استخدام التحليل البيزي لتقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة

د/ عماد عبدالجليل علي اسماعيل

المدرس بقسم التأمين -كلية التجارة – جامعة القاهرة

ملخص البحث:

يتعلق هذا البحث بتقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة The Maximum Probable Yearly (MPY)، في فرع تأمين المركبات الشامل بالمملكة العربية السعودية، باستخدام التحليل البيزي Bayesian Analysis للبيانات، في تقدير عدد المطالبات المتوقع ، وكذلك قيمة التعويضات المتوقعة، على مستوى محفظة التأمين، عن طريق تحديد كل من التوزيع الاحتمالي التنبؤي لعدد المطالبات Predictive Number of Claims Distribution، والتوزيع الاحتمالي التنبؤي لقيم المطالبات Predictive Claims Distribution، ثم تحديد التوزيع الاحتمالي التنبؤي الإجمالي لقيم المطالبات Predictive Aggregate Claims Distribution ، والذي يعطي معلومات فنية شبه كاملة عن هذا النوع من التأمين، منها تقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة.

مشكلة البحث:

يغطي تأمين المركبات الشامل بشركات التأمين السعودية، أخطار المسئولية المدنية عن الإصابات الجسدية والوفاة التي تلحق بالغير، وأخطار المسئولية المدنية عن الخسائر التي تلحق بممتلكات الغير، وكذلك الأضرار التي تلحق بسيارة المؤمن له من كافة الأخطار التي تتعرض لها فيما عدا الاستثناءات الواردة بوثيقة التأمين، فضلاً عن تغطية مصاريف العلاج الطبي الطارئ، في حالة إصابة السائق أو ركاب المركبة المؤمن عليها*، ونظراً لتعرض بعض مناطق المملكة في الفترة الأخيرة (2008 - 2011) لحوادث السيول، وخصوصاً منطقة جدة، مما كبد شركات التأمين العاملة بالمملكة خسائر فادحة، الأمر الذي اضطرت معه هذه الشركات، إلى تحصيل قسط إضافي لتغطية خطر البرد والفيضانات†، وتشير دراسة استطلاعية قام بها الباحث‡، إلى أن شركات التأمين السعودية، تفتقد إلى النماذج الكمية المستندة إلى الخبرة ، والتي تستخدم في دعم واتخاذ الكثير من القرارات المتعلقة ببعض النواحي الفنية في التأمين، كعملية تسعير الأخطار، وفرض أقساط إضافية، وغيره، حيث أن معظم القرارات الفنية في سوق التأمين السعودي، تكون غير مستندة إلى الخبرة، وتعتمد بدرجة كبيرة على معلومات محدودة وردود أفعال أنية، ومشكلة هذا البحث تتمثل في الإجابة على السؤال التالي:

ما هو النموذج الكمي المستنتج من خبرة شركات التأمين محل الدراسة، والذي سيستخدم في تقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة لفرع تأمين المركبات الشامل؟

---

* وثيقة تأمين المركبات الشامل، الشركة التعاونية للتأمين، المملكة العربية السعودية

† شركة التعاونية للتأمين تقوم بتغطية هذا الخطر مقابل قسط إضافي 50 ريالاً، خلال 2010 - 2011

‡ مقابلات شخصية مع العديد من المسئولين بشركات التأمين السعودية

هدف البحث:

يتمثل الهدف الرئيسي لهذا البحث في تقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة لفرع تأمين المركبات الشامل، عن طريق تحديد كل من:

1. التوزيع الاحتمالي التنبؤي لعدد المطالبات.
2. التوزيع الاحتمالي التنبؤي لقيم المطالبات.
3. التوزيع الاحتمالي التنبؤي لإجمالي قيم المطالبات

أهمية البحث:

يتعلق هذا البحث بتأمين المركبات، وهو ثاني أكبر فروع التأمينات العامة من حيث حجم الأقساط المكتتبة، فقد بلغ متوسط حجم الأقساط 2308.8 مليون ريال بأهمية نسبية 26.2% من إجمالي حجم أقساط التأمين المكتتبة في شركات التأمين التعاونية العاملة بالمملكة، خلال الفترة من 2005 – 2009 (مؤسسة النقد العربي السعودي، 2009)، ويأتي مباشرة بعد التأمين الطبي الذي يحتل المرتبة الأولى، حيث يبلغ متوسط حجم الأقساط المكتتبة فيه 3750.8 مليون ريال، بأهمية نسبية 37.6 % عن نفس الفترة، ومن هنا تكمن مدى الخطورة التي تتعرض لها محفظة اكتتابات الشركة، وخصوصاً فيما يتعلق بتقدير أقصى خسارة مادية محتملة نتيجة وقوع بعض الكوارث الطبيعية، وبالتالي لابد من استنتاج نماذج رياضية وإحصائية، وفقاً لخبرة الشركة الماضية والحالية، تستخدم في عمليات التقدير المختلفة، حتى تتجنب الشركة خطر التقلبات العكسية في النتائج المتوقعة.

منهجية البحث:

نعتمد في هذا البحث على أسلوب التجميع الميداني للبيانات الخاصة بتأمين المركبات الشامل، والتي تشتمل على (عدد المطالبات، قيم المطالبات، حجم الأقساط المباشرة، تاريخ المطالبة، تاريخ سداد التعويض) من إحدى شركات التأمين السعودية، وكذلك نعتمد في هذا البحث على إتباع مفهوم نظرية نماذج الخطر التجميعية Collective Risk Models حيث إنه ينظر لمحفظة التأمين على إنها مفتوحة Open Portfolio أي تسمح بدخول وخروج أي وثيقة، وهذا يتناسب مع طبيعة عمليات التأمين بعكس الحال مع نماذج الخطر الفردية التي تقوم على أن المحفظة مغلقة Closed Portfolio. [Ibrahim Morgan, 1983]

وفي ظل استخدام نماذج الخطر التجميعية Collective Risk فإنه يتم استخدام التحليل البيزي للبيانات Data Bayesian analysis في تقدير عدد المطالبات المتوقع على مستوى محفظة التأمين، وكذلك قيمة التعويضات المتوقعة، وذلك من خلال التوزيع الاحتمالي لعدد المطالبات والتوزيع الاحتمالي لقيم المطالبات، ثم إيجاد التوزيع الاحتمالي الإجمالي لقيم المطالبات Aggregate claims distribution [Dickson, 1998]

الدراسة التطبيقية

أولاً: التوزيع الاحتمالي التنبؤي لعدد مطالبات تأمين المركبات الشامل :

تم تحديد التوزيع الاحتمالي التنبؤي لعدد المطالبات Predictive Number of Claims Distribution ، في ضوء مفهوم نظرية الخطر التجميعية Collective Risk مع إتباع خطوات التحليل البيزي للبيانات Bayesian Data Analysis ، وذلك بتحديد كل من:

* التوزيع الاحتمالي لعدد المطالبات.

* التوزيع الاحتمالي القبلي لمعلمة التوزيع الاحتمالي لعدد المطالبات.

* التوزيع الاحتمالي البعدي لمعلمة التوزيع الاحتمالي لعدد المطالبات.

* التوزيع الاحتمالي التنبؤي لعدد المطالبات.

1- التوزيع الاحتمالي لعدد المطالبات The Data Distribution

بفرض أن X متغير عشوائي يعبر عن عدد مطالبات تأمين المركبات الشامل، حيث إن $(X = 0, 1, 2, \ldots..)$، وإن هذا المتغير له توزيع احتمالي نظري منفصل.

والجدول التالي يوضح توزيع عينة عشوائية حجمها 2588 وثيقة، من وثائق تأمين المركبات الشامل، حسب عدد المطالبات خلال سنة 2010، بإحدى شركات التأمين السعودية.

