تقدير دالة المخاطرة لتوزيع log-logistic بأستعمال الاحصاءات المرتبة مع تطبيق عملي

الباحثة سماح صباح حسن أ.د. انتصار عريبي فدعم
كلية الادارة والاقتصاد/ جامعة بغداد / قسم الأحصاء

المستخلص :
يعد توزيع Log-Logistic من التوزيعات الاحصائية المهمة اذ يمكن تطبيقه في العديد من التجارب الحياتية والبايلوجية والكيميائية والفيزيائية واهميته تأتي من اهمية تحديد دالة المخاطرة لتلك التجارب فأن البحث سوف يعمل على تحديد الخصائص التوزيع من خلال استعمال الاحصاءات المرتبة لتقدير معلماته بأستعمال طريقة التجزيئية (Q.E) ولتحديد الطريقة الامثل من خلال مقارنتها مع طريقة الامكان الاعظم (MLE) وفق اسلوب محاكاة بأخذ نماذج مختلفة للمعالم الاولية وحجوم عينات مختلفة وبمعايير مقارنة MSE,IMSE وكذلك تطبيقها على بيانات حقيقية لمرضى سرطان الثدي وايضاً تحديد دالة المخاطرة.
المصطلحات الرئيسة في البحث: دالة المخاطرة ,توزيع log-logistic ,طريقة الامكان الاعظم , طريقة التجزيئية ,الاحصاءات المرتبة .

Abstract :
The Log-Logistic distribution is one of the important statistical distributions, as it can be applied in many fields , biological, chemical and physical experiments, and its importance comes from the importance of determining the risk function for these experiments. The research will work to determine the characteristics of the distribution through the use of order statistics To estimate its parameters using the Quantile Method (Q.E) and to determine the optimal method by comparing it with the MLE method., according to a simulation method by taking different models for primary parameters , different sample sizes and MSE, IMSE comparison criteria, as well as applying it to real data for breast cancer patients and determining the risk function.
Keywords : risk function, log-logistic distribution, maximum likelihood method, Quantile method, ordered statistics.

1- المقدمة (Introduction)
تعد الاحصاءات المرتبة (order statistics) أحد أهم الفروع الأحصاء الرياضي حيث تلعب دوراً مهماً وأساسياً في العديد من الفروع العلوم النظرية والتطبيقية وكذلك في المشاكل اللامعلمية من الناحيتين النظرية والعملية. من الناحية النظرية ، تعتمد نظرية الاستدلال الإحصائي على الإحصاء الترتيبي ومن الناحية العملية ، يساعدنا الاستدلال الإحصائي القائم على العينات الترتيبية في التوصل الى نتائج بسيطة وقابلة للتطبيق تعتمد على جودة التوفيق و اختبارات الحياة لتوزيعات القائمة على العينات الترتيبية ، حيث أن هذه التوزيعات أكثر ملائمة لدراسة نظرية الخاصة بالقيم المتطرفة والاحتمالات ونظرية التقدير وأختبارات الفروض [9]Ragab A. and Green J.

تم في هذا البحث تقدير معلمتي توزيع log-logistic الذي يستعمل في دراسات الدخل ،ونظرية البقاء والمخاطرة .ويمكن ان يعرف بواسطة معلمة القياس (scale parameter) لتكن α ومعلمة الشكل (shape parameter) ولتكن β

.وفي هذا البحث أجرت العديد من الدراسات السابقة لهذا التوزيع ،في عام(1995) قام الباحثان(**Guo & Singh**)[10] بدراسة لتقدير معلمات التوزيع log-logistic, لقد أستعملوا مبدأ الانتروبيا العظمى (POME) لتقدير المعلمات لتوزيع log-logistic ذو المعلمتين. في عام (2005) حلل الباحث (**Fitzgerald**)[7] هطول الأمطار الغزيرة باستعمال التوزيع log-logistic، استعملوا أساليب الكمبيوتر المكثفة لفحص قيم t من التوزيع log-logistic إلى الحد الأعظم السنوي لهطول الأمطار الأيرلندية، في الدراسة ، قام بفحص خصائص L- moment باستعمال bootstrap التقليدي في البيانات. في العام (2020) م قام الباحث (**Dali Chen**)[5] واخرون تطبيق نماذج الوقت الحراري على توزيعات مختلفة أظهرت النتائج أن أنموذجًا أكثر مرونة مع التوزيع log-logistic . يهدف هذا البحث الى تقدير دالة المخاطرة بأستعمال طريقة الأمكان الاعظم وبالاضافة الى طريقة التجزيئية بأستعمال أنموذج الاحصاءات المرتبة الأعتيادية .

