مجلة البحوث المالية والتجارية
المجلد (٢٣) – العدد الرابع – أكتوبر ٢٠٢٢

استخدام النماذج الخطية المعممة في تصنيف الأخطار وتسعير تأمينات الحياة

Using Generalized Linear Models in Risk Classification and Pricing of Life Insurance

د./ محمد أحمد فؤاد عبده البرقاوي

مدرس بقسم الإحصاء التطبيقي والتأمين

كلية التجارة –جامعة المنصورة

elbarkawy@mans.edu.eg

رابط المجلة: https://jsst.journals.ekb.eg/

ملخص:

يعد تصنيف الأخطار واشتقاق أفضل نموذج يستخدم في تقدير معدلات الوفاة الخاصة بتأمينات الحياة من الأمور الهامة لشركات تأمينات الحياة حيث أن تقدير احتمالات البقاء على قيد الحياة له أهمية خاصة عند تقدير سعر التأمين ومعظم العمليات الإكتوارية في شركات التأمين ومن الملاحظ أن كل شخص له معدل الوفاة الخاص به والذي يختلف من شخص لآخر حيث يتأثر معدل الوفاة بعدة عوامل يطلق عليها عوامل خطر الوفاة (Mortality Risk Factors) وفي هذا البحث يستخدم الباحث النماذج الخطية المعممة (GLM) في تصنيف الأخطار بهدف الوصول إلى نموذج كمي يستخدم في تقدير معدلات الوفاة وسوف يستخدم الباحث في تقدير معلمات النموذج طريقة الإمكان الأعظم (MLE) ومن ثم استخدام معدلات الوفاة المقدرة باستخدام النموذج المقترح في تقدير السعر الخاص ببعض أنواع تأمينات الحياة.

كلمات مفتاحية:

(النماذج الخطية المعممة , تسعير تأمينات الحياة, تقدير معدلات الوفاة, عوامل الخطر)

Abstract

The classification of risk factors and the derivation of the best model used in estimating the mortality rates of life insurance are important matters for life insurance companies, estimating the probability of survival is particular importance when pricing insurance contracts and most of the actuarial operations in insurance companies, it is noted that each person has its mortality rate, which varies from one person to another, the death rate is affected by several factors called risk factors (Mortality Risk Factors). In this paper researcher uses generalized linear models for classifying risks to arrive to a quantitative model used in estimating mortality rates. The researcher will estimating the parameters of the model using the maximum likelihood estimation method (MLE) and then use the estimated mortality rates from the proposed model for pricing some types of life insurance.

Key words:

(Generalized linear models, pricing life insurance, estimating mortality rates, risk factors)

١. المقدمة:

حساب سعر التأمين على الحياة يتوقف على ثلاثة عناصر أساسية هي:

- **معدلات الوفاة السائدة في المجتمع.**
- **معدل الفائدة الفني المستخدم.**
- **معدل مصروفات وتحميلات القسط.**

ومن الملاحظ أن هناك صعوبة في تقدير معدلات الوفاة التي تمثل خبرة المؤمن عليهم بصورة دقيقة والتي تستمد أساساً من جداول الوفيات، وذلك بسبب ارتفاع تكلفة بناء جداول وفيات تعتمد على خبرة شركات التأمين، وكذلك الحاجة إلى فريق من الخبراء الإكتواريين لتكوين هذه الجداول وندرة بعض بيانات المؤمن عليهم وخاصة في الأعمار الكبيرة، لذا تلجأ معظم شركات تأمينات الحياة المصرية إلى استخدام جداول وفاة معدة مسبقاً اعتماداً على خبرة بعض المجتمعات الأكثر تقدماً مثل الجداول المستخدمة في سوق التأمين الخاص بالمملكة المتحدة أو الولايات المتحدة الأمريكية وذلك عند تقدير تكلفة منتجاتها التأمينية المختلفة، وهذا ما يترتب عليه أن هذه المعدلات لا تعكس واقع خبرة المؤمن عليهم محل التسعير، وبالتالي فإن تحديد السعر الخاص بتأمينات الحياة قد يختلف باختلاف الجداول المستخدمة، ويترتب على ذلك أن أي انحراف بين معدلات الوفاة المستمدة من هذه الجداول ومعدلات الوفاة الفعلية قد يؤثر تأثيراً سلبياً على سعر منتجات تأمينات الحياة، ومن هنا لابد من تقدير تكلفة تأمينات الحياة من خلال معدلات وفاة مستمدة من خبرة المؤمن عليهم بشركات تأمينات الحياة المصرية، كما يجب أن لا تكون جداول الوفيات المعيارية المستخدمة في تحديد سعر التأمين قد أعدت منذ فترة طويلة، حيث أن معدلات الوفاة تتحسن بمرور الوقت ويرجع ذلك إلى التقدم الطبي وتحسن مستوى المعيشة[محمد،أبو زيد، ٢٠٢١، ١٥٧-١٥٩].

وتزيد وطأة المشكلة عندما تضيف هذه الشركات هامش أمان على معدلات الوفاة المستمدة من جداول الخبرة الأجنبية، حتى تكون هذه الشركات بمأمن عن حدوث أي انحرافات عكسية بالنتائج، وينعكس ذلك على ارتفاع أسعار منتجات تأمينات الحياة مما قد يترتب عليه انخفاض الطلب على منتجات القطاع التأميني.

وبناء على ذلك يمكن لشركة التأمين إما أن تقوم بتسعير منتجات تأمينات الحياة باستخدام معدلات ثابتة لديها، أو تقوم بالتسعير وفقاً لمعدلات تتوافق مع المتغيرات المختلفة التي تؤثر على معدلات الوفاة، حيث يجب على شركات التأمين مراعاة العوامل الديموغرافية للمؤمن عليهم مثل (العمر، النوع، مستوى التعليم، الدخل، المهنة، الحالة الاجتماعية، الحالة الصحية)

ومما لاشك فيه أن عوامل التسعير هذه مهمة وتؤثر تأثيراً كبيراً على معدلات الوفاة [Abachi, 2018,1].

وبدراسة سوق التأمين المصري نجد أنه على مدار أكثر من ١٢٠ عاماً هي عمر قطاع التأمين المصري لم يتم إنشاء جدول حياة مصري، وكان الاعتماد على اختيار أقرب الجداول الإنجليزية توافقا مع الخبرة المصرية فى معدلات الوفاة والعجز ومن المعروف أن سعر المنتج التأميني من أهم العوامل المؤثرة في الطلب على التأمين، والذي يتم تحديده بواسطة الخبراء الإكتواريين، وبالتالي يجب مراعاة الدقة الكافية في التسعير والتي تأتي من خلال تصنيف عوامل الخطر ومدى تأثيرها على معدلات الوفاة وبالتالي سعر التأمين الصافي.

الجدول التالي يعرض تطور معدلات الوفاة لبعض شركات التأمين في سوق التأمين المصري خلال الفترة من ٢٠١٥ إلى ٢٠٢١:

جدول رقم (١)

تطور معدلات الوفاة بسوق التأمين المصري(١)

الســـــنـــة	معدلات الوفاة(في الألف)
٢٠١٥	٢,٥٠
٢٠١٦	٢,٢٤
٢٠١٧	٢,٤٨
٢٠١٨	٢,٢٥
٢٠١٩	١,٩٣
٢٠٢٠	١,٨٤
٢٠٢١	١,٧٨

المصدر: الكتاب السنوي عن نشاط التأمين أعداد مختلفة.

يتضح من الجدول السابق:

اتجاه معدلات الوفاة عن إجمالي السوق المصرية للإنخفاض التدريجي المستمر(٢) بصفة عامة من سنة ٢٠١٥ حتى سنة ٢٠٢١ ثم بدأت في ارتفاع طفيف سنة ٢٠١٧ و ٢٠١٨ مقارنة بسنة ٢٠٢١ ولكن بدأت المعدلات في الإنخفاض مرة أخرى ابتداء من سنة ٢٠١٨ حتى سنة ٢٠٢١ ، وهذا مؤشر هام لتحسن المستوى الصحي العام لجمهور حملة وثائق تأمينات الحياة،

(١) معدل الوفاة لإجمالي سوق التأمين= عدد الوثائق المنتهية بالوفاة ÷ متوسط عدد الوثائق السارية.

(٢) ص = ٢.٧٥٨ – ٠.١٢٣ س (الخطأ المعياري للتقدير ٠.١١).

وهذا التحسن لابد أن يتبعه تعديل في جداول الحياة المستخدمة مما قد يؤدي إلى إعادة النظر في سعر تأمينات الحياة في السوق المصرية.

