استعمال بعض الطرائق لتقدير معلمات ودالة البقاء لتوزيع بيركس المعمم مع تطبيق عملي[1]

Using some methods of estimating Parameters and Survival function for The Exponentiated Perks distribution with practical application

أ.م.د مهدي وهاب نعمة نصر الله
Ass.Prof.Mahdi Wahab Nea'ama
Mahdi.na2002@yahoo.com

سلوى نعيم جميل السلمان
Salwa N.ALsalman
salwa.nae166@yahoo.com

كلية الإدارة والإقتصـاد _ جامعة كربـلاء
Economics and Administration College – Karbala University

المستخلص :

جرى في هذا البحث دراسة بعض خصائص توزيع بيركس المعمم ودالة البقاء و استُعمل للتقدير طريقتين هي: طريقة الامكان الاعظم وطريقة المربعات الصغرى و اشتُقت طرائق التقدير هذه للتوصل الى صيغ مقدراتها. واجريت مقارنة بين المقدرات باستعمال اسلوب المحاكاة اذ نفذت تجارب المحاكاة باستعمال مجموعة عينات بأحجام مختلفة(10, 50,100) وكررت كل تجربة(2000) مرة لتحقيق الهدف ولثلاثة نماذج مختلفة وقد قورنت النتائج باستعمال احد اهم المقاييس الاحصائية وهو متوسط مربعات الخطأ التكاملي(IMSE) وجرى التوصل الى افضل الطرائق لكل انموذج ولكل حجم عينة . وجرى معرفة نتائج متوسط مربعات الخطأ التكاملي لدالة البقاء, واجريت المقارنة بين تلك النتائج ولجميع النماذج حيث ظهرت النتيجة بأفضلية طريقة الامكان الاعظم . اما في الجانب التطبيقي فقد اجري تطبيقا عمليا لعينة عشوائية من البيانات بحجم(200) مشاهدة الخاصة بأوقات الوفيات بالشهور لمرضى سرطان الثدي في محافظة البصرة حيث اخذت فترة بقائهم على قيد الحياة , وعن طريقها النتائج التي ظهرت تبين ان طريقة الامكان الاعظم اعطت مقدرات اكثر ملائمة للبيانات المدروسة.

Abstract :

In this research, some characteristics of the Exponiated Perks distribution and survival function were studied and Two methods of estimation were used, namely the Maximum Likelihood method and the least squares method, and these estimation methods were derived to arrive at formulas of their capabilities. A comparison was made between the capabilities using the simulation method, as the simulation experiments were carried out using a set of samples of different sizes (10, 50,100) and each experiment was repeated (2000) once to achieve the goal and for three different models. The results were compared using one of the most important statistical measures, which is the average of the integrative

[1] بحث مستل من رسالة ماجستير

error squares (IMSE). The best methods were found for each model and for each sample size. The results of the mean of squares of the integral error of the survival function were known, and a comparison was made between those results and all models where the result appeared in the best possible way. As for the applied side, a practical application was made of a random sample of data with a size of (200) Watching times of deaths per month for breast cancer patients in Basra Governorate where they took their survival period, and this sample was applied to distribute Exponiated Perks,Through the results, it became clear that the Maximum Likelihood method gave more appropriate capabilities for the data studied .

1-1 المقدمة Introduction

كان الاهتمام المتزايد لكثير من الدراسات بموضوع المعولية نتيجة للدور المهم الذي تلعبه في التعامل مع الزمن او بمعنى ادق مع الاعمار سواء للمعدات او الكائنات الحية كاحتمالات البقاء ومتوسط الحياة فكل من نظرية المعولية " Reliability Theory " ونظرية البقاء" Survival Theory " تشتركان في قياس طول الحياة, اذ الاولى للمعدات و المكائن والثانية للكائنات الحية " سواء كان الانسان ام الحيوان ام النبات ". اذ يهتم علم الاحصاء بدراسة الظواهر التي تتبع السلوك العشوائي وماتؤول اليه تلك الظواهر في المستقبل, وغالبا ماتكون تلك الظواهر ذات اهمية في حياة الفرد والمجتمع, لذلك لابد لنا من معرفة تلك الظواهر اي التوزيعات تسلكها لكي تُفسِر السلوك العشوائي الذي تسلكه تلك الظواهر، وتكون دراستنا للظواهر اما عن طريق توزيعات جاهزة او توزيعات مطورة لوصف وقياس ماتؤول اليه تلك الظاهرة من نتائج احتمالية.

1-2 مشكلة البحث Problem of Research

يعتمد تقدير دالة البقاء على البيانات المستعملة في تقدير معلمات التوزيع الاحتمالي. ومن تلك البيانات هي بيانات الاورام السرطانية التي اخذت بالانتشار بشكل كبير في الوقت الحالي وخصوصاً في محافظة البصرة , حيث لوحظ تسجيل اعداد كبيرة من الاصابات بسرطان الثدي , فكان لابد من دراسة سلوك هذا المرض والبحث عن توزيع احتمالي يصف احتمالات البقاء للمصابين به والسعي الى استعمال طرائق مختلفة لتقدير معلمات التوزيع الاحتمالي للبيانات ومن ثمّ التوصل الى تقديرات دقيقة لدالة البقاء للمصابين للتوصل الى نتائج تفيد المخططين واصحاب القرار في الحد والسيطرة على هذا المرض.

1-3 هدف البحث Aim of Research

إنَّ هدف البحث هو دراسة توزيع بيركس المعمم وبعض خصائصه وتقدير معلمات التوزيع عن طريق محاكاة مونت-كارلو باستعمال الطريقتين الكلاسيكيتين : طريقة الامكان الاعظم وطريقة المربعات الصغرى, ومن ثم استعمال تلك التقديرات في تقدير دالة البقاء للتوزيع والمقارنة بين الطرائق باستعمال المعيار الاحصائي متوسط مربعات الخطأ التكاملي (IMSE) للتوصل الى افضلها. وتطبيق افضل الطرائق التي جرى التوصل اليها على بيانات حقيقية متمثلة بمدة بقاء المصابين بسرطان الثدي لتقدير احتمالات بقاء المصابين بهذا المرض.

