مقدر دالة الإمكان الأعظم الحصين لدالة معولية توزيع ريلي

قتيبـــة نبيل نـايف القزاز* و سرمد علوان صالح**

*** قسم الإحصاء، كلية الإدارة والإقتصاد، جامعة بغداد.**

**** قسم الإحصاء، كلية الإدارة والإقتصاد، جامعة بغداد.**

الخلاصة

يهدف هذا البحث إلى تقدير دالة المعولية لتوزيع ريلي بوجود حالة التلوث في مشاهدات العينة عند الدراسة، باستخدام مقدرات الامكان الأعظم الحصينة ومقارنة النتائج مع طريقة الامكان الأعظم الاعتيادية بأسلوب المحاكاة.

1- الجانب النظري

1-1 المقدمة

أن معظم البحوث والدراسات في مجال الإحصاء وسيما في حقل المعولية تهدف للحصول على مقدرات ذات مستوى عالٍ من الكفاءة وخاصة عندما تكون بيانات العينة قيد الدراسة ملوثة ومن هنا جاء هدف البحث في الوصول إلى مقدرات كفوءة لدالة المعولية لتوزيع ريلي من خلال استخدام مقدرات الامكان الأعظم الحصين(RMLE) ومقارنة النتائج مع طريقة الامكان الأعظم (ML) باستخدام أسلوب المحاكاة وبالاعتماد على المقياسيـــن الإحصائييـــن متوســـــط مربعـــات الخطأ Mean Square Error (MSE)، ومـــتوسط مربعات الخطأ التكاملي Integral Mean square Error (IMSE) لغرض المقارنة بين النتائج.

2-1 دالة المعولية:[1]

هي احتمال بقاء الوحدة التجريبية تعمل على الأقل حتى الوقت t إذ أن $t \geq 0$ أي إذا كان T المتغير العشوائي الذي يمثل وقت البقاء للوحدة التجريبية فان R(t) تمثل احتمال بقاء المفردة تعمل لمدة قادمة أي:

$$\mathbf{R(t) = Pr(T > t) = 1 - P(T \geq t)} \quad \text{.................... (1)}$$

$$= \int_t^{\infty} \mathbf{f(u)du}$$

$$= \mathbf{1 - F_T(t)}$$

$$= \mathbf{\bar{F}_T(t)}$$

وتكون R(t) غير متزايدة ومستمرة من جهة اليسار.

3-1 توزيع ريلي:[4]

أوجد هذا التوزيـــع مـــن قبـــل العَالـِـم الإنكليزي Lord Rayleigh، ويستخدم في التحليلات المفردة وتحليلات الخطأ لمختلف الأنظمة، ويعد توزيع ريلي حالة خاصة من توزيع ويبل عندما تكون معلمة الشكل $\beta=2$، وان دالة الكثافة الاحتمالية (P.d.f.) لهذا التوزيع كما يلي:

$$\mathbf{f(t) = \frac{2}{\theta} t\, e^{-\frac{t^2}{\theta}}\ ;\ \theta > 0\ ,\ t > 0} \quad \text{...................... (2)}$$

4-1 طريقة الامكان الأعظم في التقدير:[3]

أن أسلوب هذه الطريقة يعتمد على إيجاد قيمة تقديرية للمعلمة التي تجعل الدالة في نهايتها العظمى، وان صيغة دالة الامكان كالأتي:

$$\mathbf{L = \prod_{i=1}^{n} f(t_i, \theta)}$$

وللحصول على مقدر الامكان الأعظم لتوزيع ريلي وكالاتي:

$$\mathbf{L(t_1, t_2, ..., t_n; \theta) = \left(\frac{2}{\theta}\right)^n \prod_{i=1}^{n} t_i\, e^{-\frac{\sum_{i=1}^{n} t_i^2}{\theta}}} \quad \text{....... (3)}$$

وبأخذ اللوغريتم الطبيعي للمعادلة (3) نحصل على:

$$\mathbf{L(t_1, t_2, ..., t_n; \theta) = nLn(2) - nLn(\theta) \sum_{i=1}^{n} Ln(t_i) -}$$

$$\frac{\sum_{i=1}^{n} t_i^2}{\theta} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

وبأخذ المشتقة الجزئية للمعلمة θ نحصل على ما يلي:

$$\frac{\partial Ln}{\partial \theta} L = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} t_i^2}{\theta^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

