تقدير دالة البقاء باقتراح فئات عمرية أحادية السنّ لتوزيع كاما العام باستعمال المحاكاة

أ. م. د. عمر عبدالمحسن علي   الباحثة: رغدة زياد طارق
كلية الادارة والاقتصاد / جامعة بغداد
raghda_ziad2001@yahoo.com   dromarqaisy@yahoo.com
تاريخ استلام البحث: 2016/1/31   تاريخ قبول النشر: 2016/4/14

المستخلص
يُعدّ موضوع تحليل دالتّي البقاء على قيد الحياة والمعولية من الموضوعات والطرائق الاحصائية المهمة في وقتنا الحاضر وذلك بسبب اهميتها في مختلف المجالات السكانية والطبية والصناعية والهندسية. ركز هذا البحث على توليد بيانات عشوائية للحصول على عينات من التوزيع الاحتمالي كاما العام Generalized Gamma: GG بأسلوب يعرف بطريقة "التحويل العكسي" Inverse Transformation Method: ITM ، اذ يشمل هذا التوزيع بدوره تكامل دالة كاما الناقص ضمنياً مما يزيد من صعوبة التقدير التقليدي لذا ستكون الحاجة الى اللجوء الى الاسلوب العددي للتقريب ومن ثم التقدير لدالة البقاء. وتم تقدير دالة البقاء عن طريق المحاكاة بأسلوب "مونت كارلو"، واستعملت طريقة الانتروبي لأغراض التقدير والموائمة fitting ، هذا الى جانب استعمال الطريقة التقليدية. وتم التعرّف على الطريقة الافضل بالتقدير من خلال استعمال معياريّ المقارنة: جذر متوسط مربعات الخطأ، و متوسط مطلق الخطأ النسبي . وتم اختيار حجوم عينات (n=18, 30, 50, 81) اذ تمثل توليد البيانات بحجم n=18 فئات العمر خماسية السنّ للظاهرة المبحوثة وحجم العينة n=81 تمثل فئات العمر أحادية السنّ، وبتكرار التجربة (500) مرة وأظهرت نتائج المحاكاة ان طريقة الامكان الاعظم هي الافضل في حالة العينات الصغيرة والمتوسطة. وكذلك تمّ تطبيقها على بيانات لفئات العمر خماسية السنّ التي تعاني من اضطراب وتشويش للمسح الاجتماعي والاقتصادي للأسرة في العراق 2012 (Iraq Household Socio-Economic Survey: IHSES II 2012) بينما تفوقت طريقة الانتروبي في حالة العينات الكبيرة حيث تم تطبيقها على بيانات فئات العمر أحادية السنّ الناتجة من استعمال طريقة رياضية تؤدي الى النتائج بالاستناد الى معادلة التدريج (Formula for Interpolation) وضعها العالم سبراغ (Sprague) وهذه المعاملات او ما تدعى معاملات سبراغ (Sprague multipliers) تستعمل لاستنباط اعداد الوفيات و اعداد السكان حسب احادية السن داخل فئات عمرية خماسية السنّ معطاة باستعمال اعداد الوفيات و اعداد السكان في هذه الفئات العمرية وما جاورها من فئات خماسية السنّ باستخدام برنامج Excel اذ تم استعمال بيانات فئات العمر أحادية السنّ لدقة تحسسّها لأي عُمُر يواجه خطر الفناء.

الكلمات المفتاحية : توزيع كاما العام، دالة البقاء، المحاكاة، دالة الامكان الاعظم، مبدأ اعظم دالة انتروبي.

Survival Function Estimating of Single age groups For Generalized Gamma Distribution with Simulation

Omar Abdulmohsin Ali1
dromar72@coadec.uobaghdad.edu.iq
1Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq.

Raghda Ziad Tariq2
raghda_ziad2001@yahoo.com
2Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq.

Abstract
This research covered two problems: theoretical problem represented by the integration of the incomplete gamma function implicitly in the generalized gamma (GG) distribution, and the other one was a problem of application in the presence of disturbance and confusion in Iraq Socio-Economic Survey: IHSES II (2012), resulting from the use of the five-year groups in estimating the survival function. The entropy method was therefore used for estimating and fitting purposes, in addition of using a traditional method, which is the maximum likelihood. The best method was determined by using the comparison criteria: the root of the mean square error (RMSE), and the mean absolute percentage error (MAPE). The sample size selected were (n = 18, 30, 50, 81) with (500) replicates. The simulation results showed that the maximum likelihood method is best for small and medium samples. While the entropy method excelled with large samples. The five-year age groups were converted into single-age groups on the basis of the formula for interpolation developed by the scientist (Sprague) through coefficients called Sprague multipliers.

Keywords: Generalized Gamma distribution, Survival function, Maximum Likelihood, Entropy, Sprague multipliers, Simulation.

1. المقدمة و هدف البحث

1.1: المقدمة

تبرز الحاجة الى تمثيل دالة البقاء على قيد الحياة بالضرورة تحديد احد التوزيعات الاحتمالية الملائمة ليعكس ظاهرة البقاء بشكل واضح. وغالباً ما يكون هذا التوزيع هو احد التوزيعات الأسية أو ما تدعى بتوزيعات الحياة Life distributions وتم تحديد توزيع كاما العام ذو المعلمات الثلاثة بسبب ما يميزّه عن توزيع كاما الاعتيادي بمعلمتين من دقة وحساسية عاليتين يتكيّف لصياغة دالة البقاء. الا أن لهذا التوزيع بعض الصعوبة الرياضية في اشتقاق المؤشرات الاحصائية الخاصّة به، لذا كان لابد من اللجوء الى طرائق عددية لغرض عمل التقديرات اللازمة وينطوي تحليل البقاء على النمذجة من وقت البيانات قبل وقوع حدث العطل أو الموت. في هذا السياق، يُعدّ العطل أو الموت "حدثاً" في أدبيات تحليل البقاء. في الجانب التقليدي فقط يقع حدثاً واحداً لكل وحدة تجريبية (عطل ماكنة أو موت أنسان). وأن دراسة أحداث العطل أو الموت المتكررة هي ذات الصلة في معولية الأنظمة وفي مجالات كثيرة من العلوم الاجتماعية والبحوث الطبية. وبطريقة أكثر وضوحا فأن التحليلات لدالة البقاء وتشمل نمذجة وقت البقاء[12]،ويعدّ الانتروبي مقياساً عملياً للاضطرابات disturbances التي تعتري التمثيل الاحتمالي لظاهرة معينة والذي يأتي في أطار المعلومات المرتبطة بالتوزيع الاحتمالي ذاته. وقد تم استعمال اسلوب الانتروبي في العديد من المجالات العلمية المختلفة. و كان أول من استخدم مصطلح الانتروبي العالم الالماني رودولف كلاوزيوس (Rudolf Clausius) عام (1865)[8]،وقد تم تعريف الانتروبي في التحليل الاحصائي بانها طريقة تقدير الدوال الاحتمالية ضمن قيود محدّدة يمكن التعبير عنها بتكامل الدالة الاحتمالية ذاتها، ويدعى هذا المبدأ في الانتروبي بأنه مبدأ تعظيم الانتروبي Principle of Maximizing Entropy: POME ،بينما قدّم (Shannon's) (1948) [18] الانتروبي على أنه عبارة عن نوع من قياس للتوزيع الاحتمالي بشكل تحقق فيه اعلى انتروبية مقارنة مع توزيعات احتمالية اخرى. فحين يكون لدينا معلومات يجب تحقّق بعض الشروط على توزيعها الاحتمالي دون التحديد الكامل لها. فلو كان لدينا توزيعين احتماليين وتقريبا متساويين في المعلومات فأننا نفضل التوزيع الاكثر انتروبية.

