نموذج الإنحدار البييزي اللوجستي لتحليل سرطان الثدي

Bayesian Logistic Regression Model

For Analyzing Breast Canser Data

الملخص

تتجه الدراسات العلمية في الآونه الأخيره الى المعالجة الإحصائيه لبيانات الأمراض الأكثر انتشاراً فمن أكثر الأمرض انتشاراً هو مرض سرطان الثدي. لذا أُجري البحث بهدف تحديد العوامل المؤثرة في مرض سرطان الثدي ومعرفة مدي تأثير كل متغير من هذه المتغيرات من خلال تطبيق نموذج الإنحدار اللوجستي linear logistic model حيث أن الظاهرة المدروسة ذات ظاهرة ثنائية. وقد تمت الدراسه على عينه من ١٠٠ مريض مصاب ومريض بمرض سرطان الثدي ، تتروح أعمارهم من ٣٩ الى٨١ سنة، بلغت نسبة الإصابة من الإناث (٩٦%) والذكور (٤%) جُمعت بياناتهم من مستشفي الزقازيق الجامعي في مدينة الزقازيق لمدة سنة في الفترة (٢٠١٥-٢٠١٦) م.

أستخدمت طريقة الإمكان الأكبر لتقدير معلمات النموذج لتحديد مدي تأثير المتغيرات التوضيحية وتم استخدام اختبار Wald test لمعرفة مدي معنوية المتغيرات.

المنارة للاستشارات

www.manaraa.com

١. المقدمة

يُعد القطاع الصحي أحد القطاعات الخدمية الذي يتولي مهمة توفير الخدمات الصحية لجميع أفراد المجتمع من خلال المؤسسات الصحية من أجل حماية المجتمع وتحقيق الرفاهية لأفراده. ومما لاشك فيه أن الصحة العامة من المؤشرات المهمة علي تقدم الدولة وتطورها. ومن مقومات بناء الصحة في المجتمع درء جميع الأمراض.

ويُعتبر سرطان الثدي في مقدمة هذه الأمراض السرطانية وهو من أكثر أنواع السرطانات انتشارا هذه الأيام حيث أن سرطان الثدي هو السرطان الأول عالميا ، والسبب الثاني للوفاة من السرطان عند المرأة .سرطان الثدي هو أحد أنماط الأورام الخبيثة الشائعة، وينتج عن نمو غير طبيعي لخلايا الثدي،ويُعد سرطان الثدي من أكثر أنواع الأورام التي تصيب السيدات على اختلاف أعمارهن، فقد ساعد دعم حملات التوعية بسرطان الثدي، وتمويل البحوث، على سرعة تشخيص وعلاج سرطان الثدي، فزادت معدلات البقاء على قيد الحياة، وانخفض عدد الوفيات، وذلك بفضل الله ثم الكشف المبكر للثدي والعلاجات الجديدة.

نسبة الإصابة بسرطان الثدي في تزايد مستمر في جميع دول العالم، ولكن لوحظ انخفاض في معدل الوفيات المرتبطة به منذ عام ١٩٩٠م في الدول المتقدمة بسبب إتباع برامج الكشف المبكر وتطوير علاجات أفضل.

من المتوقع حسب الدراسات وخلال العشرين سنة القادمة زيادة معدلات الإصابة به لأكثر من أربعة أضعاف في منطقة الشرق الاوسط.

هدف البحث

ان هذا البحث يهدف الى استخدام نموذج الانحدار اللوجستي في دراسة زمن البقاء للمرضى المصابين بمرض سرطان الثدي ، وتحديد المتغيرات المؤثرة في حالة المصاب.

٢. التوزيع اللوجستي

على مدى ١٥٠ سنة ماضية أو نحو ذلك، تم إجراء بحوث كبيرة فيما يتعلق بنظرية وتطبيقات التوزيع اللوجستي. في السابق استخدمت الدالة اللوجستية كمنحنى لنمو من قبل Verhulst (١٨٣٨، ١٨٤٥). حيث تم استخدام المنحنى اللوجستي لأغراض ديموغرافية

المنارة للاستشارات
www.manaraa.com

اقتصادية منذ نهاية القرن التاسع عشر. وقد سميت الدالة اللوجستية باسمها الحالي من قبل Reedو Berkson (١٩٢٩)، حيث أنهم من أول من كتب الدالة اللوجستية.

وقد استخدمت الدالة اللوجستية لتقدير نمو السكان من قبل كلٍ من: Pearl و Reed (١٩٤٠) Pearl et al. (١٩٤٠)، Schultz (١٩٣٠)، واخيرا Oliver (١٩٨٢) .

وقد استخدمت بعض تطبيقات الدوال اللوجستية في المشاكل البيولوجية من قبل كلاٍ من Pearl (١٩٤٠) Emmens (١٩٤١)، Wilson و Worcester (١٩٤٣)، Berkson (١٩٥٣،١٩٥١،١٩٤٤)، وFinney (١٩٥٢،١٩٤٧)؛ وقد استخدم Plackett (١٩٥٩) الدالة لوجستية في تحليل بيانات البقاء على قيد الحياة، وفي سنه (١٩٦١) استخدمها في إحصاءات الأخلاق.

وأهم تطبيقات التوزيع اللوجستي في المجالات الطبية و الاجتماعية كما في Dyke وPatterson (١٩٥٢) Grizzle (١٩٦١) في مجال الصحة العامة .

وأيضا يغطي كتاب Balakrishnan (١٩٩٢) التوزيع اللوجستي في العديد من الجوانب بدءا من الدراسة المرجعية إلى التطبيقات وأحدث البحوث التي أجريت في هذا الموضوع. ويخصص الفصل (٢٣) من *et al.* Johnson (١٩٩٥) للتوزيع اللوجستي: الدراسات التاريخية وطرق الاستدلال والتطبيقات والتعميم وعلاقتها بالتوزيعات الاخرى .