وعند مساواة المعادلة (5) للصفر، وحل المعادلة نحصل على مقدر الامكان الأعظم للمعلمة θ وكالأتي:

$$\hat{\theta}_{ML} = \frac{\sum_{i=1}^{n} t_i^2}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

وكما هو معروف أن مقدر الامكان الأعظم يمتاز بخاصية الثبات ، وهذه الخاصية يمكن من خلالها الحصول على مقدر دالة المعولية لتوزيع ريلي وكالاتي:

$$\because R(t) = 1 - F(t)$$

$$R(t) = 1 - \int_0^t \frac{2}{\theta} u\, e^{-\frac{u^2}{\theta}} du$$

$$R(t) = 1 - \left[1 - e^{-\frac{t^2}{\theta}}\right] = e^{-\frac{t^2}{\theta}}$$

$$\therefore \hat{R}(t)_{ML} = e^{-\frac{t^2}{\hat{\theta}_{ML}}} = e^{-\frac{nt^2}{\sum t_i^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

5-1 طريقة الامكان الأعظم الحصينة في التقدير:

قام الباحثان بأقتراح هذه الطريقة وهي عبارة عن اشتقاق مقدر دالة البقاء لتوزيع ريلي الملوث باستخدام طريقة (ML) الحصينة وكالاتي:

$$X_1, X_2, \ldots, X_{n_1} \sim f_1 \quad \Rightarrow R_1(t) = e^{-\frac{t^2}{\theta_1}}$$

$$X_{n_1+1}, \ldots, X_n \sim f_2 \quad \Rightarrow R_2(t) = e^{-\frac{t^2}{\theta_2}}$$

إذ أن:

$$\begin{bmatrix} f(t,\theta_1) = \frac{2}{\theta_1}\, t\, e^{-\frac{t^2}{\theta_1}} \\ f(t,\theta_2) = \frac{2}{\theta_2}\, t\, e^{-\frac{t^2}{\theta_2}} \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

$$\therefore f(t,\theta_1) = (1-\tau) f(t,\theta_1) + \tau f(t,\theta_2)$$

$$= (1-\tau)\frac{2}{\theta_1} t\, e^{-\frac{t^2}{\theta_1}} + \tau \frac{2}{\theta_2} t\, e^{-\frac{t^2}{\theta_2}} \quad \ldots\ldots (9)$$

إذ أن τ تمثل مستويات التلوث في مشاهدات العينة قيد الدراسة، ومن المعادلة(9) فان:

$$L = \prod_{i=1}^{n} f(t_i)$$

$$= \prod_{i=1}^{n} \left[(1-\tau)\frac{2}{\theta_1} t_i\, e^{-\frac{t_i^2}{\theta 1}} + \tau \frac{2}{\theta_2} t_i\, e^{-\frac{t_i^2}{\theta_2}} \right] (10)$$

وأن:

$$Ln(L) = \sum_{i=1}^{n} Ln \left[(1-\tau)\frac{2}{\theta_1} t_i\, e^{-\frac{t_i^2}{\theta 1}} + \tau \frac{2}{\theta 2} t_i\, e^{-\frac{t_i^2}{\theta_2}} \right]$$

$$\ldots\ldots\ldots\ldots \quad (11)$$

وبأخذ المشتقة الجزئية للمعادلة (11) نحصل على:

$$\frac{\partial}{\partial \underline{\theta}} Ln(L) = \begin{bmatrix} \frac{\partial}{\partial \tau} Ln(L) \\ \frac{\partial}{\partial \theta 1} Ln(L) \\ \frac{\partial}{\partial \theta 2} Ln(L) \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = G \ldots\ldots\ldots (12)$$

إذ أن:

$$\underline{\theta} = \begin{bmatrix} \tau \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

