توزيع (ايرلنك الخليط - بواسون) كأنموذج لبيانات العد مع تطبيق عملي

أ.م.د. ايمان محمد عبد الله الباحث مصطفى عبد الجبار جداح

كلية الادارة والاقتصاد - جامعة بغداد - قسم الاحصاء

المستخلص :

تم بناء انموذج احتمالي (توزيع ايرلنك الخليط - بواسون) لبيانات العد لمعالجة مشكلة التشتت العالي للبيانات كما تم التعرف على خصائص التوزيع والمتمثلة بالعزوم المركزية الاربعة و التباين و الالتواء و التفلطح ,كما تم تقدير معالم التوزيع بثلاثة طرق (طريقة مقدر العزوم , طريقة الامكان الاعظم , طريق مقدر التقلص) وقد تم اعتماد اسلوب المحاكاة في تحديد المقدر الجديد , كما تم اجراء اختبار مربع كاي χ^2 لحسن المطابقة على عينة من البيانات اخذت من سجلات الراقدين الموثقة في شعبة الاحصاء الصحي والحياتي في مستشفى الزهراء التعليمي (ع) المتمثلة بعدد الراقدين ولكل فترة زمنية (بالايام) لمرضى COV-19 لفترة دخول من (14-3-2020) ولغاية (10-7-2020) , حيث اظهرت نتائج الاختبار افضلية مقدر الامكان الاعظم يليه مقدر التقلص ومن ثم مقدر العزوم , كما اظهرت نتائج اختبارمربع كايχ^2 لحسن المطابقة تفوق توزيع (ايرلنك الخليط – بواسون) على كل من توزيع ثنائي الحدين السالب المركب و توزيع بواسون في نمذجة بيانات فترة رقود المرضى .

Key Words : Poisson distribution , Compound Negative Binomial distribution , Negative Binomial distribution , over-dispresion , Erlang distribution , Mixture Erlang distribution , Mixture Erlang - Poisson distribution , goodness of fit χ^2test .

Abstract

A probabilistic model (Mixture Erlang – poisson distribution) was built for the counting data to process the problem of over-dispresion of the data .The characteristics of the distribution represented by the four central moments, variance, Kurtosis and Skewness ,The parameters of the distribution were estimated in three ways (moment estimator method ,maximum likelhood estimator and the shrinkage estimator method) ,and the simulation method was relied on to determine the good estimator .

Chi-squre χ^2 test of goodness of Fit was conducted on a sample of data taken from the records of the inpatients documented in the Health and Life Statistics Division at Al-Zahraa Teaching Hospital,represented by the number of patients and for period of time (in days) for COV-19 patients from (14-3-2020) until (10-7-2020) ,Where the test results showed preference the maximum likelhood estimator followed the shrinkage estimator and then the moments estimator ,also the results of the chi-square χ^2 test for goodness of fit showed the superiority of Erlang Mixture-Poisson distribution on both Compound Negative Binomial distribution and Poisson distribution in modeling patients' inpatient period data .

1- المقدمــة : Introduction

يعد توزيع بواسون[1] توزيع احتمالي منفصل يمثل تجربة برنولية مكررة عدد لانهائي من المرات (مثال على ذلك عدد المكالمات خلال 24 ساعة و بمعدل ثابتλ) , من خصائص التوزيع التوقع يساوي التباين لكن من حيث الواقع نادرا ما يتحقق ذلك حيث نلاحظ التباين اكبر من المتوسط الحسابي اي ان البيانات تكون عالية التشتت (over-dispresion)[20], وهذا بدوره يؤدي الى تحليل غير دقيق للبيانات عندما يكون التوزيع المتبع للبيانات توزيع بواسون , لذلك لابد من اعادة النظر في نمذجة البيانات وبالتالي ايجاد توزيع جديد يكون اكثر ملائمة للبيانات , تتمثل دراستنا باجراء التعديلات على توزيع بواسون , وذلك على افتراض توزيع مستمر معين (توزيعات مدة الحياة مثلا) للمعلمةλ (استنادا الى بنية الشجرة البيانية) لنحصل على توزيع مركب جديد يكون اكثر مرونه كونه يمتلك عدد من المعلمات بدلا من معلمة واحدة وبالتالي نحصل على افضل ملائمة ,حيث اقترح الاسلوب لاول مرة من قبل عالم الاحصاء فيشر (Fisher1941) في بحثه من خلال تقديم توزيع ثنائي الحدين السالب المركب[9][13][2] , الناتج من تركيب توزيع بواسون مع توزيع كاما (على افتراض ان المعلمة λ تتبع توزيع $\text{gamma}(\alpha,\beta)$), ليأتي اسلوب تركيب مشابه لاسلوب فيشر من حيث طريقة التركيب والمتمثل بافتراض ان الحد $p = (\frac{1}{1+\beta})$ في توزيع ثنائي الحدين السالب المركب مساوي لـ $p=e^{-\lambda}$ وعلى فرض ان λ تتبع توزيع معين (توزيعات مدة الحياة مثلا) , للحصول على توزيعات متقطعة مركبة جديدة منافسة لتوزيع ثنائي الحدين السالب المركب مثل (Negative binomial –Weighted Garima distribution)[20] الا ان هذا التوزيع لاينطبق عليه شروط دالة الكثافة الاحتمالية المعروفة حيث يظهر بعض قيمها سالبة علاوة على ذلك ان مجموعها يكون اكبر من واحد صحيح في عدد من الحالات ,فمثلا عندما ($r = 1$, $\beta = 3$, $\alpha = 12$) فان مجموع الدالة عند حجم عينة (n=30) يساوي (5.1287) , كما ان عندما ($r = 14$, $\beta = 16$, $\alpha = 1$) فان مجموع الدالة عند حجم عينة (n=30) يساوي (-674.5981) . لذا انا بدوري بعد المامي بالموضوع والتعمق بالافكار واجراء الكثير من التجارب كباحث اقترحت توزيع جديد (ايرلنك الخليط) تم اشتقاقه بالاستنتاج من توزيع كاريما الموزون[18][19] (Weighted Garima distribution) ومن ثم تم تركيب توزيع بواسون مع توزيع ايرلنك الخليط لنحصل على توزيع مركب جديد(ايرلنك الخليط- بواسون) كانموذج لبيانات العد عالية التشتت , كما لايخفى ان نوع المقدر لمعلمات التوزيع له دور كبير الوصول لافضل ملائمة للبيانات , ان هدف البحث تحــديد طريقــة التقديــر المثلــى لتقدير معالم التوزيع المركب الجديد من بعض طرق التقدير .