2.1: مشكلة البحث

تظهر مشكلة البحث في جانبها التطبيقي بوجود بيانات لفئات خماسية السنّ فقط يتم التعامل معها وتحليلها لدى الجهات الرسمية المتمثلة بالجهاز المركزي للأحصاء في العراق، ولا يوجد أي بيانات تذكر حول دالة البقاء تعتمد على فئات عمرية أحادية السنّ على وجه الدقّة. هذا، فضلا عن وجود مشكلة نظرية تحليلية تعترض مسألة تقدير توزيع كاما العام Generalized Gamma: GG بثلاثة معلمات، بسبب تضمنه لتكامل كاما الناقص Incomplete gamma integration عند ايجاد الدالة الاحتمالية التراكمية CDF: Cumulative Distribution Function وبالتالي دالة البقاء وهو ما يجعل الصياغة الرياضية للعمليات الاستدلالية أكثر تعقيداً لهذا التوزيع وبهذا يكون الاسلوب التقليدي في التقدير امراً بالغ الصعوبة.

3.1: هدف البحث

يتركز اهتمام البحث على هدفين رئيسين هما:

أولاً: إقتراح إستعمال معاملات سبراغ لتحويل الفئات خماسية السنّ للسكان في العراق والمتمثلة بـ (18) فئة خماسية الى فئات عمرية أحادية السنّ والذي سينتج عنه (81) فئة أحادية سواء بالجانب التجريبي أو التطبيقي.

ثانياً: إيجاد تقدير لدالة البقاء يتغلّب على اضطراب البيانات وتذبذبها وهي طريقة الأنتروبي، الى جانب التقدير بطريقة الأمكان الأعظم، بعد تقدير معلمات توزيع كاما العام GG بثلاثة معلمات عن طريق المحاكاة بأسلوب مونت- كارلو بعد توليد البيانات بأربع حجوم عينـات: (n= 18, 30, 50, 81). هذا، الى جانب تطبيق الأجراء نفسه على بيانات حقيقية لسكان العراق. ومن ثم إستعمال جذر متوسط مربعات الخطأRoot Mean Square Error (RMSE) ومتوسط مطلق الخطأ النسبي Mean Absolute Percent Error (MAPE) كمعايير مقارنة الأفضلية.

2. الجانب النظري

1.2:المقدمة

سيتم تحديد صيغة دالة الكثافة الاحتمالية (pdf) f(t) ودالة التوزيع التراكمية (CDF) $F(t)$ ودالة البقاء $S(t)$لتوزيع كاما العام بثلاث معلمات وعرض بعض خصائصه مثل المتوسط والتباين والحالات الخاصة له، وكذلك بيان كيفية اشتقاق وتقدير معلمات ودالة البقاء باستخدام طرق مختلفة مثل الامكان الاعظم ML ومبدا اعظم دالة انتروبيPrinciple of Maximizing Entropy (POME) .

2.2: توزيع كاما العام [9],[10] (Generalized gamma: GG)

يرجع اصل توزيع كاما العام بأربع معلمات(k,β,θ,μ) الى العالم Amoroso في عام [1925] [6] ويعتبر الاصل لنموذج العمر اذ ان (β,k) تمثل معلمات الشكل و θ معلمة قياس و μ معلمة موقع او ازاحة الدالة الاحتمالية لتوزيع:

$$\text{Amoroso}(k,\beta,\theta,\mu)=\frac{1}{\Gamma(k)}\left|\frac{\beta}{\theta}\right|\ (\frac{t-\mu}{\theta})^{\beta k-1}\ e^{-\left(\frac{t-\mu}{\theta}\right)^{\beta}}\quad t,\beta,\theta,\mu\ \in R$$

$$t\ge\mu\ if\ \theta>0\ ,t\le\mu\ if\ \theta<0$$

وقد تم تقديم توزيع كاما العام بثلاث معلمات من قبل Stacy في عام [1962] [20] اذ افترض ان معلمة الموقع $\mu = 0$.

اما الدالة الاحتمالية لتوزيع GG ذي المعلمات الثلاث (k,β,θ) بالصيغة الاتية[17] :

$$f(t,\theta,\beta,k)=\frac{\beta}{\Gamma(k)\ \theta^{\beta k}}\ t^{\beta k-1}\ e^{-(\frac{t}{\theta})^{\beta}}\ \ t>0,k>0,\theta>0,\beta>0\ \ldots(1)$$

اذا ان:

$t \Leftarrow$ تمثل المتغير العشوائي (العمر)، $\theta \Leftarrow$ معلمة قياس (Scale Parameter)

$k,\beta \Leftarrow$ معلمات شكل (Shape Parameter)

بينما تكون الدالة الاحتمالية التراكمية:

$$F(t)=\frac{IG((\frac{t}{\theta})^{\beta},k)}{\Gamma(k)}\ldots(2)$$

اذ ان :

$$IG(S,K)=\frac{1}{\Gamma(k)}\int_0^S u^{k-1}e^{-u}\,du$$

$IG(S,K)$: هي دالة كاما الناقصة (incomplete gamma function)، وأن قيمة المتوسط للمتغير t من مرتبة β:

$$E\left(t^{\beta}\right)=\frac{\theta^{\beta}\Gamma(k+1)}{\Gamma(k)}=\ \theta^{\beta}k\ \ \ldots(3)$$

ومتوسط اللوغاريتم للمتغير t :

$$E(\ln t)=\frac{\Psi(k)}{\beta}+\ln\theta\ldots(4)$$

3.2: تحليل دالة البقاء[16], [14], [3]

يعدّ تحليل البقاء على قيد الحياة فرع من فروع الإحصاء الرياضي التي تتناول تحليل المدة الزمنية المستغرقة قبل ظهور واحد أو أكثر من الأحداث مثل الموت في حالة الكائنات الحيّة، والعطل في حالة المكائن، ويدعى هذا الموضوع نظرية المعّولية أو تحليل المعّولية في مجال الهندسة للمواد الصناعية غير الحية.