أما المشتقة الثانية للمعادلة (11) تكون كالأتي:

$$\frac{\partial}{\partial\underline{\theta}'\partial\underline{\theta}}\mathrm{Ln}(L)=\begin{bmatrix}\frac{\partial}{\partial\tau'}\mathrm{Ln}(L) & \frac{\partial}{\partial\tau\,\partial\theta_1}\mathrm{Ln}(L) & \frac{\partial}{\partial\tau\,\partial\theta_2}\mathrm{Ln}(L)\\ \frac{\partial}{\partial\tau\,\partial\theta_1}\mathrm{Ln}(L) & \frac{\partial}{\partial\theta_1^2}\mathrm{Ln}(L) & \frac{\partial}{\partial\theta_1\,\partial\theta_2}\mathrm{Ln}(L)\\ \frac{\partial}{\partial\theta_2\partial\tau}\mathrm{Ln}(L) & \frac{\partial}{\partial\theta_2\partial\theta_1}\mathrm{Ln}(L) & \frac{\partial}{\partial\theta_2^2}\mathrm{Ln}(L)\end{bmatrix}$$

$$\frac{\partial}{\partial\underline{\theta}'\partial\underline{\theta}}\mathrm{Ln}(L)=\begin{bmatrix}h_{11} & h_{12} & h_{13}\\ h_{21} & h_{22} & h_{23}\\ h_{31} & h_{32} & h_{33}\end{bmatrix}=H \text{ .... (13)}$$

لغرض إيجاد تقدير المعالم $(\tau,\theta1,\theta2)$ في الصيغة (9) يتم الاعتماد على طريقة نيوتن رافسون العددية لحل هذه المعادلة وبالاعتماد على الصيغ (12) (13) وكما يلي:

$$\underline{\theta}^{n+1}=\underline{\theta}^{n+1}-H^{-1}\underline{G} \text{ .... (14)}$$

إذ أن الصيغة (14) ما هي إلا مقدر الامكان الأعظم الحصين Robust ML للمعالم (τ,θ_1,θ_2) أي أن:

$$\underline{\theta}_{RML}=\underline{\theta}^{n+1} \text{ .... (15)}$$

وباستخدام خاصية الثبات فان مقدر دالة المعولية يكون كالأتي:

$$R(t)=\int_t^{\infty}f(u)du=(1-\tau)R1(t)+\tau R2(t)$$

$$=(1-\tau)\,e^{-\frac{t^2}{\theta1}}+\tau\,e^{-\frac{t^2}{\theta2}}$$

$$\hat{R}_{MLR}(t)=(1-\hat{\tau})\,e^{-\frac{t^2}{\hat{\theta}1(RML)}}+\hat{\tau}\,e^{-\frac{t^2}{\hat{\theta}2(RML)}} \text{ .... (16)}$$

2- المحاكاة

تتضمن تجارب المحاكاة المراحل التالية:

<u>**المرحلة الأولى**</u>: تعد هذه المرحلة من أهم المراحل التي تعتمد عليها باقي المراحل اللاحقة حيث يتم فيها أختيار القيم الأفتراضية وكما يلي:

1. تحديد حجوم العينات المفترضة وكما يلي:

$n = 30, 50, 100$

2. يتم أختيار القيم الأفتراضية للمعلمة θ والمعلمة τ وكما يلي:

$q_1 = 2,\ q_2 = 3.5$

$t = 10\%, 20\%$

<u>**المرحلة الثانية**</u>: يتم توليد بيانات توزيع ريلي بصورة عامة كما يلي:[2]

$$t_i=\exp\left\{\mathrm{Ln}\left\{q\,\mathrm{Ln}\left(\frac{1}{1-u_i}\right)\right\}/2\right\} \text{ .... (17)}$$

u_i : متغير عشوائي مستمر يتبع التوزيع المنتظم.

أما توزيع ريلي الملوث قيد الدراسة في هذا البحث فيتم الحصول علية حسب الصيغة التلية وبالأعتماد على الصيغة (17) وكما يلي:

$$t_i=\left[(1-t)\exp\left\{\mathrm{Ln}\left\{q_1\mathrm{Ln}\left(\frac{1}{1-u_i}\right)\right\}/2\right\}+\right.$$
$$\left.t\exp\left\{\mathrm{Ln}\left\{q_2\mathrm{Ln}\left(\frac{1}{1-u_i}\right)\right\}/2\right\}\right] \text{ .... (18)}$$

<u>**المرحلة الثالثة**</u>: في هذه المرحلة تم فيها تقدير دالة المعولية لأنموذج ريلي وفق طريقة الأمكان الأعظم ML والمتثلة بالصيغة (7) وطريقة الأمكان الأعظم الحصينة RML الصيغة (16).