2- الجانب النظري :

2-1- توزيع ايرلنك الخليط- بواسون: (Mixture Erlang - Poisson)(XEL-P) dist.

هو توزيع مركب متقطع دالة الكتلة الاحتمالية للتوزيع (p.m.f) كالتالي :

$$f(x;\alpha,\beta) = \frac{\alpha}{\alpha+\beta+1}\binom{x+\alpha-1}{x} p^{\alpha}(1-p)^{x} + \frac{\beta+1}{\alpha+\beta+1}\binom{x+\alpha-1}{x} p^{x}(1-p)^{\alpha} \ldots(1)$$

$$p = \frac{1}{1+\beta}\ ; \beta > o\ ; \quad x = 0,1,2,3,\ldots\ ;\ \alpha = 1,2,3,\ldots$$

وتم التوصل للتوزيع (صيغة (1)) حسب الخطوات :

اولاً : تم اقتراح توزيع ايرلنك الخليط (XEL) (Mixture Erlang) حيث تم اشتقاقه بالاستنتاج من توزيع كاريما الموزون (WG)[19] و كالتالي :

$$f(x;\alpha,\beta) = \frac{\beta^{\alpha}}{(\beta+\alpha+1)}\frac{x^{\alpha-1}}{\Gamma(\alpha)}(1+\beta+\beta x)e^{-\beta x} \quad .... \ (2)$$

$$f(x;\alpha,\beta) = \frac{\alpha}{\alpha+\beta+1}\ \text{gamma}(\alpha+1,\beta) + \frac{\beta+1}{\alpha+\beta+1}\ \text{gamma}(\alpha,\beta) \quad .... \ (3)$$

حيث ان :

$$\text{gamma}(\alpha,\beta) = \frac{\beta^{r}}{\Gamma r}x^{r-1}\ e^{-\beta x}$$

يمكن صياغة (3)على صيغة شجرة بيانية (Tree diagram) وكالتالي :

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{\alpha}{\alpha+\beta+1} \longrightarrow \text{gamma}(\alpha+1,\beta) \\[2ex] \dfrac{\beta+1}{\alpha+\beta+1} \longrightarrow \text{gamma}(\alpha,\beta) \end{cases} \quad .... \ (4)$$

$$x > 0 \ ; \ \alpha > 0; \ \beta > 0$$

حيث يمكن ان يمثل التوزيع عملية رمي عملة معدنية مرة واحدة احتمال النجاح $p = \frac{\alpha}{\alpha+\beta+1}$ والذي يؤدي الى ان المتغيرالعشوائي x يتبع توزيع $\text{gamma}(\alpha+1,\beta)$ و احتمال الفشل $1-p = \frac{\beta+1}{\alpha+\beta+1}$ والذي يؤدي الى ان المتغيرالعشوائي x يتبع توزيع $\text{gamma}(\alpha,\beta)$

لو اعدنا صياغة الشجرة البيانية (tree diagram) باستبدال توزيع $\text{gamma}(\alpha+1,\beta)$ بتوزيع $\text{Erlang}(\alpha,\theta)$[17] و توزيع $\text{gamma}(\alpha,\beta)$ بتوزيع $\text{Erlang}(\alpha,\beta)$ اي :

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{\alpha}{\alpha+\beta+1} \longrightarrow \text{Erlang}(\alpha,\theta) \\[2ex] \dfrac{\beta+1}{\alpha+\beta+1} \longrightarrow \text{Erlang}(\alpha,\beta) \end{cases} \quad .... \ (5)$$

$$x > 0 \ ; \ \theta = \frac{1}{\beta} \ ; \ \beta > 0 \ ; \ \alpha = 1,2,3,....$$

لنحصل بعد ذلك على دالة الكثافة الاحتمالية الاتية :

$$f(x;\alpha,\beta) = \frac{\alpha}{(\alpha+\beta+1)}\ \frac{1}{\Gamma\alpha\ \beta^{\alpha}}\lambda^{\alpha-1}\ e^{\frac{-\lambda}{\beta}} + \frac{(\beta+1)}{(\alpha+\beta+1)}\ \frac{\beta^{\alpha}}{\Gamma r}x^{\alpha-1}\ e^{-\beta x} .....(6)$$

ثانياً : يتم تركيب توزيع ايرلنك الخليط (XEL) مع توزيع بواسون و كالتالي :