جدول رقم (1)

توزيع عدد وثائق تأمين المركبات الشامل حسب عدد المطالبات

عدد المطالبات	عدد الوثائق
صفر	1859
1	615
2	97
3	14
4	3
مجموع	2588

ومن التحليل الإحصائي للبيانات نجد أن:

متوسط عدد المطالبات $E(\underline{x_1}) = .33346213$

تباين عدد المطالبات $\operatorname{var}(\underline{x_1}) = .343727$

ومن أكثر التوزيعات الاحتمالية النظرية استخداماً للتعبير عن عدد المطالبات في مجال التأمينات العامة توزيع بواسون Poisson Distribution وتوزيع ثنائي الحدين السالب Negative Binomial Distribution، ويفضل استخدام توزيع ثنائي الحدين السالب على توزيع بواسون في هذا البحث للأسباب الآتية:

ا -أن توزيع ثنائي الحدين السالب يأخذ في اعتباره عدم التجانس بين الوحدات المؤمن عليها. (ممدوح حمزة، 1990)

ب -أن توزيع ثنائي الحدين السالب ذو معلمتين، وبالتالي فهو أكثر دقة من توزيع بواسون ذو المعلمة الواحدة. [Andrew Gelman,2004]

ج -متوسط توزيع بواسون = تباينه وهذا نادراً ما يحدث في مجال التأمينات العامة.

د -إن توزيع ثنائي الحدين السالب هو توزيع مختلط Mixture Distribution، وذلك عندما يكون المتغير العشوائي يتبع توزيع بواسون ومعلمة توزيع بواسون تتبع توزيع جاما Gamma Distribution. [Andrew Gelman,2004]

وتأخذ دالة توزيع ثنائي الحدين السالب الشكل التالي:

$$Neg - bin(x / \propto, \beta) = \int poisson(x / \theta) Gamma(\theta / x, \beta) d\theta$$

$$f(x; r_1, \theta) = C_x^{r_1+x-1} \theta^{r_1} (1-\theta)^x ; X_1 = 0,1,2.....$$

حيث أن:

r_1 ، θ : معالم توزيع ثنائي الحدين السالب

$$1 \geq \theta \geq 0$$

r_1 : عدد صحيح موجب

- تقدير معالم توزيع ثنائي الحدين السالب (r_1, θ)

متوسط عدد المطالبات للتوزيع الفعلي = متوسط توزيع ثنائي الحدين السالب

تباين عدد المطالبات للتوزيع الفعلي = تباين توزيع ثنائي الحدين السالب

وبحل المعادلتين (1)، (2) نجد أن قيم معالم توزيع ثنائي الحدين السالب هي:

$$\frac{r_1(1-\theta)}{\theta} = .33346213 - (1) \qquad \frac{r_1(1-\theta)}{\theta^2} = .34372716 - (2)$$

$$r_1 = 10.83 \qquad \theta = .97$$

- اختبار جودة التوفيق* Goodness-of-Fit

الفرض العدمي: عدد مطالبات تأمين المركبات الشامل يتبع توزيع ثنائي الحدين السالب.

ونظراً لأنه توجد تكرارات فعلية ومتوقعة أقل من 5 فإنه تم استخدام اختبار كولمجروف-سيمرنوف Kolomogorov-Simnrov لاختبار ذلك الفرض لأن اختبار كا^2 يشترط أن يكون التكرار في أي خلية لا يقل عن 5 وحتى لو توافرت جميع شروط اختبار كا^2 فإنه يفضل استخدام اختبار كولمجروف-سيمرنوف لأنه أكثر قوة من اختبار كا^2 (سمير، عاشور 2005)

وقد وجد أن إحصائي الاختبار 0.00309 وهو أقل من القيمة الحدية الجدولية التي بلغت 0.026734، وبناءاً عليه تم قبول الفرض العدمي إن عدد مطالبات تأمين المركبات الشامل يتبع توزيع ثنائي الحدين السالب بدرجة ثقة 95%، أي أن:

$$X \sim NB\,(r1,\theta);\ r_1 = 10.83 \qquad \theta = 0.97$$

ووفقاً لنظرية الخطر التجميعية، فإن المعلمة θ تأخذ قيماً مختلفة، نتيجة أن شركة التأمين تسمح بدخول أو خروج الوثائق من محفظة اكتتابها. وبالتالي فإن المعلمة θ تعتبر متغيراً عشوائياً لها توزيع احتمالي (θ) π، ويسمى التوزيع الاحتمالي القبلي للمعلمة θ، أما المعلمة r_1 لها قيمة معينة ثابتة خلال فترة الخبرة وتختلف هذه القيمة باختلاف فترة الخبرة.

---

* راجع ملحق الدراسة

2 – تحديد التوزيع الاحتمالي القبلي $\pi(\theta)$ The Prior Distribution

إن التوزيع الاحتمالي القبلي $\pi(\theta)$ للمعلمة θ هو توزيع احتمالي يصف كل المعلومات والخبرات المتوفرة حول المعلمة θ قبل الحصول على العينة (الصياد جلال، 1993) أو معلومات مشابهة أو معلومات من دراسات سابقة، وسوف يتم الاعتماد فى هذا البحث على عينة من خبرة تأمين المركبات الشامل بنفس شركة التأمين، خلال سنة 2009 قبل الحصول على العينة

والتوزيع الاحتمالي القبلي $\pi(\theta)$ للمعلمة θ خلال فترة الخبرة الأساسية (سنة 2010) يعتبر هو التوزيع الاحتمالي البعدي خلال فترة الخبرة القبلية (سنة 2009) ، والذي يعتمد في تحديده على دالة الإمكان Likelihood Function خلال فترة الخبرة القبلية (سنة 2009) ودالة التوزيع الاحتمالي القبلي خلال نفس فترة الخبرة (سنة 2009) والذي يعتبر التوزيع الإحتمالى البعدي خلال الفترة القبلية ما قبل سنة 2009 والذي لا يتوافر عنها أي بيانات.

وعندما لا تتوافر أي معلومات عن المعلمة θ ، وكانت θ تتراوح قيمتها بين الصفر، الواحد فإن دالة التوزيع الاحتمالي البعدي حيث لا توجد عينة $\pi(\theta)$ تتبع التوزيع المنتظم وهو على الشكل التالي (William M. Bolstad, 2007):

$$\pi(\theta)=\frac{1}{b-a};a<\theta<b$$

وحيث إن المتغير العشوائي (عدد المطالبات) خلال فترة الخبرة يتبع توزيع ثنائي الحدين السالب وأن:

$$0<\theta<1$$

أي أن: $a = 0, \qquad b = 1$

فإن التوزيع الاحتمالي القبلي يكون على الشكل التالي:

$$\pi(\theta)=\frac{1}{b-a}=1$$

أي أن التوزيع الاحتمالي البعدي خلال فترة الخبرة ما قبل سنة 2009 مع عدم وجود أي معلومات $\pi(\theta)=1$، وهو يعتبر التوزيع الاحتمالي القبلي لفترة الخبرة (سنة 2009).

- تحديد التوزيع الاحتمالي البعدي خلال فترة الخبرة (سنة 2009)

لتحديد التوزيع الاحتمالي البعدى $\pi(\theta/\underline{X}_1)$ ، أي التوزيع الاحتمالي بعد الحصول على العينة $\underline{x}_1$ وهي حجم الخبرة خلال الفترة (سنة 2009)، يتم دمج دالة الإمكان خلال هذه الفترة مع التوزيع الإحتمالى القبلي $\pi(\theta)=1$ ، حيث أن التوزيع البعدي يتناسب مع التوزيع المشترك بينهما (Peter Congdon, 2003).

