إشتقاق دالة البقاء بإستخدام توزيع القوى الأسى "دراسة تطبيقية"

Deriving the Survival Function Using the Exponential Power Distribution

(Applied study)

د٠محمد توفيق البلقينى
أستاذ الإحصاء الإكتوارى والتأمين
كلية التجارة – جامعة المنصورة

د٠هبه محمود السجاعى
مدرس الإحصاء التطبيقى
كلية التجارة – جامعة المنصورة

الباحثة

هبه سلامه جمعه محمد الشواف
كلية التجارة – جامعة المنصورة

ملخص :

تقوم دالة البقاء بدور هام وفعال فى العديد من العلوم لما لها من أهمية كبرى فى دراسة زمن إحتمال بقاء الكائن الحى بعد مدة محددة من الزمن(t)، بالإضافة إلى دورها الأساسى فى تحليل أغلب الظواهر إعتماداَعلى البيانات والمعلومات المتاحة عن الظاهرة، لذلك لجأ العديد من الباحثين للبحث فى هذا المجال حيث قاموا بدراسة وتحليل دالة البقاء، ومن هنا تم البحث عن طرق جيدة وجديدة لإستخدامها فى دراسة وتحليل دالة البقاء، ولذلك تضمن البحث إشتقاق دالة البقاء بإستخدام توزيع القوى الأسى، كما تم تقدير معلمات التوزيع بطريقتين هما:الإمكان الأعظم ولأول مرة تم إستخدام العزوم الخطية المبتورة ، ولتحقيق هدف البحث تم إستخدام مجموعتين مختلفتين من البيانات : المجموعة الأولى عبارة عن بيانات لعدد من وفيات كوفيد-١٩فى مصر والمجموعة الثانية عبارة عن بيانات لبعض حالات الوفاة المبكرة للأشخاص المؤمن عليهم لدى شركة مصر لتأمينات الحياة .حيث أثبتت الدراسة مدى توافق ومرونة ودقة توزيع القوى الأسى مع البيانات المستخدمة، كما تم التوصل إلى أن طريقة العزوم الخطية المبتورة أفضل فى التقدير من طريقة الإمكان الأعظم وذلك من خلال المفاضلة بينهم بإستخدام : p-value –Anderson Darling test- Cramér –Von Mises test (A* - W*-P-value).

الكلمات الإفتتاحية (دالة البقاء – دالة وطأة الوفاة – توزيع القوى الأسى – العزوم الخطية المبتورة – دالة الإمكان الأعظم).

Abstract:

The Survival function has an important and effective role in many sciences because of its great importance in studying time and the probability of survival of a living organism after a specified period of time (t) in addition to its primary role in analyzing most phenomena based on the data and information available about the phenomenon. Therefore, many researchers have resorted to research in this area, where they study and analyze the survival function. The present research included the deriving of the survival function using the exponential power distribution, two methods are used to estimate the parameters of the distribution and parameters estimation is demonstrsded by two approachs maximum lilklihood and trimmed linear moments which is used for the first time in estimating the parameters. To achieve this objective, we used two different groups of data sets: the first group is data of covaid-19 deaths in Egypt and the second group is data of some premature death cases of people insured with Misr life insurance company. This study demonstrated the compatibility, flexbility and occuracy of the exponential power distribution with the data used. It has also been concluded that the trimmed linear moments is better in estimating than the maximum liklihood method through using: P-value-Anderson Darling test- Cramér –Von Mises test (p-value - A* - W*).

The Key Words :(Survival function – Hazard rate function – Exponential power distribution – Trimmed Linear Moments – Maximum Likliehood Estimation).

(١) مقدمة:

تعد دالة البقاء من الدوال الهامة التى تستخدم بواسطة كثير الباحثين، مثل الإحصاء فهى تعرف فى الإحصاء الطبى والوبائيات بتحليل البقاء وتحليل الأخطار (Hazard Analysis)، وفى الهندسة بتحليل وقت الفشل (Failure Time Analysis) وتحليل المصداقية (Reliability Analysis)، وفى علم النفس تعرف بتحليل تاريخ الحدث (Event History Analysis) وتحليل المدة (Duration Analysis)، وكذلك تعرف فى الإقتصاد بتحليل الإنتقال (Trasition Analysis).

لذلك بدأت الدراسات الجدية لها منذ خمسنيات القرن الماضى، ولقد قام العديد من الباحثين بإعداد دراسات عن دوال البقاء وأوقات الفشل ومن هذه الدراسات دراسة (Sinha and Sloon 1988) توصلت إلى أن مقدرات بيز للمعلمات ودالة البقاء تمتلك تبايناالاحقا أصغر من التباين المحاذى، ودراسة (Lye and Ryan 1993) توصلت إلى أن طريقة الإمكان الأعظم لدالة البقاء أفضل من طريقة بيزللتقديروتم قياس كفاءة المقدرات بإستخدام جذر متوسط مربع الخطأ، ودراسة (الوزير2004) توصلت إلى أنه لايمكن الإعتماد كلية على التحليل ثنائى المتغيرات فى تقييم معنوية تأثير عوامل الخطر على وقت بقاء الزارع على قيد الحياة لأن ذلك

يقودنا إلى إستنتاجات خاطئة وغيرها العديد من الدراسات التى إستخدمت توزيعات إحصائية مختلفة ولدالة البقاء دور أساسى فى تحليل معظم الظواهر إعتمادا على البيانات والمعلومات الإحصائية المتوفرة عن تلك الظاهرة.

وفى هذا البحث سوف يتم إستخدام توزيع القوى الأسى كتوزيع معلمى يتم من خلاله إشتقاق دالة البقاء وبالتالى يتم الحصول على دالة وطأة الوفاة .

(٢) الأساليب الإحصائية المستخدمة:

(٢-١) توزيع القوى الأسى Exponential Power Distribution :

يُعرف توزيع القوى الأسى (EPD) بتوزيع الخطأ المعمم(Generalized Error Distribution(GED) وهو حالة خاصة من التوزيع الطبيعى المعمم (Generalized Normal Distribution) حيث يتكون التوزيع الطبيعى المعمم من عائلتين من التوزيعات الإحتمالية ، ونجد ان كلتا العائلتين تضيف معلمة الشكل(shape parameter) إلى التوزيع الطبيعى. ويشار إلى التوزيع الطبيعى المعمم بإصدارين حيث يعد توزيع القوى الاسى الإصدار الأول من التوزيع الطبيعى المعمم وتوزيع القوى الأسى واحد من أشهر التوزيعات المستخدمة فى تحليل بيانات البقاء والموثوقية ويستخدم غالبا فى إنشاء نماذج سرعة الفشل Accelerated Failures Times(AFT) والتى تصف اعتماد توزيع البقاء على المتغيرات التفسيرية ولقد ناقش العديد من الباحثين الإستدلال الإحصائى للمعالم فى توزيع القوى الأسى باستخدام عينات مستقلة ولها نفس التوزيع.

قدم توزيع القوى الأسى (EPD) لأول مرة كنموذج لجدول الحياة بواسطة Smith & Bain(1975) ولقد تم مناقشة هذا التوزيع (EPD) بواسطة العديد من الباحثين على سبيل المثال Leemis (1986)& Rajarshi (1988) Hanagal & Dabade (2015).

ومع ذلك نجد أن توزيع القوى الأسى لم يستخدم بواسطة تحليل البقاء فقط ولكنه مرتبط ايضا بتوزيع القوى الأسى المتماثل فى الإحصاء كما ذكره Hazan etal (2003) and Delicado & Goria (2008) .