نفرض ان : $\lambda \sim XEL(\alpha,\beta)$

هذا يؤي الى :

$$P_r(\lambda)= \frac{\alpha}{(\alpha+\beta+1)}\ \frac{1}{\Gamma\alpha\,\beta^{\alpha}}\lambda^{\alpha-1}\,e^{\frac{-\lambda}{\beta}} + \frac{(\beta+1)}{(\alpha+\beta+1)}\ \frac{\beta^{\alpha}}{\Gamma r}x^{\alpha-1}\ e^{-\beta x}\ ;\ \lambda>0 .$$

$$p(x|\lambda) = \frac{\lambda^{x}\,e^{-\lambda}}{x!} \qquad ; x = 0,1,2,3,.. \quad \lambda>0$$

$$\because\ P(x,\lambda) = p(\lambda).\,p(x|\lambda)$$

$$\therefore\ P(x)=\int_0^{\infty} P(x,\lambda)\ d\lambda$$

لنحصل بعد ذلك على الصيغة (1) .

2-2 - خصائص توزيع خليط ايرلنك – بواسون :

Properties of Mixture Erlang – Poisson (XEL-P) distribution

خصائص توزيع(XEL-P) تكون كالتالي :

1- $Ex = \frac{\alpha^2\beta}{\alpha+\beta+1} + \frac{(\beta+1)\alpha}{(\alpha+\beta+1)\beta}$............. (7) (التوقع)

2- $Ex^2 = \frac{\alpha}{\alpha+\beta+1}(\alpha\beta + \alpha(\alpha+1)\beta^2) + \frac{\beta+1}{\alpha+\beta+1}(\frac{\alpha}{\beta} + \frac{\alpha(\alpha+1)}{\beta^2})$........ (8)

3- $Ex^3 = \frac{\alpha}{\alpha+\beta+1}(\ \alpha\beta + 3\ \alpha(\alpha+1)\beta^2 + \alpha(\alpha+1)(\alpha+2)\ \beta^3) + \frac{\beta+1}{\alpha+\beta+1}\ (\frac{\alpha}{\beta} + 3\frac{\alpha(\alpha+1)}{\beta^2} + \frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3})$ .... (9)

4- $Ex^4 = \frac{\alpha}{\alpha+\beta+1}(\alpha\beta + 7\ \alpha(\alpha+1)\beta^2 + 6\alpha(\alpha+1)(\alpha+2)\beta^3 + 7\alpha(\alpha+1)(\alpha+2)(\alpha+3)\beta^4) + \frac{\beta+1}{\alpha+\beta+1}(\frac{\alpha(1-p)}{p} + \frac{7\alpha(\alpha+1)(1-p)^2}{p^2} + \frac{6\alpha(\alpha+1)(\alpha+2)(1-p)^3}{p^3} + \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)(1-p)^4}{p^4})$ .... (10)

$$V(x) = \frac{\alpha}{\alpha+\beta+1}\ (\alpha\beta + \alpha(\alpha+1)\beta^2) + \frac{\beta+1}{\alpha+\beta+1}\left(\frac{\alpha}{\beta} + \frac{\alpha(\alpha+1)}{\beta^2}\right) - (\frac{\alpha^2\beta}{\alpha+\beta+1} + \frac{(\beta+1)\alpha}{(\alpha+\beta+1)\beta})^2 \text{ .... (11)}$$

5- $Skewness(x) = \frac{E(X-E\,X)^3}{(V(x))^{3/2}}$

$$= \frac{2(E\,X)^3 - 3E\,X\ E\,X^2 + E\,X^3}{(V(x))^{3/2}} \text{ .... (12)}$$

من تعويض المعادلات(7) و(8) و (9) و (11) في (12) نحصل على معامل الالتواء (Skewness) .

6- $Kurtosis\ (x) = \frac{E(X-E\,X)^4}{(V(x))^2}$

$$= \frac{(E\,X)^4 - 4EX(EX)^3 + 6E\,X^2\,(EX)^2 - 4EX^3EX + E\,X^4}{(V(x))^2} \text{ .... (13)}$$

من تعويض المعادلات(7) و(8) و (9) و(10) و(11) في (12) نحصل على معامل التفلطح (Kurtosis) .

ان الدالـــة التراكمـــية لتوزيع (XEL-P) كالتالـــي :

$$F(x;\alpha,\beta)=\frac{\alpha}{\alpha+\beta+1}\sum_{u=0}^{x}\binom{u+\alpha-1}{u}p^{\alpha}(1-p)^{u}+\frac{\beta+1}{\alpha+\beta+1}\sum_{u=0}^{x}\binom{u+\alpha-1}{u}p^{u}(1-p)^{\alpha}$$

$P=\frac{1}{\beta+1}$; $\beta>0$; $\alpha=1,2,3,\ldots$; $x=0,1,2,3,\ldots$ .... (14)

حيث ان معكوس الدالة التراكمية لايمكن ايجادها جبريا , ولتوليد بيانات التوزيع المركب (XEL-P) لغرض تقدير المعلمات اقترح الباحث الاعتماد على توليد بيانات توزيع المركب(NB) وتقدير معلماته مستنادا الى امكانية صياغة الصيغة (1) على صيغة شجرة بيانية اي ان :

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{\alpha}{\alpha+\beta+1}\longrightarrow NB(\alpha,p)\\[2ex]\dfrac{\beta+1}{\alpha+\beta+1}\longrightarrow NB(\alpha,1-p)\end{cases}\qquad\ldots\ldots\ldots(15)$$

$x=0,1,2,3,\ldots$; $p=\frac{1}{\beta+1}$; $\beta>0$; $\alpha=1,2,3,\ldots$

حيث كما يظهر من صيغة الشجرة البيانية عند تقدير معالم توزيع NB(α, p) نكون بذلك قد قدرنا معالم توزيع XEL-P .