ووجد أيضا Shao (٢٠٠٢) مقدرات دالة الأمكان الأكبر لتوزيع اللوجستي. كما قدم Zhou (٢٠٠٧) تقدير المعلمات في التوزيع اللوجستي والتوزيع اللوغارتمي اللوجستي. وكذلك قدم *et al.* Asgharzadeh (٢٠١٦) تقدير التوزيع اللوجستي بنقطة وفترة يعتمد على بيانات السجل

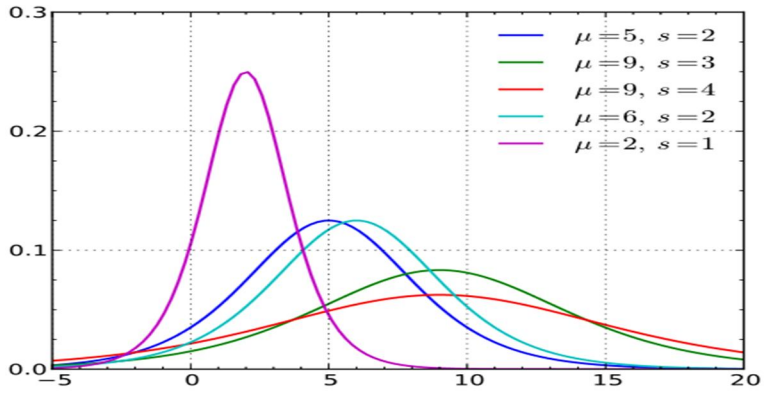
وتأخذ دالة كثافة الاحتمال للمتغير عشوائي X الذي يتبع توزيع اللوجستي كالتالي

المنارة للاستشارات
www.manaraa.com

$$f(x;\ \mu, s) = \frac{e^{-(x-\mu)/s}}{s\left(1+e^{-(x-\mu)/s}\right)^2}, -\infty < x < \infty,$$

(1)

حيث μ معلمة موقع location parameter ، و s معلمة مقياس scale parameterوكلاهما أكبر من الصفر. ويمكن رسم دالة كثافة الاحتمال للتوزيع اللوجستي بالطريقة التالية

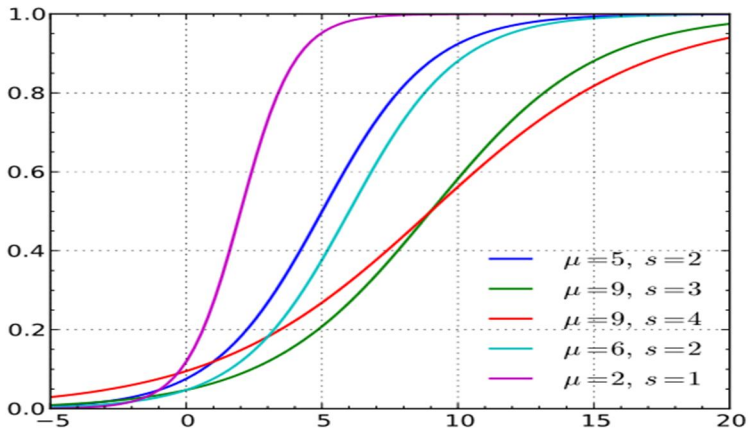


شكل ١ : دالة كثافة الاحتمال للتوزيع اللوجستي

وأيضا دالة التوزيع التراكميه للمتغير عشوائي X الذي يتبع التوزيع اللوجستي كالتالي

$$f(x) = \frac{1}{\left(1+e^{-(x-\mu)/s}\right)}, -\infty < x < \infty, \qquad (2)$$

ويمكن رسم الدالة التراكمية كالتالي



المنارة للاستشارات
www.manaraa.com

شكل ٢ : الدالة التراكميه للتوزيع اللوجستي

٣. نموج الإنحدار اللوجستي

يُستخدم تحليل الإنحدار اللوجستي في الدراسات الوبائية والطبية والذي من خلاله يتم تحديد المتغيرات المستقلة الكمية والنوعية التي تؤثر في إحتمال حدوث المتغير الناتج. عندما يتم تطبيق الإنحدار اللوجستي فلابد أن يكون متغير النتيجة ثنائي التفسير يحمل الترميز ٠ و ١.

الإنحدار اللوجستي البسيط

واحد متغير مستقل كمي أو اسمي مثل الوزن ، الطول ، الحالة الإجتماعية ، الجنس.

واحد متغير إسمي تابع ثنائي النوع مثل وجود المرض (نعم ، لا) نوع الجنس (ذكر ، أنثى .) وغيرها

الإنحدار اللوجستي المتعدد

إثنان أو أكثر من المتغيرات المستقلة الكمية أو الإسمية مثل الوزن ، الطول ، الحالة الاجتماعية ، الجنس

واحد متغير إسمي تابع ثنائي النوع مثل وجود المرض (نعم ، لا) نوع الجنس (ذكر ، أنثى) وغيرها.

استخدامات الإنحدار اللوجستي

التنبأ في إحتمال وقوع مشاهدات الدراسة في إحدى نواتج المتغير التابع (متغير النتيجة) الذي يرمز له ٠ ، ١ عن طريق معرفتنا بالمتغيرات المستقلة التفسيرية ذات النوع الكمي والإسمي وذات الدلالة الإحصائية فلو افترضنا بأن المتغير الإسمي التابع هو وجود المرض الذي يرمز له (٠= غير موجود ، ١= موجود) ونريد معرفة علاقة المتغيرات المستقلة الكمية أو النوعية في إحتمال حدوث المرض من عدمة