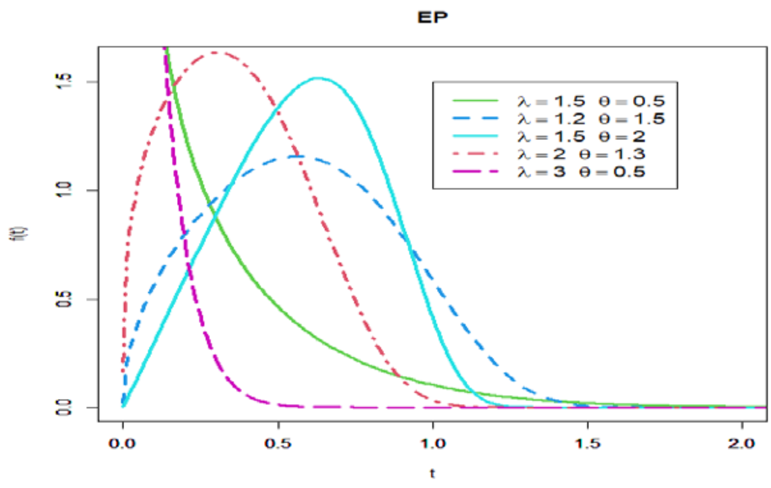
(٢-١-١) الدوال الأساسية لتوزيع القوى الأسى:

١- دالة كثافة الإحتمال لتوزيع القوى الأسى Probabilty density functiong

$$f(t) = \lambda\theta t^{\theta-1} \exp(\lambda t^{\theta}) \exp(1-\exp(\lambda t^{\theta})) \quad , t > 0 \qquad (1)$$

حيث t متغير عشوائى متصل موجب بمعلمتين (λ,θ) حيث ($\theta > 0$) وترمز إلى معلمة الشكل (shape parameter) حيث ($\lambda > 0$) وترمز إلى معلمة القياس (scale parameter).

شكل (١)



المصدر :مخرجات برنامج R

يوضح الشكل (١) دالة كثافة الإحتمال f (t) لتوزيع EP مع المتغير العشوائى (t) عند قيم مختلفة للمعلمات (λ،θ) حيث توجدعدة أشكال لتوزيع EP نتيجة إستخدام قيم مختلفة للمعلمات كما هو موضح بالشكل حيث نجد أنه عند λ⟵1.5، θ⟵ 2أصبح الشكل قريب من التوزيع الطبيعى .

٢- دالة التوزيع التجميعية لتوزيع القوى الأسى :

(The Cumulative function for exponential power distribution)

$$\mathbf{F(t) = 1 - \exp(1 - \exp(\lambda t^{\theta}))} \qquad \textbf{(2)}$$

(٢-٢) دالة البقاء على قيد الحياة The survival function:

دالة البقاء والتى يرمز بالرمز S(t) تعرف على أنها نموذج إحتمالى لمتغير عشوائى t غير سالب يمثل الوقت ولذلك فإن نماذج البقاء تهتم بالزمن الذى يسبق حدوث حدث معين مع متغير مستقل أو أكثر ولايوضع فى الإعتبار طبيعة هذه المتغيرات من حيث كونها كمية أو وصفيه أو مختلطة .

(٢-٢-١) الدوال الأساسية لدالة البقاء Basic Survival Function

١- دالة البقاء على قيد الحياة

تعرف على أنها إحتمال البقاء إلى بعد t على الأقل حتى زمن t

S(t)=P(T>t) **t>0** **(3)**

T≥0 → S (0) =1

٢- دالة التوزيع التجميعية (Cumulative Distribution Function)

دالة البقاء التجميعية هى دالة مكملة لدالة التوزيع التجميعية أى أنها دالة متمه للدالة التجميعية ، حيث أن صغر دالة البقاء يعنى كبر الدالة التجميعية وكبر دالة البقاء يعنى صغر الدالة التجميعية .

$$S(t)=1-F(t) \qquad (4)$$

٣ - دالة كثافة الإحتمال (P.d.f) (Probability Density Function)

يرمز لها بالرمز f (t) وتعرف على أنها تفاضل ل F (t)

$$f(t)=\frac{d}{dt}F(t)=\frac{-d}{dt}s(t) \quad , \qquad t\geq 0$$

وبالتالى فإن

$$F(t)=\int_0^t f(y)dy$$

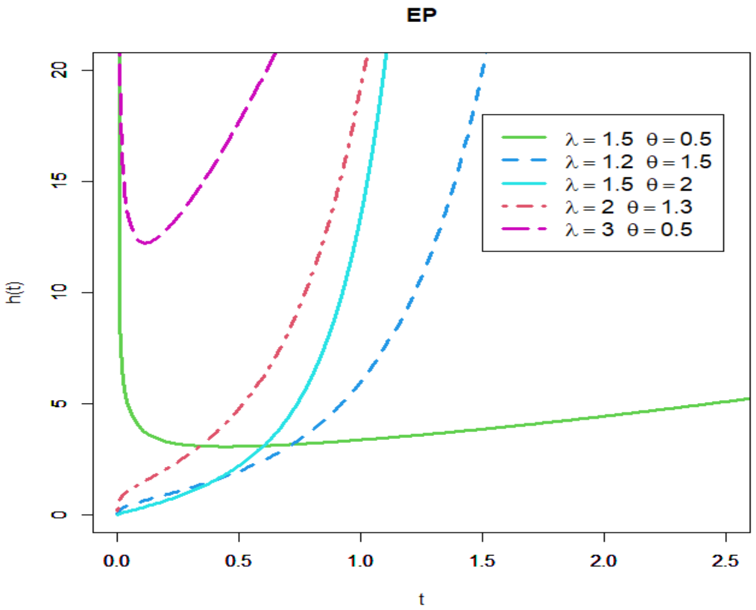
و

$$S(t)=\int_t^\infty f(y)dy$$

ومن ثم فإن S (t),F (T) إحتمالات مرتبطة بالفترات بمعنى أن دالة التوزيع التجميعية F(t) تعبر عن إحتمال تحقق الفشل قبل الوقت t، وتعبر دالة البقاء على قيد الحياة S(t) عن إحتمال تحقق الفشل بعد الوقت t وهو نفسه إحتمال البقاء على قيد الحياة عند الوقت t، بينما تعبر دالة كثافة الإحتمال f (t) عن الكثافة غير الشرطية للفشل عند الوقت t وهى مقياس لحظى وهذا يعنى أن f (t) تكون موجودة لوقت الفشل فى زمن t=0

$$S(t)=1-(1-e^{(1-e^{(\lambda t^{\theta})})}) \qquad t>0 \qquad (5)$$

شكل (٢)



المصدر: مخرجات برنامج R

يوضح شكل (٢) العلاقة بين دالة البقاء لتوزيع القوى الأسى والمتغير العشوائى (t) عند قيم مختلفة للمعلمات (θ،λ) حيث أنه عند λ⟵3، θ⟵0.5 يشبه التوزيع الأسى.