2- الاحصاءات المرتبة المعممة : (Generalized order statistics)

تسمى المتغيرات العشوائية $U(j,n,\tilde{m},s)$ أذ ان $j=1,2,\dots,n$ بالاحصاءات المرتبة المعممة المنتظمة اذا كانت دالة الكثافة الاحتمال المشتركة بالصيغة الاتية :

$$f^{U(1,n,\tilde{m},s)..U(n,n,\tilde{m},s)}(t_1 \dots . t_n) = C_{n-1}[\prod_{i=1}^{n-1}(1-t_i)^{m_i}](1-t_n)^{s-1}$$

اذ ان :

$$n \in N; n \geq 2; S \geq 1$$
$$\tilde{m} = (m_1, \dots,\ m_{n-1}) \in R^{n-1}$$
$$0 < t_1 \leq \cdots \leq t_n \leq 1$$
$$C_{n-1} = \prod_{i=1}^{n} Y_i = s\prod_{i=1}^{n} Y_i$$
$$Y_i = s + n - j + \sum_{i=j}^{n-1} m_i > 0$$

لجميع قيم $j \in \{1,2,\dots,n-1\}$

يمكن الحصول على الدالة الاحتمالية المشتركة للأحصاءات المرتبة المعممة لعدد n *من المتغيرات العشوائية* $(j,n,\tilde{m},s)$ *أذ ان* $j\in \{1,2,\dots,n\}$ T *التي تتبع أي توزيع احتمالي بدالة كثافة الاحتمالية هي* $f(x)$ *ودالة التوزيع التراكمية هي* $F(x)$ *وتأخذ الصيغة الاتية* أنظر ***kamps(1995 a)***

$$f^{t(1,n,\tilde{m},s)..t(n,n,\tilde{m},s)}(t_1 \dots . t_n) = C_{n-1}\left[\prod_{i=1}^{n-1}(1-F(t_i))^{m_i} f(t_i)\right]\left[\left(1-F(t_n)\right)^{s-1} f(t_n)\right] \dots . (1)$$

اذ ان $F^{-1}(0) < t_1 \leq \cdots \leq t_n < F^{-1}(1)$

يحتوي أنموذج الاحصاءات المرتبة المعممة على العديد من نماذج المتغيرات العشوائية المرتبة ,وبأختيار مناسب للمعلمات $(\tilde{m}, s)$ نحصل على نماذج الاحصاءات المرتبة المعممة .

3- *الاحصاءات المرتبة الاعتيادية* : (*ordinary order statistics*)

يمكن الحصول على الاحصاءات المرتبة الاعتيادية وذلك بوضع $m_i = m = 0$ لقيم *i=1,2,..,n-1* و *S=1* و $Y_i =$ $n-i+1$ في معادلة *(1)* نحصل على دالة الكثافة الاحتمال المشتركة لـ n من الاحصاءات المرتبة الاعتيادية وبالشكل الاتي :

$$f\left(t_{(1)}, t_{(2)}, \dots, t_{(n)}\right) = n! \prod_{i=1}^{n} f(t_i) \qquad (2)$$

أذ أن : $-\infty < t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)} < \infty$

4- التوزيع اللوغارتمي اللوجستي وخصائصه ((Log-Logistic distribution and its properties :

يعد التوزيع log-logistic هو أحد التوزيعات الاحتمالية المستمرة التي حازت باهتمام الباحثين Raghab A. and Green.J (1984) , تم استعماله مع نطاق واسع في أنماذج النموالسكان والمشاكل الحيوية وله بعض التطبيقات المهمة في الحل العديد من المشكلات العملية خاصة في بيانات البقاء على قيد الحياة كنوذج للأحداث كذلك عرف التوزيع log-logistic بـ (توزيع Fisk في الاقتصاد). لنفرض ان $(t_1 \le t_2 \le \cdots \le t_n)$ عينة ترتيبية بحجم n مسحوبة من مجتمع يتبع توزيع log-logistic دالته الاحتمالية كالاتي :

$$f\,(\mathrm{t_i}, \alpha, \beta) = \frac{\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\alpha}\right)^{\beta}\right]^2} \qquad t_i > 0 \quad \alpha, \beta > 0 \ , [0, \infty] \qquad (3)$$

اذ أن : α هي معلمة القياس Scale Parameter , β هي معلمة الشكلShape Parameter

- وان الدالة التوزيعية التراكمية **Cumulative Distribution Function** تعطى بالشكل الاتي:

$$F(\ t_i) = P_r(T_i < \ t_i) = \int_0^t f(x)\,dx = \frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}} \qquad (4)$$

- **دالة البقاء والمخاطرة (survival and Risk function) كالأتي:**

دالة البقاء هو احتمال بقاء المريض على قيد الحياة خلال مدة زمنية معينة تحت ظروف وعوامل خاصة، ويرمز لدالة البقاء بالرمز S (t)، ويعبر عن دالة البقاء بالمعادلة الرياضية الآتية :

$$S(t) = \Pr\,(T > t) = \int_t^{\infty} f(y)dy = \frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{\beta}} \qquad \mathbf{(5)}$$