المنارة للاستشارات
www.manaraa.com

معادلة خط الإنحدار اللوجستي

$$y = \frac{exp(b_0 + b_1 x_1 + \cdots + b_p x_p)}{\left(1 + exp(b_0 + b_1 x_1 + \cdots + b_p x_p)\right)}, \qquad (3)$$

y تعني المتغير التابع الذي يرمز له ٠ و ١

تعني الثابت وهو معامل الإنحدار اللوجستي للمتغير التابع عندما تكون قيمة المتغير المستقل

b_0 صفر

b_i تعني معامل الإنحدار اللوجستي الذي يوضح مقدار التغير في المتغير التابع في حالة التغير بوحدة واحدة للمتغير المستقل

x تعني المتغير المستقل الذي يؤثرعلى المتغير التابع

شروط الإنحدار اللوجستي

أولا : يجب أن تكون العينة كبيرة وممثلة للسكان

ثانيا : يجب أن يكون هناك علاقة ارتباط بين المتغيرات المستقلة والمتغير التابع

٤. تقدير معلمات الإنحدار اللوجستي المتعدد

توجد عدة طرق لتقدير معلمات الانحدار اللوجستى منها طريقة الإمكان الأكبر وطريقة المربعات الصغرى وطريقة بييز وسوف نقتصر في هذه الدراسة على استخدام طريقة الإمكان الأعظم وطريقة بييز

٤-١ طريقة الامكان الأعظم Maximum Likelihood Method

وهي من الطرق الشائعة لتقدير معلمات النماذج الغير خطية، وتعتبر أفضل من طريقة المربعات الصغري المستخدمة في تقدير معلمات النماذج الخطية لانها لا تفترض قيودا كثيرة مثل الخطية

المنارة للاستشارات
www.manaraa.com

وثبات التباين، وهي تقيس الاحتمالات لعدد n من المتغيرات المستقلة في العينة ، أي ان مضروب هذه الاحتمالات هو عبارة عن دالة الامكان (عبدالمرضي١٩٩٨ م).

تعطي دالة الامكان كالتالي

$$L(\underline{\beta},\underline{x})=\left(\frac{e^{\left[\left(\frac{e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}{1+e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}\right)^{y_i}\left(1-\frac{e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}{1+e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}\right)^{1-y_i}\right]}}{1+e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}\right)^{y_i}$$

$$\left(1-\frac{e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}{1+e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}\right)^{1-y_i}$$

, (4)

حيث ان $\underline{\beta}=(\beta_0,\beta_1,\beta_2,\ldots,\beta_p)$

حيث (x_i) المتغيرات المستقلة ، ودالة الامكان لعدد n من المتغيرات هي

$$L(\underline{\beta},\underline{x})=\prod_{i=1}^{n}\left[\left(\frac{e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}{1+e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}\right)^{y_i}\left(1-\frac{e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}{1+e^{\beta_0+\beta_1x_{i1}+\cdots+\beta_px_{ip}}}\right)^{1-y_i}\right],$$

(5)

وتوزيع احتمال النجاح والفشل يتبع توزيع بيرنوللي.

٢-٤ التقدير البييزي Bayesian estimation

توماس بيز (١٧٦٣) الذي أثبت حالة خاصة تسمى الأن بنظرية بيز. تحدث ريتشارد برينس عنها بورقة عنوانها " مقال نحو مشكلة في عقيدة الفرص" ثم طور بير سيمون لابلاس (١٧٧٤)

المنارة للاستشارات
www.manaraa.com

من نظرية بيز ، تم تطوير الاستدلال البيزي بواسطة هارولد جيفيريز (١٩٣٩) في كتاب نظرية الاحتمالات و يليه أرنلدو زيلنر (١٩٧١) أستخدم الاستدلال البيزي لتحليل المشاكل اقتصادية القياسية محددة واختبار الفرضيات والنماذج. طبق جيمس (١٩٧٢) الطرق البيزئية و التكرارية لتحليل المتغيرات المتعددة . ديفيد و جوناثان و جونز و كيث (١٩٩٨) قدموا أستخدم بايزي في التكنولوجيا الصحية . عمل سكوتاري (٢٠١٠) شبكات بايزي مع حزمة بنلير(٢٠١٧) طبق تقدير بيز لعدد قليل من العينات : كشف المجتمع و المشاكل ذات صلة.

التوزيع البعدي $\pi(\theta,\underline{x})$ هو توزيع احتمالي مشروط للمعلمة θ للمعلمة التوزيع البعدي بشرط الحصول علي العينة $\underline{x}$ وهو يصف درجة اعتقادنا المختلفة للمعلمة θ بعد الحصول علي العينة ويأخذ الشكل التالي

$$\pi(\theta,\underline{x}) = \frac{\pi(\theta)L(\theta|\underline{x})}{\int_{\theta}\ \pi(\theta)L(\theta|\underline{x})d\theta}, \qquad (6)$$

حيث $\pi(\theta)$ التوزيع القبلي ،$L(\theta|\underline{x})$ دالة الامكان.

و التوزيع القبلي Prior distribution هو توزيع احتمالي يصف كل المعلومات والخبرات المتوفرة حول المعلمة قبل الحصول على العينة كما يصف درجة اعتقادنا في القيم الممكنة لهذه المعلمة بحيث $f(\underline{x})$ هو التوزيع الهامشي للمتغير $\underline{x}$ ويكون أنواع التوزيع القبلي توزيعات قبليه غير معلمه Non informative prior وتعبر عن المعرفة القليلة أو النادرة وتمثل الحالة التي تكون المعلومات المتوفرة لدينا حول المعلمة قبل المعاينة لا تعتبر جوهرية بالنسبة الي المعلومات المتوقع الحصول عليها من العينة.