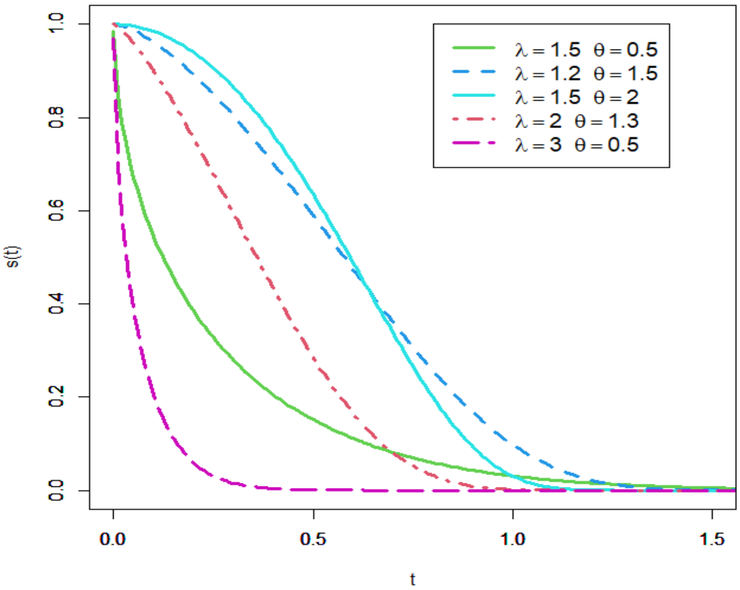
٤ـ دالة وطأة الوفاة أو دالة الوفاة اللحظى أو معدل الفشل (Hazard Rate Function) (HRF)

يرمز لها بالرمز h (t) وتعبر عن كثافة الفشل الشرطية بشرط البقاء على قيد الحياة حتى الوقت t

$$h(t) = \frac{f(t)}{s(t)} \qquad (6)$$

$$h(t)= \lambda\theta t^{\theta-1} \exp(\lambda t^{\theta}) \qquad t > 0 \qquad (7)$$

شكل (٣)



المصدر : مخرجات برنامج R

يوضح شكل (٣) العلاقة بين دالة معدل الفشل لتوزيع القوى الاسى والمتغير العشوائى t عند قيم مختلفة للمعلمات (λ،θ) .

(١-٣) تقدير معلمات التوزيع :

سوف نرمز للمتغير العشوائى t الذى يتبع توزيع القوى الأسى بالرمز (EPD) بمعلمتين هما(θ, λ) كما هو مشار إليه فى معادلة (1) وسوف يتم تقدير معلمات التوزيع (θ, λ) وذلك بإستخدام : ١- دالة الإمكان الأعظم (MLE) ٢- العزوم الخطية المبتورة (TLM).

(١-٣-١) دالة الإمكان الأعظم

بفرض أن $(T_1, T_2, \dots\dots\dots, T_n)$ متغيرات عشوائية مستقلة ومتشابهة فى التوزيع (iid) وتتبع توزيع (EP) بمعلمات مجهولة (θ,λ) ومن خلال دالة كثافة الإحتمال لتوزيع (EP) المشار إليها فى المعادلة (1) يتم التعويض عنها فى المعادلة التالية:

$$L(\theta, \lambda) = \pi_{i=1}^{n} f(t_{i,} \theta, \lambda)$$

$$L(\theta, \lambda) = \pi_{i=1}^{n} \left(\lambda \theta t^{\theta-1} \ . e^{\lambda t^{\theta}} \ . e^{1-e^{\lambda t^{\theta}}}\right)$$

$$= n\log \lambda + n \log\theta + (\theta - 1) \sum_{i=1}^{n} \log ti + \lambda \sum_{i=1}^{n} t_i^{\theta} + \sum_{i=1}^{n} (1 - e^{\lambda t_i^{\theta}})$$

ويتم التفاضل بالنسبة ل θ ومساواة الناتج بالصفر للحصول على $\widehat{\theta}$

$$\frac{\partial \log L(\theta,\lambda)}{\partial\theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\log t_i+\lambda\sum_{i=1}^{n}t_i^{\theta}\log\sum_{i=1}^{n}t_i$$
$$-\sum_{i=1}^{n}e^{\lambda t_i^{\theta}}-\sum_{i=1}^{n}e^{\lambda t_i^{\theta}}\quad .\lambda t_i^{\theta}\ \log t_i$$

ويتم التفاضل بالنسبة ل (λ) ومساواة الناتج بالصفر للحصول على $\hat{\lambda}$

$$\frac{\partial \log L(\theta,\lambda)}{\partial\lambda}=\frac{n}{\lambda}+\sum_{i=1}^{n}t_i^{\theta}-\sum_{i=1}^{n}e^{\lambda t_i^{\theta}}\quad .\quad t_i^{\theta}$$

ولا نستطيع الوصول إلى شكل نهائى للمعلمة θ ، λ لذلك نلجأ إلى الطرق الرقمية لإيجاد مقدرات لتلك المعلمات .

(١-٣-٢) العزوم الخطية المبتورة (TLM)

فى هذا الجزء سوف يتم إشتقاق العزوم الخطية المبتورة لأول مرة لتوزيع القوى الأسى فى حالة التماثل وعدم التماثل وإستخدامها فى الحصول على حالات خاصة مثل العزوم الخطية المبتورة العليا والعزوم الخطية المبتورة الدنيا والعزوم الخطية المبتورة من الدرجة الأولى والعزوم الخطية، وبإستخدام الصيغة العامة للعزوم الخطية المبتورة للمجتمع فى حالة عدم التماثل والمعادلة (١) أمكن الحصول على الصيغة العامة للعزوم الخطية المبتورة للمجتمع من الدرجة r فى حالة عدم التماثل $i_1 \neq i_2$ لتوزيع القوى الأسى كالتالى : -

$$\lambda_r^{(i_1,i_2)}=\frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}\frac{(r+i_1+i_2)!}{(r+i_1-k-1)!\,(k+i_2)!}\,(I)\qquad(8)$$

$$I=\int_0^{\infty} t\lambda\theta t^{\theta-1}.\ e^{\lambda t^{\theta}}\ .e^{1-e^{\lambda t^{\theta}}}.\left(1-e^{1-e^{\lambda t^{\theta}}}\right)^{r+i_1-k-1}.\,e^{(i_2+k)(1-e^{\lambda t^{\theta}})}dt$$

وبفرض أن

$$A = \left(1 - e^{1-e^{\lambda t^{\theta}}}\right)^{r+i_1-k-1}$$

وبإستخدام مفكوك ذو الحدين نحصل على

$$A = \sum_{s=0}^{\infty} \binom{r+i_1-k-1}{s} (-1)^s e^{s\left(1-e^{\lambda t^{\theta}}\right)}$$

$$C_1 = \sum_{s=0}^{\infty} \binom{r+i_1-k-1}{s} (-1)^s$$

$$I = C_1 \int_0^{\infty} \lambda \theta t^{\theta}.\ e^{\lambda t^{\theta}}\ .e^{(i_2+k+s+1)\left(1-e^{\lambda t^{\theta}}\right)} dt$$

وبفرض ان

$$B = e^{(i_2+k+s+1)\left(1-e^{\lambda t^{\theta}}\right)}$$

وباستخدام مفكوك exponential نحصل على

$$B = \sum_{m=0}^{\infty} \frac{(i_2+k+s+1)^m}{m!} \left[1 - e^{\lambda t^{\theta}}\right]^m$$

حيث

$$= \sum_{m=0}^{\infty} \frac{(i_2+k+s+1)^m}{m!} c_2$$

وباستخدام مفكوك ذو الحدين للمقدار

$$\left[1 - e^{\lambda t^{\theta}}\right]^m = \sum_{j=0}^{\infty} \binom{m}{j} . (-1)^j . e^{\lambda t^{\theta}}$$