2-3- طـــــــــرق التقـــــــــدير : Estimation Methods

سيتم اللجوء الى تقدير معالم ثنائي الحدين السالب المركب ($p=\frac{1}{1+\beta}$; NB(α,p)) لتقدير معالم توزيع XEL-P , وبما ان المعلمة α مجالها يتمثل بالاعداد الصحيحة الموجبة لذالك سيتم تقريب قيمة المعلمة المقدرة $\hat{\alpha}$ لاقرب عدد صحيح موجب , وان طرق التقدير كالتالي :

2-3- 1 - طريقة مقدر العزوم: Moment Estimation Methods (mom)

تعتبر من الطرق الشائعة في التقدير[10], تنطلق هذه الطريقة من مبدأ مساوات عزم المجتـــمع لعزم العينة , اذا كان ($x_1, x_2, x_3, \ldots, x_n$) تمثل مشاهدات عينة عشوائية مستـقـلـة ومتماثلة التوزيع (iid) تتبع توزيع $NB(\alpha\ ,p=\frac{1}{1+\beta})$ فـــان مقدرات طريقة العزوم لتوزيع NB تحسب من المعادلات الاتية :

$$Ex=\bar{x}\ \ \ldots\ldots\ (16)$$

$$E\,x^2=\frac{\sum_{i=1}^{n}xi^2}{n}\ \ \ldots\ldots(17)$$

نعوض المعادلتين (16) و (17) في (18)

$$v(x)=\frac{\sum_{i=1}^{n}xi^2}{n}-\bar{x}^2\ \Longrightarrow\ v(x)=S^2\qquad\ldots\ldots(18)$$

لذلك فان مقدرات العزوم للمعلمات β و α كالتالي :

$$\hat{\beta}=\frac{S^2-\bar{x}}{\bar{x}}\qquad\ldots\ldots(19)$$

$$\hat{\alpha}=\frac{(\bar{x})^2}{S^2-\bar{x}}\qquad\ldots\ldots(20)$$

المعادلتين (19) و(20) يمثلان مقدرات لتوزيع ثنائي الحدين السالب NB و توزيع (XLE-P) في ان واحد .

2-3-2- طريقة مقدر الامكان الاعظم :

Maximum Likelhood Estimation Methods (MLE)

ان مقدرات الامكان الاعظم (10)(15) لتوزيع ثنائي الحدين السالب المركب NB يتم احتسابها باعتماد برنامج الماتلاب من الايعاز nbinfit او mle لاحتساب $\hat{p}$ و $\hat{\alpha}$, و ان : $\hat{\beta}(\frac{1-\hat{p}}{\hat{p}})=$.

2-3-3- طريقة مقدر التقلص: **Shrinkage Estimation Method**

الصيغة العامة لمقدر طريقة التقلص(3)(5)(8)(11)(12)(14) :

$$\hat{\theta}_{sh}= k\,\hat{\theta} + (1-k)\,\theta_\circ \; ; \qquad 0< k <1 \;........(21)$$

$\hat{\theta}_{sh}$: يمثل مقدر طريقة التقلص .

$\hat{\theta}$: يمثل مقدرغير متحيز او مقدر اي طريقة تقدير .

$\theta_\circ$: قد تمثل مقدر طريقة تقدير اخرى او معلومات سابقة حول العينة .

k : تمثل قيمة احتمالية تحدد من قبل الباحث استنادا الى معلومات مسبقة حول العينة او كما هو معروف ان عملية تقدير هذه لابد ان يصاحبها خطأ وهو ابتعاد المقدر $\hat{\theta}_{sh}$ عن المعلمة θ بشكل عام لذلك من الممكن ان نحـدد قيمة k بحيث تجـعل متوسـط مربعات الخطــأ التجريبي MSE لل $\hat{\theta}_{sh}$ اقل ما يمكن اي ان :ــ

$$K=\frac{[\theta_\circ-\theta]^2}{[\,MSE(\hat{\theta})+[\theta_\circ-\theta]^2]} \;...................(22)$$

في الجانب التطبيقي قيمة (k 0.999 =) .

اما فيما يخص توزيع (NB) فان مقدرات التقلص :

$$\boldsymbol{\hat{\alpha}_{sh}= k1\,\hat{\alpha}_{mle} + (1-k1)\,\hat{\alpha}_{mom} \;...................(23)}$$

$$\boldsymbol{K1=\frac{[\hat{\alpha}_{mom}-\alpha]^2}{[\,MSE(\hat{\alpha}_{mle})+[\hat{\alpha}_{mom}-\alpha]^2]} \;...................(24)}$$

$$\boldsymbol{\hat{\beta}_{sh}= k2\,\hat{\beta}_{mle} + (1-k2)\,\hat{\beta}_{mom} \;...................(25)}$$

$$\boldsymbol{K2=\frac{[\hat{\beta}_{mom}-\beta]^2}{[\,MSE(\hat{\beta}_{mle})+[\hat{\beta}_{mom}-\beta]^2]} \;...................(26)}$$

4-2 - اختبار مربع كاي χ^2 لحسن المطابقة: Test of chi-square χ^2 for Goodness of fit:

اولا: فرضـــــــــية الاختبار (2) (16) (17):

فرضية العـــــــدم : البيانات تتبع التوزيع (بواسون او NB او XLE-P) .