<u>**المرحلة الرابعة**</u>: وهي المرحلة الأخيرة حيث يتم فيها المقارنة بين طرائق تقدير دالة معولية توزيع ريلي وذلك بأستخدام المعايير التالية:

1. متوسط مربعات الخطأ (MSE) وفق الصيغة التالية:

$$MSE\left(\hat{R}(t)\right)=\frac{1}{r}\sum_{i=1}^{r}\left(\hat{R}_i(t)-R_i(t)\right) \text{ .... (19)}$$

حيث r تمثل عدد المكررات لكل تجربة.

2. متوسط مربعات الخطأ التكاملي (IMSE) وفق الصيغة التالية:

حيث أن المقياس الأول هو:

$$IMSE\left(\hat{R}(t)\right)=\frac{1}{r}\sum_{i=1}^{r}\left\{\frac{1}{n_t}\sum_{j=1}^{n_t}\left[\hat{R}_i(t_j)-R_j(t_j)\right]\right\} \text{ .... (20)}$$

حيث أن المقياس الأول هو مقياس مطلق للمقارنة بين المقدرات أما المقياس الثاني فهو مقياس نسبي يستخدم لدقة القياسات في حالة كون المقدرات مختلفة وللمقارنات الدقيقة.

والجداول التالية تبين تقدير دالة المعولية، ومتوسط مربعات الخطأ MSE ومتوسط مربعات الخطأ التكاملي IMSE لتقدير دالة المعولية لجميع النماذج ولجميع حجوم العينات.

جدول رقم (1)

يوضح مقدرات R(t) و MSE و IMSE لها لجميع طرائق التقدير عندما t = 10% و n = 30.

t	R(t)			MSE	
	Real	ML	RML	ML	RML
0.1	0.99523	0.99740	0.99658	0.0000050	0.0000023
0.3	0.95786	0.97688	0.96969	0.0003787	0.0001677
0.5	0.88735	0.93713	0.91831	0.0025972	0.0010870
0.7	0.79137	0.88061	0.84681	0.0083665	0.0033530
0.9	0.67962	0.81067	0.76080	0.0181000	0.0070648
1.1	0.56224	0.73123	0.66645	0.0302210	0.0116130
1.3	0.44830	0.64640	0.56980	0.0417460	0.0159080
1.5	0.34477	0.56013	0.47612	0.0496670	0.0189330
1.7	0.25596	0.47591	0.38949	0.0522320	0.0202420
IMSE				0.0252760	0.0098476

جدول رقم (2)

يوضح مقدرات R(t) و MSE و IMSE لها لجميع طرائق التقدير عندما t = 20% و n = 30.

t	R(t)			MSE	
	Real	ML	RML	ML	RML
0.1	0.99544	0.99736	0.99665	0.00000395	0.00000242
0.3	0.95972	0.97649	0.97041	0.00030144	0.00016932
0.5	0.89221	0.93609	0.92054	0.00206810	0.00104060
0.7	0.80004	0.87869	0.85131	0.00666530	0.00310890
0.9	0.69226	0.80773	0.76815	0.01443000	0.00651950
1.1	0.57840	0.72725	0.67723	0.02411600	0.01090200
1.3	0.46705	0.64145	0.58497	0.03335100	0.01551300
1.5	0.36488	0.55437	0.49790	0.03973600	0.01973300
1.7	0.27618	0.46957	0.42380	0.04186100	0.02355700
IMSE				0.020224	0.010846

جدول رقم (3)

يوضح مقدرات R(t) و MSE و IMSE لها لجميع طرائق التقدير عندما t = 10% و n = 50.

t	R(t)			MSE	
	Real	ML	RML	ML	RML
0.1	0.99523	0.99743	0.99660	0.00000499	0.000001906
0.3	0.95786	0.97713	0.96981	0.00038150	0.000144940
0.5	0.88735	0.93778	0.91852	0.00261470	0.000986210
0.7	0.79137	0.88176	0.84696	0.00841360	0.003139900
0.9	0.67962	0.81233	0.76074	0.01817400	0.006689000
1.1	0.56224	0.73332	0.66604	0.03028200	0.010962000
1.3	0.44830	0.64876	0.56893	0.04171700	0.014817000
1.5	0.34477	0.56256	0.47469	0.04945900	0.017210000
1.7	0.25596	0.47821	0.38740	0.05178200	0.017641000
IMSE				0.025157	0.0087862

جدول رقم (4)

يوضح مقدرات R(t) و MSE و IMSE لها لجميع طرائق التقدير عندما t = 20% و n = 50.