٢. مشكلة البحث

تتعدد أنواع منتجات التأمين في سوق التأمين المصري، حيث تتوفر منتجات تأمينات الحياة، ومنتجات التأمينات العامة – تأمينات الممتلكات وتأمينات المسئولية المدنية – ولكن تختلف منتجات تأمينات الحياة عن منتجات التأمينات العامة في أنها تعتمد على جداول الوفاة في الحسابات الإكتوارية لتقدير قيمة أسعار منتجاتها التأمينية المختلفة، لذلك لابد عند تسعير منتجات تأمينات الحياة تكوين جداول وفاة تعبر عن خبرة سوق التأمين المصري وذلك للوصول إلى معدلات الوفاة المستخدمة في الحسابات الإكتوارية المختلفة ومن الملاحظ عدم اعتماد أسعار منتجات تأمينات الحياة في سوق التأمين المصري في تقديرها على خبرة السوق المصرية، حيث تعتمد على خبرة بعض الجداول التي تمثل خبرة مجتمعات أخرى.

ونظراً لوجود إختلافات معنوية بين خبرة سوق التأمين المصري والخبرة الأمريكية أو الإنجليزية وعدم مواكبة شركة التأمين المصرية بشكل مستمر للتغير في معدلات الوفاة، وهو الأمر الذي قد يؤدي إلى عدم العدالة في تسعير بعض منتجات تأمينات الحياة، حيث أن الاعتماد على مثل هذا النوع من الجداول يقدم معدلات وفاة قد تكون أعلى أو أقل من نظائرها بالمجتمع المصري، وكلا الأمرين يضر بمصلحة كل من المؤمن والمؤمن عليهم حيث قد يؤدي ارتفاع أسعار هذه الوثائق إلى انخفاض الطلب على وثائق تأمينات الحياة كما أن انخفاض أسعار هذه الوثائق عما يجب أن تكون عليه قد يؤدي إلى تكبد المؤمن خسائر قد تؤدي إلى توقفه عن ممارسة النشاط– [البحيري ،٢٠١٦].

ولذلك كان لابد من التوصل إلى بعض النماذج الجديدة التي يمكنها التنبؤ بمعدلات الوفيات مع الأخذ في الاعتبار التحسن في الرعاية الصحية والذي أدى إلى ظهور ظاهرة طول العمر بشكل أكثر دقة من النماذج المستخدمة حاليا والتي لا تعطي أي أهمية للعوامل المؤثرة في معدلات الوفاة، وتقديم أساليب جديدة لتحليل الوفيات والتنبؤ بها من خلال إدخال طرق إحصائية جديدة تقدم وجهات نظر مختلفة حول تطورات الوفيات.

ومن الملاحظ أن معدلات الوفاة خلال السنوات الأولى التي تلي إصدار وثيقة التأمين على الحياة تكون منخفضة نسبياً وذلك نتيجة لإجراءات الكشف الطبي وشروط القبول والاكتتاب، مما يترتب عليه فائض بشركات التأمين خلال السنوات الأولى للتعاقد، وكلما اختلفت معدلات الوفاة

الفعلية عن معدلات الوفاة المقدرة فإنه ينتج عنه اختلاف في قيمة هذا الفائض [عبد الباقي،متعال، ٢٠١٨، ٧٧].

ومما سبق يمكن تلخيص مشكلة البحث في:

- ✓ جداول الحياة التي تعتمد عليها شركات التأمين جداول غير محدثة، ولا تأخذ في الاعتبار العوامل المختلفة التي تؤثر على معدلات الوفاة، مما يؤثر على عدالة وكفاية سعر تأمينات الحياة.
- ✓ لا تستخدم شركات التأمين في السوق المصرية دوال رياضية لتعديل الأسعار القديمة التي تستخدمها مما يجعل المشكلة مستمرة في التأثير على سوق تأمينات الحياة.
- ✓ عدم توافر دوال رياضية أو احصائية معدلة، تمكننا من تقدير معدلات وفاة خاصة بشركة التأمين حسب خبرة سوق التأمين المصري مما يؤثر على تقدير تكلفة التأمين على الحياة وتحقيق العدالة بين طرفي عقد التأمين.

٣. هدف البحث

الهدف العام لهذه الدراسة هو المساهمة ببعض النماذج الحديثة التي تستخدم في دراسة المشكلة الأساسية السابق الاشارة إليها والمتمثلة في استخدام معظم شركات تأمينات الحياة جداول وفيات معيارية معدة مسبقا اعتماداً على خبرة بعض المجتمعات الأكثر تقدماً ومضى عليها سنوات طويلة مما جعل بيانات هذه الجداول غير محدثة بناء على بيانات الخبرة الفعلية لسوق التأمين المصري ولا تراعى التغير في معدلات الوفاة لمراحل العمر المختلفة على مدار السنوات المنقضية والناتج عن التحسن الصحي والبيئي في التوصل إلى بيانات تكون أقرب ما يمكن إلى البيانات الفعلية.

وبالتالي فإن هدف الدراسة هو التوصل إلى نموذج كمي يستخدم في تصنيف العوامل المختلفة التي تؤثر على معدل الوفاة وتحديد أكثر العوامل المؤثرة ثم استنتاج النموذج الرياضي الذي سوف يستخدم في التنبؤ بمعدلات الوفاة المختلفة من واقع خبرة المؤمن عليهم مما يساعد في تحديث البيانات اعتماداً على بيانات الخبرة الماضية، وتحديد اتجاهات الوفيات وما يرتبط بها من حسابات إكتوارية، وكيف يمكن إظهار تأثير عوامل الخطر المختلفة على معدلات الوفاة مما يجعل شركات التأمين قادرة على مواكبة التغيرات التي تحدث في معدلات الوفاة.

٤. أهمية البحث

تنبع أهمية هذا البحث من أهمية معدلات الوفاة في العديد من المجالات، وبخاصة المجالات التي تعتمد مزاياها المستقبلية على معدلات الوفاة حيث تختلف أهمية معدلات الوفاة لكل

من الدراسات الديموغرافية للمجتمع عنها لمجال التأمين، وتختلف أيضاً بالنسبة لمجال التأمين بشكل عام عنها في التأمينات الاجتماعية.

وبالتالي يمكن تلخيص أهمية البحث في النقاط التالية:

- تعتبر معدلات الوفاة حجر الأساس لبناء جداول الحياة وإذا واكبت هذه المعدلات التغيرات التي تحدث في عوامل الخطر فإن هذا قد يساعد على عدالة تقدير تكلفة التأمين.
- تحقيق العدالة في حساب القسط التأميني بما يحقق الملاءة المالية لشركات التأمين، مما يعني عدم المبالغة في زيادة الأقساط المحصلة عن الالتزامات الفعلية والذي يكون ضد مصلحة المستأمن.
- القدرة على حساب المخصصات الرياضية بشكل صحيح مما يجنب الشركة الخسارة في حال دفع تعويضات مالية تم تقديرها بشكل خاطئ، سواء مبالغ فيه أو أقل بكثير مما هو متوقع.

٥. حدود البحث:

١. تقتصر الدراسة التطبيقية على عينة عشوائية من المؤمن عليهم من شركة مصر لتأمينات الحياة حجمها ٤٠٠ مشاهدة[٣].

٢. تقتصر الحدود الزمنية على الفترة من ٢٠١٤ إلى ٢٠٢٠ م.

٦. منهجية البحث

استخدام النماذج الخطية المعممة في تصنيف الأخطار وتقدير معلمات النموذج باستخدام طريقة الإمكان الأعظم (MLE) بهدف الوصول إلى نموذج كمي يستخدم في تقدير معدلات الوفاة واستخدام معدلات الوفاة المقدرة في تقدير القسط الخاص ببعض أنواع تأمينات الحياة.

٧. فروض البحث:

١. تسعير تأمينات الحياة يعتمد بشكل أساسي على معدلات الوفاة.

٢. معدلات الوفاة المعيارية المستخدمة حالياً في شركات التأمين تحتاج إلى تحديث لكي تواكب التحسن في المستوى الصحي.

٣- بفرض أن نسبة خطأ التقدير (0.05) وبدرجة ثقة (95%) ، فإن حجم العينة يتحدد بالعلاقة:

$$n = \frac{z^2 \times P \times (1-P)}{e^2} = \frac{(1.96)^2 \times 0.5 \times 0.5}{(0.05)^2} = 384.16$$

* استخدام (P=0.5) في الصيغة السابقة يؤدى إلى الحصول على أفضل قيمة ممكنة لحجم العينة والتي تعطينا أفضل تقديرات ممكنة.

