تقارب توزيع Poisson المعدل المقترح من توزيع Binomial

An approximation of suggested modified Poisson distribution to Binomial distribution

ا.م.د.فياض عبدالله علي

قسم الاحصاء / جامعة واسط

مستخلص

يعد توزيع Poisson تقريب لتوزيع Binomial عندما يقترب حجم العينة من مالا نهاية وان احتمال النجاح P يؤول إلى الصفر ، ولكن ماذا يحدث عندما يقترب حجم العينة من مالا نهاية واحتمال النجاح P يقترب من النصف . ومن هنا جاء هدف البحث في اقتراح دالة Poisson معدلة لتكون مقاربة لتوزيع Binomial وقد تم برهنة ان دالة Poisson المعدلة هي دالة كثافة احتمالية (p.d.f) كما تم اشتقاق الدالة المولدة للعزوم (m.g.f) لهذا التوزيع المعدل ومن خلال الدالة المولدة للعزوم تم احتساب الوسط الحسابي والتباين. وقد تبين ان توزيع Poisson المعدل المقترح يتقارب إلى توزيع Binomial بشكل أفضل من توزيع Poisson عندما يكون احتمال النجاح P قريب من النصف وحجم العينة يقترب من مالا نهاية .

Abstract

Poisson Distribution is an approximation to Binomial distribution when the sample size approach to infinity and the probability of success P is a small , but the question a rise when the sample size approach to infinity but the probability of success P is close to $\frac{1}{2}$, we suggest modification on Poisson Distribution to get new function and prove that it is p.d.f and calculate mean and variance .

We explain that the suggestion modified Poisson distribution is more approximation than ordinary Poisson when the probability of success P close to $\frac{1}{2}$ and the sample size approach to infinity.

1. **مقدمة**

يعد توزيع Binomial وتوزيع Poisson من أهم التوزيعات الاحتمالية المنفصلة وقد حظي هذين التوزيعين باهتمام كبير من قبل الإحصائيين فكثرت البحوث التي تناولت تقارب هذه التوزيعات من توزيعات أخرى فقد قام Wallis[٦] بمحاولة جديدة لتقريب توزيع Binomial المتقطع من توزيع Normal المستمر تحت شروط معينة باستخدام عدد من الطرق الإحصائية والاختبارات ، فيما حاول Kanint[٣] تقريب توزيع Poisson من توزيع Beta- Negative Binomial {BN(α,β,n)} ، فيما حاول نفس الباحث Kanint[٤] [ ] وضع نظرية لتقريب توزيع Poisson من توزيع Beta-Binomial {BB((α,β,n)} .

٢. **هدف البحث**

تركزت فكرة البحث في محاولة إيجاد دالة توزيع جديدة (مشتقة من توزيع Poisson) وبرهنتها على أنها دالة كثافة احتمالية (p.d.f) وتقريبها إلى توزيع Binomial عندما يكون حجم العينة قريبا من مالا نهاية واحتمال النجاح P قريبا من $\frac{1}{2}$.

٣. **توزيع** Binomial

يسمى أيضا محاولات Bernoulli نسبة للعالم السويسري Jacob Bernoulli مكتشف التوزيع المسمى باسمه (١٧١٣) وقد اشتق توزيع Binomial من توزيع Bernoulli الذي هو عبارة عن توزيع احتمالي متقطع يأخذ القيمة ١ باحتمال P والقيمة صفر باحتمال (1-P) ويمكن توضيح الفكرة كالأتي :

لتكن X1,X2,........,Xn متغيرات عشوائية مستقلة بعضها عن البعض الآخر وكل منها لها توزيع Bernoulli باحتمال P فأن $\sum_{i=1}^{n} x_k$ تتوزع توزيع Binomial .

وحسب نظرية الاحتمال والإحصاء فأن توزيع Binomial توزيع احتمالي متقطع لعدد من حالات النجاح في تعاقب من n من تجارب تكون مخرجاتها إحدى حالتين النجاح أو الفشل .

توزيع Binomial يستخدم بكثرة لنماذج عدد النجاحات في عينة بحجم n مسحوبة بأرجاع من مجتمع بحجم N ويمتاز بما يلي :

١. التجربة مكونة من n من المحاولات المكررة .