نحاول عن طريق تحليل البقاء على قيد الحياة الإجابة على أسئلة مثل: ما هي نسبة عدد السكان والتي سوف تبقى على قيد الحياة إلى وقت معين؟ ومن تلك الفئات العمرية أيضا التي تستطيع البقاء على قيد الحياة؟ ماهي نسبة الموت؟ وكيف يمكن تحديد أسباب متعددة للموت؟ وكيف إن ظروف أو خصائص معينة تزيد أو تقلل من احتمال البقاء على قيد الحياة؟

وتعرّف دالة البقاء ببساطة على أنها احتمال ان حدث (الموت) لم يحدث حتى الان حسب الزمن t وبالتالي فان T يدل على الزمن حتى الموت فان s(t) يدل على احتمال البقاء على قيد الحياة ، أذ ان:

$$S(t)=1-F(t)=P(T>t)\qquad \text{for } t>0$$

ولنفترض ان $S(t=0)\ =1$ أي أن احتمال بقاء المصاب على قيد الحياة في الزمن (0) يساوي واحد وكذلك يجب أن تكون دالة البقاء غير متزايدة ومستمرة من الجانب الايمن:

$$S(u)\le S(t)\qquad if\,u>t$$

فاحتمال البقاء غالبا يفترض انه يقترب من الصفر كلما ازداد عمر الكائن الحي (الانسان مثلاً):

$$S(t)\to 0\quad as\ t\to\infty$$

ومن خصائص بيانات البقاء على قيد الحياة هو ان وقت البقاء على قيد الحياة لا يمكن ان يكون سالبا.

ودالة البقاء للتوزيع المستمر هي:

$$S(t)=p(T>t)=\int_t^{tmax}p(t)dt$$

فان دالة البقاء لتوزيع كاما العام تمثل حسب المعادلة (1):

$$S_{GG}(t)=1-\frac{IG((\frac{t}{\theta})^{\beta},k)}{\Gamma(k)}\ \ldots(5)$$

4.2: طرائق التقدير

سيتم هنا تقديم طريقتين لأغراض تقدير دوال البقاء على قيد الحياة لفئات العمر أحادية السنّ وهي طريقة الامكان الاعظم ML وطريقة مبدأ اعظم دالة انتروبيPOME.

1.4.2: طريقة الامكان الاعظم [5]

تعد من الطرق التقليدية والاكثر كفاءة لتقدير المعلمات في أي توزيع فضلا عن انه يواكب جميع العينات وتمتلك العديد من الخصائص الاحصائية الثبات inversion، الاتساق consistency، عدم التحيز unbiased.

بما ان مقدرات الامكان الاعظم تتصف بخاصية الثبات[4],[21]، اي اذا كانت $T=g(\theta)$ هي دالة بدلالة المتغير θ نفترض ان $\hat{\theta}_n$ هي مقدر الامكان الاعظم ل θ فان $\hat{T}_n = g(\hat{\theta}_n)$ هي مقدر الامكان الاعظم لدالة T.

دالة الامكان لتوزيع كاما العام ذي المعلمات الثلاثة (θ, k, β) يعبر عنها بالمعادلة الأتية*:

$$L=\prod_{i=1}^{n} f(t_i, k, \beta, \theta) \ldots (6)$$

$$L=\left(\frac{\beta}{\Gamma(k)\theta^{\beta k}}\right)^n \left(\prod_{i=1}^{n} t_i^{\beta k-1}\right) \left(e^{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{\theta^{\beta}}}\right) \ldots (7)$$

لتحويل دالة الامكان (7) الى الشكل الخطي يتم أخذ اللوغاريتم الطبيعي لها:

$$\text{Log } L= n \log \beta - n\beta k \log \theta - n \log \Gamma(k)+ (\beta k-1)\sum_{i=1}^{n} \log t_i - \frac{\sum_{i=1}^{n} t_i^{\beta}}{\theta^{\beta}}$$

$$= n \log \beta - n\beta k \log \theta - n \log \Gamma(k)+\beta k\sum_{i=1}^{n} \log t_i-\sum_{i=1}^{n} \log t_i - \frac{\sum_{i=1}^{n} t_i^{\beta}}{\theta^{\beta}} \quad \ldots(8)$$

لإيجاد القيم التقديرية للمعلمات الثلاث (θ, k, β) نجد المشتقات الجزئية للمعادلة (8) نسبة الى المعالم ومن ثم مساواة المشتقات الجزئية الى الصفر نحصل على القيم التقديرية للمعالم التي تجعل دالة الامكان اعظم ما يمكن:

$$\hat{k}=\frac{\hat{\theta}^{-\hat{\beta}} \sum_{i=1}^{n} t_i^{\hat{\beta}}}{n} \quad \ldots (9)$$

$$\hat{\theta}=\left(\exp\left(-\Psi(\hat{k}) + \frac{\hat{\beta}\sum_{i=1}^{n} \log t_i}{n}\right)\right)^{\frac{1}{\hat{\beta}}} \quad \ldots (10)$$

حيث ان :

$$\Psi(k) = \frac{d}{dk}\ln(\Gamma(k)) = \frac{\Gamma'(k)}{\Gamma(k)}$$

بعد تعويض القيم التقديرية ˆو ‸ يتم استخراج القيمة التقديرية ل β بطرق عددية .