٨. مراجعة الدراسات السابقة:

دراسة (Kim،2021) بعنوان:

" Applications of reserving methods for property and casualty insurance in modeling of mortality rates"

في هذه الدراسة تم استخدام نموذج (chain ladder) لتقدير معدلات الوفاة المستقبلية، وقام الباحث بتطبيق نموذج (chain ladder) كأحد النماذج التي لا تعتمد على توزيع محدد وكذلك تم التطبيق على النماذج التي تعتمد على التوزيعات الاحتمالية باستخدام نموذج التوزيع اللوغاريتمي الطبيعي لتقدير معدلات الوفيات المستقبلية، وتم عمل محاكاة لمعدلات الوفيات للتأكد من دقة المعدلات المستنتجة باستخدام بيانات من الولايات المتحدة الأمريكية وانجلترا واليابان، وقد اعتمد الباحث على بعض الأساليب الاحصائية مثل: (MAE – RMSE – MSPE) لقياس دقة التنبوء وأظهرت النتائج أن نموذج (chain ladder) يحقق أفضل أداء بشكل عام، يليه التوزيع اللوغاريتمي الطبيعي.

دراسة (Lestari،2019) بعنوان:

" Pricing life insurance premiums using Cox regression model"

أوضحت هذه الدراسة أن أهم عنصر لتقدير القسط الخاص بالتأمين على الحياة هو معدل الوفاة، وقام الباحث بتقدير معدل الوفيات من خلال نموذج انحدار كوكس والذي تضمن بعض عوامل الخطر التي يمكن أن تؤثر على معدل الوفاة مثل النوع والتدخين والحالة الاجتماعية ومستوى التعليم، واستخدم الباحث طريقة المربعات الصغرى الجزئية لتقدير معالم النموذج وتوصلت الدراسة إلى أن الأقساط الصافية للمدخنين أعلى من أقساط غير المدخنين.

دراسة (Agnesti،2018) بعنوان:

" Poisson Regression Analysis for Risk Classification and Derivation of Mortality Rate Estimation in a Life Insurance Company"

هدفت هذه الدراسة إلى تصنيف الأخطار التي تؤثر على معدل الوفاة واشتقاق أفضل نموذج يستخدم في تقدير معدل الوفيات لشركة التأمين على الحياة باستخدام نموذج انحدار بواسون، حيث قامت هذه الدراسة بنمذجة العوامل المؤثرة في خطر الوفاة مثل عدد الوفيات ، مدة الوثيقة، ومبلغ التأمين، الحالة الإجتماعية، النوع، وتوصلت الدراسة إلى أن شركات التأمين التي تعتمد على تصنيف عوامل الخطر من المتوقع أن تكون أكثر دقة في تحديد الاحتياطيات المطلوبة وتقدير أقساط التأمين بشكل أكثر عدالة.

في دراسة (عبد الباقي،متعال ،٢٠١٨) بعنوان: " استخدام نموذج التمهيد الأسى والانحدار التكيفي في تقدير معدلات الوفاة بشركات التأمين" أوضحت الدراسة أن تقديرات معدلات الوفاة من الموضوعات المهمة في مجال التأمين على الحياة بصفة عامة، وتعرض الباحثان إلى نموذج مدمج لتقدير معدلات الوفاة بشركات التأمين عن طريق الدمج بين نموذجي التمهيد الأسي والانحدار التكيفي وتوصلت الدراسة إلى أنه من أهم محددات الوفاة التي أظهرها النموذج المدمج هي توقع الحياة، ومعدل الإنفاق على الصحة. وأوصى الباحثان بضرورة اهتمام شركات التأمين بدراسة اتجاهات التغير في معدلات الوفاة في المستقبل، وتأثيرها على تحديد قسط التأمين.

وهدفت دراسة (الميه وآخرون ،٢٠١٧) بعنوان: " تحسين نموذج جومبيرتز ماكيهام لتسوية معدلات الوفاة" إلى ايجاد معدلات وفاة دقيقة لمختلف الأعمار حيث أن المعدلات الخام للوفاة المستمدة من البيانات الفعلية لشركات التأمين عادة لا تعبر عن الواقع نظراً لوجود اختلافات كبيرة في المعدلات للسنوات المختلفة، وذلك وصولاً إلى بيانات وفيات مسواه تستطيع شركات التأمين استخدامها في حساب أقساطها ومخصصاتها بالشكل المناسب لها، وأوصت الدراسة باستخدام النماذج الخطية المعممة في تسوية معدلات الوفاة.

وهدفت دراسة (البحيري ،٢٠١٢) بعنوان: " استخدام دالة الحياة في تطوير جداول الوفيات استجابة لأثر التحسن الصحي، بالتطبيق على سوق التأمين المصري" إلى بناء نموذج رياضي يمكن من خلاله دراسة التحسن الصحي والبيئي واستخدامهما في تقدير تكلفة التأمين، وقد استخدم الباحث دالة جومبرتز- ماكيهام (Gompertz- Makeham) والتي تعد من أهم الدوال العمرية المعبرة عن معدلات الوفاة، وتمكن الباحث من الوصول إلى دالة حياة مقترحة تتسم بالسهولة في التطبيق، وتأخذ أثر التحسن الصحي في الاعتبار عند التنبؤ بمعدلات الحياة والوفاة.

٩. نظرة عامة على النماذج الخطية المعممة

النماذج الخطية المعممة Generalized Linear Models (GLM) هي أسلوب لنمذجة العلاقة بين المتغير التابع وعدد من المتغيرات المستقلة، والنموذج الخطي المعمم يتبع المتغير التابع فيه أحد التوزيعات الأسية، وتعتبر النماذج الخطية المعممة أقل قيوداً من نماذج الانحدار التقليدية، وتوفر هذه النماذج طرقاً لنمذجة البيانات التي لا تتبع التوزيع الطبيعي وقدمت النماذج الخطية المعممة لأول مرة بواسطة كلاً من (Neider and Wedderburn) في عام ١٩٧٢ وذلك في محاولة لتعميم فروض نماذج الانحدار التقليدية لكي تتلائم مع الواقع العملي[Barnett, Dobson, 2008, 51].

والنماذج الخطية المعممة أقل قيوداً من نماذج الانحدار التقليدية حيث: [Evelien, 2014, 15].

- لا يشترط في النماذج الخطية المعممة أن يكون التباين ثابت أي من الممكن أن يوجد اختلاف في التباين.
- لا يشترط في النماذج الخطية المعممة أن تكون العلاقة خطية بين المتغير التابع والمتغيرات المستقلة ولكن لابد من وجود علاقة خطية بين دالة الربط ومجموعة المتغيرات المستقلة ومن هنا يمكن توفيق بعض النماذج غير الخطية باستخدام النماذج الخطية المعممة.
- في النماذج الخطية المعممة يتم تقدير معالم النموذج باستخدام طريقة الإمكان الأعظم Maximum Likelihood Estimation (MLE) كما يمكن أيضاً استخدام طريقة المربعات الصغرى التقليدية Ordinary Least Squares(OLS).
- في النماذج الخطية المعممة يتم استبدال القيمة المتوقعة لمتغير الاستجابة $g(u)$ بدالة الربط (Link function) أي استبدال $g(u)$ بـ (η) وبالتالي يكون تباين الخطأ أكثر استقراراً, وتكون دالة الربط عبارة عن تركيبة خطية من المتغيرات التفسيرية

١٠. الصورة العامة للنموذج الخطي المعمم (GLM):

الصورة العامة للنموذج الخطي المعمم هي: [Murphy,et el, 2012, 109-110]

$$Y_i = h(X\beta) + \varepsilon$$

حيث أن :

Y: المتغير التابع أو متغير الإستجابة وهو متغير عشوائي يتبع أحد التوزيعات الأسية.

X: المتغير أو المتغيرات المستقلة (التفسيرية).

h: لوغاريتم دالة الربط.

ε : خطأ التقدير.