الفرضية البديلة : البيانات لا تتبع التوزيع (بواسون او NB او XLE-P) .

: The data followed the distribution(poisson or NB or XEL) H_o

: The data not followed the distribution(poisson or NB or XEL) H_1

ثانيا : احصاءة الاختبار:

$$\chi^2=\sum_{i=1}^{k}\frac{(Oi-\mathrm{Ei})^2}{\mathrm{Ei}} \sim \chi^2_{(k-p-1,\alpha)}$$

K : يمثل عدد مستويات(fi) في عدم وجود تعديل او عدد مستويات(fi) بعد التعديل .

p : تمثل عدد معلمات التوزيع التي تم تقديرها .

Ei : يمثل التكرار النظري ويحسب كالتالي : Ei = N*p(xi)

N : تمثل مجموع التكرارات الكلي .
ان فكرة الاختبار تنطلق من ان التكرار الحقيقي (fi) يتبع توزيع بواسون(تحت شرط ان مستويات x مستقلة) وان التكرار النظري (Ei) يمثل القيمة المتوقعة للتكرار (fi) التي تكون منحازة لقيمة التكرار الحقيقي (fi) وحسب مبرهنة الغاية المركزية و تحت شرط (Ei>5 (للوصول الى التماثل لتوزيع بواسون)) يتم التوصل لصيغة الاختبار .
وفي حالة عدم تحقق الشرط (Ei<5) نلجأ اجراء تعديل للبيانات والمتمثلة بجمع التكرار النظري لمشاهدتين متتاليتين او اكثر(جمع التكرار النظري(Ei<5) مع ما يليه من التكرارات النظرية Ei وقد نلجأ الى جمع مجموع الناتج مع التكرار النظري السابق في حالة استمرار عدم تحقق الشرط) حتى نصل الى تكرار نظري جديد(Ei>5) ينوب عن التكرارات النظرية للمشاهدات التي تم جمعها, وبالمقابل يتم جمع التكرارات الحقيقية المناظرة للتكرارات النظرية التي تم جمعها, فيكون عدد التكرارات النظرية او الحقيقية كالتالي :

عدد التكرارات النظرية او الحقيقية = عدد التكرارات قبل التعديل – عدد التكرارات الداخلة في عملية الجمع 1+.

ثالثا : قاعدة القـــــرار:

نرفض فرضية العدم Ho عندما تكون القيمة المحتسبة لـχ^2 اكبر من القيمة الجدولية لـχ^2 اي ان البيانات لا تتبع التوزيع المفترض , والعكس صحيح .

5- الجانب التجريبي :

ويتمثل الجانب التجريبي بالمحاكاة (Simulation) حيث ان هذا الاسلوب يؤدي الى تحديد المقدر الجيد باعتماد بيئة افتراضية عوضا عن خاصية الكفاءة (effeciency) التي كان يعتمدها (Fisher)[9] في تحديد المقدر الجيد و نظرا لما يمتلك اسلوب المحاكاة من امكانيات مثلا السرعة والدقة , اضافتا الابتعاد عن المعادلات الجبرية الصعبة والتي من الممكن ان نواجهها عند اعتماد خاصية (الكفاءة) حيث قد ينتهي بنا المطاف بعدم الوصول الى المبتغى , للمحاكـــــاة طرائق متنوعة ومختلفة ابرزها وأكثرها اسـتعمــــالا طريقــــة مونات كارلـو (Monte-Carlo) [6][7] فيما يخص البحث تتلخص الطريقة بالخطوات الاتية :

1- توفير قيم مختلفة لمعلمات توزيع (NB) واحجام عينات مختلفة لاحظ الجدول (1) .

جدول (1): يبين احجام العينات مع قيم المعالم

n	α	β
30	5	15
50	2	2
75	3	4
100	7	10

2- ايجاد دالة الكثافة لتوزيع (XLE-P).

3- توليد بيانات توزيع (NB) وتقدير معلمات توزيع (NB) .

4- ايجاد دالة الكثافة المقدرة لتوزيع (XLE-P) .

5- تكرار التجربة(T=1000) مرة ومن ثم احتساب المقياس (Mse)(متوسط مربعات الخطأ) للمعلمات المقدرة ولدالة الكثافة المقدرة لتوزيع (XLE-P) ولكل طرق التقدير والصيغة العامة لمعيار (Mse) كالاتي :

$$\text{Mse} = \frac{\sum_{i=1}^{T}(\hat{\theta}_i - \theta)^2}{T}$$

(في حالة $\hat{\theta}_i$ تمثل المعلمة المقدرة للمعلمة θ عند المستوى i)

$$Mse = mean(\frac{\sum_{i=1}^{T}(\hat{f}_i - f)^2}{T})$$

(في حالة $\hat{f}_i$ تمثل دالة الكثافة لدالة الكثافة لتوزيع (XLE-P) , وان mean تمثل الوسط الحسابي للقيم) .

و ان المفاضلة بين المقدرات فيما يخص مجال بحثنا من الناحية التجريبية تكون على اساس ان دالة الكثافة المقدرة التي تمتلك اقل Mse يعتبر مقدرها الافضل .

5-1- مناقشة نتائج تجربة المحاكاة لتوزيع (XLE-P) :

Discuss The Results Of The Simulation Experiment of (XLE-P) dist.