وباستعمال خاصية الثبات نحصل على مقدر دالة البقاء لتوزيع كاما العام لثلاث معلمات بالصيغة الآتية:

$$\hat{S}_{GG}(t) = 1 - \frac{IG\left(\left(\frac{t}{\hat{\theta}}\right)^{\hat{\beta}}, \hat{k}\right)}{\Gamma(\hat{k})} \ldots (11)$$

2.4.2: مبدأ اعظم انتروبي [13],[15]

أن استخدام هذا المبدأ في تكوين توزيعات احتمالية ذات تحيز قليل على أساس المعلومات المحددة وغير التامة قد جاءت من قبل (Jaynes 1968)[11] و (Singh & Fiorentino 1992) [19] استخدام (POME) لاختيار $f(t,\theta)$ التي تعظم وتحقق المعلومات المتوفرة والممثلة بصورة قيود (Constraints) أي بعبارة أخرى ، الحصول على المعلومات المعينة مثل (على سبيل المثال الوسط الحسابي ، التباين ، الالتواء ، الحد الأدنى ،الحد الأعلى) فأن التوزيع الذي يتم اشتقاقه بواسطة (POME) سيكون أفضل ما يمثل المتغير (t) .

ويمكن ان يعبر عنها:

$$H(f) = -\int_{-\infty}^{\infty} f(t,\theta) \ln f(t,\theta) dt \quad \ldots (12)$$

أذ ان :

$$\int_{-\infty}^{\infty} f(t,\theta) dt = 1 \quad \ldots (13)$$

يمثل الرمز $H(f)$ دالة الانتروبي للدالة $f(t,\theta)$

يمكن كتابة دالة الانتروبي بالشكل الاتي :

$$H(f) = E - \ln f(t,\theta) \ldots (14)$$

اذ ان $H(f)$ هو متوسط القيمة $-\ln f(t,\theta)$

فان التوزيع الأقل تحيز بالنسبة للمتغير (t) هو ذلك التوزيع الذي يعمل على تعظيم دالة انتروبي طبقاً للمعلومات المعطاة ، أي انه الذي يحقق مبدأ أعظم دالة انتروبي (Principle of Maximum Entropy) حسب ما جاء به الباحث (Jaynes, 1961) .

ونتيجة لذلك فأن الحصول على معلمات التوزيع يتحقق من خلال تعظيم (f) Maximum.

* الاشتقاق من قبل الباحثين.

ادخال قيود جديدة تتعلق بالمعلمات او الحصول على دالة انتروبي جديدة مقيدة حيث يمكن توظيف مسائل الامثلية مثل مضاعفات لاكرانج عند تقدير المعلمات.

عند تعويض دالة الاحتمالية لتوزيع كاما العام كما في المعادلة (1) في المعادلة (12) يتم استخراج القيود†:

$$a_0 = \ln \Gamma(k) - \ln \beta + k\beta \ln \theta$$
$$a_1 = \frac{1}{\theta^{\beta}}$$
$$a_2 = -(k\beta - 1) = 1 - k\beta$$

بعد الاشتقاق يتم التوصل الى دالة احتمالية بدلالة القيود ومضاعفات لاكرانج :

$$f(t, a_1, a_2, \beta) = \frac{\beta\, a_1^{\frac{1-a_2}{\beta}}}{\Gamma(\frac{1-a_2}{\beta})}\; t^{-a_2}\, e^{(-a_1 t^{\beta})} \quad \dots (15)$$

نوجد المتوسط للمتغير t للمعادلة (15) المرفوع للاس β

$$E(t^{\beta}) = \frac{\frac{1-a_2}{\beta}}{a_1} \quad \dots (16)$$

نستخرج متوسط اللوغاريتم للمتغير t من المعادلة (15):

$$E(\ln t) = \frac{\Psi(\frac{1-a_2}{\beta})}{\beta} - \frac{\ln a_1}{\beta} \quad \dots (17)$$

نوجد الدالة التراكمية للمعادلة (15):

$$F(t, a_1, a_2, \beta) = \frac{IG(a_1 t^{\beta}, \frac{1-a_2}{\beta})}{\Gamma(\frac{1-a_2}{\beta})} \quad \dots (18)$$

ومن المعادلة (18) نوجد دالة البقاء ل$f(t, a_1, a_2, \beta)$:

$$S(t, a_1, a_2, \beta) = 1 - \frac{IG(a_1 t^{\beta}, \frac{1-a_2}{\beta})}{\Gamma(\frac{1-a_2}{\beta})} \quad \dots (19)$$

بعد تعويض الدالة الاحتمالية بدلالة القيود والمضاعفات في المعادلة (14) ينتج:

$$H(f) = -\ln \beta - \left(\frac{1-a_2}{\beta}\right) \ln a_1 + \ln \Gamma\left(\frac{1-a_2}{\beta}\right) + a_2 E(\ln t) + a_1 E t^{\beta} \dots (20)$$

العلاقة بين معلمات التوزيع والقيود:

$$\frac{dH(f)}{da_1} = \frac{\frac{1-\hat{a}_2}{\hat{\beta}}}{\hat{a}_1} - E(\, t^{\hat{\beta}}) = 0 \dots (21)$$
$$\frac{dH(f)}{da_2} = \frac{1}{\hat{\beta}} \ln \hat{a}_1 - \frac{1}{\hat{\beta}}\, \Psi(\frac{1-\hat{a}_2}{\hat{\beta}}) + E(\ln t) = 0 \;\dots (22)$$
$$\frac{dH(f)}{d\beta} = -\frac{1}{\hat{\beta}} + \frac{1}{\hat{\beta}^2}(1-\hat{a}_2)\ln \hat{a}_1 + \Psi\left(\frac{1-\hat{a}_2}{\hat{\beta}}\right)\left(-\frac{1}{\hat{\beta}^2}\right) + \hat{a}_1 E(\, t^{\hat{\beta}}) \ln(t) = 0 \dots (23)$$

وبسبب صعوبة حل المعادلات سيتم اللجوء الى طريقة نيوتن رافسن لتقدير a_1 من المعادلة (21) و a_2 من المعادلة (22) و β من المعادلة (23).

وعند تقدير المعلمات(a_1, a_2, β) نقدر دالة البقاء حسب الصيغة الاتية :

$$\hat{S}(t, \hat{a}_1, \hat{a}_2, \hat{\beta}) = 1 - \frac{IG(\frac{t^{\hat{\beta}}}{\hat{a}_1^{-1}}, \frac{1-\hat{a}_2}{\hat{\beta}})}{\Gamma(\frac{1-\hat{a}_2}{\hat{\beta}})} \quad \dots (24)$$

5.2: معايير المقارنة [1]

يتم استعمال بعض معايير المقارنة للتمييز بين طريقتي تقدير دالة البقاء المستعملة في البحث وهي كما في أدناه.