١٠-١ مكونات نموذج (GLM):

هناك ثلاثة مكونات رئيسية لنموذج (GLM) هي[Zhang, 2020, 24-26]:

أولاً:الجزء العشوائي

أي توزيع متغير الاستجابة (Response distribution) وهو التوزيع الذي يتبعه المتغير التابع (Y)، ومن المعروف أن متغير الاستجابة في النماذج الخطية المعممة لابد أن يتبع أحد توزيعات العائلة الأسية، ويمثل هذا المركب العنصر العشوائي(Random Component)، ويمكن عرض التوزيعات الأسية Exponential Family على النحو التالي:

بفرض أن متغير الاستجابة (Y) له معلمة واحدة (θ) يطلق عليها المعلمة الطبيعية للتوزيع(Natural Parameter)، فإن هذا التوزيع في هذه الحالة يتبع أحد التوزيعات الأسية، وهناك العديد من التوزيعات الأسية المشهورة والمستخدمة في مجال تسعير التأمينات منها:

- **توزيع بواسون Poisson distribution.**
- **توزيع ثنائي الحدين Binomial distribution.**

كما أن هناك بعض التوزيعات المستخدمة في مجال تسعير التأمين، والتي يطلق عليها عائلة التشتت الأسية Dispersion Family Exponential وهي تلك التوزيعات التي تحتوي بالإضافة إلى المعلمة الطبيعية(θ) معلمة أخرى(ϕ) وتسمى معلمة التشتت (Dispersion Parameter)، وعلى هذا الأساس فإن التوزيع يكون تابعاً لأحد توزيعات التشتت الأسية، لذلك فإنه من الممكن كتابة دالته الاحتمالية على الصورة التالية:

$$f\ \ y_i,\theta_i,\phi\ =\exp\left[\frac{y_i\theta_i - b\ \ \theta_i}{a_i\ \ \phi} + C\ \ y_i\phi\ \right]$$

ومن أهم أنواع هذه التوزيعات ما يلي:

- **التوزيع الطبيعي Normal distribution.**
- **توزيع جاما Gamma distribution.**
- **توزيع معكوس جاوس Inverse Gaussian distribution.**
- **توزيع ذي الحدين السالب Negative Binomial distribution.**

ثانياً: المتنبئ الخطي Linear Predictor:

المتنبئ الخطي Linear Predictor (η) ، وهو عبارة عن متجه المعلمات (β) ومتجه المتغيرات التفسيرية ($x_1, x_2, x_3, \ldots\ldots\ldots, x_n$) ويمثل هذا المركب العنصر المنتظم ويمكن التعبير عن المتنبئ الخطي كما يلي:

$$\eta_i = x_i^T \beta = \sum_{j=1}^{P} x_{ij}\beta_j$$

حيث:

$$(\beta_0 + \beta_1 + \beta_2 + \dots\dots + \beta_k)$$

$$x_i = x_{i1}, x_{i2}, x_{i3}, \dots\dots\dots\dots, x_{ip}{}^{T}$$

ثالثاً: دالة الربط Link Function $g(\mu_i)$:

دالة الربط link function وهي عبارة عن الدالة التي تربط بين متوسط الاستجابة (μ_i) والمتنبئ الخطي (η) وتتميز بأنها دالة رتيبة أو متفاوتة (Monotonic) كما تتميز دالة الربط بأنها دالة قابلة للتفاضل، وتستخدم دالة الربط لتوضيح العلاقة بين القيمة المتوقعة للمتغير المستقل والمتنبئ الخطي حيث:

$$g(\mu_i) = \eta_i$$

وبالتالي فإن دالة الربط تستخدم لربط القيمة المتوقعة للمتغير التابع $g(\mu_i)$ بالمتنبئ الخطي (η_i)، وتختلف دالة الربط باختلاف التوزيع الاحتمالي لخطأ التقدير ويوجد العديد من الصور لدالة الربط تستخدم كل صورة على حسب التوزيع الاحتمالي المستخدم، ومن ثم توجد عدة صور لدالة الربط نذكر منها ما يلي:

- **دالة الربط المتطابقة (Identity Link Function)**

تستخدم لربط القيمة المتوقعة للمتغير التابع $[E(Y_i) = \mu_i]$ بالمتنبئ الخطي (η_i) وذلك في حالة ما إذا كان المتغير التابع يتبع التوزيع الطبيعي $(N(0,\sigma^2)$ ~Y وفي هذه الحالة تكون دالة الربط على الصورة التالية:

$$g(\mu_i)\eta_i = x_i^T \beta\eta_i$$

$$g(\mu_i)x_i^T\beta_i = g^{-1}(x_i^T\beta_i)(\mu_i)$$

- **دالة الربط اللوغاريتمية (Log Link Function)**

تستخدم لربط القيمة المتوقعة للمتغير التابع $[E(Y_i) = \mu_i]$ بالمتنبئ الخطي (η_i) وذلك في حالة ما إذا كان المتغير التابع يتبع توزيع بواسون $(Y \sim poisson\ \lambda_i)$، وفي هذه الحالة تكون دالة الربط على الصورة:

$$\log(\mu_i)\eta_i$$

- **دالة اللوجيت (Logit Link Function)**

تستخدم لربط القيمة المتوقعة للمتغير التابع $\left[E\left(Y_i\right)=\mu_i\right]$ بالمتنبئ الخطي (η_i) وذلك في حالة ما إذا كان المتغير التابع يتبع توزيع ذي الحدين $(Y\sim Binomial(n\,,p))$ وفي هذه الحالة تكون دالة الربط على الصورة:

$$g\left(\mu_i\right)=\log\frac{\mu}{1-\mu}$$

- **دالة المقلوب (Reciprocal Link Function)**

تستخدم دالة المقلوب دائماً لربط القيمة المتوقعة للمتغير التابع $\left[E\left(Y_i\right)=\mu_i\right]$ بالمتنبئ الخطي (η_i) في حالة أن المتغير التابع يخضع لتوزيع جاما، وفي هذه الحالة تكون دالة الربط على الصورة:

$$\eta_i=\frac{1}{\mu_i}$$

١٠-٢ تقدير معلمات نموذج (GLM) [Johansson, 2010, 30-33]:

سوف يتم تقدير المعاملات $(\beta_0+\beta_1+\beta_2+........+\beta_k)$ لنموذج (GLM) باستخدام طريقة الإمكان الأعظم (MLE) حيث أن الهدف هو الوصول لتقدير المعلمة $(\boldsymbol{\beta})$ والتي تعظم دالة الإمكان من بيانات متغير الاستجابة $(\boldsymbol{Y})$.

فإذا كانت مشاهدات متغير الاستجابة $(y_i\ ,\quad i=1,2,........,n)$كلها مستقلة ولها دالة كثافة احتمالية على الشكل $f\left(Y\,,\theta,\phi\right)$ فإن دالة الإمكان لعدد $(\boldsymbol{n})$ من المشاهدات بمعلمات $(\boldsymbol{Y})$ والتي يمكن التعبير عنها بالمتجه $(y_1,y_2,..........,y_n)$ والتي يمكن التعبير عنها بالمتجه $(\boldsymbol{\beta})$ يمكن كتابتها على $(\beta_0+\beta_1+\beta_2+.......+\beta_k)$ الصورة التالية:

$$L\left(y\,;\beta\right)=\prod_{i=1}^{n}\exp\left[\frac{y_i\theta_i-b\left(\theta_i\right)}{a_i\left(\phi\right)}+c_i\left(y_i\,,\phi\right)\right] \qquad (1)$$

حيث:

(θ) : المعلمة الطبيعية.

(ϕ) : معلمة التشتت.

(ϕ) : دالة التشتت.

بأخذ اللوغاريتم لدالة الإمكان فتصبح المعادلة رقم (١) على الشكل التالي:

$$l\left(y;\beta\right)=\sum_{i=1}^{n}\left[\frac{y_i\theta_i-b\left(\theta_i\right)}{a_i\left(\phi\right)}+c_i\left(y_i,\phi\right)\right] \qquad (2)$$

وحيث أن الهدف هو تعظيم دالة الإمكان فنأخذ التفاضل الجزئي لها بالنسبة للمعلمة (β) ونساوي التفاضل بالصفر فتصبح على الصورة التالية:

$$\frac{\partial l\left(y;\beta\right)}{\partial\beta_j}=\sum_{i=1}^{n}\frac{\partial}{\partial\beta_j}\left[\frac{y_i\theta_i-b\left(\theta_i\right)}{a_i\left(\phi\right)}+c_i\left(y_i,\phi\right)\right] \qquad (3)$$

$$=\sum_{i=1}^{n}\frac{1}{a_i\left(\phi\right)}\left[y_i\frac{\partial\theta_i}{\partial\beta_j}-\frac{\partial b\left(\theta_i\right)}{\partial\beta_j}\right]$$

ونلاحظ أن المعلمة الطبيعية (θ) هي فقط دالة للمعلمات (β_j) والعلاقة بينهما معقدة إلى حد ما، وفي المقابل قيم (Y) ومعلمة التشتت (ϕ)ولتوضيح هذه العلاقة نستخدم المعادلات التالية:

$$\mu_i=b'^{(\theta_i)}$$

$$g\left(\mu_i\right)=\eta_i \qquad (4)$$

$$\eta_i=\beta_0+\beta_1x_{i1}+.........+\beta_kx_{ik}$$

وبتفاضل المعادلة السابقة باستخدام قاعدة السلسة نحصل على المعادلات التالية:

$$\frac{\partial l\left(y;\beta\right)}{\partial\beta_j}=\frac{\partial l}{\partial\theta_i}\frac{\partial\theta_i}{\partial\mu_i}\frac{\partial\mu_i}{\partial\eta_i}\frac{\partial\eta_i}{\partial\beta_j} \qquad (5)$$

$$=\sum_{i=1}^{n}\frac{\left(y_i-\mu_i\right)x_{ij}}{a_i\left(\phi\right)b''\left(\theta_i\right)g'\left(\mu_i\right)}$$