ويلاحظ أن التوزيع الاحتمالي القبلي لفترة الخبرة الأساسية (سنة 2010)يكــون على شكل دالة بيتا بمعالم* :

$$\beta(n_0 r_0 + 1, T_0 + 1)$$

3- تحديد التوزيع الاحتمالي البعدي $\pi(\theta/X)$ The Posterior Distribution

تم استنتاج التوزيع الاحتمالي البعدي لفترة الخبرة الأساسية (سنة 2010) بــدمج دالة الإمكان لفترة الخبرة الأساسية $f(X_1/\theta)$ مع التوزيع الاحتمــالي القبلــي لفترة الخبرة الأساسية $\pi(\theta/X_0)$ (Peter Congdon, 2003) ، حيــث أن دالــة الإمكان خلال فترة الخبرة الأساسية هي:

$$f(\underline{x}_1/\theta)=\prod_{i=1}^{n_1} C_{x_i}^{r_i+x_i-1}\theta^{r_i}(1-\theta)^{x_i}$$

---

* راجع ملحق الدراسة

ونجد أن المقدار $C_{x_i}^{r_1+x_i-1}$ لا يعتمد على θ فيمكن اختزاله في ثابت التناسب،

$$f(\underline{X}_1/\theta) \propto \theta^{n_1 r_1}(1-\theta)^{\sum_{i=1}^{n_1} x_1}$$

والمقدار $T_1 = \sum_{i=1}^{n1} x_i$ يمثل إحصاء كافي عبارة عن إجمالي عدد المطالبات، n1 تمثل حجم العينة خلال فترة الخبرة الأساسية.

وبالتالي فإن دالة الإمكان خلال فترة الخبرة الأساسية تكون على الشكل التالي:

$$f(\underline{X}_1/\theta) \propto \theta^{n_1 r_1}(1-\theta)^{T_1}$$

وبالتالي فإنه يتم تحديد التوزيع الاحتمالي البعدي خلال فترة الخبرة الأساسية كما يلي:

$$\pi(\theta/\underline{X}_1) \propto f(\underline{X}_1/\theta)\pi(\theta/\underline{X}_1)$$

$$\pi(\theta/\underline{X}_1) \propto \theta^{n_1 r_1}(1-\theta)^{T_1}\theta^{n_0 r_0}(1-\theta)^{T_0}$$

$$\pi(\theta/\underline{X}_1) \propto \theta^{n_0 r_0 + n_1 r_1}(1-\theta)^{T_0+T_1}$$

ويلاحظ أن التوزيع الاحتمالي البعدي خلال فترة الخبرة الأساسية يتناسب مع توزيع بيتا بمعالم $\beta\,(n_0 r_0 + n_1 r_1 + 1,\ T_0 + T_1 + 1)$

وتكون دالة التوزيع الاحتمالي البعدي خلال فترة الخبرة الأساسية هي دالة بيتا على الشكل التالي:

$$\pi(\theta/\underline{X}_1) = \frac{1}{\beta(n_0 r_0 + n_1 r_1 + 1, T_0 + T_1 + 1)}\theta^{n_0 r_0 + n_1 r_1}(1-\theta)^{T_0+T_1}$$

حيث أن:

$n_0 = 525$ $r_0 = 1.95$ $T_0 = 109$

$n_1 = 2588$ $r_1 = 10.83$ $T_1 = 863$

4- التوزيع الاحتمالي التنبؤي ($P(X_{n+1}/\underline{x}_1)$ The Predictive Distribution

تم استنتاج هذا التوزيع بإجراء التكامل المحدود لحاصـــل ضـــرب دالـــة التوزيـــع الاحتمالي الأصلي للبيانات للسنة القادمة و دالة التوزيع الاحتمالي البعدي خـــلال فترة الخبرة الأساسية (William M. Bolstad, 2007)، حيث أن:

$$P(X_{n+1}/\underline{x}_1) = \int_0^1 f(X_{n+1}/\theta)\pi(\theta/\underline{x}_1).d\theta$$

والتكامل يكون بالنسبة لـــ θ ومن صفر إلى 1 لأن قيمة θ تتراوح من صـــفر إلى 1.

و بعد إجراء الاختصارات والاستنتاجات اللازمة نجد أن دالة التوزيع الاحتمالي التنبؤي لعدد المطالبات تكون على الشكل التالي* :

$$P(X_{n+1}/\underline{x}_1) = C_{X_{n+1}}^{X_{n+1}+11-1}(.9681)^{11}(1-.9681)^{X_{n+1}}$$

* راجع ملحق الدراسة

والدالة السابقة هي دالة توزيع ثنائي الحدين السالب أي أن دالة التوزيع الاحتمالي التنبؤى لعدد مطالبات تأمين المركبات الشامل هي دالة توزيع ثنائي الحدين السالب بمعالم (11,.9681) أي أن:

$$E(X_{n+1}) = \frac{11(1-.9681)}{.9681} = .36$$

$$Var(X_{n+1}) = \frac{11(1-.9681)}{(.9681)^2} = .37$$

- **العزوم الأربعة للتوزيع الاحتمالي التنبؤي لعدد المطالبات حول الصفر:**

حيث أن التوزيع التنبؤي لعدد المطالبات يتبع توزيع ثنائي الحدين السالب فإن دالة توليد العزوم حول الصفر تأخذ الشكل التالي (الصياد جلال، 1993):

$$M'(t) = \theta^{r_1}\left[1.-(1-\theta)e^{t}\right]^{-r_1}$$

حيث إن $M'(t)$ الدالة المولدة للعزوم حول الصفر

$$\theta = .9681 \qquad r_1 = 11$$

ويمكن الحصول على العزوم الأربعة حول الصفر بالاشتقاق المتتالي بالنسبة لـ (t) ثم التعويض بعد ذلك عن t=0 بعد كل عملية اشتقاق. وباستخدام برنامج MathCAD يتم الحصول على المشتقة الأولى، ثم التعويض عن t=0 فنحصل

على العزم الثاني وهكذا. وبالتالي فإن العزوم الأربعة حول الصفر تتولد من تفاضل الدالة التالية:

$$M'_X(t) = \frac{d^x}{dt^x} M(t)$$

حيث أن: $X = 1, 2, 3, 4$ $t = 0$

العزوم الأربعة حول الصفر هي:

$$M'_1(X) = E(X_{n+1}) = .362463 \qquad M'_2(X) = .505785$$

$$M'_3(X) = .853825 \qquad M'_4(X) = 1.762411$$

حيث أن العزم الأول حول الصفر هو متوسط التوزيع الاحتمالي التنبؤي.

العزوم الأربعة للتوزيع الاحتمالي التنبؤي حول المتوسط:

إذا كانت X متغيراً عشوائياً منفصلاً، فإن دالة توليد العزوم حول المتوسط يكـــون باستخدام الدالة التالية (Lester D. Taylor, 1997):

$$E\left[(X - M)^m\right] = \Sigma (X - M)^m F(X)$$

وبالتالي فإن العزوم الأربعة للتوزيع الاحتمالي التنبؤي حول المتوسط، باســتخدام برنامج MathCAD ، هي:

$$M_1(X) = 0 \qquad M_3(X) = .39908$$

$$M_2(X) = Var(X_{n+1}) = .374415 \qquad M_4(X) = .871408$$

ثانياً: تحديد التوزيع الاحتمالي التنبؤي لقيم مطالبات تأمين المركبات الشامل

تم إتباع نفس خطوات التحليل الإحصائي البيزي المتبعة بالنسبة لعدد المطالبات، وذلك بتحديد التوزيع الاحتمالي الأصلي للبيانات والتوزيع الاحتمالي القبلي، والتوزيع الاحتمالي البعدي للمعلمة، والتوزيع الاحتمالي التنبؤي للمتغير العشوائي للسنة القادمة، على النحو التالي:

1 – تحديد التوزيع الاحتمالي لقيم مطالبات تأمين المركبات الشامل

بفرض أن y_i متغير عشوائي يعبر عن قيم مطالبات تأمين المركبات الشامل، حيث أن ($i=1, 2, \ldots\ldots, n$) ، وأن هذا المتغير له توزيع احتمالي متصل.

والجدول التالي يوضح توزيع عدد وثائق تأمين المركبات الشامل، حسب فئات قيم المطالبات خلال فترة الخبرة الأساسية (2010).