1.5.2: جذر متوسط مربعات الخطأ(RMSE)

أن الصيغة العامة لــ (RMSE) هي:

$$RMSE(\hat{s}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(s_i - \hat{s}_i)^2} \quad \dots (25)$$

2.5.2: متوسط مطلق الخطأ النسبي(MAPE)

ان الصيغة العامة لــ (MAPE) هي:

$$MAPE(\hat{s}) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{s_i - \hat{s}_i}{s_i}\right| \quad \dots (26)$$

† الاشتقاق من قبل الباحثين.

3. الجانب التجريبي

1.3:المقدمة

لغرض إعطاء صورة واضحة عن أهمية ودقة طرائق تقدير دالة البقاء فقد استخدم أسلوب المحاكاة (Simulation) من اجل المقارنة بين أفضلية المقدرات وقد تم تقييم النتائج باستخدام المقاييس الإحصائية وهي (RMSE)و(MAPE).

2.3: المحاكاة

التقديرات السنوية للسكان حسب الجنس والعمر قد تكون متاحة للفئات عمرية خماسية السنّ وليس في التوزيع أحادي السنّ الذي يغطي جميع الاعمار اي على شكل فئات العمر أحادية السنّ حيث تطلب استخدام المحاكاة لتوليد بيانات العمر على شكل فئات احادية السن لدقة تحسسها لأي عمر يواجه خطر الفناء (الموت) حيث تمثل البيانات الحقيقية التي تم الحصول عليها بطريقة رياضية (معاملات سبراغ) لتفكك الفئات الخماسية التي تم الحصول عليها من وزارة التخطيط/ الجهاز المركزي للإحصاء وتكنلوجيا المعلومات(IHSES 2012) وتحويلها الى فئات احادية السن.

وتعرف المحاكاة بأنها أسلوب رياضي لحل المشاكل المعقدة التي تظهر من خلال المعاينة لذلك تعد المحاكاة من أكثر الأساليب استعمالاً لأنها تتميز بالمرونة إذ تعطي القدرة على الاختبار والتجريب من خلال تكرار العملية لمرات عديدة وكذلك تمتاز باختصار الوقت وتبنى فكرتها الأساسية على تقليد الموقف في عالم الواقع باستعمال الأنموذج الرياضي الذي لا يؤثر على الأداء [7] .

هناك عدة طرائق للمحاكاة تعد أكثرها استخداماً هي طريقة مونت كارلو والتي تقوم على فكرة توليد العينات العشوائية من المجتمع النظري المفترض المماثل للمجتمع الحقيقي .

- مراحل بناء تجربة المحاكاة:

تتضمن بناء تجربة المحاكاة على أربع مراحل أساسية ومهمة لتقدير المعلمات (β, K, θ) لتوزيع كاما العام $f(t, \beta, k, \theta)$وهي كما في أدناه.

المرحلة الاولى (تعيين القيم الافتراضية)

هذه المرحلة من المراحل الأساسية والمهمة التي تعتمد عليها المراحل الأخرى آلاتية من برنامج بناء تجربة المحاكاة ، إذ يتم تعيين القيم الافتراضية الحقيقية وكما يأتي :

أولاً: تحديد قيم افتراضية للمعلمات (β, k, θ)

تم اختيار قيم افتراضية للمعلمات والتي تتمثل بأخذ تجارب لقيم المعلمات

$$A_1: (\theta = 30 , \beta = 1, k = 0.3), A_2: (\theta = 15 , \beta = 0.67, k = 0.36):, A_3: (\theta = 9 , \beta = 0.3, k = 0.2), A_4: (\theta = 25 , \beta = 0.77 \ \ , k = 0.32), A_5: (\theta = 5 , \beta = 0.2k = 0.3)$$

ثانياً: اختيار حجم العينة (n)

تم اختيار حجوم مختلفة للعينة بشكل يتناسب مع الظاهرة المبحوثة ومعرفة مدى تأثير حجم العينة مع دقة وكفاءة النتائج المستحصلة من طريقتي التقدير المستخدمة في هذه الدراسة فقد أخذت حجوم العينة الصغيرة وهي (n=18) وحجم عينة متوسطة وهي (n=50,30) وأخذت حجوم عينة كبيرة وهي (n=81) .

ثالثاً : اختيار عدد تكرارات التجارب (q)

اختيار عدد تكرارات عينة (q=500)

المرحلة الثانية (توليد البيانات)

في هذه المرحلة يتم توليد بيانات عشوائية حيث يمكن ايجاد بيانات عشوائية التي تستخدم اساسا للحصول على العينات من أي التوزيعات الاحتمالية :كالتوزيع المنتظم والتوزيع الثلاثي والتوزيع الطبيعي والاسلوب الاساسي الذي يستخدم لتوليد الارقام العشوائية يعرف باسم طريقة التحويل العكسي(Inverse Transformation Method :ITM) وهو الاسلوب الاسهل والاكثر شيوعا من الاساليب المستخدمة لتوليد البيانات العشوائية والذي يستخدم المعلومات التي فيها المتغير العشوائي

$$\mathrm{F(x)} = \mathrm{r}$$
$$r = f(x)$$
$$x = \overset{-1}{F}(r)$$

المرحلة الثالثة (إيجاد المقدرات)

يتم في هذه المرحلة تقدير كل من المعلمات (β, k, θ) لإيجاد دوال البقاء الناتجة من توزيع كاما العام $f(t, \beta, k, \theta)$ من خلال طريقتي التقدير المتناولة في الجانب النظري من هذا البحث.

المرحلة الرابعة (مرحلة المقارنة)

يتم في هذه المرحلة المقارنة بين دوال البقاء لتوزيع كاما العام $f(t, \beta, k, \theta)$ التي تم الحصول عليها من المرحلة الثالثة إذ تم استخدام معياران لإجراء المقارنة وهما RMSE و MAPE.

3.3: نتائج المحاكاة

يتم عرض نتائج محاكاة طريقتي التقدير MLE و Entropy . ومن ثم تحليلها وتحديد الى افضل المقدرات المستحصلة لمعلمات التوزيع $f(t,\beta,k,\theta)$ وايجاد دالة البقاء من خلال المفاضلة بين قيم هذه المقدرات بالاعتماد على معايير المقارنة المستخدمة وقد تم الحصول على هذه النتائج باعتماد برنامج Matlab . وفيما يأتي النتائج الموضحة في الجداول التي سيتم تحليلها حسب التسلسل وكما في أدناه.