دالة المخاطرة تمثل أحتمال وفاة المريض قيد الدراسة خلال الفترة $(t + \Delta t, t)$ ، مع ملاحظة أن المريض كان على قيد الحياة خلال الفترة t.

$$\mathrm{h}(\ t_i) = \lim_{\Delta t \to 0} \frac{1}{\Delta\ \ t_i} \frac{p_r(\ t_i \le T_I < \ t_i + \Delta)}{p_r(T_i \ge \ t_i)} \qquad (6)$$

$$h(t_i) = \frac{f(t_i)}{S(t_i)} = \frac{\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1}}{1+\left(\frac{t_i}{\alpha}\right)^{\beta}} \qquad (7)$$

- **خصائص التوزيع (Distribution characteristics) :**

- العزم المركزي (Central moment)

$$\mathrm{ET^k} = \int_0^{\infty} \mathrm{T^k} f(t)\,dt = \alpha^{\mathrm{K}} \mathrm{B}\left(1 - \frac{\mathrm{k}}{\beta}, 1 + \frac{\mathrm{K}}{\mathrm{B}}\right); where\ K = 1,2, \dots;$$

$$;B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$$

$$= \alpha^k (k\pi/\beta)\csc(k\pi/\beta) \quad if\ \beta > 1$$ **[Dixit Asha , 2008]**

- المتوسط (Mean)

$$M_t = \alpha(\pi\ /\ \beta)\csc(\pi/\beta)$$

- التباين (Variance)

$$\sigma_t^2 = \alpha^2[(2\pi\ /\ \beta)\csc(2\pi\ /\ \beta) - (\pi/\beta)^2 \csc^2(\pi/\beta)]$$

5- طرائق التقدير(estimation methods)

1-5 طريقة الإمكان الأعظم (Maximum likelihood Methods)

تعد هذه الطريقة من أهم الطرائق التقديرالمهمة لما تحتويه من خواص جيدة ومنها الثبات ،كفاءة عالية واتساق في بعض الأحيان . افترض أن لدينا عددًا من المشاهدات تتبع توزيع log-logistic ، يُشار إليه بـ $t_1 \le t_2 \le \ldots \le t_n$. لدالة الاحتمالية اللوغارتمية لـ β,α كالاتي:

نحصل على دالة الامكان عندما نعوض الدالة $f(t_i)$ التي سبق تعريفها في معادلة **(2)** وكالاتي: [عبدالقادر, **2018**]

$$L(\alpha,\beta\,|t) = n!\prod_{i=1}^{n} f(t_i)$$

$$=n!\prod_{i=1}^{n}\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1}\left[1+\left(\frac{t_i}{\alpha}\right)^{\beta}\right]^{-2}$$

$$L(\alpha,\beta\,|t) = n!\,\beta^n\alpha^{-n\beta}\prod_{i=1}^{n} t_i^{\beta-1}\prod_{i=1}^{n}\left[1+\left(\frac{t_i}{\alpha}\right)^{\beta}\right]^{-2} \qquad (8)$$

ويمكن كتابة دالة الامكان اللوغارتمية(log likelihood)

$$Ln\,L(\alpha,\beta|t) = ln\,n! + n\,ln\,\beta - n\beta ln\,\alpha + (\beta-1)\sum_{i=1}^{n} ln\,(t_i) - 2\sum_{i=1}^{n} ln\,\left(1+\left(\frac{t_i}{\alpha}\right)^{\beta}\right) \qquad \mathbf{(9)}$$

نشتق معادلة (**9**) بالنسبة لمعلمات α و β

$$\frac{\partial \ln L(\alpha,\beta|t)}{\partial\alpha} = -\frac{n\beta}{\alpha} + \frac{2\beta}{\alpha}\sum_{i=1}^{n}\frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}} \qquad \mathbf{(10)}$$

$$\frac{\partial \ln L(\alpha,\beta|\,t)}{\partial\beta} = \frac{n}{\beta} - n\ln(\alpha) + \sum_{i=1}^{n} ln\,(t_i) - 2\sum_{i=1}^{n}\frac{ln\left(\frac{t_i}{\alpha}\right)}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}} \qquad \mathbf{(11)}$$

ومن بعدها نساوي المعادلتين (10) و (11) بالصفر ومن ثم نحلها بأحدى الطرق العددية للحصول على مقدرات الامكان الاعظم لمعلمات α و β .