جدول رقم (4)

توزيع عدد وثائق تأمين المركبات الشامل، حسب فئات

قيم المطالبات خلال فترة الخبرة الأساسية (2010)

فئات الخسارة	عدد الوثائق
صفر -	750
10000 -	77
20000 -	21
30000 -	10
50000 -	3
80000 -100000	2
مجموع	863

ومن نتائج التحليل الإحصائي للبيانات نجد أن:

متوسط قيم المطالبات $E(\underline{y}_1) = 5203$

تباين قيم المطالبات $Var(\underline{y}_1) = 68178049$

وقد قام الباحث باختبار جودة التوفيق Goodness-of-Fit لقيم المطالبات الفعلية مع معظم التوزيعات الاحتمالية المتصلة باستخدام برنامج StatGraphics و تبين أنها تتفق مع التوزيع اللوغاريتمي الطبيعي

Lognormal Distribution؛ حيث إن دالة كثافتـــه الاحتماليـــة هـــي :

(Athanasios Papoulis, 1990)

$$f(Y/\mu) = \frac{1}{\sigma_1 y \sqrt{2\Pi}} e^{-\frac{1}{2\sigma_1^2}(\ln y - \mu)^2}$$

ومعالم التوزيع (μ, σ_1^2)

حيث أن:

$$-\infty < \mu < \infty \qquad \& \qquad y > 0$$

– تقدير معالم التوزيع اللوغاريتمي الطبيعي:

حيث إن متوسط وتباين التوزيع اللوغاريتمي الطبيعي هما:

$$E(y_1) = e^{\left(\mu + \frac{1}{2}\sigma_1^2\right)} = 5203 - (1)$$

$$Var(y_1) = e^{2\mu + \sigma_1^2}\left(e^{\sigma_1^2} - 1\right) = 68178049 \; - (2)$$

وبحل المعادلتين السابقتين نجد أن قيم معا لم التوزيع اللوغاريتمي الطبيعي هي:

$$\mu = 7.928 \qquad \sigma_1^2 = 1.258$$

ووفقاً للتحليل البيزي فإن المعلمة μ متغير عشوائي، بينما المعلمة σ_1^2 تم اعتبارها ثابتة خلال فترة الخبرة والتي تختلف باختلاف فترة الخبرة.

2 – تحديد التوزيع الاحتمالي القبلي π (μ) خلال فتـــرة الخـــبرة الأساســـية (2010):

عندما لا تتوافر أي معلومات عن المعلمة μ ، وذلك خلال فترة ما قبـــل 2009 وكانت قيمة μ تتراوح ما بين $(\infty > \mu > \infty -)$ ، فإن التوزيـــع الاحتمـــالي البعدي حيث لا توجد عينة هو π (μ) ، ويكون تقدير قيمة μ هـــو μ^* وهـــي قيمة μ قبل الحصول على العينة، حيث إن: (الصياد جلال، 1993)

$$\pi\,(\mu) \;\propto d_{\mu}$$

وهو توزيع غير كامل أو ناقص Improper، لأن تكامل هذا التوزيع على مدى المعلمة μ يعطي ∞. و على ذلك فإن تقدير قيمة μ في ظل عدم توافر أي بيانات وبفرض أن μ تتبع التوزيع الطبيعي بمعالم (a, b^2) هو:

$$\mu^* = \frac{n/\sigma^2 \bar{y} + a/b^2}{n/\sigma^2 + 1/b^2}$$

ويلاحظ من التقدير السابق لـــ μ أنه إذا كانـــت $(b^2 \;=\infty)$ أي أن تبـــاين البيانات ما قبل سنة 2009 بلغ ∞ أي لا توجد بيانات، فـــإن تقـــدير μ قبـــل الحصول على العينة هو: $\mu^* = \bar{y}$ أي أن تقدير μ سوف يكون بالاعتماد على العينة أو فترة الخبرة القبلية (سنة 2009). وبالتالي فإن التوزيع الاحتمالي القبلـــي

يعتمد على دالة الإمكان خلال هذه الفترة والذي يعتبر التوزيع البعدي؛ حيث لا توجد أي معلومات سابقة والذي يكون بدوره هو التوزيع القبلي لفترة الخبرة الأساسية 2010 كما يلي:

- تحديد دالة الإمكان خلال سنة 2009

الجدول التالي يوضح توزيع عدد وثائق تأمين المركبات الشامل، حسب فئات قيم المطالبات خلال سنة 2009

جدول رقم (5)

توزيع عدد وثائق تأمين المركبات الشامل، حسب فئات

قيم المطالبات خلال سنة 2009

فئات الخسارة	عدد الوثائق
صفر-	72
10000-	28
30000-	3
50000-	3
90000-120000	3
مجموع	109

وحيث أن قيم المطالبات يتبع التوزيع اللوغاريتمي الطبيعي خـــلال فتـــرة الخبرة الأساسية، فإن قيم المطالبات يتبع التوزيع اللوغاريتمي الطبيعي خلال فتـــرة الخبرة القبلية وبالتالي فإن:

$$E\left(\underline{y}_0\right) = 12721\ .55 = e^{\mu + \frac{1}{2}\sigma_0^2}$$

$$Var\left(\underline{y}_0\right) = 399241875 = e^{2\mu+\sigma_0^2}\left(e^{\sigma_0^2} - 1\right)$$

وبحل المعادلتين السابقتين نجد أن معالم التوزيع اللوغاريتمي الطبيعي خلال فترة الخبرة القبلية هي:

$\mu = 8.829$ $\sigma_0^2 = 1.243$

و حيث أن قيم مطالبات تأمين المركبات الشامل، تتبع التوزيع اللوغاريتمي الطبيعي بمعالم (μ, σ_0^2) ويكون شكل التوزيع هو:

$$f\left(y / \mu\right) = \frac{1}{\sigma_0 y \sqrt{2\Pi}} e^{-\frac{1}{2\sigma_0^2}|\ln y - \mu|^2}$$

فنجد أن دالة الإمكان خلال فترة الخبرة القبلية تكون على الشكل التالي:

$$f\left(\underline{y}_0 / \mu\right) = \prod_{i=1}^{n_0} \frac{1}{\sigma_{01} y_{0i} \sqrt{2\Pi}} e^{\frac{1}{2\sigma_0^2}(\ln y_{0i} - \mu)^2}$$

$$f\left(\underline{y}_0 / \mu\right) \propto e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n_0}(\ln y_i-\mu)^2}$$

وبعد إجراء الاختصارات اللازمة نجد أن دالة الإمكان تكون على الشكل التالي:

$$f\left(\underline{y}_0 / \mu\right) \propto e^{-\frac{n_0}{2\sigma_0^2}\left(\mu-\frac{T_0}{n_0}\right)^2}$$

ويتضح أن شكل دالة الإمكان تتناسب مع التوزيع الطبيعي ومعالم التوزيع $\left(\frac{T_0}{n_0},\frac{\sigma_0^2}{n_0}\right)$ باعتبار أن المتغير العشوائي هنا هو المعلمة μ وبالتالي فإن متوسط هـذا التوزيـع $\frac{T_0}{n_0}$ وتباينه $\frac{\sigma_0^2}{n_0}$.