t	R(t)			MSE	
	Real	ML	RML	ML	RML
0.1	0.99544	0.99739	0.99664	0.00000392	0.000003265
0.3	0.95972	0.97675	0.97032	0.00029911	0.000194540
0.5	0.89221	0.93675	0.92019	0.00204970	0.001034500
0.7	0.80004	0.87987	0.85043	0.00659380	0.002955300
0.9	0.69226	0.80944	0.76638	0.01423800	0.006192200
1.1	0.5784	0.7294	0.67399	0.02371400	0.010458000
1.3	0.46705	0.64389	0.57914	0.03264900	0.015263000
1.5	0.36488	0.55691	0.48703	0.03868000	0.020800000
1.7	0.27618	0.47199	0.40181	0.04046100	0.029488000
IMSE				0.019673	0.013604

جدول رقم (5)

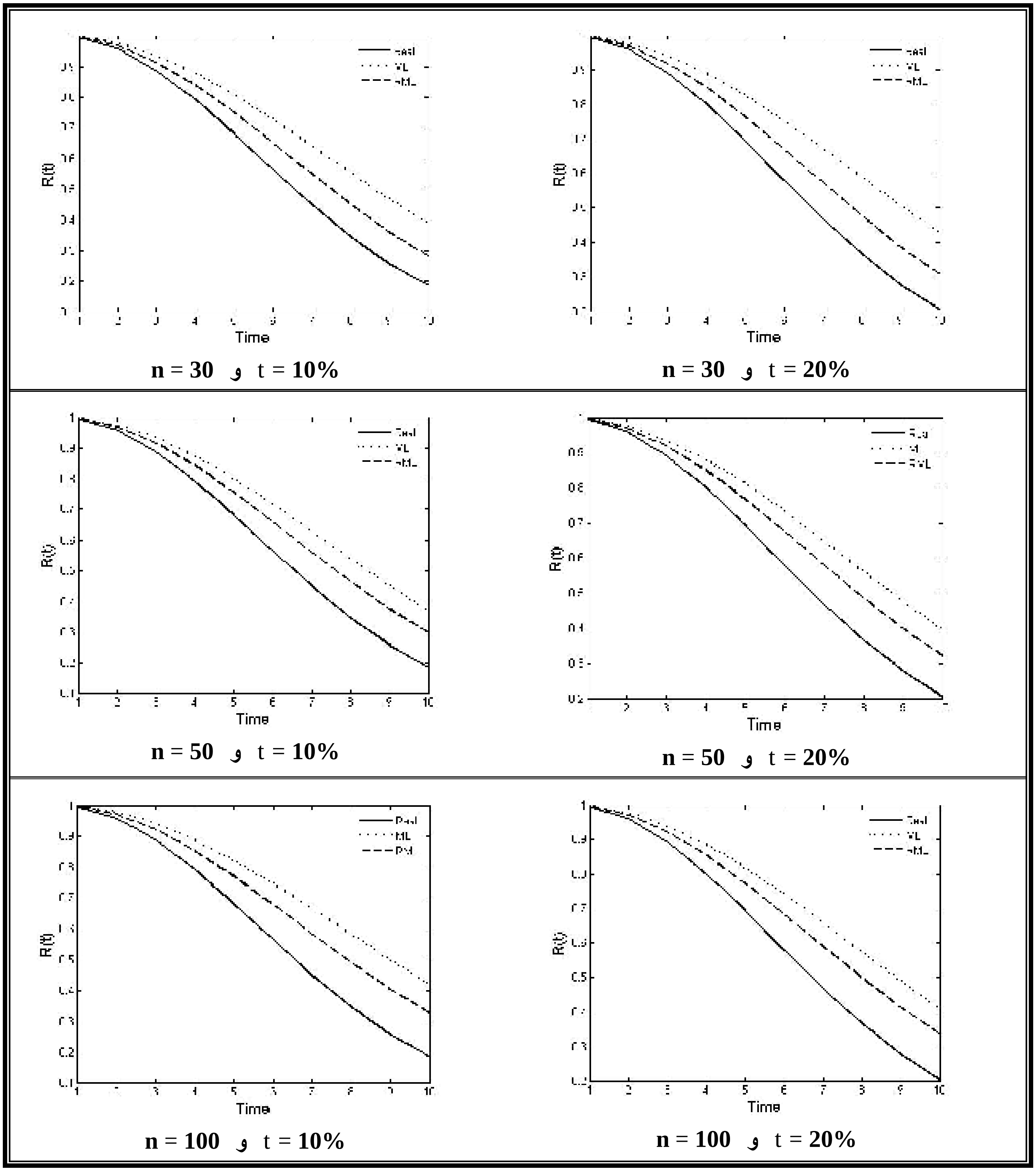
يوضح مقدرات R(t) و MSE و IMSE لها لجميع طرائق التقدير عندما t = 10% و n = 100.