والتوزيعات القبلية الغير معلمة تاخذ بعض الصور كالتالي:

١- معلومات فشر

$$\pi(\theta)\alpha\ \sqrt{I(\theta)}$$

٢- التوزيع المنتظم

المنارة للاستشارات
www.manaraa.com

إذا كانت المعلمة θ تأخذ قيما في المدى $(-\infty,\infty)$

١- إذا كانت $a < \theta < b$ فإن $\pi(\theta) = \frac{1}{b-a}$

٢- إذا كانت $-\infty < \theta < \infty$ فإن $\pi(\theta) \propto d\theta$

٣- اذا كانت $-\infty < \theta < \infty$ فإن $\pi(\theta) \propto \frac{d\theta}{\theta}$

الجدول التالي يوضح التوزيعات الاحتمالية للمتغير العشوائي X والتوزيعات القبلية المرافقة لها

الجدول ١

التوزيعات القبلية المرافقة

توزيع X	المعلمة	التوزيع القبلي المرافق
ذي الحدين	احتمال النجاح	بيتا
بواسون	المتوسط	جاما
الأسي	مقلوب المتوسط	جاما
الطبيعي	المتوسط (التباين معلوم)	الطبيعي
الطبيعي	التباين (المتوسط معلوم)	معكوس جاما

توزيع ذي الحدين هو توزيع القبلي لهذه المعالم، ويكون التوزيع القبلي للمعالم المجهولة

$\beta_0, \beta_1, \dots, \beta_p$ هو

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}, n = 0,1,2, \dots. \qquad (7)$$

المنارة للاستشارات
www.manaraa.com

حيث P احتمال النجاح، $(1-p)$ احتمال الفشل، n عدد مرات النجاح k عدد مرات اجراء النجاح.

5-2 أختبار معنوية معلمات النموذج

عند بناء النموذج لابد من معرفة هل جميع المتغيرا ت الداخلة في النموذج فاعلة في النموذج (لها تاثير معنوي علي المتغير التابع) أيٍ منها لابد أن تبقي في النموذج وأي منها سوف تحذف وهذه الاختبارات هي:

١- اختبار (Wald Test)

٢- اختبار نسبة الامكان الاعظم (Likelihood Ratio)

٣- اختبار الدرجة (Score Test)

وسوف نستخدم في هذه الدراسة اختبار Wald وهو يستخدم لاختبار معنوية تاثير المتغير التوضيحي علي المتغير التابع في النماذج الاحصائية الغير خطية، ولان في النموذج متغير ثنائي تابع مع متغير مستقل واح أو أكثر، ويرافق كل متغير مستقل متغير معلمة واحدة فقط.

ويستخدم اختبار والد لاختبار كل معلمة من معلمات المتغيرات المستقلة الداخلة في بناء النموذج، ونقوم باختبار معنوية المعلمات كالتالي:

١- الفروض

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

٢- احصاءة الاختبار

$$W = \left(\frac{\hat{\beta}}{s(\hat{\beta})}\right)^2$$

المنارة للاستشارات
www.manaraa.com

حيث ان :

$\hat{\beta}$ هي مقدر المعلمة ، $s(\hat{\beta})$ هي الانحراف المعياري لمقدر المعلمة.

٣- القيمة الجدولية للاختبار هي قيمة χ^2، الجدولية عند درجات حرية تساوي الواحد الصحيح .

٤- قاعدة القرار: اذا كانت احصاءة الاختبار ذات دلالة احصائية فأن ذلك يعني رفض الفرض العدمي القائل بان فيمة معامل الإنحدار تساوي صفر، وبالتالي المتغير المستقل Xسيكون له تاثير في المتغير التابع Y. أما اذا اذا كانت احصاءة الاختبار ليست ذات دلالة احصائية فأن ذلك يعني عدم رفض الفرض العدمي القائل بان فيمة معامل الإنحدار تساوي صفر، وبالتالي المتغير المستقل X ليس له تاثير في المتغير التابع Y.

٦- العوامل المؤثرة في سرطان الثدي

خطر الاصابة بسرطان الثدي يزداد مع العمر ، و تكون المرأة عرضة للإصابة بسرطان الثدي في الستينيات من عمرها أعلى بما يفوق ١٠٠ مرة منها في العشرينيات. السمنة وعدم ممارسة التمارين الرياضة، يزيد الخطر لدى النساء لتعرض بسرطان الثدي .الحمل والإنجاب والارضاع الطبيعي في سن مبكر و عدم إنجاب الاطفال يضاعف خطر بالإصابة ٣ مرات . أن الرجل أقل عرضة بكثير لخطر الاصابة بسرطان الثدي من النساء.

٦-١وصف البيانات

تم الحصول علي هذه البيانات من مستشفي الزقازيق الجامعي. تم سحب عينة عشوائية مكونة من ١٠٠ مفردة من المرضي المصابين بسرطان الثدي في الفترة من ٢٠١٥-٢٠١٦م.

وقد حددت العوامل المؤثرة في زيادة حدوث المرض وهي:

١-العمر (Age)

٢- الجنس(Gender) وتم التصنيف علي اساس

١: للذكر (Male) -٢: للأنثي (Female)

المنارة للاستشارات
www.manaraa.com

٣- الحالة الاجتماعية (Social Status)

١: متزوج(Married)-٢:مطلق(Divorced)-٣: أرمل (Wideo)

٤- الرضاعة الطبيعة

أ.تمت الرضاعة الطبيعيه ب.لم تتم الرضاعة الطبيعيه.

٥- السمنة (obesity)

١: نحيف – ٢ :مثالي – ٣: مثالي جدا – ٤: بدين – ٥: بدين جدا

أما المتغير المعتمد أو متغير الاستجابة ويمثل حالة المصاب أو غير مصاب أعطيت القيمة

(١) في حالة المصاب والقيمة (٠) في حالة الغير مصاب.

٦-٢ التحليل الاحصائي

تم تحليل البيانات باستخدام برنامج SPSS من خلال خطوات متعددة

١- اختبار معنوية المتغيرات بتطبيق اختبار والد حيث اعطي النتائج التالية

جدول ٢

اختبار معنوية العوامل المؤثرة في النموذج اللوجستي

المتغيرات	Wald	Df	Sig
العمر	٢.٣٠	١	٠.٠
الجنس	٢٧.٠٦	١	٠.٠
السمنة	١١.٩١	١	٠.٠
الحالة الاجتماعية	١٦.٥٣	١	٠.٠
الرضاعة الطبيعية	١٩.٨٤	١	٠.٠

المنارة للاستشارات
www.manaraa.com

يتضح من الجدول ٢ ان جميع المتغيرات الداخلة في بناء النموذج معنوية لان قيمة p-value أقل من ٠.٠٥

جدول ٣

العلاقة بين السمنة والاصابة بالسرطان

مريض السرطان	السمنة					المجموع
	نحيف	مثالي	وزن زائد	بدين	بدين جدا	
غير مصاب	٠	٠	٠	٧	١٧	٢٤
مصاب	١	١	٢	٣٦	٣٦	٦٧

نلاحظ من جدول ٣ وجود علاقة قوية بين السمنة والاصابة بالسرطان كلما كان الشخص بدين كانت فرصة الاصابة بالسرطان أكبر.