ودالة الإمكان السابقة تعتبر هي دالة التوزيع الاحتمالي البعدي خلال فترة الخـبرة القبلية وهي في نفس الوقت تعتبر دالة التوزيع الاحتمالي القبلـي لفتـرة الخـبرة الأساسية، وتكون هذه الدالة باعتبارها التوزيع القبلي لفترة الخبرة الأساسية علـى الشكل التالي:

$$\pi\left(\mu / \underline{y}_0\right)=\frac{1}{\sqrt{2\Pi}\,\frac{\sigma_0}{\sqrt{n_0}}}e^{-\frac{1}{2\sigma_0^2/n_0}\left(\mu-\frac{T_0}{n_0}\right)^2}$$

حيث إن:

$n_0 = 109 \qquad \sigma_0^2 = 1.243$

$$T_0 = \sum_{i=1}^{n_0} \ln y = 936.9$$

3– تحديد التوزيع الاحتمالي البعدي $\pi(\mu/\underline{y}_1)$ خلال فترة الخبرة الأساسية

تم استنتاج التوزيع الاحتمالي البعدي لفترة الخبرة الأساسية بدمج دالـــة الإمكـــان لفترة الخبرة الأساسية مع التوزيع الاحتمالي القبلي لفترة الخبرة الأساسية، ووجـــدنا أنه توزيع طبيعي، أى أن:

$$\pi\left(\mu / \underline{y}_1\right) = \frac{1}{\sqrt{2\pi}\sqrt{\dfrac{\sigma_0^2\sigma_1^2}{n_1\sigma_0^2 + n_0\sigma_1^2}}} e^{-\frac{1}{2\sigma_0^2\sigma_1^2/(n_1\sigma_0^2+n_0\sigma_1^2)}\left(\mu - \frac{n_1\sigma_0^2\frac{T_1}{n_1} + n_0\sigma_1^2\frac{T_0}{n_0}}{n_1\sigma_0^2+n_0\sigma_1^2}\right)^2}$$

ومعالم هذا التوزيع هي:

$$\left(\frac{\sigma_0^2 T_1 + \sigma_1^2 T_0}{n_1\sigma_0^2 + n_{01}\sigma_1^2}, \frac{\sigma_0^2\sigma_1^2}{n_1\sigma_0^2 + n_0\sigma_1^2}\right)$$

حيث أن:

$n_1 = 863 \qquad \sigma_1^2 = 1.258$

$$T_1 = \sum_{i=1}^{n_1} \ln y = 6792.76$$

4- تحديد التوزيع الاحتمالي التنبؤى $P(y_{n+1}/\underline{y}_1)$ خلال فترة الخبرة الأساسية:

$$P\left(y_{n+1}/\underline{y}_1\right)=\int_{-\infty}^{\infty} f\left(y_{n+1}/\mu\right)\pi\left(\mu/\underline{y}_1\right)d\mu$$

حيث أن:

$$f\left(y_{n+1}/\mu\right)=\frac{1}{\sqrt{2\Pi}\,y_{n+1}\sigma_1}e^{-\frac{1}{2\sigma_1^2}(\ln y_{n+1}-\mu)^2}$$

$$\pi\left(\mu/\underline{y}_1\right)=\frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma_0^2\sigma_1^2}{n_1\sigma_0^2+n_0\sigma_1^2}}}e^{-\frac{1}{2\sigma_0^2\sigma_1^2/(n_1\sigma_0^2+n_0\sigma_1^2)}\left(\mu-\frac{n_1\sigma_0^2\frac{T_1}{n_1}+n_0\sigma_1^2\frac{T_0}{n_0}}{n_1\sigma_0^2+n_0\sigma_1^2}\right)^2}$$

وبعد إجراء الاختصارات والاستنتاجات اللازمة، وجدنا أن التوزيع الإحتمالى التنبؤي لقيم المطالبات يتبع التوزيع اللوغاريتمى الطبيعي، وتكون دالة كثافته الاحتمالية على الشكل التالي:

$$P\left(y_{n+1}/\underline{y}_1\right)=\frac{1}{\sqrt{2\Pi}\,Y_{n+1}\frac{\sigma_1 d}{G}}e^{-\frac{1}{2\frac{\sigma_1^2d^2}{G^2}}(\ln y_{n+1}-C)^2}$$

ومعالم التوزيع هى: $\left(C, \frac{\sigma_1^2 d^2}{G^2}\right)$

حيث أن:

$$\mu_p = C = \frac{\sigma_0^2 T_1 + \sigma_1^2 T_0}{n_1 \sigma_0^2 + n_0 \sigma_1^2}$$

$$\sigma_p^2 = \frac{\sigma_2^2 d^2}{G^2} = \sigma_1^2 . \frac{\sigma_0^2 \sigma_1^2}{n_1 \sigma_0^2 + n_0 \sigma_1^2} \left(\frac{1}{\sigma_1^2} + \frac{n_1 \sigma_0^2 + n_0 \sigma_1^2}{\sigma_0^2 \sigma_1^2} \right)$$

وبالتالي فإن قيم معالم التوزيع اللوغاريتمي الطبيعي هي:

$\mu_p = 7.953194$ $\sigma_p^2 = 1.259292$

حيث إن σ_p^2 , μp هما المتوسط والتباين للتوزيع اللوغاريتمى الطبيعى. وبالتالي فإنه يمكن تحديد المتوسط والتباين للمتغير العشوائى (قيم المطالبات) كما يلى:

$$E(y_{n+1}) = e^{\left(\mu_p + \frac{1}{2}\sigma_p^2\right)} = 5339$$

$$Var(y_{n+1}) = e^{2\mu_p + \sigma_p^2}\left(e^{\sigma_p^2} - 1\right) = 7.1922484753 \times 10^7$$

وبالتالي فإن متوسط قيمة مطالبة تأمين المركبات الشامل العام القادم 5339 ريال والتباين 7.1922484753×10^7 ريال.

-العزوم الأربعة حول الصفر، للتوزيع الاحتمالي التنبؤى لقيم مطالبات:

تم تحديد العزوم الأربعة حول الصفر وفقاً لدالة توليد العزوم للتوزيـــع الاحتمـــالي التنبؤي (التوزيع اللوغاريتمي الطبيعي)، حيث إن دالة توليد العزوم حول الصـــفر هي:

$$M'_r(y) = e^{r\mu_p + \frac{1}{2}r^2\sigma_p^2}$$

وبالتالي فإن:

$$M'_1(y) = 5339 = E(Y_{n+1})$$

$$M'_2(y) = 1.0043 \times 10^8$$

$$M'_3(y) = 6.65504 \times 10^{12}$$

$$M'_4(y) = 1.55361 \times 10^{18}$$

-العزوم الأربعة حول المتوسط الحسابي للتوزيـــع الاحتمـــالي التنبـــؤي لقـــيم المطالبات:

$$M_1(y) = 0$$

$$M_2(y) = e^{2\mu} e^{\sigma^2}\left(e^{\sigma^2} - 1\right) = Var(y_{n+1})$$
$$= 7.192248 \times 10^7$$

$$M_3(y) = \left(e^{\sigma^2}\right)^{3/2}\left[e^{\sigma^2} - 1\right]\left[e^{\sigma^2} + 2\right]e^{3\mu}$$

$$= 2.12086 \times 10^{12}$$

$$= 1.4262169 \times 10^{18}$$

$$M_4(y) = \left(e^{\sigma^2}\right)^2 \left[e^{\sigma^2} - 1\right]^2 \left[\left(e^{\sigma^2}\right)^4 + 2\left(e^{\sigma^2}\right)^3 + 3\left(e^{\sigma^2}\right)^2 - 3\right] e^{4\mu}$$

ثالثاً: تحديد التوزيع الاحتمالي التنبؤي الإجمالي لقيم مطالبات تأمين المركبات الشامل:

إن أهم ما تشير إليه النظرية الإحصائية ما يلي: (الديب علي، 1996)

1 - أن الخصائص المشاهدة لشكل التوزيع الاحتمالي تتلخص في أربع قيم وهي الموقع Location والتشتت Dispersion والالتواء Skewness والتفرطح Kurtosis.

2 - أن الخصائص الأربعة لشكل التوزيع ترتبط بشكل كبير جداً بالعزوم الأربعة للمتغير العشوائي.