٢. كل محاولة ينتج عنها حالتين فقط تسمى الأولى نجاح باحتمال P والثانية فشل باحتمال (1-P) .

٣. احتمال النجاح P ثابت في كل محاولة من المحاولات .

٤. المحاولات مستقلة بعضها عن البعض الآخر .

دالة الكثافة الاحتمالية لهذا التوزيع هي :

$$p(k) = C_k^n \theta^k (1-\theta)^{n-k}$$

حيث ان kεΩ وان Ω = 0,1,2,........n

وان الوسط الحسابي للتوزيع يساوي $n\theta$ بينما التباين يساوي $n\theta(1-\theta)$.

الدالة المولدة للعزوم (m.g.f) تساوي $(1+\theta+\theta.e^t)^n$.

مجالات استخدام التوزيع كثيرة تشمل الإحصاءات والأسئلة التي تعتمد فيها الإجابة نعم أو لا .

٤.توزيع Poisson

سمي هذا التوزيع توزيع Poisson نسبة للعالم الفرنسي Simeon Denis Poisson عندما نشر بحثه عن التوزيع عام ١٨٣٧ حيث اشتق كغاية لتوزيع Binomial عندما يؤول احتمال النجاح P إلى الصفر.

ويمكن اعتماد توزيع Poisson كتقريب لتوزيع Binomial إذا تحققت الشروط التالية [٢] :

١. حجم العينة كبير جدا يقترب إلى مالا نهاية .

٢. احتمال حدوث الحدث صغير جدا .

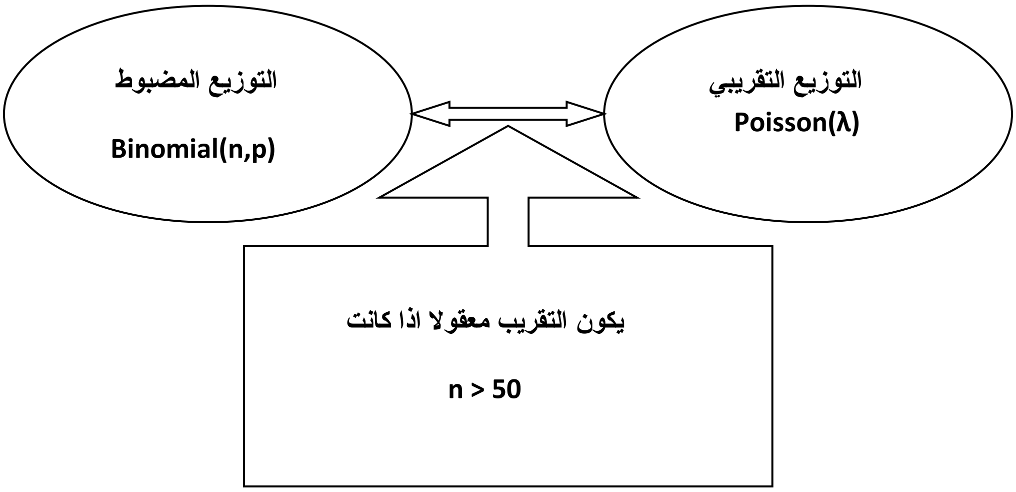
٣. الوسط الحسابي (λ) اقل من ٥ .

الدالة الاحتمالية لتوزيع Poisson هي

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

K ε Ω , Ω=0,1,2,........

والمخطط التالي يوضح متى وكيف ان توزيع Poisson ممكن استخدامه كتقريب لتوزيع Binomial .



الوسط الحسابي للتوزيع يساوي λ وكذلك التباين يساوي λ .

الدالة المولدة للعزوم (m.g.f.) للتوزيع تساوي λ(et-1) .

وقد استخدم التوزيع في مجالات عديدة مثل تكنولوجيا الاتصالات ومراقبة الجودة الإحصائية وفي الرصد الجوي وعلم الأحياء وبعض الظواهر التابعة لميدان التفكيك النووي المشع .