ومقدر الامكان الاعظم لدالة المخاطرة h(t) حيث t>0 *بالاعتماد على الاحصاءات المرتبة على التوالي:*

$$\hat{h}(t_i)_{MLO} = \frac{\frac{\hat{\beta}_{MLO}}{\hat{\alpha}_{MLO}}\left(\frac{t_i}{\hat{\alpha}_{MLO}}\right)^{\hat{\beta}_{MLO}-1}}{1+\left(\frac{t_i}{\hat{\alpha}_{MLO}}\right)^{\hat{\beta}_{MLO}}} \qquad \mathbf{(12)}$$

2-5 الطريقة التجزيئية Quantile Method (Q.E)

أقترح الباحثان Enrique and Hadi في عام (1997) طريقة التجزيئية , يمكن تلخيصها على النحو التالي: بأفتراض أن α =$(\alpha_1,\alpha_2\cdots\alpha_n)$ تمثل المعلمات المراد تقديرها و المتغيرات t_1 , t_2 , …,t_n تمثل احصاءات مرتبة مسحوبة من عينة عشوائية تتبع توزيع log-logistic ويمكن تعبير عنه بالصيغة التالية $F(t_{1:n},\alpha) = \boldsymbol{P_{i:n}}$: [**عبدالاحد, 2012**] **[Afify , 2004,pp2]**

وان دالة التوزيع التراكمية لتوزيع log-logistic ستكون كالآتي :

$$F(t) = \frac{1}{1+\left(\frac{t}{\alpha}\right)^{-\beta}}$$

ويمكن كتابتها بالصيغة التالية :-

$$p_i = \frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}}$$

$$p_i + p_i\left(\frac{t_i}{\alpha}\right)^{-\beta} = 1$$

$$\left[p_i\left(\frac{t_i}{\alpha}\right)^{-\beta} = 1-\rho_i\right] * \frac{1}{p_i}$$

$$\left(\frac{t_i}{\alpha}\right)^{-\beta} = \frac{1-\rho_i}{\rho_i}$$

$$\frac{t_i}{\alpha} = \left(\frac{1-\rho_i}{\rho_i}\right)^{-\frac{1}{\beta}}$$

$$t_i = \alpha\left(\frac{1-\rho_i}{\rho_i}\right)^{-\frac{1}{\beta}}$$

بما انه لدينا معلمتين ,اذن سوف يكون لدينا معادلتين

$$t_i = \alpha\left(\frac{1-\rho_i}{\rho_i}\right)^{-\frac{1}{\beta}} \qquad (13)$$

$$t_j = \alpha\left(\frac{1-\rho_j}{\rho_j}\right)^{-\frac{1}{\beta}} \qquad (14)$$

وعندما $i < J$ ، وبحل المعادلتين (13) و (14) لتقدير كل من α و β :

$$\hat{\alpha}_{i,j} = \frac{t_i - tj}{\left(\frac{1-\rho_i}{\rho_i}\right)^{-\frac{1}{\beta_{i,j}}} - \left(\frac{1-\rho_j}{\rho_j}\right)^{-\frac{1}{\beta_{i,j}}}} \qquad (15)$$

$$t_i = \alpha_{i,j}\left(\frac{1-\rho_i}{\rho_i}\right)^{\frac{-1}{\beta_{i,j}}}$$

$$\ln t_i = \frac{-1}{\beta_{i,j}} \ln\left(\alpha_{i,j}\left(\frac{1-P_i}{p_i}\right)\right)$$

$$-\ln\left(\alpha_{ij}\left(\frac{1-P_i}{\rho_i}\right)\right) = \beta_{ij}\ln t_i$$

$$\hat{\beta}_{i,j} = \frac{-\ln\left(\hat{\alpha}_{i,j}\left(\frac{1-p_i}{\rho_i}\right)\right)}{\ln t_i} \qquad (16)$$

واذا كان لدينا r من المعلمات للتقدير فان $C_2^{\,n}$ سيكون عدد التقديرات المختلفة ، $p_i \ ,i = 1,2,\ldots,n$ وعليـة المقدرات التجزيئية ستكون كالآتي :-

$$\hat{\alpha}_{Q.E} = median(\hat{\alpha}_{i,j})$$

$$\hat{\beta}_{Q.E} = median(\hat{\beta}_{i,j})$$

حيث ان (Median) يمثل الوسيط للمقدرات بدلالة (i , j) واما مقدر دالة المخاطرة h(t) للطريقة التجزيئية كالآتي :

$$\hat{h}(t_i)_{Q.E} = \frac{\frac{\hat{\beta}_{Q.E}}{\hat{\alpha}_{Q.E}}\left(\frac{t_i}{\hat{\alpha}_{Q.E}}\right)^{\hat{\beta}_{Q.E}-1}}{1+\left(\frac{t_i}{\hat{\alpha}_{Q.E}}\right)^{\hat{\beta}_{Q.E}}} \quad (17)$$

6- الجانب التجريبي : (The Simulation Approach)

تم استعمال اسلوب المحاكاة بطريقة (Monte Carlo) للمقارنة بين طرائق التقدير المختلفة حيث يتميز هذا الاسلوب بالمرونة ويوفر الكثير من التكاليف عن طريق اخذ بنظر الاعتبار حجوم العينات المختلفة والقيم المختلفة لمعلمات التوزيع وتكرار التجربة في كل مرة ويتم في هذا الاسلوب توليد البيانات دون اللجوء الى البيانات الحقيقية مع عدم الاخلال بالدقة المطلوبة وتتلخص هذه الطريقة بالخطوات التالية :