t	*R(t)*			*MSE*	
	Real	***ML***	***RML***	***ML***	***RML***
0.1	0.99523	0.99747	0.9966	0.00000507	0.000001907
0.3	0.95786	0.97742	0.96986	0.00038725	0.000144960
0.5	0.88735	0.93854	0.91866	0.00265290	0.000985980
0.7	0.79137	0.88313	0.84721	0.00853070	0.003137100
0.9	0.67962	0.81435	0.76108	0.01840800	0.006677300
1.1	0.56224	0.73592	0.66644	0.03062900	0.010930000
1.3	0.44830	0.65178	0.56933	0.04211400	0.014752000
1.5	0.34477	0.56580	0.47504	0.04980300	0.017103000
1.7	0.25596	0.48144	0.38766	0.05196800	0.017493000
IMSE				0.02532	0.0087316

جدول رقم (6)

يوضح مقدرات R(t) و MSE و IMSE لها لجميع طرائق التقدير عندما t = 20% و n = 100.

t	*R(t)*			*MSE*	
	Real	***ML***	***RML***	***ML***	***RML***
0.1	0.99544	0.99744	0.9967	0.00000407	0.00000160
0.3	0.95972	0.97723	0.97073	0.00031089	0.00012202
0.5	0.89221	0.93804	0.92097	0.00213040	0.00083257
0.7	0.80004	0.8822	0.85146	0.00685290	0.00266160
0.9	0.69226	0.81293	0.76752	0.01479500	0.00570150
1.1	0.5784	0.73399	0.67507	0.02463300	0.00940760
1.3	0.46705	0.64939	0.57988	0.03389600	0.01282100
1.5	0.36488	0.56302	0.48703	0.04012000	0.01503100
1.7	0.27618	0.47839	0.40047	0.04190600	0.01556700
IMSE				0.020397	0.007666



شكل رقم (1) تغير دالة المعولية مع الزمن لجميع حجوم العينات ونسب التلوث.

3- تفسيـــر النتائج

يتضح من النتائج في الجداول (1)، (2)، (3)، (4)، (5) و (6) المذكورة أنفاً الأتي:

- لجميع النماذج ولجميع حجوم العينات ونسب التلوث نلاحظ تناقص MSE و MAPE عند زيادة حجم العينة.
- نلاحظ ان طريقة الأمكان الأعظم الحصينة RML كانت الأفضل ولجميع الحالات من طريقة الأمكان الأعظم ML.

4- الأستنتاجات

يلاحظ مما سبق أفضلية طريقة الأمكان الأعظم الحصينة المقترحةRML من طريقة الأمكان الأعظم ML الأعتيادية في التقدير، أذ أن الطريقة المقترحة RML استطاعت ان تجد مقدر يحتوي على جميع معلومات العينة الملوثة، أي ان هذا المقدر استطاع ان يعبر عن المجتمعين الذين تكونت منهما العينة قيد الدراسة.

5- المصادر

[1] Charles, E. E. (1997). "*An Introduction to Reliability and Maintainability Engineering*" the Mc-GrawHial companies, Inc. New York.

[2] Crowder, M. J., Kimber, A. C., Smith, R. L., and Sweeting, T. J. (1991). "*Statistical Analysis of Reliability Data*". Chapman and Hall. Great Britain.

[3] Lewis, E. E. (1994). "*Introduction of Reliability engineering*" 2nd edition, wiley, New York.

[4] Sinha, S. K. and Kate, B. K. (1980). "*Life Testing and Reliability*" wiley Eastern limited

Abstract

This paper is goaling to estimate the Reliebilty functions for Rayleigh distribution when observation of a sample is contaminated by using robust maximum likelihood estimeters (RMLE) and compare the results with maximum likelihood (ML) method by using simulation.