وقد تم التوصل في هذا المبحث إلى العزوم الأربعة الأولى حول المتوسط لعدد المطالبات، ولكن بالنسبة لوثيقة واحدة، ويمكن تحديد هذه العزوم لعدد 2588 وثيقة كما يلي:

$M_1(X) = M_1(x) * 2588 = 938.054$

$M_2(X) = M_2(x) * 2588 = 968.962$

$M_3(X) = M_3(x) * 2588 = 1033$

$M_4(X) = 2588[M_4(x) - 3M_2(x)] + 3(4856)^2 [M_2(x)]^2 = 2.816009 \times 10^6$

وتحسب العزوم المركزية الأربعة الأولى للتوزيـــع الاحتمـــالي التنبـــؤي الإجمـــالي $P(A_{n+1}/\underline{y}_1)$ لقيم مطالبات تأمين المركبات الشامل كما يلي:

$$M'_1(A) = M_1(x) * M_1(y) = 5.00827036\times 10^{6}$$

$$M'_2(A) = M_2(x) * [M_1(y)]^2 + M_1(x) * M_2(y) = 9.5087355\times 10^{10}$$

$$M'_3(A) = M_3(x) \,.\, [M_1(y)]^3 + M_1(x) \,.\, M_3(y) + 3M_1(y).\, M_2(y) \,.\, M_2(x)$$

$$= 3.26291831 \times 10^{15}$$

$$M'_4(A) = M_4(x).\, [M_1(y)]^4 + M_1(x) \,.\, M_4(y) + 4M_1(y).\, M_3(y).\, M_2(x) + 6\, [M_1(y)]^2 \,.\, M_2(y)\, [M_1(x).\, M_2(x) + M_3(x)] + 3\, [M_2(y)]^2\, [M_1(x)]^2 - M_1(x) + M_2(x)] = 2.851922697\times 10^{22}$$

حيث إن A متغير عشوائي يعبر عن إجمالي قيم مطالبات تأمين المركبات الشـــامل، ويتم استخدام هذه العزوم المركزية لحساب معامل الالتـــواء ومعامـــل التفـــرطح لاستخدامها في تحديد قيمة المعادلة التفاضلية لبيرسون، وذلك لتحديـــد التوزيـــع الاحتمالي الإجمالي التنبؤى لقيم المطالبات، كما يلي:

- معامل الالتواء B_1

$$B_1 = \frac{M_3(A)}{[M_2(A)]^{3/2}} = 0.11128126$$

- معامل التفرطح B_2

$$B_2 = \frac{M_4(A)}{[M_2(A)]^{2}} = 3.154222$$

-تحديد قيمة المعادلة التفاضلية لبيرسون (إبراهيم رأفت، 1996)

$$P = \frac{B_1(B_2+3)^2}{4(4B_2-3B_1)(2B_2-3B_1-6)} = -3.377$$

وحيث إن قيمة معادلة بيرسون أقل من الصفر، فإن منحنى التوزيع الاحتمالي الإجمالي التنبؤي لقيم مطالبات تأمين المركبات الشامل، يكون من النوع الأول ، وفقاً لمنحنيات بيرسون وتكون دالة الكثافة الاحتمالية pdf على الشكل التالي:

$$f(A) = K(A - a_1)^{m1}(a_2 - A)^{m2}$$

حيث إن:

A متغير عشوائي يمثل إجمالي قيم مطالبات تأمين المركبات الشامل.

$$a1 < A < a2$$

a1 تمثل الحد الأدنى لقيمة المتغير العشوائي = صفر

a2 تمثل أقصى قيمة لمطالبات تأمين المركبات.

m_1, m_2 معالم التوزيع الاحتمالي.

K مقدار ثابت

وباستخدام تحويلة معينة، توصلنا إلى شكل توزيع احتمالى معروف، على الشكل التالى:

$$f(A) = \frac{1}{\beta(m_1+1, m_2+1)} \left(\frac{A}{a_2}\right)^{m_1} \left(1 - \frac{A}{a_2}\right)^{m_2} * \frac{1}{a_2}$$

حيث أن: $0 < A_2 < a_2$

$(m_1+1), (m_2+1)$ معالم التوزيع

وقد قمنا بتقدير معالم التوزيع الإحتمالى السابق باستخدام برنامج MathCAD ، وتبين أن:

$$m1=130.174322 \qquad m2=129.741095$$

و دالة الاحتمالات التراكمية CDF هي:

$$G(A) = \int_0^A f(A).dA$$

- **تقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة لـتأمين المركبات الشامل**

إن الطريقة الأكثر دقة في تقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة The Maximum Probable yearly (MPY) هو توفيق توزيع احتمالي لإجمالي الخسائر السنوية لإيجاد قيمة L^* التي تجعل $f(L^*) = .99$. تمثل دالة توزيع تراكمي مناسب، ومن ثم تصبح (L^*) هي تقدير لقيمة MPY (الديب علي، 1996)وبالتالي فإن:

$$G(A) = \int_0^A f(A).dA = 0.99$$

وباستخدام برنامج MathCAD نجد أن الحد الأقصى لإجمالي الخسائر السنوية المحتملة لعدد 2588 وثيقة هو 5.726669×10^7 ريال

النتائج والتوصيات

أولاً: النتائج

1- عدد مطالبات تأمين المركبات الشامل تتفق مع توزيع ثنائي الحدين السالب، أى أن:

$$X \sim NB\,(r1,\theta);\; r1 = 10.83 \qquad \theta = 0.97$$

2- دالة التوزيع الاحتمالي التنبؤي لعدد المطالبات تكون على الشكل الآتي:

$$P(X_{n+1}/\underline{x}_1) = C_{X_{n+1}}^{X_{n+1}+11-1}(.9681)^{11}(1-.9681)^{X_{n+1}}$$

و هي دالة توزيع ثنائي الحدين السالب بمعالم (11,.9681)

3- دالة التوزيع الاحتمالي لقيم مطالبات تأمين المركبات الشامل، تتفق مع التوزيع اللوغاريتمي الطبيعي، بعالم:

$$\mu = 7.928 \qquad\qquad \sigma_1^{\,2} = 1.258$$

4- التوزيع الإحتمالى التنبؤي لقيم المطالبات يتبع التوزيع اللوغاريتمى الطبيعي، وتكون دالة كثافته الاحتمالية على الشكل التالي:

$$P\left(y_{n+1}/\underline{y}_1\right) = \frac{1}{\sqrt{2\Pi}\,Y_{n+1}\,\dfrac{\sigma_1 d}{G}}\, e^{-\frac{1}{2\frac{\sigma_1^{\,2} d^2}{G^2}}(\ln y_{n+1} - C)^2}$$

ومعالم التوزيع هي: $\left(C, \frac{\sigma_1^{\,2} d^2}{G^2}\right)$

حيث أن:

$$\mu_p = C = \frac{\sigma_0^{\,2} T_1 + \sigma_1^{\,2} T_0}{n_1 \sigma_0^{\,2} + n_0 \sigma_1^{\,2}}$$

$$\sigma_p^2 = \frac{\sigma_2^{\,2} d^2}{G^2} = \sigma_1^{\,2} . \frac{\sigma_0^{\,2} \sigma_1^{\,2}}{n_1 \sigma_0^{\,2} + n_0 \sigma_1^{\,2}} \left(\frac{1}{\sigma_1^{\,2}} + \frac{n_1 \sigma_0^{\,2} + n_0 \sigma_1^{\,2}}{\sigma_0^{\,2} \sigma_1^{\,2}} \right)$$

و معالم التوزيع اللوغاريتمي الطبيعي هي:

$\mu_p = 7.953194$ $\sigma_p^2 = 1.259292$

5- التوزيع الاحتمالي الإجمالي التنبؤي لقيم المطالبات يكون من النــــوع الأول ، وفقاً لمنحنيات بيرسون وتكون دالة الكثافة الاحتماليةpdf على الشكل الأتي:

$$f(A) = \frac{1}{\beta(m_1+1, m_2+1)} \left(\frac{A}{a_2}\right)^{m_1} \left(1 - \frac{A}{a_2}\right)^{m_2} * \frac{1}{a_2}$$

حيث أن: $0 < A_2 < a_2$

معالم التوزيع $(m_1+1), (m_2+1)$

وتبين أن:

$m1=130.174322 \qquad m2=129.741095$

6- دالة الاحتمالات التراكمية CDF

$$G(A)=\int_0^A f(A).dA$$

7- الحد الأقصى لإجمالي الخسائر السنوية المحتملة لعدد 2588 وثيقة هو 5.726669×10^7 ريال

ثانياً: التوصيات

تتمثل أهم توصيات البحث فيما يلي:

1- استخدام النموذج المقترح في فروع أخرى في التأمين

2- الاعتماد على الأساليب الكمية التي تجمع بين عدد وقيم المطالبات

3- استخدام أساليب رياضية أو إحصائية تتلائم مع الواقع العملي

المراجع

أولاً: المراجع العربية:

- جلال مصطفى الصياد، *الاستدلال الإحصائي*، دار المريخ، 1993
- رأفت أحمد علي إبراهيم، *نحو نظام علمي لإدارة أخطار الكوارث الطبيعية في جمهورية مصر العربية*، رسالة دكتوراه، كلية التجارة، جامعة القاهرة، 1996
- سمير كامل عاشور، سامية أبو الفتوح سالم، *مقدمة في التحليل الإحصائي*، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة، 2005
- علي السيد عبده الديب، "استخدام التوزيعات الاحتمالية (منحنيات بيرسون) في تقدير الحد الأقصى لإجمالي الخسائر السنوية المحتملة التي تتعرض لها شركة التأمين"، المجلة المصرية للدراسات التجارية، كلية التجارة – جامعة المنصورة، المجلد العشرون، العدد الثاني، 1996 .
- ممدوح حمزة أحمد، *استخدام التوزيعات الاحتمالية فى تسعير التأمين مع التطبيق على تأمين السطو/ محلات تجارية*، رسالة دكتوراه، كلية التجارة- جامعة القاهرة، 1990

ثانياً: المراجع الأجنبية:

- Andrew Gelman, et.al, *Bayesian Data Analysis*, Second edition Chapman & Hall/CRC, USA, 2004
- Athanasios Papoulis, *Probability Statistics*, Prentice-Hall, Inc., USA, 1990

- Dickson, David, "Predictive Aggregate claims distribution", Journal of Risk & Insurance, Dec., 98, Vol.65, Issue 4
- Hossack, I.B., *Introductory Statistics with Application in general insurance*, British library cataloguing in Publication, Cambridge University Pkss, 1985
- Ibrahim Mohamed Morgan, *Credibility theory under the collective Risk Model*, PhD dissertation, university of wisconson, U.S.A, 1983
- Lester D. Taylor, *Probability and Mathematical Statistics*, Harpecs & Row, Publishers, Inc., New York, 1997
- Peter Congdon, *Applied Bayesian Modelling*, John Wiley & Sons, Ltd., USA, 2003
- William M. Bolstad, *Introduction to Bayesian Statistics* , Second edition, John Wiley & Sons, Inc., Publication, USA, 2007

ملحق الدراسة

- اختبار جودة التوفيق لتوزيع عدد المطالبات مع توزيع ثنائي الحدين السالب:

نظراً لأن قيمة المعلمة r_1 بها كسر، فإنه يصعب استخدام دالة توزيع ثنائي الحدين السالب للحصول على الاحتمالات النظرية إلا عندما يكون عدد المطالبات = صفر، حيث إن:

أما بالنسبة لحساب باقي الاحتمالات النظرية $f(1), f(2), f(3), f(4)$ ، فإنه يتم استخدام القاعدة التالية (Hossack, 1985)

$$f(0;r_1,\theta) = C_0^{r_1+0-1}\theta^{r_1}(1-\theta)^0 = \theta^{r_1} = .719014886$$

$$f(X;r_1,\theta) = f(X-1)\times(1-\theta)\times\left(\frac{r_1+x-1}{x}\right)$$

$$f(1;r_1,\theta) = f(0)\times(1-.97)\times\left(\frac{10.83+1-1}{1}\right)$$

$$f(1) = .232939285$$

$$f(2) = .041156592$$

$$f(3) = .0052575$$

$$f(4) = .00054296$$

وبالتالي فإنه يمكن تحديد التكرارات النظرية كما يلي:

$F(0) = f(0) \times 2588 = 1863$

$F(1) = f(1) \times 2588 = 603$

$F(2) = f(2) \times 2588 = 106$

$F(3) = f(3) \times 2588 = 14$

$F(4) = f(4) \times 2588 = 2$

والجدول التالي يوضح الاحتمالات التراكمية الفعلية والاحتمالات التراكمية النظرية المقابلة لكل قيمة من قيم عدد مطالبات تأمين المركبات الشامل.

جدول رقم (2)

الاحتمالات التراكمية الفعلية والاحتمالات التراكمية النظرية

المقابلة لعدد مطالبات تأمين المركبات الشامل

عدد المطالبات	التكرارات الفعلية	التكرارات النظرية	الاحتمالات التراكمية الفعلية	الاحتمالات التراكمية النظرية	الفرق المطلق
صفر	1859	1863	0.718315	0.719861	0.000155
1	615	603	0.95595	0.9528594	0.003091
2	97	106	0.99343	0.9938176	0.000388
3	14	14	0.998841	0.999227	0.000386
4	3	2	1	1	صفر
مجموع	2588	2588			

يلاحظ أن أكبر فرق مطلق = د $_{\text{المحسوبة}}$ = 0.003091.

ويتم الحصول على د $_{\text{الجدولية}}$ عند مستوى معنوية 5% وحجم عينة 2588 كما يلي:

د $_{\text{الجدولية}}$ = $\frac{1.36}{\sqrt{n}}$ = 0.026734

وحيث إن د المحسوبة أقل من د الجدولية ، فإننا نقبل الفرض العدمي إن عدد مطالبات تأمين المركبات الشامل تتبع توزيع ثنائي الحدين السالب بدرجة ثقة 95%، أي أن:

$$X \sim NB\,(r1,\theta);\; r1= 10.83 \qquad \theta= 0.97$$

– تحديد دالة الإمكان خلال الفترة سنة 2009

الجدول التالي يوضح توزيع عدد وثائق تأمين المركبات الشامل حسب عدد المطالبات خلال الفترة القبلية سنة 2009.

جدول رقم (3)

توزيع عدد وثائق تأمين المركبات الشامل

حسب عدد المطالبات خلال الفترة القبلية سنة 2009

عدد المطالبات	عدد الوثائق
صفر	431
1	81
2	11
3	2
مجموع	525

$$E(\underline{X}_0)=.2076190$$

$$\operatorname{var}(\underline{X}_0)=.22971283$$

وحيث إن المتغير العشوائي (عدد المطالبات) يتبع توزيع ثنائي الحدين السالب خلال فترة الخبرة الأساسية، فإن المتغير العشوائي (عدد المطالبات) يتبع توزيع ثنائي الحدين السالب خلال فترة الخبرة القبلي، وبالتالي فإن:

متوسط عدد المطالبات الفعلية خلال هذه الفترة = متوسط توزيع ثنائي الحدين السالب

تباين عدد المطالبات الفعلي خلال فترة الخبرة القبلية = تباين توزيع ثنائي الحدين السالب

$$\frac{r_1(1-\theta)}{\theta}=.20761904$$

$$\frac{r_1(1-\theta)}{\theta^2} = .22971283$$

وبحل المعادلتين السابقتين نجد أن:

$$r_0 = 1.95 \qquad \theta = .903819956$$

ونجد أن دالة الإمكان تكون على الشكل التالي:

$$f(X_0/\theta) = \prod_{i=1}^{n_0} C_{X_i}^{X_i+r_0-1} \theta^{r_0} (1-\theta)^{X_i}$$

$$f(\underline{X}_0/\theta) \propto \theta^{n_0 r_0} (1-\theta)^{\sum_{i=1}^{n_0} x_i}$$