في هذه الفقرة يتم عرض نتائج المحاكاة والمتمثلة بتقدير المعلمات وحساب متوسط مربعات الخطأ لدالة الكثافة لتوزيع (ELX-P), و كما في الجداول من (2) الى (5) :

جدول(2) : يبين قيم مقدرات المعلمات (ELX-P) وقيم (MSE) للمعلمات و لدوال الكثافة لكافة طرائق التقدير وحجوم العينات عندما $\alpha = 5$, $\beta = 15$

n	Method		α	β	P.d.f
30	MOM	Est.	6	14.687910	
		Mse	3.342	20.538758	0.001155
	MLE	Est.	6	14.416756	
		Mse	3.346	17.861707	0.001114
	SHRINK	Est.	6	14.515592	
		Mse	3.143953	17.050866	0.00104
	Best		SHRINK	SHRINK	SHRINK
50	MOM	Est.	5	14.744519	
		Mse	1.899	12.514231	0.000434
	MLE	Est.	5	14.586207	
		Mse	1.7390	10.773151	0.000402
	SHRIK	Est.	5	14.640025	
		Mse	1.617332	10.336855	0.000393
	Best		SHRINK	SHRINK	SHRINK
75	MOM	Est.	5	14.882938	
		Mse	1.209	9.005142	0.000205
	MLE	Est.	5	14.76856	
		Mse	1.063	7.417563	0.0001761
	SHRINK	Est.	5	14.807788	
		Mse	0.991517	7.280665	0.0001757
	Best		SHRINK	SHRINK	SHRINK
100	MOM	Est.	5	14.942446	
		Mse	0.813	6.738881	0.0001095
	MLE	Est.	5	14.819352	
		Mse	0.697	5.545577	0.000094
	SHRINK	Est.	5	14.86381	
		Mse	0.615589	5.406483	0.000101
	Best		SHRINK	SHRINK	MLE

جدول(3) : يبين قيم مقدرات المعلمات (ELX-P) وقيم (MSE) للمعلمات و لدوال الكثافة لكافة طرائق التقدير وحجوم العينات عندما $\beta=2$, $\alpha=2$.

n	Method		α	β	P.d.f
30	MOM	Est.	3	1.98664	
		Mse	3.495	0.910634	0.001869
	MLE	Est.	3	1.974214	
		Mse	4.024	0.824138	0.001835
	SHRINK	Est.	3	1.953518	
		Mse	3.904812	0.749163	0.001541
	Best		MOM	SHRINK	SHRINK
50	MOM	Est.	2	1.98869	
		Mse	1.084	0.565939	0.000834
	MLE	Est.	2	1.982397	
		Mse	0.984	0.490287	0.000781
	SHRINK	Est.	2	1.965027	
		Mse	0.907715	0.459312	0.000685
	Best		SHRINK	SHRINK	SHRINK
75	MOM	Est.	2	1.985813	
		Mse	0.585	0.382407	0.000411
	MLE	Est.	2	1.984521	
		Mse	0.499	0.323689	0.000359
	SHRINK	Est.	2	1.968978	
		Mse	0.441229	0.30666	0.000319
	Best		SHRINK	SHRINK	SHRINK
100	MOM	Est.	2	1.983062	
		Mse	0.444	0.291860	0.000248
	MLE	Est.	2	1.984928	
		Mse	0.369	0.244074	0.000213
	SHRINK	Est.	2	1.971945	
		Mse	0.308856	0.233869	0.000198
	Best		SHRINK	SHRINK	SHRINK

جدول (4) : يبين قيم مقدرات المعلمات (ELX-P) وقيم (MSE) للمعلمات و لدوال الكثافة لكافة طرائق التقدير وحجوم العينات عندما $\alpha=3$, $\beta=4$

n	Method		α	β	P.d.f
30	MOM	Est.	3	3.974843	
		Mse	2.647	2.290483	0.001483
	MLE	Est.	4	3.908948	
		Mse	2.687	2.027475	0.001385
	SHRINK	Est.	3	3.906503	

		Mse	2.538522	1.901138	0.253037
	Best		SHRINK	SHRINK	MLE
50	MOM	Est.	3	3.944464	
		Mse	1.194	1.420479	0.000607
	MLE	Est.	3	3.899363	
		Mse	1.093	1.221773	0.000555
	SHRIK	Est.	3	3.902565	
		Mse	0.991105	1.158035	0.000521
	Best		SHRINK	SHRINK	SHRINK
75	MOM	Est.	3	3.938177	
		Mse	0.78	0.968208	0.000305
	MLE	Est.	3	3.907224	
		Mse	0.647	0.766096	0.000250
	SHRINK	Est.	3	3.903672	
		Mse	0.586057	0.739955	0.000241
	Best		SHRINK	SHRINK	SHRINK
100	MOM	Est.	3	3.970668	
		Mse	0.539	0.734837	0.000179
	MLE	Est.	3	3.945	
		Mse	0.465	0.589074	0.001593
	SHRINK	Est.	3	3.943777	
		Mse	0.403006	0.570555	0.000192
	Best		SHRINK	SHRINK	SHRINK