يمكن اعادة كتابة معادلة التفاضل حسب قاعدة السلسلة السابقة باستخدام دالة التباين $\operatorname{var}\left[y_i\right]=\left(\phi/w_i\right)v\left(\mu_i\right)$لنحصل على معادلة دالة الامكان التالية:

$$\frac{\partial l\left(y;\beta\right)}{\partial\beta_j}=\sum_{i=1}^{n}\frac{w_i\left(y_i-\mu_i\right)x_{ij}}{\phi V\left(\mu_i\right)g'\left(\mu_i\right)}=0 \qquad (6)$$

للقيم $\left(j=0,1,2,.......k\right)$ نلاحظ أن $\left(x_{i0}=1\right)$وذلك لأن المعلمة $\left(\beta_0\right)$ نقطة التقاصع مع محور الصادات حيث:

$$\mu_i=g^{-1}\left(\beta_0+\beta_1x_{i1}+\beta_2x_{i2}+.......+\beta_kx_{ik}\right) \qquad (7)$$

في المعادلة رقم (٧) يوجد عدد $(k+1)$ من المعادلات وعدد $(k+1)$ من المجاهيل وبحل هذه المعادلة نحصل على مقدر الإمكان الأعظم للمعلمات $\left(\beta_0+\beta_1+\beta_2+.......+\beta_k\right)$ وبالتالي لابد من حل هذه المعادلة رقمياً باحدى طرق التحليل العددي حيث اقترح العديد من الباحثين استخدام طريقة جوارزمية (Newton-Raphson) أو

استخدام خوارزمية (Fisher scoring) لحل معادلة الإمكان الأعظم والحصول على تقديرات معالم النموذج الخطي المعمم.

١٠-٣ نمذجة الوفيات باستخدام نموذج (GLM) [Lestari, 2019, 3-5]:

تم اختيار نموذج (GLM) لنمذجة وتقدير احتمال الوفاة $q_{(it)}$ بناء على عوامل الخطر التي تؤثر في هذا المعدل لأنه في نموذج (GLM) متغير الاستجابة متغير ثنائي (binary) يأخذ قيمتين فقط (1,0).

يمكن التعبير عن احتمال الوفاة بالصورة التالية:

$$q_{(it)} = pr\left[T_i - t \middle| T_i \geq t, X_i\right] \qquad (8)$$

أي أن احتمال الوفاة يمثل الاحتمال الشرطي للشخص (i) الذي يعيش عدد من السنوات ثم يموت عند العمر (t) حيث يعيش هذا الشخص في أول الفترة ويتمتع بمجموعة من الخصائص (عوامل الخطر) (X_i) وتمثل (T_i) متغير عشوائي متقطع لوقت الوفاة.

وبما أن متغير الاستجابة ثنائي وبالتالي يتبع توزيع ذي الحدين لذلك يتم استخدام دالة اللوجيت (Logit Link Function) كدالة للربط بين معدل الوفاة (q_{it}) والمتنبئ الخطي $(X_i\beta)$ وبفرض أن $\left(q_{(it)} = \left[Y_{it} \middle| X_i\right]\right)$

تأخذ دالة اللوجيت Logit Link Function الشكل التالي:

$$g\left(\mu_i\right) = \log\frac{\mu}{1-\mu} \qquad (9)$$

$$\left(X_i\beta\right) = \log\frac{q_{it}}{1-q_{it}} \qquad (10)$$

$$\log(\exp\left(X_i\beta\right)) = \log\frac{q_{it}}{1-q_{it}} \qquad (11)$$

$$\exp\left(X_i\beta\right) = \frac{q_{it}}{1-q_{it}} \qquad (12)$$

$$\exp\left(X_i\beta\right)\left(1-q_{it}\right) = q_{it} \qquad (13)$$

$$\exp\left(X_i\beta\right) - \exp\left(X_i\beta\right)\left(q_{it}\right) = q_{it} \qquad (14)$$

$$\exp\left(X_i\beta\right) = q_{it} + \exp\left(X_i\beta\right)\left(q_{it}\right) \qquad (15)$$

$$\exp\left(X_i\beta\right) = q_{it} + \left[1+\exp\left(X_i\beta\right)\ \right] \qquad (16)$$

$$q_{it} = \frac{\exp\left(X_i\beta\right)}{\left[1+\exp\left(X_i\beta\right)\ \right]} \qquad (17)$$

١٠ – ٤ النموذج المقترح لتسعير تأمينات الحياة وفقاً لنموذج (GLM):

أولاً:القسط الوحيد الصافي

تأمين الوقفية البحتة:

$$b_{k+1}=1 \quad , k=0,1,2,.........$$

$$v_{k+1}=\begin{cases} v^{k+1} , k=0,1,2,.......... \\ v^{n}, \quad k=n,n+1,....... \end{cases}$$

$$A_{\overset{1}{t:\overline{n|}}}=\sum_{k=0}^{n-1} v^{k+1}\,{}_{k}p_{t}\;q_{t+k}+v^{n}p_{t}$$

$$A_{\overset{1}{t:\overline{n|}}}=\sum_{k=0}^{n-1} v^{k+1}\,{}_{k}p_{t}\;\frac{\exp\left(X_{t}\beta\right)}{\left[1+\exp\left(X_{t}\beta\right)\;\right]}+v^{n}p_{t}$$

تأمين الوفاة المؤقت:

$$b_{k+1}=1 \quad , k=0,1,2,........,\mathrm{n}-1$$

$$A_{\overset{1}{t:\overline{n|}}}=\sum_{k=0}^{n-1} v^{k+1}\,{}_{k}p_{t}\;q_{t+k}$$

$$A_{\overset{1}{t:\overline{n|}}}=\sum_{k=0}^{n-1} v^{k+1}\,{}_{k}p_{t}\;\frac{\exp\left(X_{t}\beta\right)}{\left[1+\exp\left(X_{t}\beta\right)\;\right]}$$

تأمين مدى الحياة

$$A_{t}=\sum_{k=0}^{n-1} v^{k+1}\,{}_{k}p_{t}\;q_{t+k}$$

$$A_{t}=\sum_{k=0}^{n-1} v^{k+1}\,{}_{k}p_{t}\;\frac{\exp\left(X_{t}\beta\right)}{\left[1+\exp\left(X_{t}\beta\right)\;\right]}$$

ثانياً: القسط السنوي الصافي:

تأمين الوقفية البحتة	$P_{\overset{1}{t:\overline{n|}}}=\dfrac{A_{\overset{1}{t:\overline{n|}}}}{\ddot{a}_{t:\overline{n|}}}$
تأمين الوفاة المؤقت	$P_{t:\overline{n|}}=\dfrac{A_{\overset{1}{t:\overline{n|}}}}{\ddot{a}_{t:\overline{n|}}}$
تأمين مدى الحياة	$P_{t}=\dfrac{A_{t}}{\ddot{a}_{t}}$

١١. التطبيق العملي لنموذج (GLM) المقترح[٤]

عوامل الخطر ((Risk Factors(R.F) المقترحة التي تؤثر على احتمالات الوفاة والتي تمثل المتغيرات المستقلة المؤثرة على متغير الاستجابة (الوفيات) هي:

- العمر عند إصدار الوثيقة ويرمز له بالرمز X_1 .
- النوع ويرمز له بالرمز X_2 ويأخذ القيمة(1 للذكر و صفر للأنثى).
- التدخين ويرمز له بالرمز X_3 ويأخذ القيمة(1إذا كان الشخص مدخن والقيمة صفر إذا كان الشخص غير مدخن).
- الحالة الاجتماعية ويرمز لها بالرمز X_4 وتأخذ القيمة(1 للمتزوج و صفر لغير المتزوج).
- الحالة التعليمية ويرمز لها بالرمز X_5 وتأخذ القيمة (1 في حالة التعليم الجامعي والقيمة 0 في حالة التعليم الأقل من الجامعي).