الطرائق الاحصائية

1-2 توزيع بيركس المعمم The exponentiated Perks distribution

يعدُّ توزيع بيركس من التوزيعات المستمرة المهمة واحد نماذج الفشل ينسب هذا التوزيع الى العالم الفيزيائي (Perks 1932) و اشتُقت توزيع بيركس من العلاقة بين توزيعين هما Gompertz–Makehm وان توزيع بيركس له استعمالات عدة في الحقول المختلفة منها في دراسات المعولية وكذلك في الدراسات السكانية المتمثلة بتوقعات الحياة في جدول وكذلك في موضوع الرقابة على الجودة وله عدة انواع او صيغ حسب عدد المعالم في التوزيع منها توزيع بيركس ذو معلمه واحدة وتوزيع بيركس ذي المعلمتين

الدالة الاحتمالية لبيركس ذي المعلمتين هي:

$$f(x,\alpha,\lambda) = \alpha\lambda\, e^{\lambda x}\frac{(1+\alpha)}{(1+\alpha e^{\lambda x})^2} \qquad \ldots \quad (1)$$

حيث ان $\alpha > 0\,, \lambda > 0\,, x \geq 0$

ان دالة التوزيع التراكمية هي :

$$F(x,\alpha,\lambda) = 1 - \frac{1+\alpha}{1+\alpha e^{\lambda x}} \qquad \ldots \quad (2)$$

كما ان دالة البقاء هي:

$$s_{(t)} = \frac{1+\alpha}{1+\alpha e^{\lambda x}} \qquad \ldots \quad (3)$$

اما في هذه الرسالة سيكون الاهتمام على توزيع بيركس ذي ثلاث معالم ويدعى توزيع بيركس المعمم The exponentiated Perks distribution

العلاقة بين توزيع بيركس المعمم وبيركس ذي المعلمتين هي إضافة معلمة شكل β الى $F(x)$ بيركس ذي المعلمتين:

إنّ دالة الكثافة الاحتمالية لتوزيع بيركس المعمم تكون كما يأتي

$$f(x,\alpha,\lambda,\beta) = \alpha\lambda\beta\ \frac{(1+\alpha)e^{\lambda x}}{(1+\alpha e^{\lambda x})^2}\left[1 - \frac{1+\alpha}{1+\alpha e^{\lambda x}}\right]^{\beta-1} \qquad \ldots \quad (4)$$

حيث ان $(\alpha,\lambda,\beta) > 0$ and $x \geq 0$

λ معلمة القياس Scale parameter

α معلمة شكل Shape Parameter

β معلمة الشكل Shape Parameter

كما ان الدالة التراكمية لتوزيع بيركس المعمم تكون كالاتي

$$F(x,\alpha,\lambda) = \left(1 - \frac{1+\alpha}{1+\alpha e^{\lambda x}}\right)^{\beta} \quad . \quad . \quad . \quad (5)$$

اما دالة البقاء لتوزيع بيركس المعمم فتعطى بالشكل الاتي :

$$s_{(t)} = 1 - \left[1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x}}\right)\right]^{\beta} \quad . \quad . \quad . \quad (6)$$

ودالة المخاطرة لهذا التوزيع ستكون

$$h_{(t)} = \frac{f_{((t,\alpha,\lambda,\beta))}}{s_{(t)}}$$

$$h_{(t)} = \frac{\alpha\beta\lambda e^{\lambda x}(1+\alpha)\left[1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x}}\right)\right]^{\beta-1}}{(1+\alpha\, e^{\lambda x})^2\left(1 - \left[1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x}}\right)\right]^{\beta}\right)} \quad . \quad . \quad . \quad (7)$$

$$h_{(t)} = \frac{\beta\,\lambda\ \alpha^{\beta}(1+\alpha)e^{\lambda x}\left[e^{\lambda x}-1\right]^{\beta-1}}{(1+\alpha\, e^{\lambda x})[(1+\alpha\, e^{\lambda x})^{\beta} - \alpha^{\beta}(e^{\lambda x}-1)^{\beta}]}$$

2-1-1 الوسط الحسابي والتباين والعزم الرائي

الوسط الحسابي لتوزيع بيركس المعمم ينتج منايجاد التكامل للصيغة الاتية:

$$\mu(x) = E(x) = \int_0^{\infty} x\alpha\beta\lambda \frac{(1+\alpha)e^{\lambda x}}{(1-\alpha e^{\lambda x})^2}\left(1 - \left(\frac{1+\alpha}{1+\alpha e^{\lambda x}}\right)\right)^{\beta-1} dx \quad . \quad . \quad . \quad (8)$$

إنَّ نتيجة التكامل انف الذكر هي (∞) اي ان توزيع بيركس المعمم من التوزيعات التي ليس لها وسط حسابي ,كذلك فان $\left(E(x^2)\right)$ هو (∞) اي ان هذا التوزيع ليس لديه تباين, وعموما ان توزيع بيركس المعمم من التوزيعات التي لاتمتلك عزم.

2-1-2 الربيعيات Quantil

تمثل الربيعيات احدى المقاييس الاحصائية والتي تستخدم لوصف البيانات حيث تقسم الربيعيات مشاهدات العينة الى اربعة اقسام ،ويمكن الحصول على من المعادلة الاتية :

$$Q = \int_0^{\infty} \alpha\beta\lambda \frac{(1+\alpha)e^{\lambda x}}{(1-\alpha e^{\lambda x})^2}\left(1 - \left(\frac{1+\alpha}{1+\alpha e^{\lambda x}}\right)\right)^{\beta-1} dx \quad . \quad . \quad . \quad (9)$$

إن الربيعي الاول يمثل قيمة المتغير العشوائي(x) عند مساواة المعادلة (9) مع (0.25), و الوسيط لتوزيع بيركس المعمم هو قيمة(x) عند مساواة المعادلة(9) مع $(x = .5)$اما الربيعي الثالث لتوزيع بيركس المعمم فهو قيمة(x) عند مساواة المعادلة(9) مع $(x = .75)$

2-1-3 الاحصاءات المرتبةOrder Statistics

تأثر الاحصاءات المرتبة تأثير كبير في دراسة دالة البقاء(survival) ودراسة المعولية(reliability) لوصف دالة البقاء او العمر بالنسبة للأنظمة المتوالية والمتوازية ونظام (K OFn) للحصول على اكبر او اصغر احصاءة ،فاذا توفرت عينة من المشاهدات تتوزع حسب توزيع بيركس المعمم وتحقق شرط (iid) فان دالة الكثافة الاحتمالية للاحصاءات المرتبة (x_k) يمكن الحصول عليها من العلاقة الاتية:

$$g_k(x_{(k)}) = \frac{n!}{(k-1)!\,(n-k)!}\left(F(x_{(x)})\right)^{k-1}\left(1-F(x_{(k)})\right)^{n-k} f(x_{(k)}) \qquad \dots(10)$$