٥. توزيع Poisson المعدل المقترح

لتكن دالة Poisson المعدلة المقترحة تأخذ الشكل التالي

$$f(x)=\frac{e^{-\left\lceil \frac{\lambda}{2}\right\rceil}.\left\lceil \frac{\lambda}{2}\right\rceil^{x-\left\lceil \frac{\lambda}{2}\right\rceil}}{(x-\left\lceil \frac{\lambda}{2}\right\rceil)!}\ldots.x\geq\left\lceil \frac{\lambda}{2}\right\rceil$$

حيث ان x يأخذ قيم صحيحة موجبة

$\left\lceil \frac{\lambda}{2} \right\rceil$ أقل عدد صحيح اكبر أو يساوي $\frac{\lambda}{2}$

فان f(x) هي دالة كثافة احتمالية

البرهان

$$\sum_x f(x) = \sum_{x=\left\lceil \frac{\lambda}{2} \right\rceil}^{\infty} \frac{e^{-\left\lceil \frac{\lambda}{2} \right\rceil} \cdot \left\lceil \frac{\lambda}{2} \right\rceil^{x-\left\lceil \frac{\lambda}{2} \right\rceil}}{(x-\left\lceil \frac{\lambda}{2} \right\rceil)!} \ldots . x \geq \left\lceil \frac{\lambda}{2} \right\rceil$$

اجعل $\theta = \left\lceil \frac{\lambda}{2} \right\rceil$ وان $y = x - \left\lceil \frac{\lambda}{2} \right\rceil$ وبالتالي فان $y = x - \theta$

فان

$$\sum_y f(y) = \sum_{y=0}^{\infty} \frac{e^{-\theta} \theta^y}{y!} \ldots\ldots y = 0,1,2,\ldots\ldots$$

$$\sum_y f(y) = e^{-\theta} \sum_{y=0}^{\infty} \frac{\theta^y}{y!} \ldots\ldots y = 0,1,2,\ldots\ldots$$

$$\sum_y f(y) = e^{-\theta} . e^{\theta} = 1$$

وهذا يعني أن f(x) تحقق شروط دالة كثافة احتمالية للمتغير المتقطع x . المتغير العشوائي الذي يمتلك هذا التوزيع أطلقنا عليه تسمية modified-Poisson أو توزيع Poisson المعدل .

<u>ملاحظة</u> نستخدم دالة Poisson المعدلة في حالة كون المتغير العشوائي يرمز لحالات النجاح في تجارب يكون الاحتمال قريب من النصف وان حجم العينة كبير

الدالة المولدة للعزوم moment generating function (m.g.f) لتوزيع Poisson المعدل يعطى حسب الصيغة التالية [١]:

$$M(t)=\sum_{x} e^{tx}$$

$$M(t)=\sum_{x=\frac{\lambda}{2}}^{\infty} e^{tx}.\frac{e^{-\left\lceil\frac{\lambda}{2}\right\rceil}.\left\lceil\frac{\lambda}{2}\right\rceil^{x-\left\lceil\frac{\lambda}{2}\right\rceil}}{(x-\left\lceil\frac{\lambda}{2}\right\rceil)!}$$

اجعل $\theta=\left\lceil\frac{\lambda}{2}\right\rceil$ وأن $y=x-\left\lceil\frac{\lambda}{2}\right\rceil$ وبالتالي $y=x-\theta$ فأن

$$M(t)=\sum_{y=0}^{\infty} e^{t(y+\theta)}.\frac{e^{-\theta}.\theta^{y}}{y!}$$

$$M(t)=e^{\theta(t-1)}.\sum_{y=0}^{\infty}\frac{(\theta.e^{t})^{y}}{y!}$$

$$M(t)=e^{\theta(t-1)}.e^{\theta.e^{t}}$$

$$M(t)=e^{\theta(e^{t}+t-1)}$$

$$M'(t)=e^{\theta(e^{t}+t-1)}.\theta(e^{t}+1)$$

$$M'(t)=\theta.e^{\theta(e^{t}+t-1)+t}+\theta.e^{\theta(e^{t}+t-1)}$$

$$M'(0)=\theta+\theta=2\theta$$

$$M'(0)=\lambda$$

$$M''(t)=\theta(e^{\theta(e^t+t-1)+t}.\theta(e^t+1)+1)+\theta(e^{\theta(e^t+t-1)}.\theta(e^t+1))$$
$$M''(0)=\theta(2\theta+1)+\theta(2\theta)$$
$$M''(0)=4\theta^2+\theta$$
$$v(x)=M''(0)-(M'(0))^2$$
$$v(x)=\theta$$
$$v(x)=\frac{\lambda}{2}$$