متغير الاستجابة (الوفيات) ويرمز له بالرمز (Y) ويأخذ القيمة (1 في حالة الوفاة والقيمة صفر في حالة البقاء على قيد الحياة) ومتغير الاستجابة يتبع توزيع ذي الحدين وحيث أن تحديد النموذج المناسب هو أساس نمذجة الانحدار، فيمكن صياغة نموذج (GLM) عن طريق المعادلة التالية[Valente, 2020, 7]:

$$Y_i = g\left(X_i \beta_i\right) + \varepsilon$$

$$g\left(\mu\right) = \eta$$

$$g\left(\mu\right) = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}$$

$$\eta_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}$$

حيث (η) هي تركيبة خطية من المتغيرات التفسيرية حيث:

$$\eta_i = \left(X\beta\right)$$

حيث $\left(\eta_i\right)$ هي دالة الربط (Link Function)، والتي تفسر كيفية ارتباط المتوسط (μ) بالمتغيرات، وحيث أن متغير الاستجابة يتبع توزيع ذي الحدين فإن دالة الربط المستخدمة في نموذج (GLM) هي دالة اللوجيت والتي تأخذ الشكل الآتي:

٤ـ البيانات المستخدمة في البحث عبارة عن عينة من بيانات المؤمن عليهم المستخرجة من سجلات شركة مصر لتأمينات الحياة.

$$g\left(\mu_i\right)=\log\frac{\mu}{1-\mu}$$

الخطوة الأولى:

لتطبيق نموذج (GLM) المقترح لابد من ترميز البيانات الخاصة بالمؤمن عليهم كما يلي:

جدول رقم (٢)

جزء من بيانات المؤمن عليهم وفقاً لعوامل الخطر (R.F)

رقم الوثيقة	تاريخ الاصدار	السن	المدة	النوع	التدخين	الحالة الاجتماعية	مستوى التعليم	انتهاء الوثيقة
133941	01/12/17	50	19	M	N	M	H	وفاة
134257	01/12/16	37	20	M	N	M	H	
153532	01/05/14	17	30	M	N	M	L	
153533	01/05/17	18	30	M	N	M	L	
153730	01/06/19	33	12	M	N	M	L	
153757	01/06/15	33	20	M	N	S	L	وفاة
153772	01/06/17	44	12	F	N	S	L	وفاة
153848	01/06/17	37	20	M	N	S	H	وفاة
153849	01/06/15	32	20	F	N	M	H	وفاة
153858	01/04/17	42	18	M	S	S	H	
……..	……..	……..	……..	……..	……..	……..		
……..	……..	……..	……..	……..	……..	……..		
155384	01/09/17	20	20	M	N	M	L	
155404	01/09/15	41	15	F	N	M	L	

ملاحظات على الجدول السابق : (ذكر=M ، انثى=F) ، (مدخن=S ، غير مدخن = N) الحالة الاجتماعية (متزوج =M ، غير متزوج =S) ، مستوى التعليم (جامعي =H ، أقل من جامعي =L)

جدول رقم (٣)

ترميز بيانات المؤمن عليهم وفقاً لعوامل الخطر (R.F)

رقم الوثيقة	Y (الوفاة)	X1 (السن)	X2 (النوع)	X3 (التدخين)	X4(الحالة الاجتماعية)	X5 (التعليم)
133941	1	50	1	1	1	1
134257	0	37	1	1	1	1
153532	0	17	1	1	1	0
153533	0	18	1	1	1	0
153730	0	33	1	1	1	0
153757	1	33	1	1	0	0
153772	1	44	0	1	0	0
……..	……..	……..	……..	……..	……..	
……..	……..	……..	……..	……..	……..	
153858	0	42	1	0	0	1
153912	0	17	1	1	0	0

الخطوة الثانية:

تقدير معلمات النموذج باستخدام طريقة الإمكان الأعظم (MLE) كما هو موضح في الجدول رقم (٤) وللحصول على أكثر عوامل الخطر (R.F) تأثيراً على المتغير التابع (الوفاة) تم تطبيق طريقة (backward elimination method) واستخدام معيار (AIC) وذلك لاستبعاد العوامل أو المتغيرات التي لا تؤثر تأثيراً معنوياً على متغير الاستجابة (الوفاة).

جدول(٤)

تقدير معلمات نموذج (GLM) باستخدام طريقة (MLE)

Variables	Estimate	Std.Error	Z value	p-value
(Intercept)	-11. 5192	3.983692	4.476	0.0000535
X_1	2.4361	0.082421	2.621	0.04631
X_2	3.1851	1.426225	1.248	0.025530
X_3	-1.3116	1.232473	3.876	0.031403
X_4	-3.0301	1.025747	2.954	0.060691
X_5	-0.33328	1.273278	0.262	0.079351

يتضح من الجدول رقم (٤) ما يلي:

- ✓ **قيمة p-value (مستوى الدلالة) لمتغير السن (x_1) تساوي(0.04631) وهي أقل من مستوى المعنوية ٥٪ وبالتالي فإن السن له تأثير معنوي ذو دلالة احصائية على متغير الاستجابة عند مستوى معنوية ٥٪.**
- ✓ **قيمة p-value لمتغير النوع (x_2) تساوي (0.025530) وهي أقل من مستوى المعنوية ٥٪ وبالتالي فإن النوع له تأثير معنوي ذو دلالة احصائية على متغير الاستجابة عند مستوى معنوية ٥٪.**
- ✓ **قيمة p-value لمتغير التدخين (x_3) تساوي (0.031403) وهي أقل من مستوى المعنوية ٥٪ وبالتالي فإن التدخين له تأثير معنوي ذو دلالة احصائية على متغير الاستجابة عند مستوى معنوية ٥٪.**
- ✓ **قيمة p-value لمتغير الحالة الاجتماعية (x_4) تساوي (0.060691) وهي أكبر من مستوى المعنوية ٥٪ وبالتالي فإن متغير الحالة الاجتماعية ليس له تأثير معنوي ذو دلالة احصائية على متغير الاستجابة عند مستوى معنوية ٥٪.**

✓ قيمة p-value لمتغير المستوى التعليمي (x_5) تساوي (0.079351) وهي أكبر من مستوى المعنوية ٥٪ وبالتالي هذا المتغير ليس له تأثير معنوي ذو دلالة احصائية على متغير الاستجابة عند مستوى معنوية ٥٪.

وبالتالي يتم استبعاد الحالة الاجتماعية والمستوى التعليمي من النموذج المقترح لعدم تأثيرهما احصائياً على متغير الاستجابة.

مقاييس دقة التنبؤ بنموذج (GLM) المقترح

لتحديد كفاءة نموذج (GLM) المقترح في التنبؤ، يوجد ثلاثة مقاييس لتقدير الخطأ والتي تقارن القيم المتوقعة (E_i) التي نحصل عليها من النموذج بالقيم الفعلية (O_i) فإذا اقترب الفرق $(A_i - F_i)$ من الصفر فإن هذا يدل على دقة النموذج في التنبؤ، ومن هذه المقاييس ما يلي(Evans,2017:308):

١- المتوسط المطلق للانحراف MAD (The mean absolute deviation):

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|O_i - E_i\right|$$

٢- متوسط مربع الخطأ MSE (Mean square error):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(O_i - E_i\right)^2$$

٣- جذر متوسط مربع الخطأ RMSE (Root mean square error):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(O_i - E_i\right)^2}$$

دقة التنبؤ باستخدام نموذج (GLM) المقترح:

جدول(٥)

دقة التنبؤ باستخدام نموذج (GLM) المقترح

Model	Mean
MSE	0.0375
MAD	0.0375
RMSE	0.1609

يتضح من الجدول رقم (٥) أن:

جميع مقاييس دقة التنبؤ (MSE) ، (RMSE) ، (MAD) تقترب من الصفر وهذا يدل على دقة النموذج المقترح في التنبؤ.

جدول التصنيف The Classification Table

يتم من خلال هذه الجدول التعرف على كفاءة النموذج في التصنيف من خلال النسبة المئوية للحالات التي صنفت تصنيفاً صحيحا للمجموعة الأولى وكذلك للمجموعة الثانية والنسبة المئوية العامة للتصنيف باستخدام النموذج المقترح، كلما زادت هذه النسبة كلما كان النموذج المقدر أكثر كفاءة في التصنيف. (محمود مشعال،٢٠١١: ٦٧).

جدول (٦)

مصفوفة التصنيف

Observed		Predicted y: 0	Predicted y: 1	Percentage Correct
y	0	384	1	99.7
	1	13	3	18.8
Overall Percentage				96.5

يتضح من الجدول رقم (٦) ما يلي:

نسبة التصنيف الصحيح أو نسبة الكفاءة (Overall Percentage) تساوي (96.5%) وهي نسبة مرتفعة مما يدل على أن النموذج يمثل البيانات تمثيلاً جيداً وتدل أيضًا على كفاءة النموذج في التصنيف.