حيث

$$= \sum_{j=0}^{\infty} \binom{m}{j} . (-1)^j c_3$$

$$I = C_1 . C_2 . C_2 . \lambda \theta \int_0^{\infty} t^{\theta}\, e^{-(-j-1)\lambda t^{\theta}} dt$$

وبفرض $z = (-j-1)\lambda t^{\theta}$

$$t = \left[\frac{z}{(-j-1)\lambda}\right]^{\frac{1}{\theta}}$$

$$|j| = \frac{dt}{dz} = \left[\frac{1}{\theta(-j-1)^{\frac{1}{\theta}}\lambda^{\frac{1}{\theta}}}\right].Z^{\left(\frac{1-\theta}{\theta}\right)}$$

$$I = C_1.C_2.C_3.\lambda\theta\int_0^{\infty}\frac{Z}{(-j-1)\lambda}.e^{-z}.\left[\frac{1}{\theta(-j-1)^{\frac{1}{\theta}}\lambda^{\frac{1}{\theta}}}\right].Z^{\frac{1}{\theta}}\,dz$$

$$I = C_1.C_2.C_3.\lambda\theta.\left[\frac{1}{\theta(-j-1)^{\frac{1}{\theta}}\lambda^{\frac{1}{\theta}}}\right]$$

$$.\frac{1}{(-j-1)\lambda}\int_0^{\infty} e^{-z}.Z^{\frac{1}{\theta}}\,dz$$

وبإستخدام تكامل جاما

$$I = C_1.C_2.C_3.\lambda\theta.\left[\frac{1}{\theta(-j-1)^{\frac{1}{\theta}}\lambda^{\frac{1}{\theta}}}\right].\frac{1}{(-j-1)\lambda}.\Gamma\left(\frac{1}{\theta}+1\right)$$

وبالتالى فان الصورة العامة للعزوم الخطية المبتورة $\lambda_r^{(i_1,i_2)}$ تاخذ الشكل التالى :

$$\lambda_r^{(i_1,i_2)} = \frac{1}{r}\sum_{k=0}^{\infty}\sum_{s=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\binom{r-1}{k}.\binom{r+i_1-k-1}{s}.\binom{m}{j}$$

$$.(-1)^{k+s+j}.\frac{(r+i_1+i_2)!\,(i_2+k+s+1)^m}{m!\,(r+i_1-k-1)!\,(i_2+k)!}.\frac{\Gamma\left(\frac{1}{\theta}+1\right)}{(-j-1)^{1+\frac{1}{\theta}}.\lambda^{\frac{1}{\theta}}} \qquad (9)$$

ويمكن الحصول على العزوم الخطية المبتورة الدنيا (TLL-moments)بالتعويض $i_{1=0}$ وكذلك يمكن الحصول على العزوم الخطية المبتورة العليا (TLH-moments) بالتعويض عن $i_2 = 0$ فى المعادلة (9).

العزوم الخطية المبتورة من الدرجة r فى حالة التماثل :

بالتعويض عن $(i_1 = i_2 = i)$ فى المعادلة (9) يمكننا الحصول على العزوم الخطية المبتورة من الدرجة r فى حالة التماثل وتأخذ الشكل التالى :

$$\lambda_r^{(i)} = \frac{1}{r}\sum_{k=0}^{\infty}\sum_{s=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\binom{r-1}{k}.\binom{r+i-k-1}{s}.\binom{m}{j}$$

$$.(-1)^{k+s+j}.\frac{(r+2i)!\,(i+k+s+1)^m}{m!\,(r+i-k-1)!\,(i+k)!}.\frac{\Gamma(\frac{1}{\theta}+1)}{(-j-1)^{1+\frac{1}{\theta}}.\lambda^{\frac{1}{\theta}}} \qquad (10)$$

(٤-١) التطبيق العملى لتوزيع (EP)

فى هذا الجزء سوف يتم إثبات مرونة توزيع (EP) بإستخدام مجموعتين من البيانات الحقيقية, لكل مجموعة من البيانات تم فى البداية تقدير معلمات توزيع (EP) عدديا بإستخدام كلا من طريقة الإمكان الأعظم وطريقة العزوم الخطية المبتورة بإستخدام برنامج (R) ثم تم إستخدام تلك التقديرات فى التأكد من جودة توفيق التوزيع للبيانات بإستخدام إختبار كولموجروف –سميرنوف وحساب قيمة (P-Value) وتم المفاضلة بين طريقتى تقدير المعلمات بإستخدام بعض مقاييس جودة التوفيق (W*، A*).

تتضمن المجموعة الأولى بيانات 52 يوم لعدد الوفيات الناتجة عن المصابين بفيروس كورونا (CoVID-19) فى الفترة من 1/7/ 2020 الى 2020/8/21 داخل جمهورية مصر العربية حيث نهدف فى هذه الدراسة إلى تحديد التوزيع الأمثل لتوفيقه مع بيانات المجموعة الأولى ويعد هذا ذو أهمية كبرى إلى كل من يهتم بدراسة كورونا داخل جمهورية مصر العربية لحدث حالات الوفاة .

تعرض المجموعة الثانية بيانات 100مشاهدة لبعض حالات الوفاة المبكرة للأشخاص المؤمن عليهم لدى شركة مصر لتامينات الحياة حيث تعرف الوفاة المبكرة بأنها الوفاة التى تحدث للمؤمن عليهم خلال الثلاث سنوات الأولى لعقد وثيقة التامين طويلة الأجل .

ولذلك تم إختيار وثائق التأمين على الحياة طويلة الأجل التى تشمل مدتها من 10 إلى 30 سنة .والتى تمت فى الفترة من 2009/1/1 إلى 2019/1/1 مع إختيار أعمار المؤمن عليهم بحيث لاتتعدى أعمارهم 40 سنة عند بدء التأمين وذلك وفقا لشروط وهى خضوعهم للكشف الطبى . والهدف من هذه الدراسة هو تحديد طريقة التقدير المثلى من بين طريقتى MLE-TLM لتوزيع القوى الأسى ويعد هذا ذو أهمية كبرى لدى شركة التأمين لما يترتب عليه من سهولة ودقة التنبؤ بمتوسط أعمار الوفاة المبكرة التى ستحدث فى المستقبل وبالتالى سوف يتم أخذ الاحتياطات اللازمة لتوفير مبالغ التأمين وغيرها.

تحليل بيانات المجموعة الأولى:

يعرض جدول (١) تقديرات الإمكان الأعظم والعزوم الخطية المبتورة لتوزيع (EP) وبعض مقاييس جودة التوفيق المستخدمة لإختبار جودة توفيق التوزيع للبيانات والمفاضلة بين طريقتى تقدير المعلمات بإستخدام بعض مقاييس جودة التوفيق (W*،A*) وإحتمالات الوفيات عند قيم مختلفة ذلك وفقا للمجموعة الأولى من البيانات.

جدول (١)

Covid19		MLE					TIM				
		lambda	theta	KS (P-Value)	W*	A*	lambda	theta	KS (P-Value)	W*	A*
Estimate		0.00202	1.46937	0.10619 (0.6007)	0.1399	0.9837	0.005533	1.216	0.10561 (0.6078)	0.1293	0.9261
T=15	Survival	0.892161					0.8516039				
	Hazard rate	0.01179266					0.01401				
T=20	Survival	0.8358742					0.79029				
	Hazard rate	0.01428707					0.0159				
T=25	Survival	0.7732131					0.7265				
	Hazard rate	0.01691293					0.01778				
T=75	Survival	0.1155401					0.15405				
	Hazard rate	0.07115296					0.04907				

المصدر: مخرجات برنامج R

يوضح جدول (١) تقديرات الإمكان الأعظم والعزوم الخطية المبتورة لمعلمات توزيع (EP) ومقاييس جودة توفيق التوزيع للمجموعة الأولى من بيانات COVID -19 وإحتمالات الوفيات عند قيم مختلفة.

