكون من جزأين (متغير وصفي ثنائي القيمة بالإصابة إلى أن متغيرالزمن الذي يسبق حدوث الحدث) .

- متغيرات مستقلة بغض النظرعن طبيعتها هل هي كمية أو وصفية أومختلطة ومن المتوقع أن يكون لها تأثير على الظاهرة المعنية بالدراسة .

والنموذج الاصلي لانحدار كوكس اقترح من قبل العالم Cox عام ١٩٧٢ . فإذا كان T متغير عشوائي متصل فإن الشكل الاساسي للنموذج هو :

$$h(t \setminus xi) = h\circ * \exp(\beta\ x`i)$$

$h(t / x`i)$: هو الخطر الشرطي لوقوع الحدث عند الزمن t بالنسبة للمفردات التي لديها متجه المتغيرات المفسرة $x`$

$h_0(t)$: دالة الخطر الاساسية Function Hazard Baseline)(التي تعتمد على الزمن والمناظرة لمتجه المتغيرات المفسرة $X_i=0$.

B :هي متجه عمودي 1 × P من المعالم المجهولة وتستخدم طريقة الامكان الجزئية لتقدير المعالم المجهولة .

Xi :هو متجه وصفي 1 × P من المتغيرات المفسرة .

$\exp(\beta\ x`i)$: هو الخطر النسبي الذي يعتمد على الزمن أي أن تأثير المتغيرات المفسرة بزيادة أو لتغير نقطة زمنية

نقصان الخطر يكون ثابت التغير تبعاً لـ (t) وإن النسبة بين أي معدلين للخطر أيضا تكون ثابتة وغير معتمدة على الزمن .

فروض نموذج انحدار كوكس

هناك بعض الفروض لنموذج Cox كما يأتي :

١ –إن أهم فروض النموذج هي فرضية المخاطرة النسبية ، والتي تفترض بأن معدل الخطورة هو ثابت من شخص إلى آخر ضمن الدراسة والتي تعتمد على الوقت .

٢ –أن نموذج cox يفترض عدم ثبات دالة المخاطرة الاساسية ويتم تقديرها فقط بالاعتماد على الوقت (وقت حصول الوفاة) وهذا هو الاختلاف بين نموذج Cox وبقية نماذج البقاء .

٣ –يجب أن تكون المتغيرات المستقلة ثابتة مع الوقت أي التتغير مع مرور الزمن .

٤ – يفترض توزيع أولي للبيانات على عكس نماذج البقاء .

٥ –يفترض أن معدل الخطورة يزداد بشكل خطي مع الزمن .

٦ –يفترض أن المتغيرات التوضيحية ترتبط بشكل خطى مع لوغاريثم دالة الخطورة.

تقدير معالم نموذج انحدار كوكس

من شروط ايجاد دالة الامكان الاعظم يجب أن يكون التوزيع معلوم .

تعرف داله الامكان بأنها حاصل لضرب الدالة المشتركة لجميع المشاهدات داخل العينة تحت فرضية معينة أي لإيجاد دالة مشتركة لجميع المشاهدات.

وفي حالة استخدام نموذج Cox فإن توزيع t يكون معلوم ، ال يمكن إيجاد دالة الامكان الاعظم L(B) ، وفي النماذج المعلمية ، أما في نموذج Cox الشبه معلمي نقوم بإيجاد دالة الإمكان بالاعتماد على ترتيب الحدث مثلا (الوفاة) بدال من توزيع الحدث الوفاة .

اختبار معنوية معالم النموذج اختبار والد (Wald)

يتم من خلال اختبار والد الكشف عن معنوية تأثير متغير مستقل معين .