1- تحديد القيم الافتراضية : تم استعمال قيم افتراضية للمعلمتين بناءاً على قيم المقدرة للجانب التطبيقي واختيارثلاث حجوم عينات وهي (50,100,250) وكما يأتي :

2- **توليد البيانات:**

1- اذ تم توليد المتغير العشوائي بطريقة التحويل العكسي كمايلي :

$$t_i = \alpha\left(\frac{1-u}{u}\right)^{-\frac{1}{\beta}} i = 1,2,3,\dots,n \quad (18)$$

2- ترتيب البيانات التي تم توليدها ترتيباً تصاعدياً $t_1 < t_2 < \cdots < t_i$

3- حل المعادلات التي تم الوصول اليها بالطرق العددية

4- تم تحديد الطريقة الافضل عن طريق مقياس المقارنة ($IMSE$) في حالة تقدير دالة التوزيع الاحتمالية ودالة المخاطرة

$$IMSE(\hat{h}(t)) = \frac{1}{r}\sum_{i=1}^{r}\left[\frac{1}{n_t}\sum_{j=1}^{n_t}\left(\hat{h}_i(t_j) - h(t_j)\right)^2\right] \quad (19)$$

$$IMSE(\hat{f}(t)) = \frac{1}{r}\sum_{i=1}^{r}\left[\frac{1}{n_t}\sum_{j=1}^{n_t}\left(\hat{f}_i(t_j) - f(t_j)\right)^2\right] \quad (20)$$

اذ أن :

r : عدد تكرارات التجربة (1000) مرة

n_t : عدد البيانات المولدة لكل عينة

$\hat{f}_i(t_j), \hat{h}_i(t_j)$: مقدر دالة التوزيع الاحتمالية ودالة المخاطرة على التوالي .

$f(t_j)$, $h(t_j)$: دالة التوزيع الاحتمالية والمخاطرة حسب القيم الابتدائية وعلى التوالي .

5- حساب متوسط مربعات الخطأ (MSE) لكل قيمة من قيم المتغير (t_i) لمعالم التوزيع α وβ .

$$MSE(\hat{\alpha}_i) = \frac{\sum_{i=1}^{r}(\hat{\alpha}_i-\alpha_i)^2}{r} \quad (21)$$

$$MSE(\hat{\beta}_i) = \frac{\sum_{i=1}^{r}(\hat{\beta}_i-\beta_i)^2}{r} \quad (22)$$

اذ أن :

r : عدد تكرارات التجربة (1000) مرة

$\hat{\alpha}_i, \hat{\beta}_i$: مقدرات معالم التوزيع وعلى التوالي .

α, β: تمثل معالم التوزيع حسب القيم الابتدائية وعلى التوالي .

- وقد تم الحصول على النتائج المتمثلة بالجداول الاتية :

- **الجدول رقم (1) يبين قيم معالم الافتراضية والمقدرة لـ α و β وحجوم العينات ولجميع التجارب**

		$\hat{\alpha}$		$\hat{\beta}$	
model	n	MLE	Q.E	MLE	Q.E
α=2.0,β=2.0	50	2.05412	2.054105	2.021349	2.023975
	100	1.088246	1.088559	1.97942	1.983844
	250	1.037706	1.036679	1.040964	**1**.039501
α=2.5,β=2.5	50	2.490443	2.58866	2.49015	2.49868
	100	2.484991	2.487163	1.411661	1.418568
	250	3400532.	2.236698	1.124394	1.225606
α=2.7,β=4.6	50	2.709721	2.719773	4.691514	4.697993
	100	2.681577	2.685855	4.590082	4.592206
	250	2.518995	2.618505	4.435429	4.538372

- **الجدول رقم (2) يبين قيم MSE لتقدير معلمة (α)و(β) بجميع الطرائق وحجوم العينات ولجميع التجارب**

		$\hat{\alpha}$			$\hat{\beta}$		
model	n	MLE	Q.E	Best	MLE	Q.E	Best
α=2.0,β=2.0	50	2.72E-02	2.75E-02	MLE	1.51E-02	1.59E-02	MLE
	100	2.00E-02	1.57E-02	Q.E	0.011276	1.46E-03	Q.E
	250	1.40E-03	1.40E-03	MLE+Q.E	0.0006247	6.36E-04	MLE
α=2.5,β=2.5	50	4.08E-02	4.10E-02	MLE	2.97E-02	2.99E-02	MLE
	100	3.22E-03	3.15E-03	Q.E	1.05E-03	2.75E-02	MLE
	250	1.38E-03	2.86E-03	MLE	1.00E-03	1.23E-03	MLE
α=2.7,β=4.6	50	1.52E-02	1.53E-02	MLE	0.04609	0.460089	MLE
	100	1.30E-03	0.001238	Q.E	0.009538	0.019511	MLE
	250	2.51E-04	2.80E-04	MLE	8.07E-03	8.22E-03	MLE

نلاحظ من جدول رقم (2) عند تقدير المعلمة (α) بصورة عامة وعند حجم عينة n=50,250 ظهرت طريقة MLE يالأفضلية اما عند حجم عينة n=100 ظهرت طريقة Q.E بالأفضلية
وعند تقدير المعلمة (β) وبصورة عامة ظهرت طريقة MLE بالأفضلية لجميع حجوم العينات ولجميع التجارب اما في التجربة الاولى وعند حجم عينة n=100 ظهرت طريقة Q.E بالافضلية .