(١-٤-١) تقدير الإمكان الأعظم والعزوم الخطية المبتورة لمعلمات توزيع (EP):

من خلال جدول (١) تم تقدير معلمات توزيع (EP) عدديا بإستخدام طريقة الإمكان الأعظم والعزوم الخطية المبتورة بإستخدام برنامج (R) وذلك إعتمادا على المجموعة الأولى من البيانات حيث تم الحصول على قيم المعلمات (θ،λ) للامكان الأعظم والعزوم الخطية المبتورةوهى على الترتيب (1.469, 0.00202) ، (1.216 -0.005533).

(١-٤-٢) إختبار جودة توفيق توزيع(EP):

تتعدد الإختبارات التى تستخدم فى إختبار جودة توفيق التوزيعات الاحتمالية بما يتناسب مع البيانات ومن بين هذه الإختبارات إختبار كولمجروف –سميرنوف وإختبار كاى x^2 وغيرها من الإختبارات ،ولكن فى هذه الدراسة سوف يتم إستخدام إختبار كولمجروف –سميرنوف وإختبار(A^*) وإختبار(W^*)لإختبار جودة توفيق توزيع (EP) للبيانات محل الدراسة والمفاضلة بين طرق التقدير (MLE –TLM).

إختبار جودة توفيق المجموعة الاولى من البيانات (COVID -19):

اولا :بإستخدام تقديرات الإمكان الأعظم

تم إختبار جودة توفيق توزيع (EP)مع البيانات عن طريق حساب إحصاء كولموجروف –سميرنوف (K.S$_{CAL}$) وحساب قيمة (P-Value) بإستخدام تقديرات الإمكان الأعظم (١-٤-١) بإستخدام برنامج (R)كما هو واضح فى جدول (١) ويتم حساب إختبار(K.S) كالتالى :

- إختبار كولمجروف – سميرنوف لتوزيع (EP)عند $(\alpha = 0.05)$

١ - الفرض العدمى : توزيع البيانات لايختلف جوهريا عن توزيع (EP)

٢ - الفرض البديل : توزيع البيانات يختلف جوهريا عن توزيع (EP)

٣- إحصاءة الإختبار

$$K.s_{cal} = \max\left\{\max\left[\frac{i}{n} - F(x)\right], \max\left[F(x) - \frac{i-1}{n}\right]\right\} = 0.10619$$

٤– حساب P-Value

$$\text{p-value}=p(ks_{\alpha} \geq ks_{cal}) = 0.6007$$

٥ - القرار : بما ان (p-value=0.6007) وهى قيمة أكبرمن α إذا القرار قبول الفرض العدمى، أى أن توزيع البيانات لايختلف جوهريا عن توزيع (EP) وذلك بدرجة 95%.

وتم حساب إحتمالات الوفيات عند قيم مختلفة وذلك بإستخدام دالة Survivalودالة hazard

عند t=15

S(t)= P(t≥15) =0.892

وهذا يدل أن معدل البقاءعند النقطة 15=0.892أو بمعنى أخر إحتمال أن عدد الوفيات يساوى 15 هو 0.892

h(t)=P(t=15/t≤15)=0.011

وهذا يعنى أن معدل الفشل عند (t=15)هو 0.011بشرط معلومية عدد الوفيات أقل من أويساوى 15
وعند t=20

$S(t)=P(t\geq 20)=0.8358742$

وهذا يدل أن معدل البقاء عند النقطة (t=20)هو 0.835أو إحتمال أن عدد الوفيات يساوى 20 هو0.835

$h(t)=P(t=20/t\leq 20)=0.01428707$

وهذا يعنى أن معدل الفشل عند(t=20)هو0.0143 بشرط معلومية عدد الوفيات أقل من أويساوى 20
وعند t=25

$S(t)=P(t\geq 25)=0.7732131$

وهذا يدل أن معدل البقاء عند النقطة (t=25)هو 0.773 أو إحتمال أن عدد الوفيات يساوى 25 هو0.773

$h(t)=P(t=25/t\leq 25)=0.01691293$

وهذا يعنى أن معدل الفشل عند(t=25)هو0.0169 بشرط معلومية عدد الوفيات أقل من أويساوى 25

وعند t=75

$S(t)=P(t\geq 75) = 0.1155401$

وهذا يدل أن معدل البقاء عند النقطة (t=75) هو 0.115 أو إحتمال أن عدد الوفيات يساوى75 هو0.1155

$h(t)=P(t=75/t\leq 75)=0.07115$

وهذا يعنى أن معدل الفشل عند(t=75)هو0.07115 بشرط معلومية عدد الوفيات أقل من اويساوى 75

ثانيا : بإستخدام العزوم الخطية المبتورة

تم إختبار جودة توفيق توزيع (EP)مع البيانات عن طريق حساب إحصاء كولموجروف –سميرنوف ($K.S_{CAL}$) وحساب قيمة (P-Value) بإستخدام تقديرات العزوم الخطية المبتورة(١-٤-٢) بإستخدام برنامج (R)كما هو موضح فى جدول (١) ويتم حساب إختبار(K.S) كالاتى

- إختبار كولمجروف –سميرنوف لتوزيع (EP)عند ($\alpha = 0.05$)

١ - الفرض العدمى : توزيع البيانات لايختلف جوهريا عن توزيع (EP)

٢ - الفرض البديل : توزيع البيانات يختلف جوهريا عن توزيع (EP)

٣ - إحصاءة الاختبار

$$K.s_{cal} = \max\left\{\max\left[\frac{i}{n} - F(x)\right], \max\left[F(x) - \frac{i-1}{n}\right]\right\} = 0.10561$$

٤ – حساب P-Value

p-value=p($ks_{\alpha} \geq ks_{cal}) = 0.6078$

٥ - القرار : قبول الفرض العدمى، أى أن توزيع البيانات لايختلف جوهريا عن توزيع (EP) وذلك بدرجة 95%.

وتم حساب إحتمالات الوفيات عند قيم مختلفة وذلك بإستخدام دالة Survivalودالة hazard

عند t=15

S(t)= P(t≥15) =0.852

وهذا يدل أن معدل البقاءعند النقطة 15=0.852 أو إحتمال أن عدد الوفيات يساوى 15 هو 0.852

h(t)=P(t=15/t≤15)=0.014

وهذا يعنى أن معدل الفشل عند (t=15)هو 0.014 بشرط معلومية عدد الوفيات أقل من أويساوى 15

وعند t=20

S(t)=P(t≥20)=0.790

وهذا يدل أن معدل البقاء عند النقطة (t=20)هو 0.790 أو إحتمال أن عدد الوفيات يساوى 20 هو0.790

h(t)=P(t=20/t≤20)=0.0159

وهذا يعنى أن معدل الفشل عند(t=20)هو0.0159 بشرط معلومية عدد الوفيات أقل من أويساوى 20