جدول(5) : يبين قيم مقدرات المعلمات (ELX-P) وقيم (MSE) للمعلمات و لدوال الكثافة لكافة طرائق التقدير وحجوم العينات عندما α =7 , β =10

n	Method		α	β	P.d.f
30	MOM	Est.	8	9.878572	
		Mse	8.09	9.495185	0.00106
	MLE	Est.	8	9.593481	
		Mse	8.527	8.059947	0.000977
	SHRINK	Est.	8	9.706206	
		Mse	8.186646	7.83286	0.27542
	Best		MOM	SHRINK	MLE
50	MOM	Est.	7	9.928887	
		Mse	3.843	5.929546	0.000431
	MLE	Est.	8	9.758753	
		Mse	3.769	5.331918	0.000404
	SHRINK	Est.	8	9.845069	
		Mse	3.558503	5.149475	0.000394
	Best		SHRINK	SHRINK	SHRINK
75	MOM	Est.	7	9.938672	

		Mse	2.261	4.101668	0.000202
	MLE	Est.	7	9.808162	
		Mse	2.217	3.610367	0.000189
	SHRINK	Est.	7	9.861247	
		Mse	2.039303	3.484746	0.000182
	Best		SHRINK	SHRINK	SHRINK
100	MOM	Est.	7	9.9746	
		Mse	1.604	3.047656	0.000117
	MLE	Est.	7	9.890991	
		Mse	1.518	2.648862	0.000105
	SHRINK	Est.	7	9.920558	
		Mse	1.40219	2.581459	0.000109
	Best		SHRINK	SHRINK	MLE

يظهر من الجداول ما يلي :

1- عندما ($\alpha = 5,\ \beta = 15$) وعند ($n = 30 , 50 , 75$) تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر التقلص و عند($n = 100$) تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر الامكان الاعظم .

2- عندما ($\alpha = 2,\ \beta = 2$) وعند كل قيم n تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر التقلص .

3- عندما ($\alpha = 3 ,\ \beta = 4$) وعند ($n = 30$) تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر الامكان الاعظم و عند ($n = 50 , 75 , 100$) تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر التقلص .

4- عندما ($\alpha = 7,\ \beta = 10$) وعند ($n = 30 , 100$) تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر الامكان الاعظم و عند ($n = 50 , 75$) تكون قيمة (mse) لدالة (pdf) المقدرة الاقل لمقدر التقلص .

اما نسبة الافضلية لطرائق تقدير المعلمات لدالة الكثافة الاحتمالية لتوزيع (XLE-P) كما في جدول (6) :

جدول (6) : يبين عـــدد مــــرات الافضلية ونسبة الافضلية لكـــل طريقة تقديـــر

الطريقة	عدد مرات الافضلية	نسبة الافضلية
مقدر التقلص	12	$\frac{12}{16}$
مقدر الامكان الاعظم	4	$\frac{4}{16}$

6- الجانب التطبيقي :

يتمثل بتطبيق ماتم ذكره في الجانب النظري من المقدرات لتوزيع (XLE-P) و اجراء تحليل للبيانات بتطبيق اختبار مربع كاي لحسن المطابقة (goodness of fit χ^2test) للتوصل الى نتائج من الواقع الحقيقي مع ملاحظة فيما اذا تطابق نتائج الواقع الحقيقي مع نتائج البيئة الافتراضية (المحاكاة) من حيث تحديد المقدر الجيد لتوزيع (XLE-P) .

البيانات التالية تمثل عـدد الاشخاص المصابين (fi) بمرض (COV-19) ولكل مدة رقود (xi) في مستشفى الزهراء التعليمي(ع) في محافظة واسط لفترة دخول من (14-3 ولغاية 10-7) 2020 و كما في الجدول الاتي :

جدول (7): يمثل عينة من البيانات الحقيقية لعدد المرضى الراقدين في المستشفى ولكل فترة زمنية

xi :	0	1	2	3	4	5	6	7	8
fi :	360	360	270	180	112	68	39	23	13
	9	10	11	12	13	14 - 31			
	7	4	2	1	1	0			

وان المدة (xi) تمثل المدة بالايام وكالتالي :

(xi =0) : مغادرة المستشفى في نفس اليوم (حالة اشتباه) .

(xi =1) : رقود المريض مدة يوم واحد .

(xi =2) : رقود المريض مدة يومين .

وهكـــــــــــــــــــــــذا

ان نتائج الاختبار للمقارنة بين انموذج توزيع (XLE-P) عند طرق التقدير المختلفة ونتائج توزيع بواسون و توزيع (NB) باعتماد مقدر (MLE) تكون كما في جدول (4) :

جدول (8) : يمثل نتائج الاختبار مربع كاي لحسن المطابقة

Dist.	h	estimator	p-value	stat
XLE-P	0	MOM	0.99999997703302	0.054794548188083
	0	MLE	0.999999979734956	0.053097572690724
	0	shrinkage	0.999999979733564	0.053098489525517
Poisson	1	MLE	6.043805184354027e-114	5.424314456114412e+02
NB	0	MLE	0.999999970752690	0.058227999300140

حيث ان :

h : تمثل قرار قبول فرضية العدم عندما قيمتها تساوي (1) كما تمثل قرار رفض فرضية العدم عندما تكون قيمتها مساوية (0) .

p-value : تمثل احتمال قبول فرضة العدم وهي صحيحة .

Stat : تمثل احصاءات اختبار مربع كاي لحسن المطابقة .