2- 2طرائق التقدير Estimation of Methods

2-2-1 طريقة الامكان الاعظم Maximum Likelihood Method (MLE)

تهدف هذه الطريقة الى جعل دالة الإمكان للمتغيرات العشوائية اعظم ما يمكن وان هذه الطريقة تستخدم غالبا لتقدير معالم التوزيع وذلك لأنها تمتلك خواص جيدة منها انها غير متحيزة وتكون كافية وتملك اقل تباين ممكن واهم خاصية عدم التغير(Invariant) او الثبات وتكون اكثر دقة من طرق التقدير الأخرى خصوصا عند زيادة حجم العينة و اول من صاغ هذه الطريقة هو العالم C.F.Gauss وقام بتطبيقها لأول مرة الباحث R.A.Fisher , وخطوات هذه الطريقة هي كالاتي :

$$\mathbf{f}_{(\underline{x},\alpha,\lambda,\beta)} = \alpha^n\lambda^n\beta^n(1+\alpha)^n e^{\lambda\sum_{i=1}^n x_i}\prod_{i=1}^n\frac{1}{(1+\alpha\, e^{\lambda x_i})^2}\prod_{i=1}^n\left[1-\left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right]^{\beta-1} \quad \dots \quad (11)$$

بأخذ اللوغاريتم الطبيعي لطرفي المعادلة (11) نحصل على :

$$\mathbf{Ln}\, f_{(\underline{x},\alpha,\lambda,\beta)} = n\ln\alpha + n\ln\lambda + n\ln\beta + n\ln(1+\alpha) + \lambda\sum_{i=1}^n x_i - 2\sum_{i=1}^n \ln\left(1+\alpha\, e^{\lambda x_i}\right) + (\beta-1)\sum_{i=1}^n \ln\left[1-\left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right] \quad \dots \quad (12)$$

وبإيجاد المشتقة الجزئية بالنسبة للمعلمات (α,λ,β)للمعادلة (12)نحصل على:

$$\frac{d\ln L}{d\alpha} = \frac{n}{\alpha} + \frac{n}{1+\alpha} - 2\sum_{i=1}^n \frac{e^{\lambda x_i}}{(1+\alpha\, e^{\lambda x_i})} + (\beta-1)\sum_{i=1}^n \frac{1}{\alpha(1+\alpha e^{\lambda x_i})} \quad \dots (13)$$

$$\frac{d\ln L}{d\beta} = \frac{n}{\beta} + \sum_{i=1}^n ln\left(\frac{\alpha(-1+e^{\lambda x_i})}{(1+\alpha e^{\lambda x_i})}\right) \quad \dots \quad (14)$$

$$\frac{d\ln L}{d\lambda} = \frac{n}{\lambda} + \sum_{i=1}^n x_i - 2\sum_{i=1}^n \frac{\alpha\, x_i e^{\lambda x_i}}{(1+\alpha\, e^{\lambda x_i})} + (\beta-1)\sum_{i=1}^n \frac{(1+\alpha)\, x_i e^{\lambda x_i}}{(e^{\lambda x_i}-1)(1+\alpha e^{\lambda x_i})} \quad \dots (15)$$

وللحصول على تقدير للمعلمات المجهول يجري مساواة المشتقة انفة الذكر بالصفر. حيث ينتج منظومة من المعادلات لا خطية اذ لا يمكن استعمال الطرائق الاعتيادية في التقدير ولابد من استعمال الطرائق العددية لحلها.وان تقدير دالة البقاء سيكون :

$$\hat{s}_{(t)} = 1 - \left[1 - \left(\frac{1+\hat{\alpha}}{1+\hat{\alpha}e^{\hat{\lambda}x}}\right)\right]^{\hat{\beta}} \quad . \ . \ . \quad (16)$$

2-2-2 طريقة المربعات الصغرى "L.S" Least Squares Method

تعد طريقة المربعات الصغرى الاعتيادية من الطرائق المهمة في عملية التقدير اذ تتميز بخصائص عده مما يجعلها من أفضل الطرائق واوسعها استعمالا وتستند هذه الطريقة على تصغير مجموع مربعات الاخطاء , ويمكن صياغتها بالشكل الاتي :

$$L.S = \sum_{j=1}^{n}\left[F_{(x)} - \left(\frac{j}{n+1}\right)\right]^2 \quad . \ . \ . \quad (17)$$

علما ان $\left(\frac{i}{n+1}\right)$ هو مقدر لا معلمي وهو تقدير للدالة التراكمية لتوزيع الاحتمالي بيركس المعمم:

$$L.S = \sum_{i=1}^{n}\left\{\left[1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right]^{\beta} - \left(\frac{i}{n+1}\right)\right\}^2 \quad . \ . \ . \quad (18)$$

وللحصول على مقدرات طريقة المربعات الصغرى لتوزيع بيركس المعمم يتم بايجاد المشتقة الاولى للمعادلة (18) وللمعلمات $(\alpha;\beta;\lambda)$ نحصل على الاتي:

$$\frac{d\,L.S}{d\,\alpha} = 2\sum_{i=1}^{n}\left(\left(1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right)^{\beta} - \left(\frac{i}{n+1}\right)\right)\left(\frac{\beta\alpha^{\beta-1}\left(e^{\lambda x_i}-1\right)^{\beta}}{(1+\alpha\, e^{\lambda x_i})^{\beta-1}}\right) \quad . \ . \ . \quad (19)$$

$$\frac{d\,L.S}{d\,\beta} = 2\sum_{i=1}^{n}\left\{\left(1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right)^{\beta} - \left(\frac{i}{n+1}\right)\right\}\left(1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right)^{\beta} \quad ln\left(1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right) \quad . \ . \ . \quad (20)$$

$$\frac{d\,L.S}{d\,\lambda} = 2\sum_{i=1}^{n}\left(\left(1 - \left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)\right)^{\beta} - \left(\frac{i}{n+1}\right)\right)\left(\frac{\beta\alpha^{\beta}(1+\alpha)x_i e^{\lambda x_i}}{\left(\frac{1+\alpha}{1+\alpha\, e^{\lambda x_i}}\right)^{\beta+1}}\right) \quad . \ . \ . \quad (21)$$

وبمساواة المعادلات انفة الذكر ينتج منظومة من المعادلات غير الخطية لا يمكن حلها بالطرائق الرياضية الاعتيادية حيث سيجري اللجوء الى الطرائق العددية لحلها وان تقدير دالة البقاء سيكون كما مشار اليه في المعادلة (16)