وعند t=25

S(t)=P(t≥25)=0.7265

وهذا يدل أن معدل البقاء عند النقطة (T=25)هو 0.7265أو إحتمال أن عدد الوفيات يساوى 25 هو0.7265

h(t)=P(t=25/t≤25)=0.01778

وهذا يعنى أن معدل الفشل عند (t=25) هو0.01778 بشرط معلومية عدد الوفيات أقل من أويساوى 25

وعند t=75

$S(t)=P(t\geq75)= 0.15405$

وهذا يدل أن معدل البقاء عند النقطة (t=75)هو 0.154 أو إحتمال أن عدد الوفيات يساوى75 هو0.15405

$h(t)=P(t=75/t\leq75)=0.04907$

وهذا يعنى أن معدل الفشل عند(t=75) هو0.04907 بشرط معلومية عدد الوفيات أقل من أويساوى 75

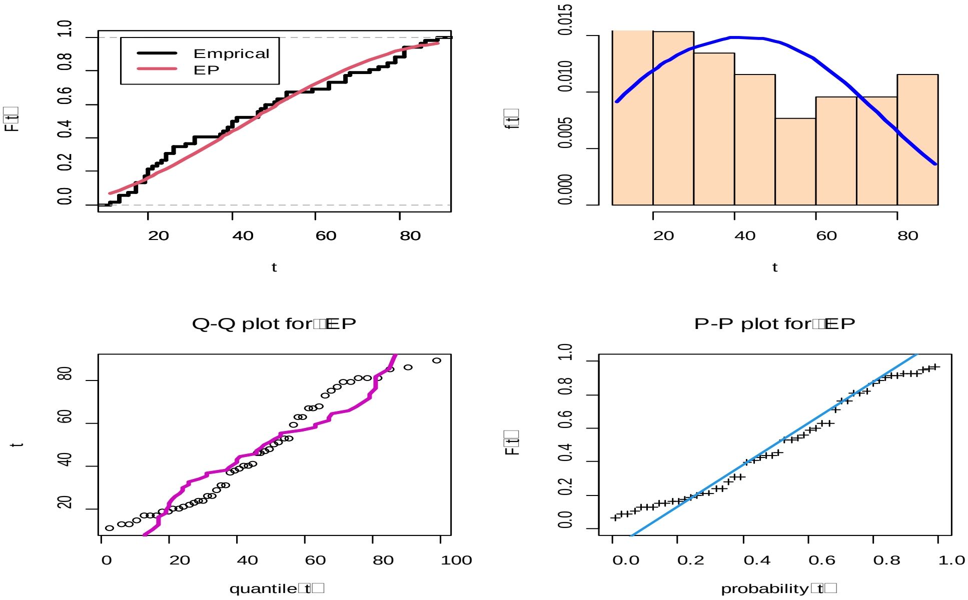
يتضح من جدول(١)

- أن تقدير المعلمات بطريقة (TLM)أفضل من تقدير المعلمات بطريقة (MLE) وذلك وفقا لإختبار(K.S) و(P- Value)
- عند المفاضلة بين طريقة (MLE)وطريقة (TLM) بإستخدام إختبارات (A* - W*) يتضح أن طريقة (TLM) أفضل من طريقة (MLE) فى حساب survival و Hazard rate حيث تتم المفاضلة على أساس النسبة الأقل لكل من (A* - W*) وتبين أن نسبتهم اقل فى طريقةTLM

ويتضح ذلك أيضا من خلال الأشكال (٢،١)

شكل (١)

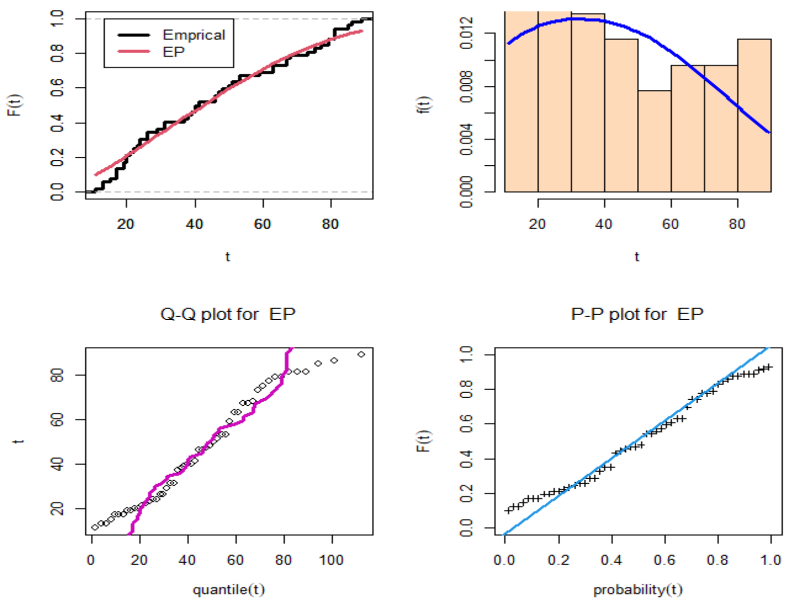
MLE



يوضح شكل (١) (emprical - histogram -P-Pplot-Q-Qplot) مدى ملائمة المجموعة الأولى من بيانات (COVID-19) مع توزيع (EP) بإستخدام (MLE)

شكل (٢)

TLM



يوضح شكل (٢) (emprical –histogram –P-Pplot-Q-Qplot) مدى ملائمة المجموعة الأولى من بيانات (COVID-19) مع توزيع (EP) بإستخدام (TLM)

وبمقارنة الأشكال(١) و(٢) توضح أفضلية العزوم الخطية المبتورة عن الإمكان الأعظم

تحليل بيانات المجموعة الثانية :

يعرض جدول (٢) تقديرات الإمكان الأعظم والعزوم الخطية المبتورة لتوزيع (EP) وبعض مقاييس جودة التوفيق المستخدمة لإختبار جودة توفيق التوزيع للبيانات والمفاضلة بين طريقتى تقدير المعلمات بإستخدام بعض مقاييس جودة التوفيق (W*، A*) وإحتمالات الوفيات عند قيم مختلفة للعمر.

جدول (٢)

Insurance		MLE					TLM				
		lambda	theta	KS (P-Value)	W*	A*	lambda	theta	KS (P-Value)	W*	A*
Estimate		0.36251	1.4713	0.074799 (0.6306)	0.0673	0.5346	0.37925	1.314	0.0665 (0.7674)	0.0646	0.5335
T=0.5	Survival	0.8696403					0.84807				
	Hazard rate	0.4384635					0.4669				
T=0.75	Survival	0.7649351					0.74322				
	Hazard rate	0.5905425					0.5904				
T=1.5	Survival	0.3939764					0.40328				
	Hazard rate	1.247115					1.07998				
T=2.5	Survival	0.04792547					0.07884				
	Hazard rate	3.317075					2.35241				

المصدر : مخرجات برنامج R

ويتضح من جدول (٢) مدى قدرة (EP) فى توفيق بيانات المجموعة الثانية ويتضح ذلك من خلال مقاييس جودة التوفيق المستخدمة (K.S- P-Value) كما يوضح افضلية طريقة (TLM) عن طريقة (MLE)وذلك بإستخدام بعض مقاييس جودة التوفيق (W*،A*)

(١-٤-٣) تقدير الإمكان الأعظم والعزوم الخطية المبتورة لمعلمات توزيع (EP)

من خلال جدول (٢) تم تقدير معلمات توزيع (EP)عدديا بإستخدام ١- طريقة الإمكان الأعظم

٢- طريقة العزوم الخطية المبتورة بإستخدام برنامج (R) وذلك إعتمادا على المجموعة الثانية من البيانات المستخدمة حيث تم الحصول على المعلمات (λ،θ) وهى على الترتيب (0.3625154- 1.4713007)، (0.37925 - 1.314) .