اختبار (AIC) لحذف المتغيرات المستقلة غير المعنوية

جدول(٧)

اختبار (AIC) لحذف المتغيرات المستقلة غير المعنوية

	Deviance	AIC	LRT	Pr(>Chi)
X_1	98.456	101.151	3.763	0.001
X_2	134.162	97.456	5.068	0.024374
X_3	96.877	94.877	2.489	0.114648
X_4	28.714	89.714	3.326	0.068194
X_5	45.714	85.714	1.326	0.03625
Deviance= 69.278			AIC: 91.278	

يتضح من الجدول رقم (٧) ما يلي:

- بالنسبة لعمود الانحراف يتبين أن المتغيرات التي تحتوي على أكبر القيم سواء كانت (Deviance) أو (AIC) هي أكثر المتغيرات تأثيرا على النموذج وإذا تم حذفها ستؤدي إلى زيادة قيمة الانحراف (Deviance) عن القيمة (69.278) أو زيادة قيمة معيار أكاكي (AIC) عن القيمة (91.278) مما يقلل من كفاءة نموذج (GLM) المقترح ، ومن

مخرجات الجدول السابق يتضح أن المتغيرات(X_2) هو أكثر المتغيرات تأثيرا ولا يجب حذفه من النموذج ويليه المتغير(X_1) ولا يجب حذفه أيضاً من النموذج وأخيراً المتغير(X_3) حسب قيم كلاِ الانحراف (Deviance) و معيار(AIC).

تقدير معلمات نموذج (GLM) المقترح بعد حذف المتغيرات غير المعنوية من النموذج.

جدول(٨)

تقدير معلمات نموذج (GLM) باستخدام طريقة (MLE) بعد التعديل

Variables	Estimate	Std.Error	Z value	p-value
(Intercept)	-16.865	0.225	2.274	0.0016
X_1	0.223	0.078436	-4.369	0.01307
X_2	3.125	1.237	1.598	0.0275
X_3	0.371	0.492	2.492	0.000

جدول(٩)

دقة التنبؤ باستخدام نموذج (GLM) المقترح بعد التعديل

Model	Mean
MSE	0.0184
MAD	0.0184
RMSE	0.1465

نلاحظ من الجدول رقم (٩) أن جميع مقاييس دقة التنبؤ تقترب من الصفر وهذا يدل على دقة النموذج في التنبؤ كما يتضح أيضًا أن قيمة جميع مقاييس دقة التنبؤ قد انخفضت بعد حذف المتغيرات غير المعنوية وغير المؤثرة في النموذج وهذا يدل على أن دقة النموذج في التنبؤ تزداد في حالة حذف المتغيرات غير المعنوية وإدخال المتغيرات المعنوية المؤثرة فقط نموذج (GLM) المقترح.

جدول(١٠)

جدول التصنيف بعد تعديل النموذج

<table>
<tr><th rowspan="3"></th><th rowspan="3" colspan="2">Observed</th><th colspan="3">Predicted</th></tr>
<tr><th colspan="2">y</th><th rowspan="2">Percentage Correct</th></tr>
<tr><th>0</th><th>1</th></tr>
<tr><td rowspan="3">Step 1</td><td rowspan="2">y</td><td>0</td><td>384</td><td>2</td><td>99.7</td></tr>
<tr><td>1</td><td>13</td><td>1</td><td>17.8</td></tr>
<tr><td colspan="2">Overall Percentage</td><td></td><td></td><td>96.7</td></tr>
<tr><td colspan="6">a. The cut value is .500</td></tr>
</table>

- نلاحظ من الجدول رقم (١٠) زيادة نسبة الكفاءة (Overall Percentage) إلى (96.7%) مما يدل على أن النموذج يمثل البيانات تمثيلاً جيداً وتدل أيضاً على زيادة كفاءة النموذج في التصنيف وذلك بعد حذف المتغيرات غير المؤثرة في النموذج.

جدول(١١)

اختبار (AIC) بعد حذف المتغيرات المستقلة غير المعنوية

	Deviance	AIC	LRT	Pr(>Chi)
X_1	97.120	95.354	5.727	0.016711
X_2	130.250	97.456	3.578	0.058541
X_3	136.389	93.021	7.939	0.004837
Deviance= 74.485				AIC: 92.485

بناء على التحليل السابق فإن النموذج المقترح لتقدير معدل الوفاة بناء على عوامل الخطر(R.F) هو:

$$q_{it} = \frac{\exp\left(X_i\beta\right)}{\left[1+\exp\left(X_i\beta\right)\ \right]}$$

$$q_{it} = \frac{\exp\left(\alpha+\beta_1X_1+\beta_2X_2+\beta_3X_3\right)}{\left[1+\exp\left(\alpha+\beta_1X_1+\beta_2X_2+\beta_3X_3\right)\ \right]}$$

$$q_{it} = \frac{\exp\left(-16.865+0.223X_1-3.125X_2+0.371X_3\right)}{\left[1+\exp\left(-16.865+0.223X_1-3.125X_2+0.371X_3\right)\ \right]}$$

فمثلاً لتقدير معدل الوفاة لشخص عمره ٣٥ عام أي ($x_1=35$) نوعه ذكر أي ($x_2=1$) ويدخن أي ($x_3=1$) فإن:

$$q_{35,1,1} = \frac{\exp\left(-16.865+0.223(35)-3.125(1)+0.371(1)\right)}{\left[1+\exp\left(-16.865+0.223(35)-3.125(1)+0.371(1)\right)\right]} = 0.00140$$

تقدير معدل الوفاة لشخص عمره ٣٥ عام أي ($x_1=35$) نوعه ذكر أي ($x_2=1$) و لا يدخن أي ($x_3=0$) فإن:

$$q_{35,1,0} = \frac{\exp\left(-16.865+0.223(35)-3.125(1)+0.371(0)\right)}{\left[1+\exp\left(-16.865+0.223(35)-3.125(1)+0.371(0)\right)\right]} = 0.00097$$

تقدير معدل الوفاة لشخص عمره ٤٥ عام أي ($x_1 = 45$) نوعه ذكر أي ($x_2 = 1$) ويدخن أي ($x_3 = 1$) فإن:

$$q_{45,1,1} = \frac{\exp\left(-16.865 + 0.223(45) - 3.125(1) + 0.371(1)\right)}{\left[1 + \exp\left(-16.865 + 0.223(45) - 3.125(1) + 0.371(1)\right)\right]} = 0.0129$$

تقدير معدل الوفاة لشخص عمره ٤٥ عام أي ($x_1 = 45$) نوعه ذكر أي ($x_2 = 1$) ولا يدخن أي ($x_3 = 0$) فإن:

$$q_{45,1,0} = \frac{\exp\left(-16.865 + 0.223(45) - 3.125(1) + 0.371(0)\right)}{\left[1 + \exp\left(-16.865 + 0.223(45) - 3.125(1) + 0.371(0)\right)\right]} = 0.00896$$

تقدير معدل الوفاة لشخص عمره ٤٥ عام أي ($x_1 = 45$) نوعه أنثى أي ($x_2 = 0$) وتدخن أي ($x_3 = 1$) فإن:

$$q_{45,0,1} = \frac{\exp\left(-16.865 + 0.223(45) - 3.125(0) + 0.371(1)\right)}{\left[1 + \exp\left(-16.865 + 0.223(45) - 3.125(0) + 0.371(1)\right)\right]} = 0.00156$$

تقدير معدل الوفاة لشخص عمره ٤٥ عام أي ($x_1 = 45$) نوعه أنثى أي ($x_2 = 0$) ولا تدخن أي ($x_3 = 0$) فإن:

$$q_{45,0,0} = \frac{\exp\left(-16.865 + 0.223(45) - 3.125(0) + 0.371(0)\right)}{\left[1 + \exp\left(-16.865 + 0.223(45) - 3.125(0) + 0.371(0)\right)\right]} = 0.00107$$

وبالمثل يمكن تقدير معدل الوفاة عند أى عمر ترغب شركة التأمين في حسابه باستخدام النموذج المقترح وهو:

$$q_{it} = \frac{\exp\left(-16.865 + 0.223X_1 - 3.125X_2 + 0.371X_3\right)}{\left[1 + \exp\left(-16.865 + 0.223X_1 - 3.125X_2 + 0.371X_3\right)\right]}$$

١٢. تطبيق نموذج (GLM) المقترح في تسعير أحد أنواع تأمينات الحياة:

تسعير التأمين المؤقت[٥]:

القسط الوحيد الصافي لوحدة النقود:

- بفرض أن شــخص ذكر عمره ٣٥ عام مدخن ويريد شــراء تأمين مؤقت مدته ٥ سنوات فيمكن حساب القسط الوحيد الصافي كما يلي:
- احتمال الوفاة بالنموذج المقترح لشخص عمره ٣٥ سنة ذكر ومدخن سبق تقديره بالنموذج المقترح $\left(q_{(35,1,1)} = 0.00140\right)$ وبالتالي فإن احتمال حياة هذا الشخص يساوي $\left(p_{(35,1,1)} = 1 - 0.00140 = 0.9986\right)$.
- معدل الخصم المستخدم ٤.٢٥٪

$$A^{1}_{t:\overline{n|}} = \sum_{k=0}^{n-1} v^{k+1}\ {}_{k}p_t\ q_{t+k}$$

$$A^{1}_{t:\overline{n|}} = \sum_{k=0}^{n-1} v^{k+1}\ {}_{k}p_t\ \frac{\exp\left(X_t\beta\right)}{\left[1+\exp\left(X_t\beta\right)\ \right]}$$

$$A^{1}_{35:\overline{5|}} = \sum_{k=0}^{5-1} v^{k+1}\ {}_{k}p_t\ \frac{\exp\left(X_t\beta\right)}{\left[1+\exp\left(X_t\beta\right)\ \right]}$$

$$A^{1}_{35:\overline{5|}} = \sum_{k=0}^{5-1} \left(1-0.0425\right)^{k+1} \left(0.9986\right)\left(0.00140\right)$$

$$A^{1}_{35:\overline{5|}} = 0.006148$$

القسط السنوي الصافي لوحدة النقود[٦]:

$$P^{1}_{35:\overline{5|}} = \frac{A^{1}_{35:\overline{5|}}}{\ddot{a}_{35:\overline{5|}}} = \frac{0.006148}{\sum_{k=0}^{5-1} \left(1-0.0425\right)^{(k+1)} \left(1-0.00140\right)} = \frac{0.006148}{4.391} = 0.0014$$

٥- تم التطبيق على التأمين المؤقت كمثال على تطبيق النموذج المقترح في التسعير ويمكن استخدام النموذج المقترح في تسعير أى نوع من أنواع تأمينات الحياة بتكرار نفس الخطوات.

6- $\ddot{a}_{t:\overline{n|}} = \sum_{k=0}^{n-1} \left(1-d\right)^{k+1} \left(1 - \frac{\exp\left(X_t\beta\right)}{\left[1+\exp\left(X_t\beta\right)\ \right]}\right)$

- **بفرض أن شخص ذكر عمره ٣٥ عام ولا يدخن ويريد شراء تأمين مؤقت مدته ٥ سنوات فيمكن حساب القسط الوحيد الصافي لوحدة النقود كما يلي :**

- **سبق تقدير معدل الوفاة بالنموذج المقترح لشخص عمره ٣٥ سنة ذكر وغير مدخن $\left(q_{(35,1,0)} = 0.00097\right)$ وبالتالي فإن احتمال حياة هذا الشخص يساوي $\left(p_{(35,1,0)} = 1 - 0.00097 = 0.99903\right)$**
- **معدل الخصم المستخدم ٤.٢٥٪**

$$A^{\;1}_{35:\overline{5|}} = \sum_{k=0}^{5-1} v^{k+1}\, {}_k p_t \, \frac{\exp\left(X_t\beta\right)}{\left[1+\exp\left(X_t\beta\right)\;\right]}$$

$$= \sum_{k=0}^{5-1}\left(1-0.0425\right)^{(k+1)}\left(0.99903\right)\left(0.00097\right)$$

$$A^{\;1}_{35:\overline{5|}} = 0.004261$$

القسط السنوي الصافي لوحدة النقود:

$$P^{\;1}_{35:\overline{5|}} = \frac{A^{\;1}_{35:\overline{5|}}}{\ddot{a}_{35:\overline{5|}}} = \frac{0.004261}{\sum_{k=0}^{5-1}(1-0.0425)^{(k+1)}(1-0.00097)} = \frac{0.004261}{4.391} = 0.00119$$

النتائج والتوصيات

أولاً: نتائج البحث

١- توجد علاقة معنوية ذات دلالة احصائية بين احتمال الوفاة وعوامل الخطر التالية: العمر ، النوع ، التدخين.

٢- معنوية النموذج المقترح لتقدير احتمالات الوفاة باستخدام النماذج الخطية المعممة الذي تم توفيقه باستخدام عوامل الخطر المختلفة وصلاحيته للتنبؤ.

٣- النموذج المقترح لتقدير احتمالات الوفاة وفقاً لعوامل الخطر هو:

$$q_{it} = \frac{\exp\left(-16.865 + 0.223X_1 - 3.125X_2 + 0.371X_3\right)}{\left[1 + \exp\left(-16.865 + 0.223X_1 - 3.125X_2 + 0.371X_3\right)\ \right]}$$

١- النموذج المقترح لتسعير تأمينات الحياة (تم التطبيق على التأمين المؤقت) وفقاً لنموذج GLM ووفقاً لعوامل الخطر هو:

$$A^{\ 1}_{t:\overline{n}|} = \sum_{k=0}^{n-1} v^{k+1}\ {}_k p_t\ \frac{\exp\left(X_t\beta\right)}{\left[1 + \exp\left(X_t\beta\right)\ \right]}$$

٤- القسط الوحيد الصافي للتأمين المؤقت إذا كان الشخص مدخن يكون أعلى إذا الشخص غير مدخن.

ثانيا : توصيات البحث

في ضوء ما توصل إليه البحث من نتائج يمكن الخروج بالتوصيات التالية:

١- يوصي الباحث بضرورة تحديث معدلات الوفاة المستخدمة في سوق التأمين المصري باستخدام النموذج المقترح.

٢- يوصي الباحث بتسعير تأمينات الحياة في ضوء عوامل الخطر التي تؤثر في احتمالات الوفاة، وذلك من أجل تحقيق العدالة بين حملة الوثائق بما يضمن تناسب الأقساط التأمينية مع درجات الخطر المختلفة.

٣- يوصي الباحث بأن تقوم هيئات الرقابة بوضع نموذج موحد لوثائق التأمين وملزم لجميع شركات التأمين، بما يضمن تجميع بيانات بشكل تفصيلي عن عوامل الخطر المختلفة، وبما يضمن المساعدة في دراسة المزيد من عوامل الخطر التي قد تؤثر في تقدير احتمالات الوفاة.

٤- يوصي الباحث مستقبلاً باستخدام طريقة (Chain Ladder) في تقدير احتمالات الوفاة.

مراجع البحث:

البحيري، مصطفى يسري عبد اللطيف" نموذج للتنبؤ بجداول الحياة المستقبلية باستخدام دالة ماكيهام"، المجلة العلمية للدراسات التجارية والبيئية، جامعة قناة السويس – كلية التجارة، المجلد السابع، العدد الرابع، ٢٠١٦.

الحصري, محمد حسن سيد, استخدام النماذج الخطية المعممة في تسعير تأمين السيارات التكميلي, مجلة البحوث الإدارية, مركز الإستشارات والبحوث والتطوير, أكاديمية السادات للعلوم الإدارية, ٢٠١٧.

محمد، أحمد محمد فرحان ، أبو زيد، محمد أحمد محمود " نموذج اكتواري مقترح لتسعير تأمين الحماية والادخار بالتطبيق على قطاع التأمين بالمملكة العربية السعودية"، مجلة البحوث المالية والتجارية، جامعة بورسعيد – كلية التجارة ، العدد الثاني، ٢٠٢١.

عبد الباقي، رضا صالح، متعال، محمود عبدالعال" استخدام نموذج التمهيد الأسى والانحدار التكيفى فى تقدير معدلات الوفاة بشركات التأمين"، مجلة كلية التجارة للبحوث العلمية، جامعة أسيوط – كلية التجارة ، العدد ٦٥، ٢٠١٨.

Abachi, jerusha," factors that influence pricing of life insurance products ", master thesis, united states, 2018.

Annette J. Dobson, Adrian G. Barnett" An Introduction to Generalized Linear Models ", Mathematics and Economics, Chapman & Hall, 3rd Edition 2008.

Björn Johansson, " The Basics of Pricing with GLMs ", Mathematics and Economics, Chapman & Hall, 2010.

D. Lestari," Generalized linear model (GLM) to determine life insurance premiums", AIP Conference Proceedings 2168, 020036 ,2019.

Evelien Brisard," Pricing of Car Insurance with Generalized Linear Models", Ph.D. thesis, University of Manchester, 2014.

Karl P. Murphy, Michael J. Brockman" Using Generalized Linear Models to Build Dynamic Pricing Systems ", Mathematics and Economics, Chapman & Hall, 3rd Edition 2008.

Kwon, Jones," The impact of the determinants of mortality on life insurance and annuities ", Mathematics and Economics, vol.38, 2006.

Shakil, M., Kibria, B, . M,. G,. and Singh, . J, . N, . " A New Family of Distribution Based on the Generalized Pearson Differential Equation with Some Applications " , Austrin Journal of statistics , Vol 39, 2010.

Valente, Ana," Health Insurance Pricing with Generalized Linear Models ", master thesis, university of Lisbon, 2020.

Yuqing Zhang," Dynamic Pricing with Application to Insurance", MSc. thesis, Brussel University, 2020.