الجانب التجريبي

3-1مفهوم المحاكاة Simulation

تُعّرف المحاكاة بأنها تقنية عددية تستخدم الحاسب الالكتروني لتكوين تجارب بوجود أنموذج رياضي أو منطقي لوصف سلوك نظام ما . هذا في الاطار العام اما في الاطار الضيق تسمى المحاكاة العشوائية أو محاكاة مونت كارلو. وتُعرّف بأنها تقنية عددية تستخدم أنموذجاً رياضياً لتوليد الارقام العشوائية المطلوبة من توزيع احتمالي معين. جرى توليد البيانات العشوائية بما يتلائم والتوزيع الاحصائي بيركس المعمم لغرض المقارنة والمفاضلة بين طرائق التقدير المذكورة , وجرى تنفيذ تجارب المحاكاة لتوليد عدد من العينات جرى توليد متغيرات تتبع توزيع منتظم $u\sim U(0,1)$بالاستعانة بالايعاز (Rand) وجرى توليد بيانات تتبع توزيع بيركس المعمم باستعمال الصيغة

$$x=\frac{1}{\lambda}\log\left(\frac{1}{\alpha}\left(\frac{1+\alpha}{1-u^{1/\beta}}-1\right)\right) \qquad \ldots (22)$$

جرى تنفيذ تجارب المحاكاة لتوليد عدد من العينات بالحجوم(10،100،50) و بتكرار(2000) مرة لكل تجربة محاكاة لغرض الحصول على اكبر تجانس(Homogenous) .

3-2 المقارنة بين طرائق التقدير

بعد ان جرى تقدير(Estimate) معلمات توزيع بيركس المعمم تأتي مرحلة المقارنة بين هذه التقديرات ، سيجري استعمال معيار متوسط الخطأ التكاملي(IMSE) (Integral Mean Square Error) يمثل تكامل للمساحة الكلية (t_i) واختزالها بقيمة واحدة وتعد عامة للزمن, أو معبرة عن الزمن الكلي. وصيغتة كما يأتي:

$$IMSE[\hat{R}(t)]=\frac{1}{r}\sum_{i=1}^{r}\left\{\frac{1}{n_t}\sum_{j=1}^{n_t}[\hat{R}_i(t_j)-R(t_j)]^2\right\} \qquad \ldots (23)$$

$$=\frac{1}{n_t}\sum_{j=1}^{n_t}MSE[\hat{R}(t_i)]\ , \qquad \ldots (24)$$

$$i=1,\ldots..,r$$

اذ ان R(Replication) تمثل عدد التكرارات لكل تجربة و الذي كان مساوي الى (2000)

n_t: معبرة عن حدود المتغير (t_i) أي من الحد الأدنى (Lower Bound) إلى الحد الأعلى (Upper Bound) كما اختيرت عدد من القيم الافتراضية لمعلمات دالة الكثافة الاحتمالية للتوزيع وكما هوموضح في الجدول (1).

جدول رقم (1) القيم الافتراضية لتجارب المحاكاة

Model	α	β	λ
1	2	1	10
2	10	4	3
3	16	1	20

3-3 تحليل تجارب المحاكاة experiments Analysis of simulation

لغرض التوصل الى افضل الطرائق دقة في تقدير دالة البقاء للتوزيع, تم عرض نتائج تجارب المحاكاة كما في الجداول الاتية:

جدول رقم (2) تقديرات دالة البقاء عند طرائق التقدير ومتوسط مربعات الخطأ التكاملي IMSE عند كل احجام العينات المفترضة للأنموذج الأول

N	*Real*	ML	LS
		α=2.37	α=2.31
		β=2.43	β=1.36
		λ=10.78	λ=10.65
10	0.98744	0.98823	0.99185
	0.93644	0.94042	0.94115
	0.86683	0.87639	0.85302
	0.78366	0.80405	0.73834
	0.68783	0.72820	0.61014
	0.58097	0.65219	0.48159
	0.46970	0.57837	0.36374
	0.36636	0.50838	0.26383
	0.28250	0.44326	0.18476
	0.22080	0.38364	0.12580
	IMSE	0.11596	0.13432
N	*Real*	ML	LS
		α=2.24	α=2.27
		β=2.22	β=1.11
		λ=10.66	λ=10.34
50	0.98744	0.99323	0.98312
	0.93644	0.94309	0.91974

	0.86683	0.85158	0.83651
	0.78366	0.73130	0.74404
	0.68783	0.59720	0.64940
	0.58097	0.46381	0.55764
	0.46970	0.34268	0.47221
	0.36636	0.24104	0.39521
	0.28250	0.16160	0.32761
	0.22080	0.10343	0.26952
	IMSE	0.01086	0.0331
N	*Real*	ML	LS
		α=2.10	α=2.23
		β=1.21	β=1.14
		λ=10.22	λ=10.30
100	0.98744	0.98386	0.98392
	0.93644	0.92075	0.92090
	0.86683	0.83778	0.83798
	0.78366	0.74589	0.74612
	0.68783	0.65201	0.65225
	0.58097	0.56088	0.56111
	0.46970	0.47571	0.47590
	0.36636	0.39843	0.39856
	0.28250	0.33001	0.33007
	0.22080	0.27068	0.27065
	IMSE	0.001001	0.001145

جدول رقم (3) تقديرات دالة البقاء عند طرائق التقدير ومتوسط مربعات الخطأ التكاملي IMSE عند كل احجام العينات المفترضة للأنموذج الثاني

N	*Real*	ML	LS
		α=10.85	α=10.28
		β=4.41	β=4.55
		λ=3.46	λ=3.32
10	0.99203	0.97241	0.97084
	0.95123	0.94815	0.94794
	0.87985	0.85432	0.85420
	0.78506	0.71626	0.71718
	0.67570	0.59844	0.60108
	0.59930	0.53776	0.54182
	0.52309	0.48950	0.49473
	0.44933	0.45093	0.45690
	0.37984	0.41967	0.42599
	0.34715	0.40620	0.41259
	IMSE	0.07944	0.08577
N	*Real*	ML	LS
		α=10.22	α=10.09
		β=4.50	β=4.20
		λ=3.33	λ=3.08
50	0.99203	0.99044	0.99208
	0.95123	0.95048	0.95185
	0.87985	0.87650	0.88162
	0.78506	0.77287	0.78835
	0.67570	0.65437	0.68065
	0.59930	0.57642	0.60529