7- الاســتنتاجات : Conclutions

اظهرت نتائج تحليل البيانات الحقيقية والمتمثلة باجراء اختبار مربع كاي χ^2 لحسن المطابقة الاتي :

- اظهرت نتائج المحاكات تفوق مقدر التقلص يلية مقدر الامكان الاعظم ومن ثم مقدر العزوم .
- ان مقدر الامكان الاعظم للبيانات الحقيقية (بيانات الراقدين) يكون المقدر الافضل يليه مقدر التقلص ومن ثم مقدر العزوم .
- ان توزيع (ايرلنك الخليط – بواسون) (XLE-P) قد تفوق على كل من توزيع ثنائي الحدين السالب المركب (NB) و توزيع بواسون .

8- التوصـــيات : Recommendations

- نوصي باعتماد توزيع (ايرلنك الخليط ـ بواسون) (XLE-P) لنمذجة مدة اقامة مرضى كورونا COV-19 .
- نوصي باعتماد مقدر الامكان الاعظم في تقدير معالم توزيع (ايرلنك الخليط ـ بواسون) (XLE-P) الا ان اختيار مقدر الامكان لا يقلل من اهمية مقدر التقلص حيث اظهرت نتائج المحاكات نسبة افضلية مقدر الامكان $(\frac{4}{16})$ في حين ان نسبة الافضلية لمقدر التقلص $(\frac{14}{16})$.

المصـــــادر :

اولا : المصادر العربية : **Arabic References**

1. صالح , بو عبد الله (2005 -2006) " محاظرات الاحصاء الرياضي لطلبة كلية العلوم الاقتصادية" https://drive.google.com/uc?export=download&id=1EAfr9rQRMEijB9zIWY0efhUp6OVYw8nU .
2. المشهداني , محمود حسن . هرمز , امير حنا (1989) " الاحصاء " , مديرية مطبعة التعليم العالي , الموصل , العراق .
3. هرمز , امير حنا (1990) "الاحصاء الرياضي" , مديرية مطبعة التعليم العالي , الموصل, العراق .
4. الجاسم ,صباح هادي . عبد , فراس صدام (2004) " مقارنة بين طريقة التقلص و طريقة الامكان الاعظم لتقدير المعلمات و دالة المعولية لتوزيع ويبل بثلاث معلمات بأستخدام المحاكاة " مجلة العلوم الاقتصادية و الادارية , المجلد 19 , العدد 73 .
5. نصار , عامر فاضل (2010)" اقتراح دالة التقلص لحالة المشاهدة الواحدة في مشكلة $N(\theta,1)$" مجلة جامعة الانبار للعلوم الصرفة المجلد (4), العدد (1) .
6. نعيمة , علي بندر (2016)" تقدير دوال الفشل للتوزيع الناتج من دمج توزيع بواسون ليندلي مع توزيعات اخرى " اطروحة مقدمة الى كلية الادارة والاقتصاد/ جامعة بغداد للحصول على درجة الدكتوراه " فلسفة في الاحصاء ".
7. محمد , نور اياد (2017) " تقدير معلمات توزيع بواسون المركب مع تطبيق عملي" رسالة مقدمة الى كلية الادارة والاقتصاد/ جامعة بغداد للحصول على درجة الماجستير في علوم الاحصاء .
8. الباقر, زينب محمد باقر صادق (2017) " تقدير دالة المعولية لتوزيع بواسون مع التطبيق عملي" رسالة مقدمة الى كلية الادارة والاقتصاد/ جامعة بغداد للحصول على درجة الماجستير في "علوم الاحصاء" .

ثانيا : المصادر الاجنبية : **Foreign References**

9. Fisher,R . A .,& S, F. R.(1941). The Negative Binomial Distribution. Annals of Eugenics ,Vol.11 , No.1 , PP.182-187.
10. Mood, A. M. , Graybill, F. A., & Boes , D.C. (1974) . Intrduction to the Theory of Statistics . MC Graw – Hill International Book Company – Singapore.
11. Pandey,M. & S.K. Upadhyay.(1988).Bayes shrinkage Estimation of Weibull Parameters, IEEE Transactions on Reliability ,Vol.,96, PP .1619 -1626 .
12. Tiejun,T .,Yuedong ,W.(2005).Optimal Shrinkage Estimation of Variance with Applications to Microarray Data Analysis .
13. Best,Donald.John.,Rayner,John.C.W., Olivier Thas (2009) . Anscombe's Tests of Fit for the Negative Binomial Distribution . Journal of Statistical Theory and Practice,Vol.3 , No.3 , PP.555-565.
14. Peter,H. (2013).Shrinkage estimators.
15. Hogg,Robert.V.,McKean,Joseph.W.,Craig,Allen.T(2013). Introduction to Mathmatical Statistics . Pearson Education Inc.
16. Balakrishnan,N.,Voinov,Vassilly.,Nikulin,M.S.(2013).Chi-Squared Goodness of Fit Test with Applications.Academic Press.
17. Kongrod,Siriporn.,Bodhisuwan,Winai.,Payakkapong,Prasit .(2014). The Negative Binomial-Erlang Distrbution with Applications. International Journal of Pure & Applied Mathematics. Vol.92,No.3, PP.389-401.
18. Shanker,Rama.(2016) .Garima Distribution and its Application to Model Behavioral Science Data.Biom Biostat Int J ,Vol.4 , No.7 ,PP.1-9 .

19. Shanker,Rama.(2018) . A two-parameter weighted Garima distribution with properties and application. Biom Biostat Int J,Vol.7 , No.3 , PP.234-242.
20. Bodhisuwan,Winai.,Seangthong,Pornpop.(2020) .The Negative Binomial – Weighted Garima Distribution : Model,properties & Applications . Bakistan journal of statisticsand Operation Research. Vol.16 ,No.1, PP.1-10 .