جدول ٤

العلاقة بين العمر والاصابة بالسرطان

مريض السرطان	العمر					المجموع
	٣٩-	٤٩-	٥٩-	٦٩-	٧٩-	
غير مصاب	٩	٩	٤	١	١	٢٤
مصاب	٢٧	٢١	٢١	٦	١	٦٧

نلاحظ من جدول (٤) العلاقة بين العمر والاصابة ان كلما كبر الشخص في العمر كلما قلت الاصابة بالسرطان.

المنارة للاستشارات
www.manaraa.com

جدول ٥

العلاقة بين الرضاعة الطبيعية والاصابة بالسرطان

مريض السرطان	الرضاعة الطبيعية		المجموع
	نعم	لا	
غير مصاب	١٤	١٠	٢٤
مصاب	٣١	٣٤	٦٧

نلاحظ من جدول ٥ أن الاشخاص الذين لا يقومون بالرضاعة الطبيعية أكثر عرضة بالاصابة بسرطان الثدي .

الجدول ٦

العلاقة بين الجنس والاصابة بالسرطان

مريض السرطان	الجنس		المجموع
	ذكر	أنثي	
غير مصاب	٢	٢٢	٢٤
مصاب	٣	٦٤	٦٧

المنارة للاستشارات
www.manaraa.com

نلاحظ من جدول ٦ أن الأناث اكثر عرضة للاصابة بمرض سرطان الثدي من الذكور.

التحليل البييزي

خطوات التحليل الاحصائي للنموذج البيزي اللوجستي:

تم تحليل البيانات العينة باستخدام برنامج Winbugs 14 كالتالي:

١- إدخال بيانات الخاصة لسرطان الثدي لعدد 100 شخص

٢- الحصول على مقدر الامكان الأكبر

٣- تحديد التوزيع القبلي المتوافق لهذه البيانات وتم تحديد توزيع ذي الحدين.

٤- تم تحدد ثلاث قيم مبدأيه لحل النموذج وهي (0,1,0,0,0)، (0.1,0.1,0.1,0.1,0.1)، (0.15,0.15,0.15,0.15,0.15).

٥- تم عمل تكرار (300000) لهذه التجربة للتأكد من ثبات النتائج.

الجدول ٧

خصائص مقدرات العوامل المؤثرة في الاصابة بسرطان الثدي ويوضح جدول ٧ mean، standard error، MC error حيث يوضح الفرق بين مقدر التوزيع البعدي ومقدر باستخدام دالة الامكان.

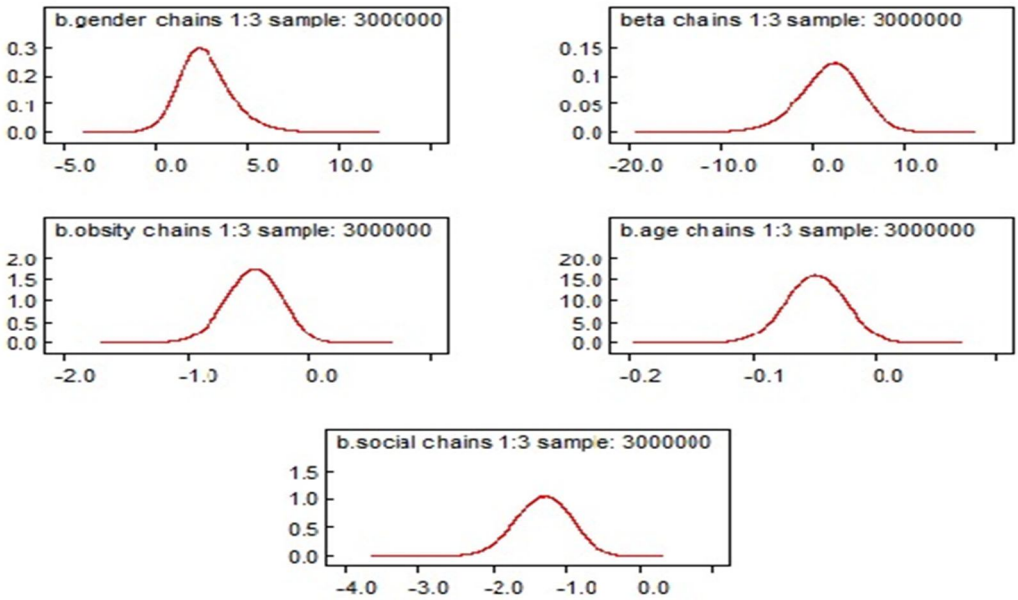
الخصائص	Intercept	age	gender	obesity	social	Breast feeding
mean	١.٧٨	٠.٠٤٨٥	٢.٦٨٨	-٠.٤٦٠٦	١.٣٢٣	٢.٨٩٥
Standard error	٣.٤٨٧	٠.٠٢٥١٣	١.٤٧٦	٠.٠٢٣٠٢	٠.٣٨٨١	٠.٠٥٩٤٦
MC erro	٠.٠٣٤٢	٠.٠٠٠١٣١	٠.٠١٣٩٥	٠.٠٠٠٩٣٥	٠.٠٠١٠٧	٠.٠٠٨٢٦

المنارة للاستشارات
www.manaraa.com

الجدول ٨ يبين Quantiles للمقدرات

الخصائص	Intercept	age	gender	obesity	social	Breast feeding
First	−5.687	−0.099	0.182	−0.929	−2.116	−٢.٠١٧
second	1.981	−0.048	2.544	−0.454	−1.313	−2.090
Third	8.138	−5.597	6.039	−0.027	−0.592	−3.012