1.3.3: نتائج تقديرات دالة البقاء باستعمال طريقتي التقدير

من خلال النتائج المبينة في الجدول رقم (1) نلاحظ ان طريقة الانتروبي هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الاول $A_1: (\theta = 30\ , \beta = 1, k = 0.3)$ عند حجوم العينات (n=18,30,50,81)

من خلال النتائج المبينة في الجدول رقم (2) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الأخرى لتقدير دالة البقاء في حالة الانموذج الثاني $A_2: (\theta = 15\ , \beta = 0.67, k = 0.36)$ عند حجوم العينات (n=18,30,50) ان طريقة الانتروبي هي الافضل لتقدير دالة البقاء عند حجوم العينات (n=81).

من خلال النتائج المبينة في الجدول رقم (3) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الثالث $A_3: (\theta = 9\ , \beta = 0.3, k = 0.2)$ عند حجوم العينات (n=18,30,50,81)

من خلال النتائج المبينة في الجدول رقم (4) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الرابع $A_4: (\theta = 25\ , \beta = 0.77k = 0.32)$ عند حجم العينة(n=18,30) ان طريقة الانتروبي هي الافضل لتقدير دالة البقاء عند حجوم العينات (n=50,81).

من خلال النتائج المبينة في الجدول رقم (5) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الخامس $A_5: (\theta = 5\ , \beta = 0.2k = 0.3)$ عند حجم العينة (n=18) ان طريقة الانتروبي هي الافضل لتقدير دالة البقاء عند حجوم العينات (n=30,50,81).

من خلال النتائج المبينة في الجدول رقم (6) نلاحظ ان طريقة الانتروبي هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الاول $A_1: (\theta = 30\ , \beta = 1, k = 0.3)$ عند حجوم العينات (n=,18,30,50,81).

من خلال النتائج المبينة في الجدول رقم (7) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الثاني $A_2: (\theta = 15\ , \beta = 0.67, k = 0.36)$ عند حجوم العينات (n=18,30) ان طريقة الانتروبي هي الافضل لتقدير دالة البقاء عند حجوم العينات (n=50,81).

من خلال النتائج المبينة في الجدول رقم (8) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الأخرى لتقدير دالة البقاء في حالة الانموذج الثالث $A_3: (\theta = 9\ , \beta = 0.3, k = 0.2)$ عند حجوم العينات (n=18,30) ان طريقة الانتروبي هي الافضل لتقدير دالة البقاء عند حجوم العينات (n=50,81).

من خلال النتائج المبينة في الجدول رقم (9) نلاحظ ان طريقة الامكان الاعظم هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الرابع $A_4: (\theta = 25\ , \beta = 0.77k = 0.32)$ عند حجم العينة(n=18,30,50,81)

من خلال النتائج المبينة في الجدول رقم (10) نلاحظ ان طريقة الانتروبي هي الافضل من الطريقة الاخرى لتقدير دالة البقاء في حالة الانموذج الخامس $A_5: (\theta = 5\ , \beta = 0.2k = 0.3)$ عند حجم العينة (n=18,30,50,81).

4. الجانب التطبيقي

1.4: المقدمة

من خلال الجانب النظري يبين كيفية تقدير دالة البقاء من خلال طريقتين وهي (Entropy،ML)، وأخيرا يتم حساب معايير المقارنة للطريقتين الذكور والاناث معا.

2.4: وصف البيانات

تم جمع البيانات من الجهاز المركزي للإحصاء حول الاوضاع المعيشية للأسرة العراقية (المسح الاجتماعي الاقتصادي للأسر)(IHSES II 2012)[2] الذي يوفر منظومة معلومات متكاملة تفصيلية، واجمالية عن الوضع

المعيشي للفرد والأسرة العراقية، ليقدم صورة رقمية موضوعية عن تطور الأوضاع المعيشية في المجالات المختلفة (الإنفاق والدخل، الوضع الديموغرافي، الصحة، السكن، التشغيل وغيرها). كما إن هذا الملف سيخدم شرائح واسعة من مستخدمي هذا النوع من البيانات على المستويين القطاعي في تنوع جوانب التغطية الإحصائية للمؤشرات، والمكاني في مستوى الشمول والقدرة على توفير مؤشرات على مستوى المحافظات والأقضية. بلغ حجم العينة (25488) أسرة، في جميع محافظات العراق بواقع (216) أسرة لكل قضاء من أقضية العراق ال (118) كما بلغ عدد العناقيد (2832) عنقود ويشمل كل عنقود (9) أسر موزعة على الأقضية والمحافظات وللبيئتين الحضرية والريفية.

لم يتم جمع اعداد الوفيات للمسح (IHSES II 2012) حسب الفئات العمرية أحادية السنّ حيث وجدت بيانات على شكل فئات عمرية خماسية السنّ لأعداد الوفيات وبما ان لباحث يحتاج في الرسالة الى اعداد الوفيات واعداد السكان وذلك بقسمة عدد الوفيات على عدد السكان لكل فئة مما اضطر الى اللجوء لاستعمال طريقة رياضية تؤدي الى النتائج بالاستناد الى معادلة التدريج (Formula for Interpolation) وضعها سبراغ (Sprague) وهذه المعاملات او ما تدعى معاملات سبراغ (Sprague multipliers) تستعمل لاستنباط اعداد الوفيات او اعداد السكان حسب احادية السن داخل فئات عمرية خماسية السنّ معطاة باستعمال اعداد الوفيات او اعداد السكان في هذه الفئات العمرية وما جاورها من فئات خماسية باستعمال برنامج Excel . اذ سيتم تقدير دالة البقاء باستعمال بيانات احادية الفئة.

3.4: نتائج التطبيق

حسب نتائج المحاكاة فان طريقة مبدأ اعظم دالة انتروبي افضل في حالة حجوم العينات الكبيرة هي الافضل اذ تمثل الفئات احادية السن التي تتكون من 81 فئة .

بعد ملاحظة نتائج الجدول رقم (11) أعلاه يتبين سلوك تقدير دالة البقاء عند استعمال الطريقة الانتروبي بأن تبدأ احتمال البقاء بالهبوط عن 0.5 لدالة الحقيقية عند العمر 51 ، اما طريقة الانتروبي تبدأ بالهبوط عند العمر 70 ويبدأ احتمال البقاء للذكور والاناث معا- حضر وريف بالهبوط عن 0.9 عند العمر 53.