الجدول رقم (3) يبين قيم IMSE لتقدير دالة المخاطرة بجميع الطرائق وحجوم العينات ولجميع التجارب

model	n	MLE	Q.E	best
α=2.0,β=2.0	50	0.249907	0.250635	MLE
	100	0.027158	0.025109	Q.E
	250	0.026278	0.026354	MLE

α=2.5,β=2.5	50	0.198167	0.198540	MLE
	100	0.102561	0.101220	Q.E
	250	0.020805	0.022782	MLE
α=2.7,β=4.6	50	0.171774	0.175303	MLE
	100	0.150202	0.146483	Q.E
	250	0.018078	0.121432	MLE

نلاحظ من جدول رقم (3) لتقدير دالة المخاطرة لجميع التجارب وعند حجم عينة n=50,250 ظهرت طريقة MLE بالافضلية وعند حجم عينة n=100 ظهرت طريقة Q.E بالافضلية .

الجدول رقم (4) يبين قيم IMSE لتقدير دالة التوزيع الاحتمالية بجميع الطرائق وحجوم العينات ولجميع التجارب

model	n	MLE	Q.E	best
α=2.0,β=2.0	50	3.05E-02	3.05E-02	MLE+Q.E
	100	4.10E-04	4.74E-04	MLE
	250	4.44E-05	4.44E-05	MLE+Q.E
α=2.5,β=2.5	50	4.80E-02	4.80E-02	MLE+Q.E
	100	4.25E-03	4.55E-03	MLE
	250	6.02E-04	6.02E-04	MLE+Q.E
α=2.7,β=4.6	50	4.89E-02	4.89E-02	MLE+Q.E
	100	0.032048	3.60E-02	MLE
	250	3.60E-03	3.60E-03	MLE+Q.E

نلاحظ من جدول رقم (4) لتقدير دالة التوزيع الاحتمالية لجميع التجارب وعند حجم عينة n=50,100,250 ظهرت طريقة MLE بالافضلية وعند حجم عينة n=50,250 ظهرت طريقة Q.E بالافضلية .

7- الجانب التطبيقي : (the practical Approach)

تم جمع بيانات حقيقية بحجم (50)لمرضى سرطان الثدي من المركز الوطني الريادي لبحوث السرطان / جامعة بغداد لعام 2020م اذ تم تسجيل وقت دخول المريض لمركز لحين الخروج وان جميعهم في حالة وفاة عند الخروج وتعتبر هذه البيانات هي بيانات كاملة (complete data)

جدول (5)يبين البيانات الحقيقية

t_i	2.1 7.2 1.2 3.4 8.6 2.1 6.4 3 8.12 5.4 5.2 3.4 2.1 2.1 3.4 6.4 6.4
t_i	3.2 3.2 3.2 6.2 12.4 1.9 5.6 4.1 2.1 3.4 9.8 3.1 2.1 2.1 12.4 11.4
t_i	3.4 3.4 6.5 3.2 12.1 5.4 6.1 6.1 6.1 2.4 8.4 10.5 10.5 8.5 9.1 9.1 5.3

والجدول (5) يمثل وقات الفشل لمرضى سرطان الثدي بالاشهر وان الرقم الكسري يمثل اوقات الفشل بالايام .

1-7 اختبار ملائمة البيانات Data fitting

لبيان مدى ملائمة البيانات الحقيقية لتوزيع log-logistic تم اجراء اختبار لبيانات اعلاه بأستعمال برنامج 5.5 Easy Fit Standard لاختبار الفرضية الاحصائية التالية عند مستوى معنوية 0.05

Ho: log-logistic البيانات تتبع توزيع

H1: log-logistic البيانات لاتتبع توزيع

وكانت نتائج الاختبار كما في الجدول (6) الاتي :

جدول (6) يبين اختبارات ملائمة البيانات التطبيقية

Test	Sig
Kolmogorov-smirnov	0.13623
Anderson Darling	0.9763
Chi-squared	8.977

ونلاحظ من جدول (6) ان قيمة sig لجميع الاختبارات اكبر من 0.05 لذلك نقبل فرضية العدم Ho اي ان البيانات الحقيقية المختارة تتبع توزيع log-logistic

جدول (7) يبين قيم معالم المقدرة التوزيع log-logistic لطرائق التقدير

$\hat{\alpha}_{MLE}$ $\hat{\beta}_{MLE}$	$\hat{\alpha}_{Q.E}$ $\hat{\beta}_{Q.E}$
3.279519	2.714018
4.889562	5.360724