في نموذج Cox الانحدار هنا كمتغير ثنائي تابع واحد مع متغير مستقيل واحد أو عدة متغيرات مستقلة و يرافق كل متغير مستقل معلمه واحدة ، وأن اختبار والد يضع فرضية العدم مساويه للصفر

$$H° : \beta i = 0$$

هذا يعني أن اختبار تأثير المتغيرات المستقلة لوقت البقاء في النموذج وأن قيمة اختبار والد هي عبارة عن مربع قيمة اختبار t وتكون صيغة اختبار والد كما يلي :

$$wj = (= \frac{\beta j}{s.E\beta j})^2$$

احصائية الانحدار تتوزع حسب توزيع مربع كاي بدرجه حرية واحد ، عند عملية بناء النموذج فإن الاختبار يقودنا لمعرفة المتغيرات الغير معنوية والتي يجب استبعادها من النموذج ، وأن فترات الثقة للمعلمة المقدرة لنموذج كوكس تعطي بالصيغة الاتية :

$$\boldsymbol{\beta} \hat{} \pm \boldsymbol{Z\alpha}\ \boldsymbol{2/S}\ \hat{}\boldsymbol{E(\beta} \hat{} \boldsymbol{)}$$

•اختبار معنوية النموذج ككل (اختبارنسبة الامكان الاعظم)

إن اختبار النموذج هو من الاختبارات المهمة أي نموذج ملائم باستخدام اختبار نسبة الامكان الاعظم حيت يتم تنفيد هذا الاختبار بتقدير النموذجين للبيانات ثم مقارنة النموذج الاول مع النموذج الثاني من خلال اللوغاريتمات

لدالة الامكان للنموذجين وأن النموذج الذي له أقل قيمة يكون هو الافضل ، يجب معرفة أن هذا الفرق هو معنوي باستخدام الصيغة :

$$LR = -2 \log (= \frac{\mathrm{Lm}}{\mathrm{L}\circ}) = 2 \log L\circ - 2 \log Lm$$

$L\circ$: هو دالة الامكان الاعظم عندما تكون دالة المخاطرة الاساسية فقط موجودة داخل النموذج .

Lm: دالة الامكان الاعظم عند وجود m من المتغيرات داخل النموذج وأن نتيجة احصائية الاختبار لها تتوزع مربع كاي بدرجة حرية مساوي لعدد المعلمات الموجودة داخل النموذج.

الجانب التطبيقي[٣]

(تحليل الانحدار اللوجستي)

تم الحصول على المخرجات الاتية من خلال اتباع المسار Analyze→Regression→Binary Logisticو تم وضع المتغير المعتمد (Y) في حقل Dependent و وضع بقية المتغيرات (المستقلة) في الحقل Covariates واختيار Enter ثم OK

جدول رقم (١) يبين المتغيرات المستقلة المدروسة ورموزها في هذا البحث وفي التحليل الإحصائي

تسلسل المتغير	رمز المتغير	عنوان المتغير
المتغير الأول	X1	العمر
المتغير الثاني	X2	الجنس
المتغير الثالث	X3	السمنة المفرطة
المتغير الرابع	X4	الوراثة
المتغير الخامس	X5	الكحول
المتغير السادس	X6	علاج شعاعي لسرطان سابق
المتغير السابع	X7	التدخين
المتغير الثامن	X8	الأمراض المزمنة
المتغير التاسع	X9	أمراض معوية التهابية
المتغير العاشر	X10	الزوائد ألغديه

الجدول رقم (٢) يبين عدد المشاهدات المفقودة

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	130	100.0
	Missing Cases	0	.0
	Total	130	100.0
Unselected Cases		0	.0
Total		130	100.0

يتضح لنا من الجدول اعلاه ان حجم العينة المدروسة ١٣٠ مشاهدة وعدد المشاهدات المفقودة يساوي صفر

الجدول رقم (٣) يمثل نوع المتغير المعتمد

Dependent Variable

Encoding

Original Value	Internal Value
نعم	0
كلا	1

نعم تمثل الإصابة وسيرمز لها بالرقم (٠) وكلا تمثل عدم الإصابة وسيرمز لها بالرقم (١)

الجدول رقم (٤) يوضح اعداد المشاهدات المصابة وغير المصابة ضمن العينة المدروسة .