	0.52309	0.50434	0.53002
	0.44933	0.44033	0.45701
	0.37984	0.38494	0.38808
	0.34715	0.36036	0.35559
	IMSE	0.02588	0.00087
N	*Real*	ML	LS
		α=10.07	α=10.38
		β=4.09	β=4.23
		λ=3.11	λ=3.12
100	0.99203	0.99202	0.99974
	0.95123	0.95136	0.99318
	0.87985	0.88035	0.96437
	0.78506	0.78611	0.89375
	0.67570	0.67742	0.76775
	0.59930	0.60145	0.65419
	0.52309	0.52566	0.52508
	0.44933	0.45226	0.39239
	0.37984	0.38307	0.26966
	0.34715	0.35049	0.21578
	IMSE	0.00042	0.00689

جدول رقم (4) تقديرات دالة البقاء عند طرائق التقدير و متوسط مربعات الخطأ التكاملي IMSE عند كل احجام العينات المفترضة للأنموذج الثالث

N	*Real*	ML	LS
		α=16.54	α=16.40
		β=1.36	β=1.42
		λ=20.63	λ=20.52
10	0.99150	0.98632	0.97218

	0.94811	0.92189	0.94826
	0.87266	0.81707	0.86333
	0.77336	0.68882	0.73307
	0.66016	0.55427	0.61519
	0.58197	0.46818	0.55312
	0.50474	0.38854	0.50285
	0.43076	0.31722	0.46206
	0.36178	0.25516	0.42875
	0.32958	0.22766	0.41439
	IMSE	0.07146	0.02442
N	*Real*	ML	LS
		α=16.21	α=16.23
		β=1.19	β=1.32
		λ=20.20	λ=20.32
50		0.99150	0.99147	0.99973
		0.94811	0.94814	0.99313
		0.87266	0.87275	0.96405
		0.77336	0.77331	0.89276
		0.66016	0.65966	0.76575
		0.58197	0.58100	0.65155
		0.50474	0.50324	0.52203
		0.43076	0.42872	0.38948
		0.36178	0.35927	0.26734
		0.32958	0.32687	0.21388
		IMSE	0.00127	0.00745
N	*Real*	ML	LS

		α=16.19	α=16.22
		β=1.08	β=1.24
		λ=20.09	λ=20.30
100	0.99150	0.99137	0.99142
	0.94811	0.94738	0.95110
	0.87266	0.87087	0.87838
	0.77336	0.77008	0.77896
	0.66016	0.65511	0.66417
	0.58197	0.57569	0.58592
	0.50474	0.49733	0.51062
	0.43076	0.42239	0.44091
	0.36178	0.35272	0.37834
	0.32958	0.32029	0.34994
	IMSE	0.00102	0.01396

يتضح من نتائج المحاكاة للنماذج الثلاثة ان طريقة الامكان الاعظم كانت افضل من طريقة المربعات الصغرى و لجميع احجام العينات الكبيرة والصغيرة.

وجدول (5) يبين عدد مرات افضلية طرائق التقدير وعند كل حجم من احجام العينات المفترضة في كل تجارب المحاكاة .

جدول رقم (5) عدد مرات افضلية طرائق التقدير عند كل حجم من احجام العينات المفترضة في كل تجارب المحاكاة

الطريقة	حجم العينة			عدد مرات الأفضلية
	$n_1 = 10$	$n_2 = 50$	$n_3 = 100$	
ML	2	2	3	7
LS	1	1	0	2

الجانب التطبيقي

4-1 وصف البيانات Description of the data

جرى اخذ البيانات الخاصة بالدراسة من المركز التخصصي للاورام وامراض الدم التابع لمستشفى الصدر التعليمي في محافظة البصرة وجرت عملية جمع البيانات الخاصة بوفيات مرض سرطان الثدي اذ جرى جمع 200 مشاهدة لعدد الوفيات بسبب هذا المرض من سجلات القسم وجرى تبويب هذه البيانات للحصول على أوقات الحياة للمصابين بهذا المرض وذلك عن طريق طرح تاريخ الإصابة من تاريخ الوفاة . والجدول الاتي يمثل اوقات الحياة للمصابين بسرطان الثدي .

جدول (6) يمثل اوقات الحياة لمرضى سرطان الثدي بالشهور

27.67361	41.80321	13.98888	15.74118	17.52831
33.43643	44.43346	24.0036	28.82578	10.94639
16.73129	25.79361	40.14268	28.32003	24.79808
45.67381	21.01278	33.9471	54.37103	41.31727
14.37032	29.99307	31.33361	18.03201	22.52475
20.97183	45.39975	27.13051	30.70509	32.31069
27.21915	30.73538	17.36873	18.91115	34.05623
18.76197	15.5715	21.83442	20.67509	43.17171
21.58317	64.47916	26.38144	14.40276	33.43204
43.59613	22.58518	27.50706	35.46944	30.97481
28.13736	21.29798	29.64965	21.71226	31.98895
39.77207	24.6079	26.53215	35.96631	33.82708
27.78258	21.79262	25.23172	30.00842	42.73277
39.2202	20.92807	27.53864	14.0523	46.43188
21.2578	24.21202	19.17518	43.7473	25.85252
17.95303	19.65168	21.35194	24.34303	23.83226
31.0714	35.22171	20.20719	25.23825	46.06243
17.68173	29.30978	32.7188	45.30921	49.54992
22.35576	23.02155	28.38207	26.20349	16.21248
17.36769	43.23819	17.47152	52.56529	23.62316
22.37284	42.68294	23.7518	20.25372	12.52384
44.00892	25.70927	41.52993	35.95247	24.8371
27.55426	16.29816	17.94704	38.56468	29.88207
21.87907	31.00705	40.51121	35.37063	18.87255
23.63873	38.56759	33.22711	24.62181	
27.01929	43.18517	19.47413	48.56043	
36.23938	30.29664	41.45411	28.20355	
21.69199	25.48724	16.57269	28.36965	
30.25591	11.14641	32.82181	22.61264	
70.12289	53.20924	31.74407	34.4685	
12.4928	27.96514	32.32969	18.03053	
26.25786	23.60049	61.29501	19.41016	
27.72515	29.52663	15.83818	63.73456	
41.04673	42.11701	30.3385	16.15579	