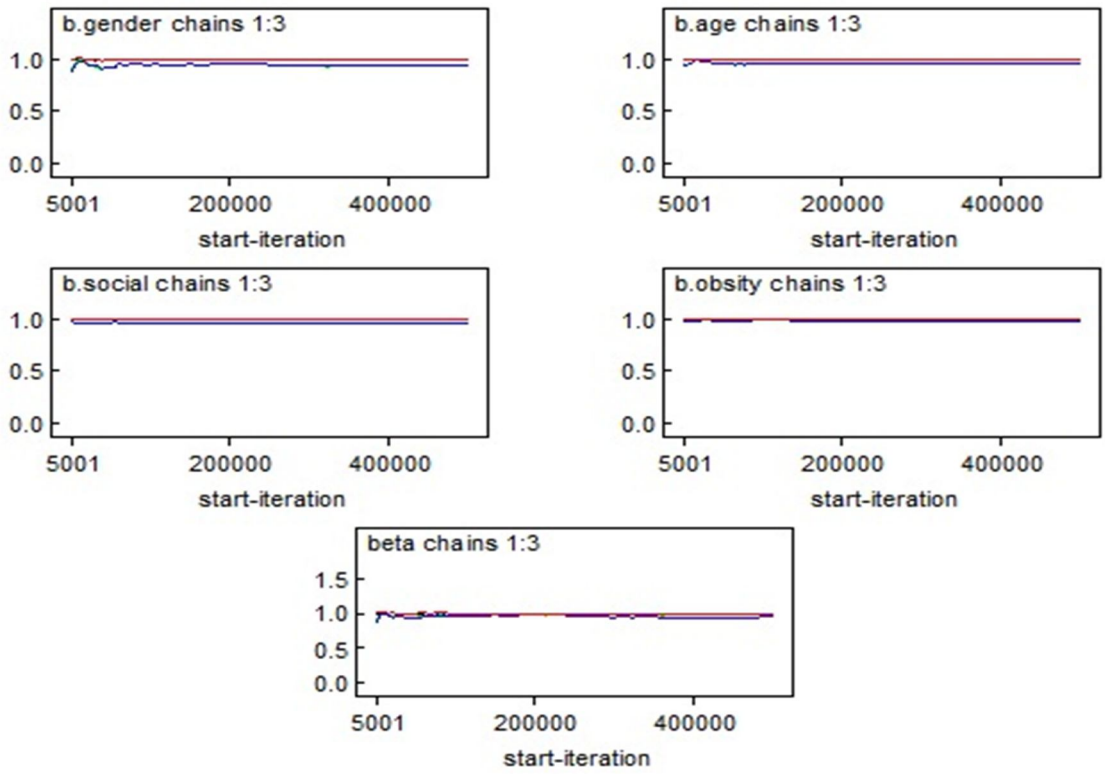
ولتوضيح ثبات تحليل البيانات الخاصة بمريض سرطان الثدي تم رسم دوال كثافة الاحتمال للتوزيع البعدي كالتالي:



شكل ٧ يوضح دالة كثافة الاحتمال للتوزيع البعدي. ويلاحظ ان التوزيع البعدي لجميع المقدرات تتبع توزيعًا طبيعيًا

ولتوضيح ثبات تتابع البيانات الخاصة لمريض سرطان الثدي تم استنتاج مقياس للتقارب ثبات حيث يكون استنتاج النتائج مبني علي ان تكون السلاسل الزمنية تكاد Gelman Rubinهو تصل الي الواحد الصحيح وهذا ما يوضحه الشكل التالي:

المنارة للاستشارات
www.manaraa.com

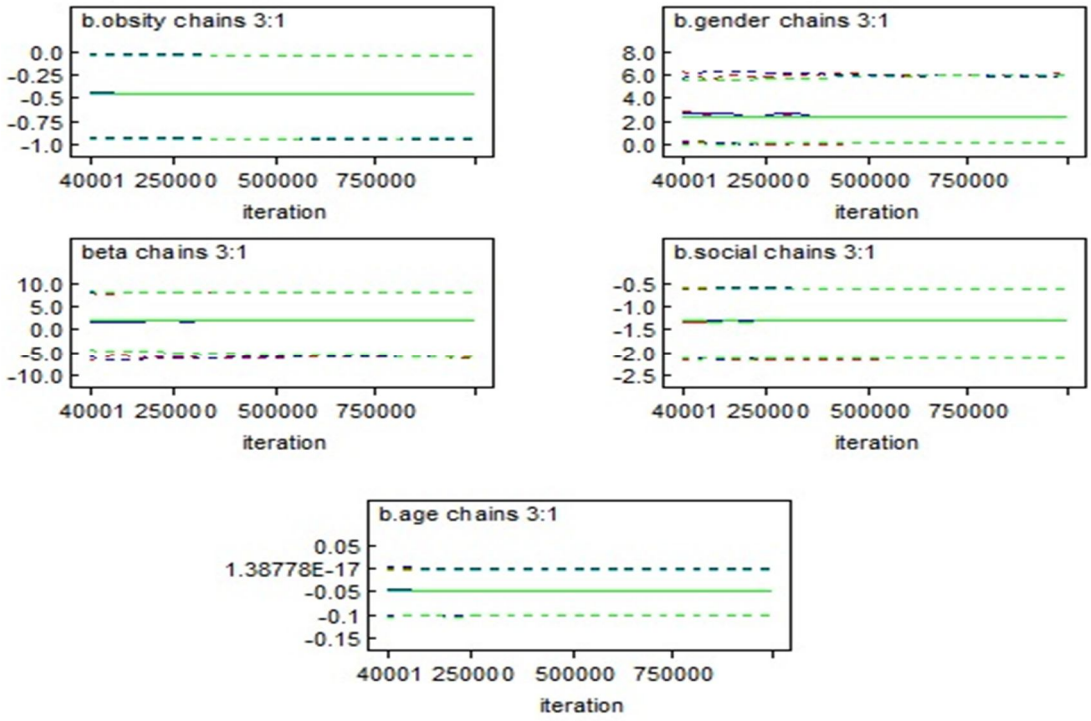


شكل ٨ يوضح مقياس التقارب بين القيم المبدئية للحل،ونجد انها تكاد

تقترب من الواحد الصحيح

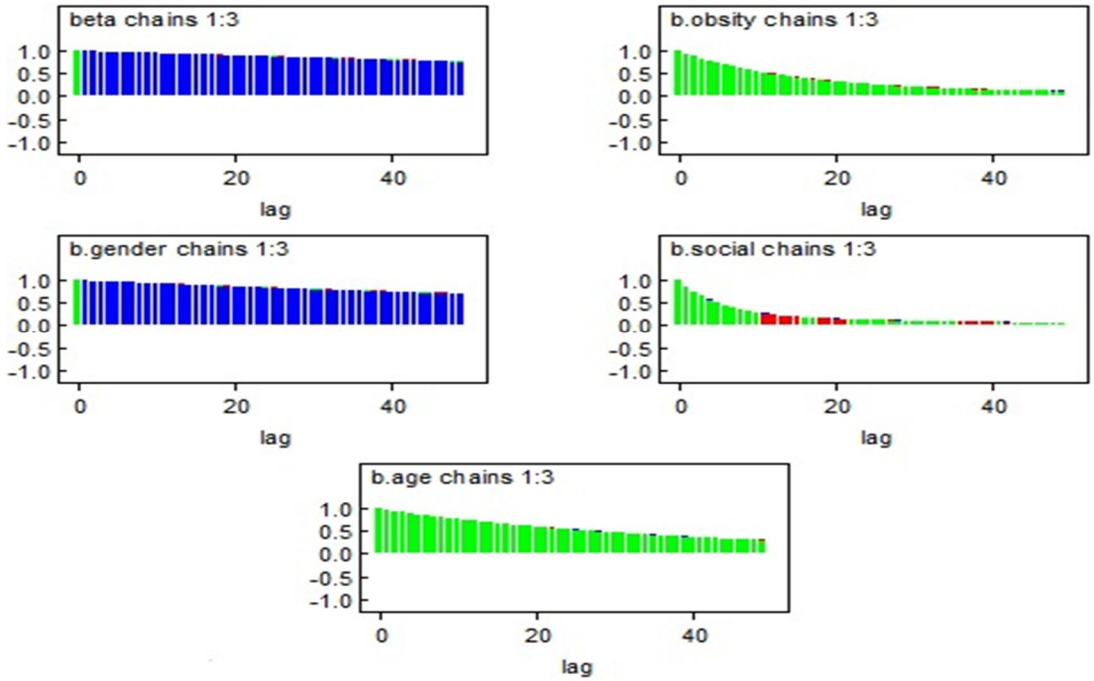
الشكل التالي يوضح الربيع الاول ، الربيع الثاني، الربيع الثالث

المنارة للاستشارات
www.manaraa.com



شكل ٩ توضح دوال الربيعات لمقدرات العوامل المؤثرة على مريض سرطان الثدي

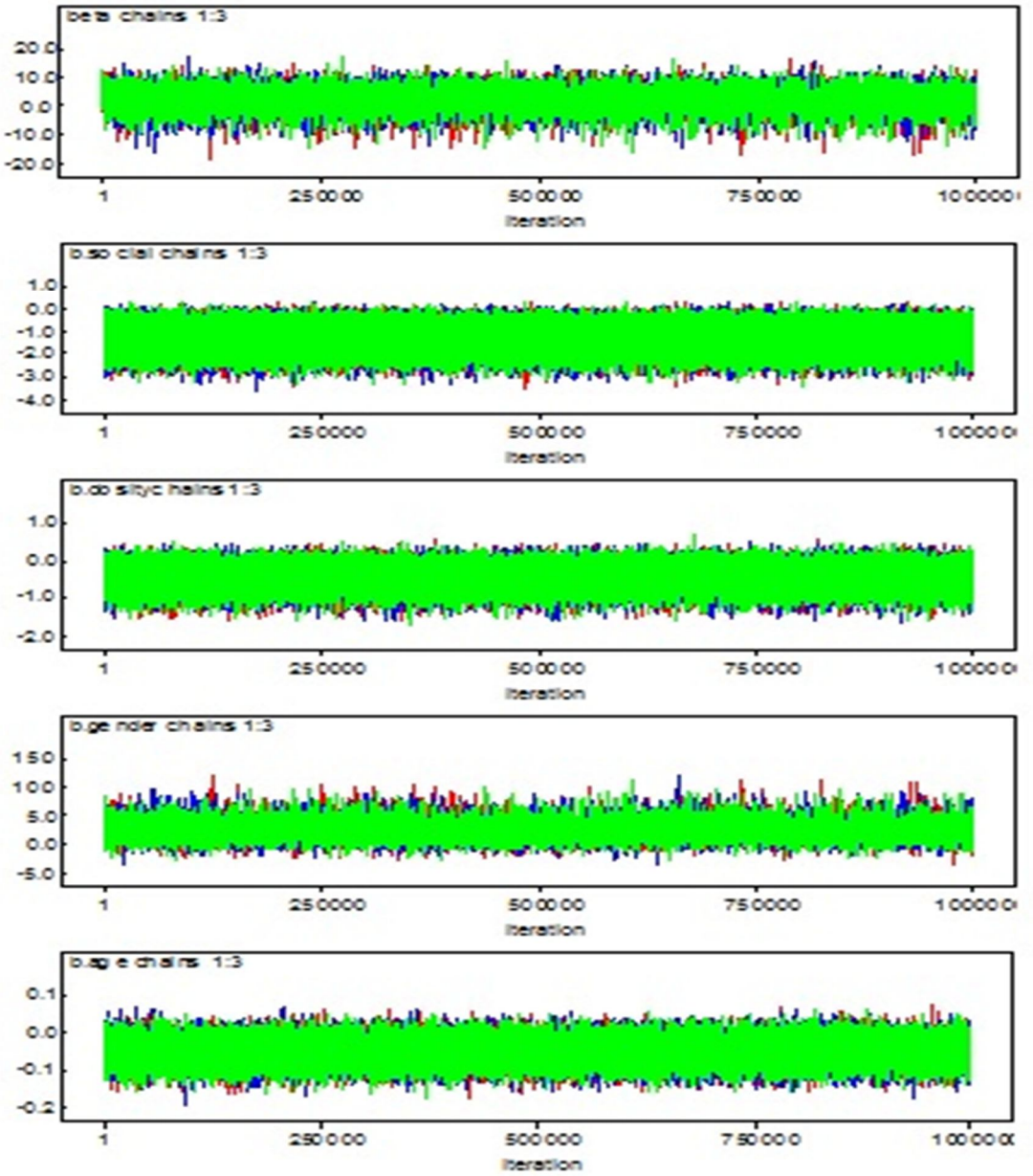
الشكل التالي يوضح معاملات الارتباط بين معاملات العوامل المؤثرة في مرضي السرطان