Classification Table[a,b]

	Observed		Predicted		
			القولون سرطان		Percentage Correct
			نعم	كلا	
Step 0	سرطان القولون	نعم	100	0	100.0
		كلا	30	0	.0
	Overall Percentage				76.9

الجدول اعلاه يبين ان 100 مشاهدة تؤثربها المتغيرات المستقلة بشكل مباشر وهي مصابة بمرض سرطان القولون و30 مشاهدة لم تتأثر بالمتغيرات المستقلة وهي غير مصابة بمرض سرطان القولون

الجدول رقم (٥) بين قيمة اختبار (WALD)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.204-	.208	33.451	1	.000	.300

من خلال مقارنة قيمة wald (33.451) مع قيمة كاي سكوير الجدولية عند مستوى معنوية 0.05 ودرجة حرية (K-1=9) والبالغة 32.36 ، وبمان ان القيمة الجدولية اقل من القيمة المحسوبة (التحليلة) لذا نرفض فرضية العدم ونقبل الفرضية البديلة بمعنى ان هناك تأثير مباشر من قبل المتغيرات المستقلة على المتغير المعتمد .

الجدول رقم (٦) يوضح قيمة اختبار (COX) وهو اختبار مناضر لاختبار مربع كاي .

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	68.685[a]	.424	.642

اي ان المتغيرات المستقلة تفسر ما قيمته (0.424) من نموذج الانحدار اللوجستي .

والباقي يكون ضمن متغيرات اخرى وتدرج ضمن حد الخطأ اي ان العوامل المدروسة فسرت ماقيمته (٠,٤٢٤) من العوامل والمسببات للمرض والباقي يجب البحث حول متغيرات اضافية تكون ذات تأثير على تشخيص المرض ومعرفة المسببات و لم تدخل ضمن متغيرات الانحدار اللوجستي .

الجدول رقم (٧) يوضح قيم متوسط الخطأ للمتغيرات المدروسة .

Variables in the Equation

		B	S.E.	Wald	Df	Sig.	Exp(B)
Step 1[a]	x1	-.048-	.028	3.004	1	.083	.953
	x2	.805	.787	1.046	1	.307	2.237
	x3	-2.986-	1.034	8.346	1	.004	.050
	x4	-1.593-	1.011	2.481	1	.115	.203
	x5	-1.815-	.880	4.257	1	.039	.163
	x6	-2.911-	.883	10.875	1	.001	.054
	x7	-1.569-	.773	4.123	1	.042	.208
	x8	-1.020-	.744	1.879	1	.170	.361
	x9	-2.915-	.842	11.992	1	.001	.054
	x10	-.968-	1.014	.911	1	.340	.380
	Constant	10.458	4.490	5.426	1	.020	34820.073

كلما زاد حد متوسط الخطأ (s.e) دل ذلك على ان حد التأثير ضمن هذا المتغير يكون اكبر من غيره من المتغيرات المدروسة وبالتالي فإن ثاثير المتغيرات على المتغير المعتد يكون ترتيبها كالتالي ويتضح ان المتغير الثالث والذي يمثل السمنة المفرطه ذو تأثير اكبر من بقية المتغيرات المدروسة بمتوسط خطأ (1.034) ، يليه 1.014 يمثل متغير العاشر (x10) الزوائد الغدية يليه المتغير ، 1.011 يمثل متغير الرابع (4) عامل الوراثة و يليه ، 0.883 يمثل متغير السادس (x6) علاج شعاعي لسرطان سابق ثم يليه المتغير ، 0.880 يمثل متغير الخامس (x5) الكحول ثم يليه المتغير ، 0.842 يمثل متغير التاسع (x9) امراض معوية التهابية و يليه المتغير ، 0.787 يمثل متغير الثاني (x2) الجنس ثم يليه المتغير ، 0.773 يمثل متغير السابع (x7) التدخين و يليه المتغير ، 0.774 يمثل متغير الثامن (x8) الامراض المزمنة ثم يليه المتغير ، 0.028 يمثل متغير الاول (1x) العمر نلاحظ ان متوسط الخطأ للمتغير x3 يؤثر على مرض سرطان القولون بشكل كبير ثم x10 وثم x4 وثم x6 وثم x5 وثم x9 وثم x2 وثم x7 وثم x8 وثم x1 وهو اقل المتغيرات تأثيرا على المتغير المعتمد الذي يمثل الإصابة بمرض سرطان القولون .