19.0197	26.136	26.90965	34.83015	
19.93766	32.52878	22.79154	25.66051	
45.937	44.25927	18.86968	18.7393	
35.89904	17.80985	49.83939	15.32794	
19.25358	20.89739	50.84991	52.08334	
46.82634	26.94311	40.79788	21.01671	
16.06956	36.53221	34.50587	24.63528	
24.48828	25.85246	24.47428	18.92686	
32.45757	19.52505	30.23361	29.24512	
13.92127	41.47532	33.0695	25.66409	

(4-2) معايير حسن المطابقة Goodness of Fit

اقترحت العديد من المعايير او الاختبارات التي يمكن أن تستخدم لمعرفة افضلية طرائق التقدير المستعملة اذ جرى تطبيق هذه الاختبارات على بيانات جرى الحصول عليها من المركز التخصصي للأورام السرطانية في محافظة البصرة ومن هذه الاختبارات الآتي:

(4-2-1)معياراكايكي(AIC)Akaike's Information Criterion

إنَّ الصيغة العامة لهذا الاختبار كما يأتي:

$$\mathrm{AIC} = -2\mathrm{L}(\hat{\theta} \setminus \mathrm{x}) + 2\mathrm{p} \qquad \dots (25)$$

p: تمثل عدد المعلمات في دالة التوزيع الاحتمالية النظرية

$L(\hat{\theta} \setminus x)$: تمثل لوغاريتم دالة الترجيح(Log Likelihood Function) لمشاهدات بيانات العينة.

(4-2-2)معياربيزاكايكي (BIC) Bayesian Information Criterion

معيار يستخدم كاختبار حسن المطابقة (GOF) ويرمز له اختصارا ب (BIC)

ان الصيغة لهذا الاختبار هي:

$$\mathrm{BIC} = -2\mathrm{L}(\hat{\theta} \setminus \mathrm{x}) + \mathrm{plog}(\mathrm{n}) \qquad \dots (26)$$

$L(\hat{\theta} \setminus x)$: تمثل لوغاريتم دالة الترجيح(Log Likelihood Function) لمشاهدات بيانات العينة.

p: تمثل عدد المعلمات في دالة التوزيع الاحتمالية النظرية.

n تمثل حجم العينة.

(4-2-3)معياراكايكي المتسقConsistent Akaike Information Criterion(CAIC)

إنَّ الصيغة لاختبار حسن المطابقة اكايكي المتسق (CAIC) هي كما يأتي

$$CAIC = -2L(\hat{\theta} \setminus x) + \frac{2np}{n-p-1} \qquad \dots(27)$$

و ان n تمثل حجم العينة

4-3 تقدير المعلمات Estimate of Parameters

يوضح الجدول (7) قيم المعلمات لتوزيع بيركس المعمم وقد جرى استخدام طريقة الامكان الاعظم وطريقة المربعات الصغرى .

جدول رقم (7)يمثل تقديرات لمعلمات توزيع بيركس المعمم لبيانات دالة البقاء

Method	α	β	λ
Max	2.020392	12.00378	0.150786
LS	2.026702	12.00426	0.06964

وجرى استخدام اختبارات حسن المطابقة (GOF) التي جرى ذكرها في الفقرة (4-2) بهدف ايجاد أكثر افضلية طرائق التقدير لتمثيل دالة البقاء لبيانات المرضى المصابين بسرطان الثدي في محافظة البصرة ، وقد ادرجت النتائج في جدول يمثل نتائج الاختبارات وكما مبين في الجدول الاتي:

جدول رقم (8) اختبارات حسن المطابقة لدالة البقاء لمرضى سرطان الثدي

Method	AIC	BIC	CAIC
Max	1547.068	1529.173	18.00627
LS	4081.557	4063.662	19.95692

يوضح جدول (8) قيم اختبارات حسن المطابقة (GOF)لبيانات دالة البقاء للمرضى المصابين بأمراض سرطان الثدي في محافظة البصرة حيث نلاحظ ان طريقة الامكان الاعظم سجلت اقل قيمة حسابية عند جميع الاختبارات أي ان طريقة الإمكان الأعظم تعطي مقدرات أكثر ملائمة لبيانات دالة البقاء وهذا يتناسب مع ماتم التوصل اليه في الجانب التجريبي.

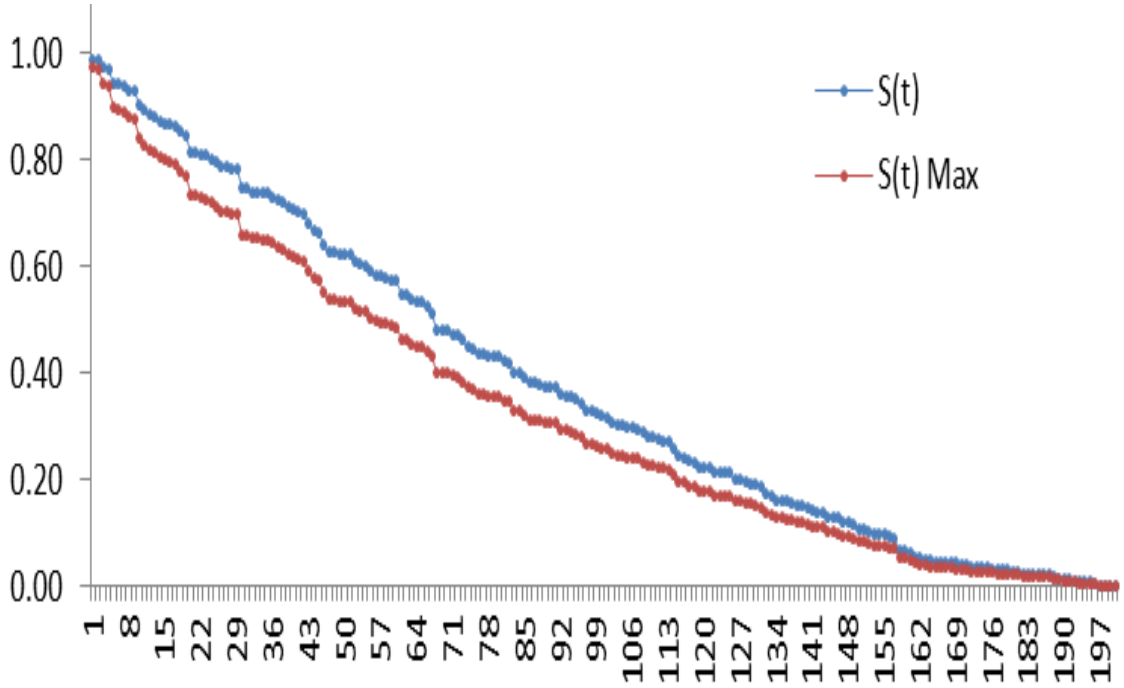
بينت نتائج تجارب المحاكاة ان افضل طريقة لتقدير دالة لبقاء لتوزيع بيركس المعمم هي الامكان الاعظم , لذلك سيجري تطبيق هذه الطريقة على البيانات الحقيقية لقياس احتمالات البقاء (دالة البقاء) للمصابين بسرطان الثدي , وباستعمال برنامج Matlab جرى انشاء برنامج خاص لتطبيق الطريقة وكانت النتائج كما في جدول (9) :