شكل ١٠ : يوضح مدى الارتباط بين العوامل المؤثرة على المريض

الشكل التالي يوضح مدي تقارب مقدرات في مرضي سرطان الثدي

المنارة للاستشارات
www.manaraa.com



شكل١١ يوضح مدى تقارب مقدرات العوامل المؤثرة

نتائج الدراسة

من أهم الاستنتجات التي حصلنا عليها من خلال دراسة تاثير العمر والجنس والسمنة والرضاعة الطبيعية والحالة الاجتماعية لعينة محل الدراسة هي:

١- عدد المصابين الاناث بمرض سرطان الثدي أكبر من فرصة إصابة الرجل

المنارة للاستشارات
www.manaraa.com

٢- وجود علاقة قوية بين السمنة ومرض السرطان كلما كان الشخص بدين كانت فرصة الإصابة بمرض السرطان أكبر .

٣- وجود علاقة عكسية بين فرصة الإصابة بالمرض والعمر بمعني أن فرصة الإصابة للعمر ٣٩ أكبر من العمر ٧٩ كلما كبر الشخص قلت فرصة إصابة بمرض السرطان الثدي.

٤- فرصة الإصابة بسرطان الثدي أكبر لدى الأشخاص الذين لايقومون بالرضاعة الطبيعية.

٥- المتزوجين هم أكثر عرضة للإصابة بالسرطان ثم الأرامل ثم المطلقين

.

وتم تأكيد هذه النتائج أيضا من خلال ادماج المعلومات القبلية مع المعلومات الحالية من خلال استخدام برنامج Winbugs 14 لتقدير نموذج الإنحدار البييزي اللوجستي.

التوصيات

١- يمكن التوصية بتحري الدقة في تسجيل البيانات، من خلال وضع برامج احصائية متخصصة ومتقدمة في كل مستشفى.

٢- يمكن التوصية بإجراء دراسات مماثلة لأنواع مختلفة من الأورام، ومعرفة العوامل المؤثرة علي كل نوع من هذه الأورام.

٣- يمكن التوصية بإضافة متغيرات أخري الي هذه الدراسة مثل جينات الوراثة، نوع الأغذية ، المنطقة التي يسكن بها المريض، عدد مرات الإنجاب.

المنارة للاستشارات
www.manaraa.com

References

Asgharzadeh, A. Valiollahi,R. And Abdi, M. (2016). Point and interval estimation for the logistic distribution based on record data, P.89-112.

Balakrishnan, N.(Ed.) (1992). *Handbook of the logistic distribution.*Marcel Dekker, New York.

Balakrishnan, N. and Chan, P.S. (1992). Estimation for the scaled half-logistic distribution under Type II censoring. *Computational Statistics and Data Analysis,*Vol.13, pp.123-141.

Grizzle, J. E.(1961). A new method of testing hypotheses and estimating parameters for the logistic model, *Biometrics,* Vol.17, pp.372-385.

Johnson, N.L., Kotz ,S. and Balakrishnan, N.(1995). *Continuous univariate distributions*,Vol.2, 2nd ed., John Wiley, New York.

Oliver, F.R.(1982). Notes on the logistic curve for human

المنارة للاستشارات
www.manaraa.com

populations. *J. Roy. Statist. Soc.,* A,Vol.145, pp. 359-363.

Plackett, R. L. (1959). The analysis of life test data. *Technometrics*,Vol.1, pp.9-19.

Shao, Q. (2002). maximum likelihood estimation for generalized logistic distribution, *Communication in Statistics, Theory and Methods*, V.31, pp.1687-1700.

Zhou,Y. (2007). Estimation of parameters in logistic and log-logistic distribution with grouped data, US National Library of Medicine, pp.9-421.

Bayes.T. (1763). An Essay towards solving a problem in the Doctrine of chances, pp. 370-481.

Stephen. M. (1986). The history of statistic, Harvard University.

Jayens .E.T.(2003). Probability theory: the logic of science, Cambridge University.

Zellner, A. (1971). An Introduction to Bayesian Inference in

Scutari M. (2010): Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software??*

Horny, M.(2014). *Bayseian Networks*, Boston University.

المنارة للاستشارات

www.manaraa.com

Hopkin, S.B. and Teurer.(2017). Bayesian estimation from few samples: community detection and related problems, *Computer Science*, pp.1-76.

المراجع باللغة العربية

-عزام وعبد المرضي (١٩٩٨) م: التحليل الاحصائي للمتغيرات المتعددة، دار المريخ للطباعة والنشر، الرياض.

- التحليل الاحصائي للمتغيرات المتعددة دار المريخ للطباعة والنشر الرياض.

المنارة للاستشارات
www.manaraa.com