جدول (8) يبين تقدير دالة المخاطرة لبيانات التطبيقية

t_i	$\hat{h}(t)_{MLE}$	$\hat{h}(t)_{Q.E}$	t_i	$\hat{h}(t)_{MLE}$	$\hat{h}(t)_{Q.E}$
1.2	2.96E-02	5.55E-02	5.4	0.912776	0.978493
1.9	0.166841	0.363434	5.4	0.912776	0.978493
2.1	0.236563	0.515179	5.6	0.933676	0.993796
2.1	0.236563	0.515179	6.1	0.93478	0.997514
2.1	0.236563	0.515179	6.1	0.93478	0.997514
2.1	0.236563	0.515179	6.1	0.93478	0.997514
2.1	0.236563	0.515179	6.2	0.955093	0.994438
2.1	0.236563	0.515179	6.4	0.965998	0.999267

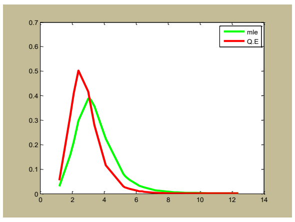
2.1	0.236563	0.515179	6.4	0.965998	0.999267
2.4	0.363627	0.761509	6.4	0.965998	0.999267
3	0.640197	0.777764	6.5	0.996619	1.787159
3.1	0.680781	0.780381	7.2	1.664887	1.790581
3.2	0.718202	0.885136	8.12	1.895094	1.858338
3.2	0.718202	0.885136	8.4	1.896291	1.88669
3.2	0.718202	0.885136	8.5	1.96983	1.90929
3.2	0.718202	0.885136	8.6	1.963499	1.962055
3.4	0.782313	0.893963	9.1	1.963683	1.988193
3.4	0.782313	0.893963	9.1	1.963683	1.988193
3.4	0.782313	0.893963	9.8	1.976583	1.986452
3.4	0.782313	0.893963	10.5	1.984104	1.990184
3.4	0.782313	0.893963	10.5	1.986104	1.990184
3.4	0.782313	0.893963	11.4	1.987941	1.990195
4.1	0.892904	0.958427	12.1	1.993415	1.992888

5.2	0.890959	0.960268	12.1	1.993415	1.992888
5.3	0.892018	0.964234	12.4	1.993729	1.997191

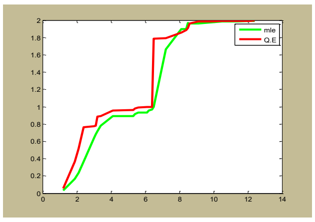
جدول (9) يبين تقدير دالة التوزيع الاحتمالية لبيانات التطبيقية

t_i	$\hat{f}(t)_{MLE}$	$\hat{f}(t)_{Q.E}$	t_i	$\hat{f}(t)_{MLE}$	$\hat{f}(t)_{Q.E}$
1.2	2.94E-02	0.054849	5.4	6.69E-02	2.36E-02
1.9	0.156025	0.316619	5.4	6.69E-02	2.36E-02
2.1	0.212528	0.411208	5.6	5.54E-02	1.89E-02
2.1	0.212528	0.411208	6.1	3.51E-02	0.011148
2.1	0.212528	0.411208	6.1	3.51E-02	0.011148
2.1	0.212528	0.411208	6.1	3.51E-02	0.011148
2.1	0.212528	0.411208	6.2	3.21E-02	1.01E-02
2.1	0.212528	0.411208	6.4	0.02697	8.26E-03
2.1	0.212528	0.411208	6.4	0.02697	8.26E-03
2.4	0.298726	0.501889	6.4	0.02697	8.26E-03
3	0.388731	0.416003	6.5	2.47E-02	7.50E-03
3.1	0.386944	0.381736	7.2	1.39E-02	3.94E-03

3.2	0.380625	0.346714	8.12	6.99E-03	1.84E-03
3.2	0.380625	0.346714	8.4	5.74E-03	1.49E-03
3.2	0.380625	0.346714	8.5	5.36E-03	1.38E-03
3.2	0.380625	0.346714	8.6	5.01E-03	1.28E-03
3.4	0.356744	0.279275	9.1	3.61E-03	8.96E-04
3.4	0.356744	0.279275	9.1	3.61E-03	8.96E-04
3.4	0.356744	0.279275	9.8	2.34E-03	5.60E-04
3.4	0.356744	0.279275	10.5	1.56E-03	3.61E-04
3.4	0.356744	0.279275	10.5	1.56E-03	3.61E-04
3.4	0.356744	0.279275	11.4	9.65E-04	2.14E-04
4.1	0.22437	0.116326	12.1	6.80E-04	1.47E-04
5.2	8.09E-02	2.97E-02	12.1	6.80E-04	1.47E-04
5.3	7.35E-02	2.65E-02	12.4	5.89E-04	1.25E-04