جدول (9) دالة البقاء الحقيقية ودالة البقاء المقدرة بطريقة الامكان الاعظم لمرضى سرطان الثدي

T	Real S(t)	MLE S(t)	T	Real S(t)	MLE S(t)	T	Real S(t)	MLE S(t)	T	Real S(t)	MLE S(t)	T	Real S(t)	MLE S(t)
10.95	0.99	0.97	18.93	0.74	0.65	23.62	0.48	0.40	27.54	0.30	0.25	32.53	0.16	0.12
11.15	0.99	0.97	19.02	0.73	0.64	23.64	0.48	0.40	27.55	0.30	0.25	32.72	0.15	0.12
12.49	0.97	0.94	19.18	0.72	0.64	23.75	0.47	0.40	27.67	0.30	0.24	32.82	0.15	0.12
12.52	0.97	0.94	19.25	0.72	0.63	23.83	0.47	0.39	27.73	0.30	0.24	33.07	0.15	0.12
13.92	0.94	0.90	19.41	0.71	0.62	24.00	0.46	0.38	27.78	0.30	0.24	33.23	0.14	0.11
13.99	0.94	0.89	19.47	0.71	0.62	24.21	0.45	0.37	27.97	0.29	0.23	33.43	0.14	0.11
14.05	0.94	0.89	19.53	0.70	0.62	24.34	0.44	0.37	28.14	0.28	0.23	33.44	0.14	0.11
14.37	0.93	0.88	19.65	0.70	0.61	24.47	0.44	0.36	28.20	0.28	0.23	33.83	0.13	0.10
14.40	0.93	0.88	19.94	0.68	0.59	24.49	0.44	0.36	28.32	0.28	0.22	33.95	0.13	0.10
15.33	0.90	0.84	20.21	0.67	0.58	24.61	0.43	0.36	28.37	0.27	0.22	34.06	0.13	0.10
15.57	0.89	0.83	20.25	0.66	0.58	24.62	0.43	0.36	28.38	0.27	0.22	34.47	0.12	0.09
15.74	0.89	0.82	20.68	0.64	0.55	24.64	0.43	0.36	28.83	0.26	0.21	34.51	0.12	0.09
15.84	0.88	0.82	20.90	0.63	0.54	24.80	0.42	0.35	29.25	0.24	0.20	34.83	0.11	0.09
16.07	0.87	0.80	20.93	0.63	0.54	24.84	0.42	0.35	29.31	0.24	0.19	35.22	0.11	0.08
16.16	0.87	0.80	20.97	0.62	0.54	25.23	0.40	0.33	29.53	0.24	0.19	35.37	0.11	0.08
16.21	0.87	0.80	21.01	0.62	0.53	25.24	0.40	0.33	29.65	0.23	0.19	35.47	0.10	0.08
16.30	0.86	0.79	21.02	0.62	0.53	25.49	0.39	0.32	29.88	0.23	0.18	35.90	0.10	0.08
16.57	0.85	0.78	21.26	0.61	0.52	25.66	0.38	0.31	29.99	0.22	0.18	35.95	0.10	0.08
16.73	0.85	0.77	21.30	0.61	0.52	25.66	0.38	0.31	30.01	0.22	0.18	35.97	0.10	0.08
17.37	0.82	0.74	21.35	0.60	0.52	25.71	0.38	0.31	30.23	0.21	0.17	36.24	0.09	0.07
17.37	0.82	0.74	21.58	0.59	0.50	25.79	0.38	0.31	30.26	0.21	0.17	36.53	0.09	0.07
17.47	0.81	0.73	21.69	0.58	0.50	25.85	0.37	0.31	30.30	0.21	0.17	38.56	0.07	0.05
17.53	0.81	0.73	21.71	0.58	0.50	25.85	0.37	0.31	30.34	0.21	0.17	38.57	0.07	0.05
17.68	0.80	0.72	21.79	0.58	0.49	26.14	0.36	0.30	30.71	0.20	0.16	39.22	0.06	0.05
17.81	0.79	0.71	21.83	0.58	0.49	26.20	0.36	0.29	30.74	0.20	0.16	39.77	0.06	0.04
17.95	0.79	0.70	21.88	0.57	0.49	26.26	0.36	0.29	30.97	0.19	0.16	40.14	0.05	0.04
17.95	0.79	0.70	22.36	0.55	0.46	26.38	0.35	0.29	31.01	0.19	0.15	40.51	0.05	0.04
18.03	0.78	0.70	22.37	0.55	0.46	26.53	0.34	0.28	31.07	0.19	0.15	40.80	0.05	0.04
18.03	0.78	0.70	22.52	0.54	0.45	26.91	0.33	0.27	31.33	0.19	0.15	41.05	0.05	0.04
18.74	0.75	0.66	22.59	0.54	0.45	26.94	0.33	0.27	31.74	0.18	0.14	41.32	0.04	0.03
18.76	0.75	0.66	22.61	0.53	0.45	27.02	0.32	0.26	31.99	0.17	0.13	41.45	0.04	0.03
18.87	0.74	0.65	22.79	0.52	0.44	27.13	0.32	0.26	32.31	0.16	0.13	41.48	0.04	0.03
18.87	0.74	0.65	23.02	0.51	0.43	27.22	0.32	0.26	32.33	0.16	0.13	41.53	0.04	0.03
18.91	0.74	0.65	23.60	0.48	0.40	27.51	0.31	0.25	32.46	0.16	0.13	41.80	0.04	0.03

نلاحظ من جدول (9) تقارب قيم دالة البقاء المقدرة بطريقة الامكان الاعظم مع قيم دالة البقاء الحقيقية (عند البيانات التطبيقية) , ونلاحظ ايضاً ان مدة الاصابة متزايدة مع الزمن اي كلما زادت مدة الاصابة زاد احتمال موت المريض وان

دالة البقاء متناقصة مع الزمن فاذا كانت مدة بقاء المريض على قيد الحياة (50) شهرا فأكثر فان احتمالية بقاءه على قيد الحياة تقترب من (الصفر) واذا كان مدة بقاء المريض المصاب بسرطان الثدي (16) شهرا فأقل فان احتمالية بقاءه على قيد الحياة تزيد عن (78%) وكما مبين في شكل (1)



شكل(1) منحنى دالة البقاء الحقيقية والمقدرة بطريقة الامكان الاعظم عند البيانات التطبيقية

الاستنتاجات والتوصيات

الاستنتاجات Conclusions

بعرض نتائج البحث أمكننا الوصول إلى جملة من الاستنتاجات وأهمها :

1- أظهرت النتائج بان طريقة الامكان الاعظم أفضل من طريقة المربعات الصغرى لأنها حققت اقل متوسط مربعات الخطأ التكاملي IMSE ولجميع النماذج و احجام العينات .