(١-٤-٤) إختبارات جودة توفيق توزيع(EP) للمجموعة الثانية من البيانات (التأمين)

كما سبق أن ذكرنا فى (١-٤-٢) تتعدد الأنواع التى تستخدم فى إختبارات جودة التوفيق بمايتناسب مع البيانات وفيما يلى إختبارات جودة التوفيق بإستخدام تقديرات (TLM-MLE).

اولا : إختبار جودة التوفيق باستخدام تقديرات الإمكان الأعظم

تم إختبار جودة توفيق بيانات المجموعة الثانية مع توزيع (EP) وذلك بإستخدام إختبار كولموجروف – سمير نوف و(P-Value) وفيما يلى خطوات الإختبار :

إختبار جودة توفيق المجموعة الثانية من البيانات :

تم إختبار توفيق توزيع (EP) عن طريق حساب إحصاء كولموحروف –سميرنوف ($k.s_{cal}$) وحساب قيمة (P-Value) بإستخدام تقديرات الإمكان الأعظم (١-٤-٣) بإستخدام برنامج (R) كما هو موضح بجدول(٢) كالتالى :

- إختبار كولموجروف –سميرنوف لتوزيع (EP) عند ($\alpha = 0.05$)

١- الفرض العدمى : توزيع البيانات لايختلف جوهريا عن توزيع (EP)

٢- الفرض البديل : توزيع البيانات يختلف جوهريا عن توزيع (EP)

٣- إحصاءة الاختبار

$$K.s_{cal} = \max\left\{\max\left[\frac{i}{n} - F(x)\right], max\left[F(x) - \frac{i-1}{n}\right]\right\} = 0.074799$$

٤ – حساب P-Value

p-value=p($ks_{\alpha} \geq ks_{cal}) = 0.6306$

٥ -القرار : قبول الفرض العدمى، أى أن توزيع البيانات لايختلف جوهريا عن توزيع (EP) وذلك بدرجة 95%

يتم حساب دالة survival ,hazard لبيانات التأمين كالأتى

عندt=0.5

S(t)=P(t≥0.5)=0.869

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى 0.5 سنة هو 0.869

h(t)=P(t=0.5/t≤0.5)=0.438

وهذا يعنى أن معدل الفشل عند t=0.5هو 0.438بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من أويساوى 0.5

عندt=0.75

S(t)=P(t≥0.75)=0.7649

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى 0.75 سنة هو 0.7649

h(t)=P(t=0.75/t≤0.75)=0.5905

وهذا يعنى أن معدل الفشل عند t=0.75هو 0.5905بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من أويساوى 0.75

عند$t=1.5$

$S(t)=P(t\geq 1.5)=0.3939$

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى 1.5 سنة هو 0.3939

$h(t)=P(t=1.5/t\leq 1.5)=1.247$

وهذا يعنى أن معدل الفشل عند $t=1.5$هو 1.247بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من اويساوى 1.5.

وعند $t=2.5$

$S(t)=P(t\geq 2.5)=0.0479$

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى2.5 سنة هو 0.0479

$h(t)=P(t=2.5/t\leq 2.5)=3.317$

وهذا يعنى أن معدل الفشل عند $t=2.5$ هو 3.317بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من أويساوى 2.5 .

ثانيا : إختبار جودة التوفيق بإستخدام تقديرات العزوم الخطية المبتورة

تم إختبار جودة توفيق بيانات المجموعة الثانية مع توزيع (EP) وذلك بإستخدام إختبار كولموجروف – سمير نوف و(P-Value)وفيما يلى خطوات الإختبار :

*إختبار جودة توفيق المجموعة الثانية من البيانات :

تم إختبار توفيق توزيع (EP) عن طريق حساب إحصاء كولموحروف –سميرنوف ($k.s_{cal}$) وحساب قيمة (P-Value) بإستخدام تقديرات العزوم الخطية المبتورة (١-٤-٣) بإستخدام برنامج (R) كما هو موضح بجدول(٢) كالتالى :

-إختبار كولموجروف –سميرنوف لتوزيع (EP) عند ($\alpha = 0.05$)

١- الفرض العدمى : توزيع البيانات لايختلف جوهريا عن توزيع (EP)

٢- الفرض البديل : توزيع البيانات يختلف جوهريا عن توزيع (EP)

٣- إحصاءة الاختبار

$$K.s_{cal} = \max\left\{\max\left[\frac{i}{n} - F(x)\right], max\left[F(x) - \frac{i-1}{n}\right]\right\} = 0.0665$$

٤ – حساب P-Value

$$\text{p-value}=p(ks_{\alpha} \geq ks_{cal}) = 0.7674$$

٥- القرار : قبول الفرض العدمى، أى أن توزيع البيانات لايختلف جوهريا عن توزيع (EP) وذلك بدرجة 95%

يتم حساب دالة survival,hazard لبيانات التأمين كالأتى

عند$t=0.5$

$$S(t)=P(t\geq 0.5)=0.848$$

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى 0.5 سنة 0.848

$$h(t)=P(t=0.5/t\leq 0.5)=0.4669$$

وهذا يعنى أن معدل الفشل عند $t=0.5$هو 0.4669بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من أويساوى 0.5

عند$t=0.75$

$$S(t)=P(t\geq 0.75)=0.74322$$

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى 0.75 سنة هو 0.74322

$$h(t)=P(t=0.75/t\leq 0.75)=0.5904$$

وهذا يعنى أن معدل الفشل عند $t=0.75$هو 0.5904بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من اويساوى 0.7

عند$t=1.5$

$$S(t)=P(t\geq 1.5)=0.4032$$

وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى 1.5 سنة هو 0.4032

$$h(t)=P(t=1.5/t\leq 1.5)=1.0799$$

وهذا يعنى أن معدل الفشل عند $t=1.5$هو 1.0799بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من أويساوى 1.5

عند $t=2.5$

$$S(t)=P(t\geq 2.5)=0.0788$$

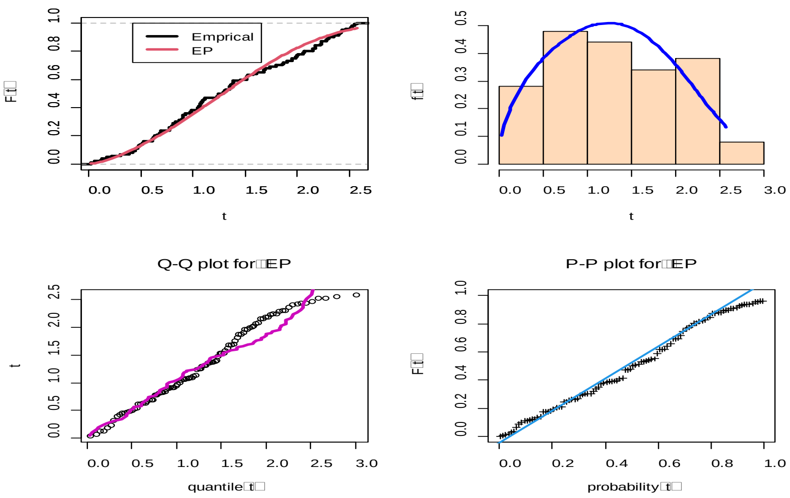
وهذا يعنى إحتمال بقاء الشخص على قيد الحياة حتى2.5 سنة هو 0.0788

$$h(t)=P(t=2.5/t\leq 2.5)=2.352$$

وهذا يعنى أن معدل الفشل عند $t=2.5$هو 2.352بشرط معلومية الفترة التى يعيشها المؤمن عليه أقل من أويساوى 2.5

يتضح من جدول(٢)
- أن تقدير المعلمات بطريقة (TLM) أفضل من تقدير المعلمات بطريقة (MLE) وذلك وفقا لإختبار(K.S) و(P- Value)
- عند المفاضلة بين طريقة (MLE) وطريقة (TLM) بإستخدام إختبارات (A* - W*) يتضح أن طريقة (TLM) أفضل من طريقة (MLE) فى حساب survival وHazard rate
ويتضح ذلك ايضا من خلال الأشكال (٤,٣) .