جدول رقم (12) يبين المقاييس الاحصائية (RMSE,MAPE) لتقدير دالة البقاء للفئات الاحادية للسكان الذكور والاناث معا بطريقة الانتروبي .

5. الاستنتاجات والتوصيات

1.5: الاستنتاجات

عند استخدام الجانب التجريبي لتقدير دوال البقاء والمقارنة بينها تم استنتاج ان مقدرات مبدأ اعظم دالة انتروبي (POME)هي الافضل في الجداول (1) لجميع حجوم العينات المستخدمة(n=18,30,50,81)وجدول رقم (2)عند حجم العينة (n=81)وجدول رقم(4) عند حجم العينة (n=50,81) اضافة الى الجدول رقم (5)عند حجوم العينات (n=30,50,81) ثم جاءت نتائج الامكان الاعظم هو الافضل في الجدول رقم (2)لحجوم العينات(n=18,30,50) وجدول رقم (3)لجميع حجوم العينات وجدول رقم (4)عند حجم العينة(n=18,30)وجدول رقم (5)عند حجم العينة (n=18) (باستخدام مقياس RMSE).

أما باستخدام مقياس MAPE فكانت النتائج مبدأ اعظم دالة انتروبي هي الافضل في حالة الجدول (6),(9),(10)لجميع حجوم العينات وجدول (7),(8)عند حجوم العينات (n=50,81) وجاءت نتائج الامكان الاعظم افضل (7),(8)عند حجوم العينات (n=18,30).تم تطبيق طريقة الانتروبي في حالة البيانات فئات العمر أحادية السنّ.

2.5: التوصيات

1. توسيع نطاق البحث في الموضوع يمكن ادخال قيود جديدة تتعلق بالمعلمات والحصول على دالة انتروبي جديدة مقيدة لتوزيعات اخرى وعندئذ يمكن توظيف مسائل الامثلية مثل مضاعفات لاكرانج عند تقدير المعلمات ومن ثم تقدير دوال البقاء على قيد الحياة او المعولية.
2. نوصي باستخدام طريقة الامكان الاعظم في حالة البيانات لفئات العمر خماسية السنّ وطريقة مبدأ اعظم دالة انتروبي في حالة البيانات فئات العمر أحادية السنّ.

المصادر

المصادر العربية:

1. ابراهيم، زينب علاوي،(2013)،" الدقة في تخطيط كمية انتاج مادة السمنت في معامل الشركة العامة للسمنت العراقية (دراسة اختبارية لطرق التنبؤ باستخدام معايير الخطأ) "،مجلة ديالى للعلوم الهندسية ،المجلد(7)،العدد(1)، ص:39-59 .
2. الجهاز المركزي للإحصاء، وزارة التخطيط ، (2014) ،"المسح الاجتماعي والاقتصادي للأسرة في العراق IHSES II 2012 تقرير الجداول "، طبعت في مطبعة الجهاز المركزي للإحصاء، العراق.
3. فرحان،ابتهال حسين، (2015) ،"تقدير دالة البقاء لبيانات حقيقية كاملة لمرض سرطان الرئة"، رسالة ماجستير في علوم الرياضيات ،كلية التربية ابن الهيثم، جامعة بغداد.

4. محمد، مكي اكرم صالح، (2006) ،"محاكاة طرائق تقدير معلمة القياس ودالة المعولية لتوزيع ويبل ذي المعلمتين "، اطروحة دكتوراه في علوم الرياضيات، كلية التربية، الجامعة المستنصرية.

المصادر الاجنبية:

5. Ahmed, A.; Mir, K.A.; and Reshi, J.A,(2014),"Some Important Statistical Properties, Information Measures and Estimation of Sized Biased Generalized Gamma Distribution", Journal of Reliability and Stati stical Studies, Vol. 7, No. Issue 2, pp.161-179.
6. Amoroso, L., (1925) ," Ricerche intorno alla curve dei redditi ", Annali di Matematica Puraed Applicants ,vol.2(1), pp.123–159
7. Arsham,H. ,(2007),"System Simulation :The shortest Route to Application", http://home.ubalt.edu/ntsbarsh/simulation/sim.htm#rintroduction .
8. Cropper, W. H.,(2004),"The Road to Entropy Rudolf Clausius", Great Physicists, The Life and Times of Leading Physicists from Galileo to Hawking, Oxford University Press, ISBN 978-0-19-517324-6,pp. 93–105.
9. De Ves, E.; Benavent, X.; Ruedin, A.; Acevedo, D.; and Seijas, L. ,(2010),"Wavelet-based Texture Retrieval Modeling the Magnitudes of Wavelet Detail Coefficients with a Generalized Gamma Distribution", IEEE, In: Proc. ICPR, pp. 221 – 224.
10. Hirukawa, M.and Sakudo, M.,(2013),"Family of Generalized Gamma Kernels: A Unified Approach to the Asymptotics on Asymmetric Kernels", Development Bank of Japan.
11. Jaynes, E.T.,(1968),"prior probabilities",*IEEE*, Transactions On Systems Science and Cybernetics,vol. SSC-4,No.3,pp.227-241.
12. Kaplan, E.L. ; Meier, P. ,(1958)," Non Parametric Estimation from Incomplete Observations", Journal of the American Statistical Association, vol. 53, No. 282 ,pp. 457-481.
13. Karmeshu;pal,N.R.,(2003)," Entropy Measures, Maximum Entropy Principle and Emerging Applications", Springer, Verlag, Berlin ,Heidelberg.
14. Lawless, J.F.,(1982)," Statistical Models And Methods for Lifetime Data", John Wiley & Sons, Inc., New York.
15. Levine, R.D.; Tribus, M.,(1979) ,"The Maximum Entropy Formalism" ,MIT Press, Cambridge ,Massachusetts, USE.
16. Qamruz, Z.; and Karl, P.,(2011),"Survival Analysis Medical Research", http://interstat.statjournals.net/YEAR/2011/abstracts/1105005.php .
17. Schutz, A.; Bombrum, L.; Berthoumieu, Y.; and Najim, M. ,(2013), "Centroid Based Texture Classification Using the Generalized Gamma Distribution", In: Proc. EUSIPCO Marrakech, Morocco .
18. Shannon,C.E., (1948) ," System Technical", PP.379-623.
19. Singh,V.P.; Fiorentino ,M., (1992), "A Historical Perspective Of Entropy Applications In Water Resources", vol.9, Physics . Earth Sciences ; Geograph , pp.21-61.
20. Stacy,. E., (1962)," A Generalization of the Gamma Distribution",The Annals of Mathematical Statistics, Vol. (33), No.(3) ,pp.1187–1192.
21. Tan, P.; Drossos, C.,(1975)," Invariance Properties of Maximum Likelihood Estimators", Mathematics Magazine, Vol. 48, No. 1 pp.37-41.