2- قيم متوسط مربعات الخطأ (IMSE) تتناقص بزيادة حجم العينة ولجميع المقدرات ، وهذا ما ينسجم مع النظرية الإحصائية .

3- عن طريق نتائج التطبيق العملي تبين ان دالة البقاء متناقصة مع الزمن ونستنتج من ذلك انه كلما زاد وقت الاصابة زاد احتمال موت المصاب بهذا المرض

التوصيات : Recommendations

في ضوء ما توصل اليه الباحث في هذا البحث من استنتاجات يوصي بالاتي:

1- اجراء بحوث ودراسات مستقبلية في تقدير دالة البقاء لتوزيع بيركس المعمم في حالة البيانات المراقبة.

2- يرجى الاخذ بنظر الاعتبار من قبل المؤسسة الصحية نسبة انخفاض متوسط حجم Hb للمرضى بمرور الزمن نتيجة اعطاء الجرع وبفروق معنوية ولذا يتطلب الامر القيام بدراسات اخرى وعلى ادوية اخرى وبجرعات مختلفة حتى يؤدي الى تقليل المضاعفات للمرضى.

المصادر References

1. تهاني مهدي و سميرة مزهر وقتيبة نبيل.(2008). "مقارنة مقدرات بيز مع مقدر الأمكان الأعظم في تقدير دالة البقاء للتوزيع الطبيعي اللوغاريتمي باستخدام بيانات مراقبة من النوع الثاني". مجلة جامعة النهرين. 1, 11 (2008) , 58 – 70.
2. جاسم , خضر نصيف (2012) " مقارنة تقدير دالة المعولية للتوزيع الاسي المختلط مع تطبيق عملي " اطروحة دكتوراة , كلية الادارة والاقتصاد بغداد .
3. هرمز، أمير حنا، (1990)، "الإحصاء الرياضي"، دار الكتب للطباعة والنشر ـ جامعة الموصل.
4. Al-Fawzan, M., (2000), "Methods for Estimating the Parameters of the Weibull Distribution", Email. mfawzan@kacst.edu.sa, kingAbdulaziz city for science and Technology .
5. Al-Sutany ,Shurooq A.and Mohammed , Sahar A. (2018)"Comparison between Bayesian and Maximum Likelihood Methods for parameters and The Reliability function of Perks Distribution "Iraqi Journal of Scicnce Vol.59, No.2 PP: 369-376 .
6. Chaudhary , Arunkumar and Kumar , Vijay (2013) " A Bayesian Analysis of perks Distribution Vie Markov Chain Monte Carlo Simulation " Nepal Journal of Science and Technology Vol. 14 .No.1 pp:153-166 .
7. Chaudhary ,A.K. and Campus, Nepal Commerce (2018) "A study of Perks – II Distribution via Bayesian Paradigm " Pravaha Journal .
8. Chen , Qming and Egan , David Mo (2006) " A Bayesian Method for Transformer Life Estimation Using Perks Hazard Function " IEEE Transaction on power systems Vol.21,No.4, November.
9. Dzages Shireen : ahmed aulay man (2009):"A comparative Study of the immunological effect of two different medical Regimens on the breast cancer " Doctoral thisis (dept of microbiology , the haeler medical university erbil Iraq
10. James Dx and Nancy ED (2006) "estrogen carcinogenesis in Breast cancer "the new England journal of medicine 354:270-282
11. Mood, Alexander &Geraybill , Frawklina (1972) "Introduction to Theory of Statistics ",International Student Edition, Third Edition .
12. Nadaraja,S and Bakar,S.A.A (2013) "Anew R package for actuarial survival models" computational statistics 28,p. 2139-2160 .

13. Nassar,M. and Mode ,N.(2013) "A new Gonzalez on of the pareto –Geometric Distribution " Journal of the Egyption Mathematical Socicty , Vol.21, PP:148-155.

14. Oguntunde,pelumiE.and others (2018)"Introducing the kumaraswamy perks distribution " proceedings of the world congress on engineering and computer science 2018 Vol.II

15. Richards,S.J (2010) "Ahandbook of parametric survival models for actuarial use" CrossRef citations to date Altmetric p.p. 233-257 .

16. Shabri , ANI, (2007)" L.Q. Moments for Statistical analysis of Extreme Events" Journal of modern Applied Statistical Method Vol. 6 , ISSUE 1, article 21 , PP:228-238.

17. Singh,Bhupendra (2016) "An Additive Perks-Weibull Model with Bathtub-Shaped Hazard rate function"

18. Singh,Bhupendra (2015) "Classical and Bayesian estimation using type-II hybrid censored data from perks distribution" Om parkash,pp. 493-510 .

19. Talacko , Joseph (2016) "A note about a family of Perks distribution " spring Indian statistical institute are collaborating with Jstor to digitize , preserve and extend access to sankhya : the Indian Journal of statistics Vol.20 No.3/4 .

20. Teimouri , Mahdi and Gupta ,Arjum,K. (2013) " On the three parameter weibull distribution shape parameter estimation " Journal of data science , Vol.11 P.P : 403-414 .

21. Tyagi,Abhishek ;Choudhary,Neha and Singh,Bhupendra (2019) "Discrete additive perks – weibull distribution :properties and applications" life cycle reliability and safety engineering 8(3), p.p.183-199 .

22. Zeng ,Hongtao.Lan,Tian and Chen,Qiming (2016) "Five and four parameter lifetime distributions for bathtub-shaped failure rate using perks mortality equation" Reliability Engineering and system safety 152,pp 307-315 .