التوصيات والاستنتاجات

الاستنتاجات

١- من خلال النتائج التي تم الحصول عليها بعد التطبيق من الجدول رقم ((5) Variables in the Equation)) فإن تأثيرالمتغيرات المستقلة مباشر على مرض سرطان القولون واكثرها تأثيراً المتغير التاسع (x3) السمنة المفرطة وبقيمة متوسط الخطأ تساوي (١,٠٣٤) ، واقلها تأثير المتغير الاول الذي يمثل عمر الشخص بمتوسط خطأ (٠,٠٢٨)

2- نستنتج ان العلاقة معنوية بين المتغيرات المدروسة والمتغير المعتمد لان قيمة wald اكبر من قيمة الاختبار المحسوبة من جداول معنوية مع درجة معنوية 0.05

3 - نستنتج من خلال قيمة كوكس و التي = (0.424) تفسير المتغيرات المدروسة لظاهرة مرض سرطان القولون وهي نسبة جيدة جداً ، اما المتبقي يدرج ضمن حد الخطأ اي توجد متغيرات لم تدخل دراستها ضمن الانحدار اللوجستي .

التوصيات

١- نوصي بمقارنة نتائج البحوث العلمية الطبية التي تستخدم الانحدار اللوجستي مع التحليل المميز من خلال معرفة العوامل المسببة للامراض المدروسة في البحوث اللاحقة كونها عوامل تمييزية .

٢- نوصي بدراسة نماذج تحليلية أخرى لتحليل مرض سرطان القولون مثل التحليل العاملي .

٣- نوصي بدراسة برامج إحصائية أخرى للوصول الى نتائج البحث العلمي مثل البرنامج الاحصائي (matlab) (max stat)(aca stat)

المراجع (reference) المصادر

المصادر العربية

١- البياتي،هبه براهيم صالح ، (٢٠٠٥) تحميل المسار في إنموذج الانحدار اللوجستي مع تطبيق ، رسالة ماجستير كلية الادارة والاقتصاد ، جامعة بغداد .

٢- الشاروط ،أ.د. محمد حبيب ،محيسن ، أ.م. اميرة جابر، (٢٠١٣) الانحدار اللوجستي لدراسة الاصابة بسرطان الثدي ، مجلة القادسية لعلوم الجاسوب والرياضيات .

٣- البياتي ا.د.محمود مهدي حسن ، ابو الشعير ا.د.محمود جواد ،(٢٠١٢) ، البرنامج spss تطبيق عملي لتحليل البيانات الاحصائية ، كتاب .

٤- شمار م.م. اياد حيبي، جيجان م.م.علي محمد علي ، (٢٠١٧) تحليل اثر بعض المتغيرات لاستجابة علاج مرض عجز القلب بإستعمال أنموذج الانحدار اللوجستي ، بحث مجلة الكوت الجامعة .

٥- حسن ا.م.د. رعد فاضل ، صالح م . عائدة هادي ، (٢٠١٤) ، استخدام نموذج انحدار كوكس (COX -regression) لاوقات البقاء المرضى المصابين بمرض سرطان الدماغ في العراق ، بحث مجلة الجامعة المستنصرية .

المصادر الاجنبية

6- GRAMER, J.S.(1991) THE LOGIT MODEL , an introduction for economists , Edward arnold, London .

7- SHAEFER , R.L. (1979) , multicollinearty and logistic regression , ph.d dissertation , university of Michigan , usa .

8- Kalbfleisch and Prentice (1980) , The Statistical Analysis of Failure Time Data , John Wiley , New York .