شكل رقم(1) سلوك مقدرات دالة المخاطرة



شكل رقم(2) سلوك مقدرات دالة pdf

- نلاحظ من جدول رقم (8) والشكل 1 ظهرت دالة المخاطرة بشكل متزايد وهذا يتطابق مع نضرية الدالة
- نلاحظ من جدول رقم (9) وشكل 2 ظهرت دالة التوزيع الاحتمالية بشكل متزايد ثم الانخفاض التدريجي
-

8- الاستنتاجات :(Conclusion)

1- من خلال نتائج تجارب المحاكاة تبين تفوق طريقة الامكان الاعظم على طريقة التجزيئية للنماذج الثلاث وعند حجم عينة n=50 ,250 أما في حالة حجم عينة n=100 ظهرت طريقة التجزيئية بالأفضلية في عملية تقدير دالة المخاطرة للتويع log-logistic .

2- في تقدير دالة التوزيع الاحتمالية ظهرت طريقة الامكان الاعظم في جميع التجارب وفي جميع حجوم العينات ذات كفاءة عالية اما في حالة حجم عينة n=50,250 ظهرت الطريقة التجزيئية بالافضلية .

3- من استنتاجات الجانب التطبيقي اتضح ان قيم دالة المخاطرة هي متزايدة بازدياد وقت الاصابة بالنسبة لمجموعة لمرضى سرطان الثدي قيد البحث وهذا يتطابق مع الخصائص النظرية لهذه الدالة كونها دالة رتيبة متزايدة (monotone increasing function)

4- ظهرت قيم دالة التوزيع الاحتمالية بشكل متزايد ثم الانخفاض التدريجي كلما زادت قيم t

9 - التوصيات : (Recommendations)

1- يوصي الباحث بأستعمال الطريقة التجزيئية في حالة حجوم العينات المتوسطة وطريقة الامكان الاعظم في حالة حجوم العينات الصغيرة والكبيرة في تقدير دالة المخاطرة لتوزيع Log-Logistic.

2- يوصي الباحث أستعمال نماذج اخرى لأحصاءات المرتبة ومنها (الاحصاءات المرتبة ذات القيم المسجلة العليا والدنيا والاحصاءات المرتبة المعممة المرافقة)وغيرها من النماذج لتقدير دالة المخاطرة لتوزيع Log-Logistic .

3- يوصي الباحث بضرورة اعتماد وزارة الصحة لمقدرات دالة المخاطرة ولدالة التوزيع الاحتمالية في الجانب المحاكاة والتطبيقي لمرض سرطان الثدي والافادة في وضع الخطط العلاجية المستقبلية ومتابعة المرضى .

4- يوصي الباحث أستعمال مبدأ مجموعة المعاينة المرتبة (Rank Set Sampling) لتقدير المخاطرة لتوزيع Log-Logistic ومقارنته مع طرائق التقدير الاعتيادية .

المصادر: (Reference)

1- **الحريري**, رانيا بن محمد (2009) **"دراسة بعض خواص الأحصاءات المرتبة المعممة"** رسالة ماجستير في الرياضيات مقدمة الى كلية العلوم التطبيقية في جامعة ام القرى, مملكة العربية السعودية.

2- **عبد الأحد**, عطاف اداور(2007), **"تقديرات المعولية للتوزيع الأسي بمعلمتين دراسة مقارنة"**, رسالة ماجستير في بحوث العمليات مقدمة الى كلية الأدارة والأقتصاد في جامعة بغداد .

3- **عبد القادر** , ايناس غانم (2018) **"تقدير معلمات توزيع Lomax بأستخدام الأحصاءات المرتبة المعممة"** رسالة ماجستير في الاحصاء مقدمة الى كلية العلوم الحاسوب والرياضيات في جامعة الموصل

4- **Afify, E.E.** (2004) **"Linear and Nonlinear Reqression of Exponential Distribution"**

5- **Dali Chen, et.al** (2020)**," Estimation of thermal time model parameters for seed germination in 15 species: the importance of distribution function"** Science Research 1–8.

6- **Dixit, Asha** (2008), **"Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions",** A Thesis Submitted to the Graduate Faculty of Auburn University

7- **Fitzgerald, D. L.** (2005). **"Extreme Rainfall Using the Log-Logistic Distribution"**, Stochastic Environmental Research and Risk Assessment, 19, 249-257.

8- **Kamps, U**. (1995). "**A concept of generalized order statistics**". Journal of Statistical Planning and Inference, Vol. 48: 1-23.

9- **Ragab, A. and Green J.** (1984). **"On order statistics from the log-logistic distribution and their properties",** Communications in Statistics-Theory and Methods, 13, 2713{2724.

10- **Singh, V. P. and Guo**, H. (1995). **"Parameter Estimation For 2-parameter Log-logistic Distribution by Maximum Entropy"**, Civil Engineering System, **12**, 343-357.