حيث أن n_0 تمثل حجم عينة فترة الخبرة القبلية (سنة 2009)،أي أن:

$$n_0 = 525$$

حيث إن $T_0 = \sum_{i=1}^{n_0} x_i$ تسمى بالإحصاء الكافي والمقدار $C_{x_i}^{x_i+r_0-1}$ لا يعتمد في حسابه على θ فيختزل في ثابت التناسب وقيمة T_0 تعبر عن إجمالي عدد المطالبات خلال فترة الخبرة القبلية(سنة 2009) أي أن:

$$T_0 = 109$$

وبالتالي فإن دالة الإمكان هي:

$$f(\underline{X}_0/\theta) \propto \theta^{n_0 r_0} (1-\theta)^{T_0}$$

وهي شكل دالة بيتا بمعالم

$$\beta(n_0 r_0 + 1, T_0 + 1)$$

وحيث إن التوزيع الاحتمالي القبلي خلال فترة الخبرة القبلية (سنة 2009) هو عبارة عن التوزيع الاحتمالي البعدي خلال فترة الخبرة ما قبل سنة 2009 وهو $\pi(\theta)$ حيث إن:

$$\pi(\theta) = 1$$

ويتم الحصول على التوزيع الاحتمالي البعدي خلال فترة الخبرة القبلية (سنة 2009)، ويرمز له بالرمز $\pi(\theta/X_1)$ ، أي تحديد التوزيع الاحتمالي للمعلمة θ بعد الحصول على العينة من فترة الخبرة القبلية (سنة 2009) كما يلي:

$$\pi(\theta/\underline{X}_0) \propto f(\underline{X}_0/\theta)\pi(\theta)$$

وحيث أن $\pi(\theta) = 1$ فإن:

$$\pi(\theta/\underline{X}_0) \propto f(\underline{X}_0/\theta)$$

$$\pi(\theta/\underline{X}_0) \propto \theta^{n_0 r_0}(1-\theta)^{y_0}$$

أي أن التوزيع الاحتمالي البعدي خلال فترة الخبرة القبلية يتناسب مع دالة الإمكان، وبالتالي فإن التوزيع الاحتمالي البعدي خلال فترة الخبرة القبلية، هو عبارة عن دالة الإمكان خلال فترة الخبرة القبلية فقط. حيث أن التوزيع الاحتمالي القبلي خلال فترة الخبرة القبلية $\pi(\theta)$ لم يؤثر حيث أنه مستنتج من فترة خبرة ما قبل سنة 2009 ولا يتوافر أي معلومات عن هذه الفترة. ويكون الاعتماد في حساب أي تقديرات على بيانات الخبرة القبلية (سنة 2009).

- **التوزيع الاحتمالي التنبؤي The Predictive Distribution $P(X_{n+1}/\underline{x}_1)$**

تم استنتاج هذا التوزيع بإجراء التكامل المحدود لحاصل ضرب دالة التوزيع الاحتمالي الأصلي للبيانات للسنة القادمة و دالة التوزيع الاحتمالي البعدي خلال فترة الخبرة الأساسية (William M. 2007 .(Bolstad,

$$P(X_{n+1}/\underline{x}_1) = \int_0^1 f(X_{n+1}/\theta)\pi(\theta/\underline{x}_1).d\theta$$

والتكامل يكون بالنسبة لـ θ ومن صفر إلى 1 لأن قيمة θ تتراوح من صفر إلى 1.

ويتم الحصول على التوزيع الاحتمالي لعدد المطالبات للسنة القادمة بالتعويض في دالة التوزيع الاحتمالي خـلال فترة الخبرة الأساسية عن X بـ X_{n+1} وبالتالي فإن:

$$f\left(x_{n+1} / \theta\right) = C_{x_{n+1}}^{x_{n+1}+r_1-1} \theta^{r_1} (1-\theta)^{x_{n+1}} \quad --(1)$$

$$\pi\left(\theta / \underline{X}_1\right) = \frac{1}{\beta\left(n_0 r_0 + n_1 r_1 + 1, T_0 + T_1 + 1\right)} \theta^{n_0 r_0 + n_1 r_1} (1-\theta)^{T_0+T_1} \quad -(2)$$

بضرب معادلة (1) × (2) مع إجراء التكامل المحدود

$$\mathrm{P}\left(x_{n+1} / \underline{X}_1\right) = \frac{C_{x_{n+1}}^{x_{n+1}+r_1-1}}{\beta\left(n_0 r_0 + n_1 r_1 + 1, T_0 + T_1 + 1\right)} \int_0^1 \theta^{r_1 + n_0 r_0 + n_1 r_1} (1-\theta)^{x_{n+1}+T_0+T_1} . d\theta$$

و بعد إجراء الاختصارات والاستنتاجات اللازمة نجد أن دالة التوزيع الاحتمالي التنبؤي لعدد المطالبات تكون على الشكل التالي:

$$P\left(X_{n+1} / \underline{x}_1\right) = C_{X_{n+1}}^{X_{n+1}+11-1} (.9681)^{11} (1-.9681)^{X_{n+1}}$$

والدالة السابقة هي دالة توزيع ثنائي الحدين السالب أي أن دالة التوزيع الاحتمالي التنبؤى لعدد مطالبات تأمين المركبات الشامل هي دالة توزيع ثنائي الحدين السالب بمعالم (11,.9681) أي أن:

$$E\left(X_{n+1}\right) = \frac{11(1-.9681)}{.9681} = .36 \qquad Var\left(X_{n+1}\right) = \frac{11(1-.9681)}{(.9681)^2} = .37$$

- **العزوم الأربعة للتوزيع الاحتمالي التنبؤي لعدد المطالبات حول الصفر:**

حيث أن التوزيع التنبؤي يتبع توزيع ثنائي الحدين السالب فإن دالة توليد العزوم حول الصفر تأخذ الشـــكل التالي:

$$M^{/}(t) = \theta^{r_1}\left[1. - (1-\theta)e^{t}\right]^{-r_1}$$

حيث إن $M^{/}(t)$ الدالة المولدة للعزوم حول الصفر

$\theta = .9681$ $\qquad$ $r_1 = 11$

ويمكن الحصول على العزوم الأربعة حول الصفر بالاشتقاق المتتالي بالنسبة لـــ (t) ثم التعويض بعد ذلك عـــن t=0 بعد كل عملية اشتقاق. وباستخدام برنامج Mathcad يتم الحصول على المشتقة الأولى، ثم التعويض عن t=0 فنحصل على العزم الثاني وهكذا. وبالتالي فإن العزوم الأربعة حول الصفر تتولد من تفاضل الدالـــة التالية:

$$M_X^{/}(t) = \frac{d^x}{dt^x} M(t)$$

حيث أن: $X = 1, 2, 3, 4$ $\qquad$ $t = 0$

العزوم الأربعة حول الصفر هي:

$$M_1^{/}(X) = E(X_{n+1}) = .362463 \qquad M_2^{/}(X) = .505785$$

$$M_3^{/}(X) = .853825 \qquad M_4^{/}(X) = 1.762411$$

حيث أن العزم الأول حول الصفر هو متوسط التوزيع الاحتمالي التنبؤي.

العزوم الأربعة للتوزيع الاحتمالي التنبؤي حول المتوسط (Lester D. Taylor, 1974)

إذا كانت X متغيراً عشوائياً منفصلاً، فإن دالة توليد العزوم حول المتوسط يكون باستخدام الدالة التالية:

حيث أن:

$$E\left[(X-M)^m\right]=\Sigma(X-M)^m F(X)$$

- العزم الأول حول المتوسط = صفر أي أن: $M_1(X) = 0$

- العزم الثاني حول المتوسط:

$$E\,[(X-M)^2] = E\,(X^2 - 2Mx + M^2)$$
$$M_2 = E\,(X^2) - 2M\,E(x) + M^2$$
$$= E\,(X^2) - 2M^2 + M^2$$
$$= E\,(X^2) - M^2$$

حيث إن: $E\,(X^2)$ تمثل العزم الثاني حول الصفر

$$E\,(X^2) = M_2(X) = .505785$$
$$M = M_1(X) = .362463$$
$$M^2 = .13137$$

وبالتالي فإن العزم الثاني حول المتوسط M_2

$$M_2 = .374415$$
$$M_3 = E\,[(X-M)^3] = M_3 - 3M_2\,M_1 + 2M_1^2$$
$$= .39908$$
$$M_4 = E\,[(X-M)^4] = M_4 - 4M_3\,M_1 + 6M_2\,m_1^2 - 3M_1^4$$
$$= .871408$$

وبالتالي فإن العزوم الأربعة للتوزيع الاحتمالي التنبؤي حول المتوسط هي:

$$M_1(X)=0 \qquad M_3(X)=.39908$$

$$M_2(X)=Var(X_{n+1})=.374415 \qquad M_4(X)=.871408$$