جدول رقم (1) يمثل قيم جذر متوسط مربعات الخطأ (RMSE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 30 \;, \beta = 1 \;, k = 0.3$

RMSE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.2315	0.2222
30	0.2328	0.2299

50	0.2423	0.2346
81	0.2446	0.2369

جدول رقم (2) يمثل قيم جذر متوسط مربعات الخطأ (RMSE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 15 , \beta = 0.67, k = 0.36$

RMSE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.1731	0.1909
30	0.1893	0.2014
50	0.2058	0.2061
81	0.2246	0.2111

جدول رقم (3) يمثل قيم جذر متوسط مربعات الخطأ (RMSE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 9 \ , \beta = 0.3 \ , k = 0.2$

RMSE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.0454	0.1038
30	0.0754	0.1135
50	0.1158	0.1226
81	0.1265	0.1279

جدول رقم (4) يمثل قيم جذر متوسط مربعات الخطأ (RMSE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 25 , \beta = 0.77 \ , k = 0.32$

RMSE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.1990	0.2033
30	0.2018	0.2119
50	0.2199	0.2180
81	0.2278	0.2212

جدول رقم (5) يمثل قيم جذر متوسط مربعات الخطأ (RMSE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 5 \ , \beta = 0.2 \ , k = 0.3$

RMSE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.1033	0.1108
30	0.1588	0.1180
50	0.1703	0.1232
81	0.1554	0.1261

جدول رقم (6) يمثل قيم متوسط مطلق الخطأ النسبي (MAPE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 30 \ , \beta = 1 , k = 0.3$

MAPE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.3149	0.2848
30	0.3151	0.2908
50	0.3273	0.2957
81	0.3303	0.2972

جدول رقم (7) يمثل قيم متوسط مطلق الخطأ النسبي (MAPE) لمقدرات دالة البقاء لأحجام

عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 15, \beta = 0.67, k = 0.36$

MAPE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.2327	0.2446
30	0.2543	0.2550
50	0.2751	0.2586
81	0.3007	0.2660

جدول رقم (8) يمثل قيم متوسط مطلق الخطأ النسبي (MAPE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 9 \ , \beta = 0.3 \ , k = 0.2$

MAPE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.0393	0.0573
30	0.0554	0.0597
50	0.0890	0.0623
81	0.0983	0.0645

جدول رقم (9) يمثل قيم متوسط مطلق الخطأ النسبي (MAPE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 25 \ , \beta = 0.77 \ , k = 0.32$

MAPE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.2644	0.2565
30	0.2666	0.2627
50	0.2894	0.2677
81	0.2999	0.2722

جدول رقم (10) يمثل قيم متوسط مطلق الخطأ النسبي (MAPE) لمقدرات دالة البقاء لأحجام عينات مختلفة وللطريقتين المستعملتين وللقيم الافتراضية: $\theta = 5 \ , \beta = 0.2 \ , k = 0.3$

MAPE		
n	$\hat{S}_{MLE}$	$\hat{S}_{Entropy}$
18	0.1003	0.0725
30	0.1483	0.0738
50	0.1597	0.0751
81	0.1463	0.0751

جدول رقم (11) تقدير دالة البقاء للفئات العمر أحادية السنّ للذكور واناث معا بطريقة الانتروبي

$\hat{S}_{ent}$	S_{real}	Age	$\hat{S}_{ent}$	S_{real}	Age
0.966901	0.637997	40	0.999512	0.916673	0
0.965675	0.633466	41	0.999206	0.901195	1
0.965465	0.632702	42	0.999105	0.896943	2
0.96409	0.627766	43	0.997979	0.862914	3
0.962801	0.62326	44	0.9976	0.854411	4
0.958731	0.609686	45	0.996052	0.826757	5
0.949137	0.580972	46	0.996045	0.826654	6
0.944694	0.568933	47	0.995766	0.822472	7
0.933958	0.542385	48	0.995648	0.820768	8
0.916484	0.505008	49	0.995054	0.812578	9
0.914827	0.501762	50	0.995047	0.812483	10

0.905366	0.48404	51	0.994341	0.803559	11
0.90529	0.483903	52	0.993755	0.796695	12
0.892318	0.461574	53	0.993271	0.791339	13
0.88712	0.453172	54	0.992704	0.785377	14
0.884415	0.44891	55	0.992083	0.779184	15
0.881042	0.443697	56	0.991481	0.773482	16
0.879816	0.441828	57	0.990909	0.76831	17
0.866532	0.422434	58	0.990345	0.763408	18
0.865736	0.421318	59	0.990046	0.760887	19
0.844307	0.392969	60	0.989744	0.758395	20
0.799316	0.341948	61	0.989431	0.755857	21
0.742394	0.288959	62	0.989115	0.753349	22
0.693002	0.250281	63	0.988476	0.748412	23
0.669978	0.234042	64	0.987882	0.74399	24
0.666639	0.231769	65	0.987327	0.739986	25
0.658627	0.226398	66	0.986747	0.735917	26
0.656003	0.224664	67	0.986091	0.731461	27
0.628893	0.207404	68	0.985594	0.728173	28
0.555251	0.165843	69	0.985338	0.726513	29
0.456074	0.119605	70	0.98518	0.725497	30
0.448584	0.116488	71	0.984982	0.724234	31
0.374071	0.08795	72	0.984893	0.723667	32
0.306591	0.065586	72	0.984022	0.718256	33
0.251088	0.049403	74	0.981835	0.705482	34
0.250697	0.049296	75	0.979972	0.695388	35
0.209605	0.038526	76	0.97875	0.689106	36
0.188682	0.033422	77	0.975519	0.673579	37
0.183247	0.032138	78	0.97181	0.657368	38
0.174	0.029992	79	0.968744	0.645026	39
0.087107	0.012293	80+			

جدول رقم (12)

Method	RMSE	MAPE
Entropy	0.307479	1.117784