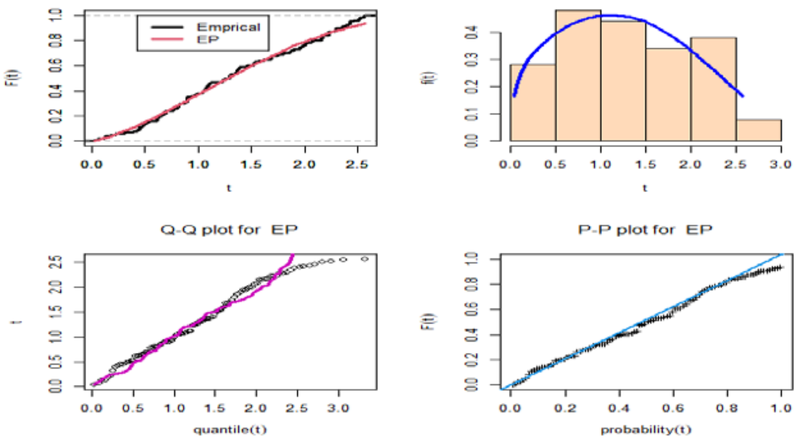
شكل (٣)
MLE



يوضح شكل (٣) (emprical –histogram –P-Pplot-Q-Qplot) مدى ملائمة المجموعة الثانية من بيانات (Insurance) مع توزيع (EP) بإستخدام (MLE)

شكل(٤)

TLM



يوضح شكل (٤) (emprical –histogram –p-pplot-Q-Qplot) مدى ملائمة المجموعة الثانية من بيانات (Insurance) مع توزيع (EP) بإستخدام (TLM) وبمقارنة الأشكال (٣) و(٤) يظهر أن طريقة العزوم الخطية المبتورة أفضل من طريقة الإمكان الأعظم حيث يظهرمدى تطابق خط البيانات الفعلية مع خط بيانات التوزيع فى طريقة TLMعن MLE.

٣- النتائج والتوصيات

(٣-١) النتائج

1- إثبات قدرة توزيع القوى الأسى (EPD) فى توفيق مجموعتين مختلفتين من البيانات الأولى عبارة عن بيانات لعدد من وفيات (كوفد-١٩) وذلك بنسبة 60%من خلال P-Value والثانية عبارة عن بيانات لبعض حالات الوفاة المبكرة للأشخاص المؤمن عليهم لدى شركة مصر لتأمينات الحياة وذلك بنسبة 60%من خلال P-Value .

٢- تم إثبات كفاءة توزيع القوى الأسى فى توفيق مجموعتين مختلفتين من البيانات وذلك بإستخدام

(K.S - P-Value).

٣- تمت المفاضلة بين TLM-MLEبإستخدام مقاييس جودة التوفيق (W^* وA^*)وتوصلنا إلى:

طريقة (TLM) أفضل من طريقة (MLE) فى تقديرات معلمات توزيع (EP) وحساب Surviva ، Hazard لمجموعتين مختلفتين من البيانات.

(٣-٢) التوصيات

يوصى الباحث دراسة توزيعات أخرى خلاف توزيع القوى الأسى مثل عائلة Lindy للتوزيعات وإدماج توزيع القوى الأسى بها.

(٤) المراجع

(٤-١) المرجع العربية

١- البلقينى، محمد توفيق (2015)" أسس النظرية الاحصائية " ، كلية التجارة – جامعة المنصورة .

٢- العلوى، أميرة عبدالغنى ،(2015)،"تقدير وتحليل مؤشرات البقاء بإستخدام بعض نماذج سلاسل ماركوف المختلفة "، رسالة ماجيستير، كلية التجارة – جامعة المنصورة .

٣ -المغاورى، سالى أبو العنين، (2014)، "إستخدام العزوم الخطية المبتورة لتقدير معلمات بعض التوزيعات الإحتمالية "، رسالة ماجيستير، كلية التجارة – جامعة المنصورة .

٤- الوزير، رزق السيد، (2004)، "تحليل البقاء على قيد الحياة قبل وبعد زراعة الكلى بإستخدام النماذج الديناميكية"، رسالة دكتوراة ، كلية التجارة – جامعة المنصورة .

(٤-٢) المراجع الأجنبية

1- Anderson, T. W. and Darling, D.A. (1954)" A Test of Goodness-of –Fit "Journal of the American Statistical Association ,49(268), 765-769.

2- Bayazit, M.and ÖnÖz, B. (2002)," LL-Moments for Estimating Low Flow Quantiles ", Hydrological Sciences, 47(5), PP 707-720.

3- Darling, D.A. (1957)," The Kolomogorov-Smirnov, Cramer-Von Mises Test". Ann. Math . Statist, 28(4), 823-838.

4- David, H.A. and Nagaraja, H.N. (2003)," Order Statistics", 3rd edition, New York, Wiley.

5- Deliado, P. and Goria, M.N, (2008), "Asmall sample comparison of maximum liklihood, moments and L-moments methods for the asymmetric exponential power distribution". Computational Statistics and Data Analysis, 52.

6- Elamir , E.A.H. and Seheult, A.H, (2003)," Trimmed L-Moments", Computional Statistics and Data Analysis, 43, PP299-314.

7- Hosking, J.R.M. (1990), ”L-Moments:analysis and estimation of distribution using linear combinations of order statistics “, Journal of the Royal Statistical Society, Serues B52: PP 105-124.

8- Hosking, J.R.M. (2007), ”Some Theory and Practical Uses of Trimmed L-Moments”, Journal of Statistical Planning and Inference, 137, PP3024-3039.

9- Hazan , A., Lansman , Z. and Makov , U.E. (2003) . ”Robustness Via a mixture of exponential power distributions”. Computational Statistics and Data Analysis, 42(1-2): 111-121.

10-Leemis, L.M. (1986). ”Lifetime distribution identities”. IEEE Transactions on Reliability, 35(2).

11- Maillet , B. and M`edecin , J. (2009) ”Extreme Volatilities and L-Moment Estimations of Tail Indexes”, Electronic copy of this paper is available at: http:/ssrn.com/abstract=1288661. 12-Smith, R.M. and Bain , L.J. (1975). ”An exponential power life-testing distribution “.Communications in Statistics-Theory and Methods, 58:469-481.

13-Tianchen, z. (2017). ”Maximum Liklihood Estimation of parameters in Exponential Power distribution”. Science in Statictics, master, Miami-Florida.

14- Wang , Q.J. (1997), ”LH-Moments for Statistical analysis of extreme events”, Water Resour Res, 33(12),2841-2849.