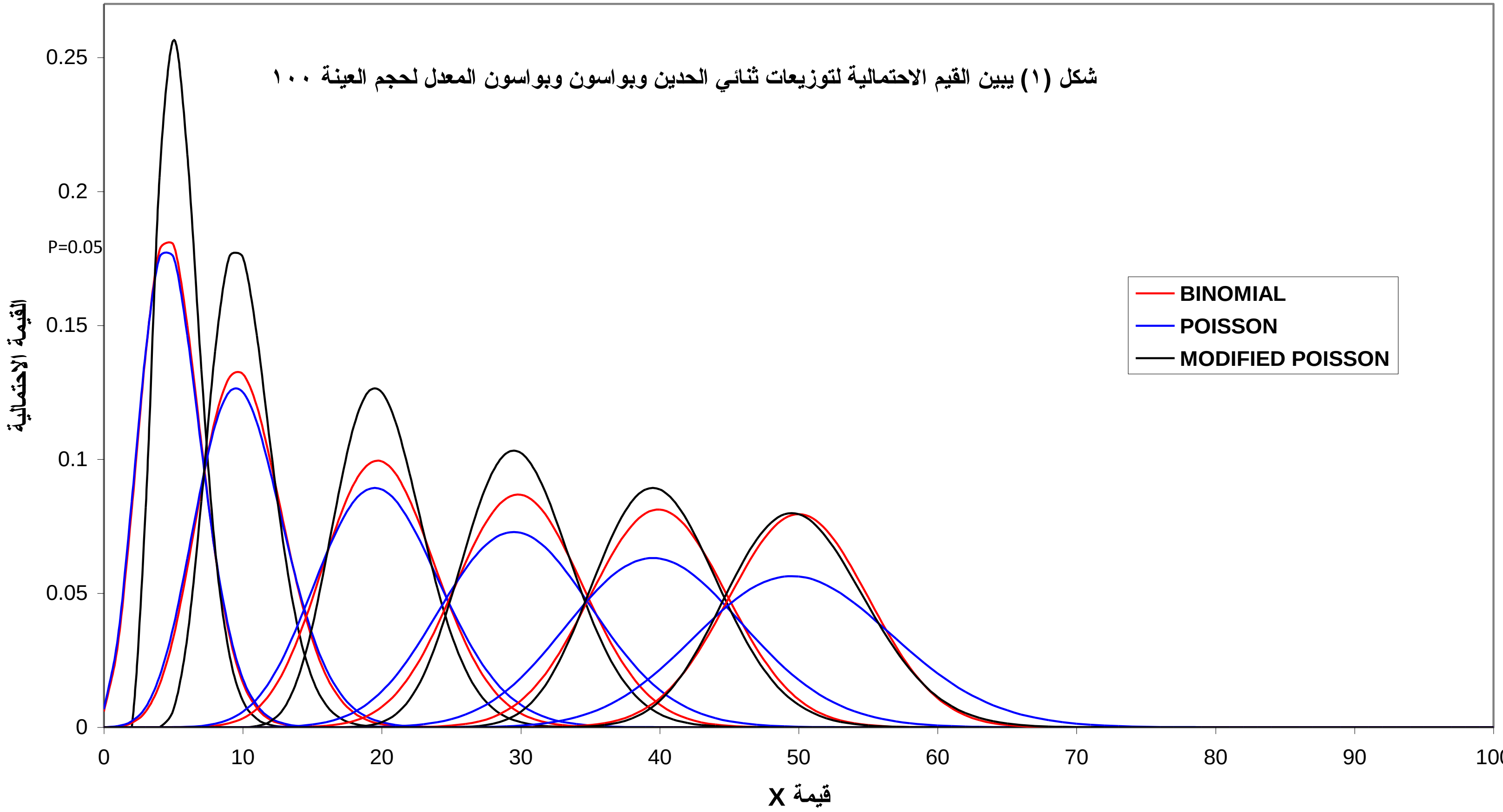
وهذا يعني ان توزيع Poisson المعدل وسطه الحسابي λ وتباينه $\frac{\lambda}{2}$.

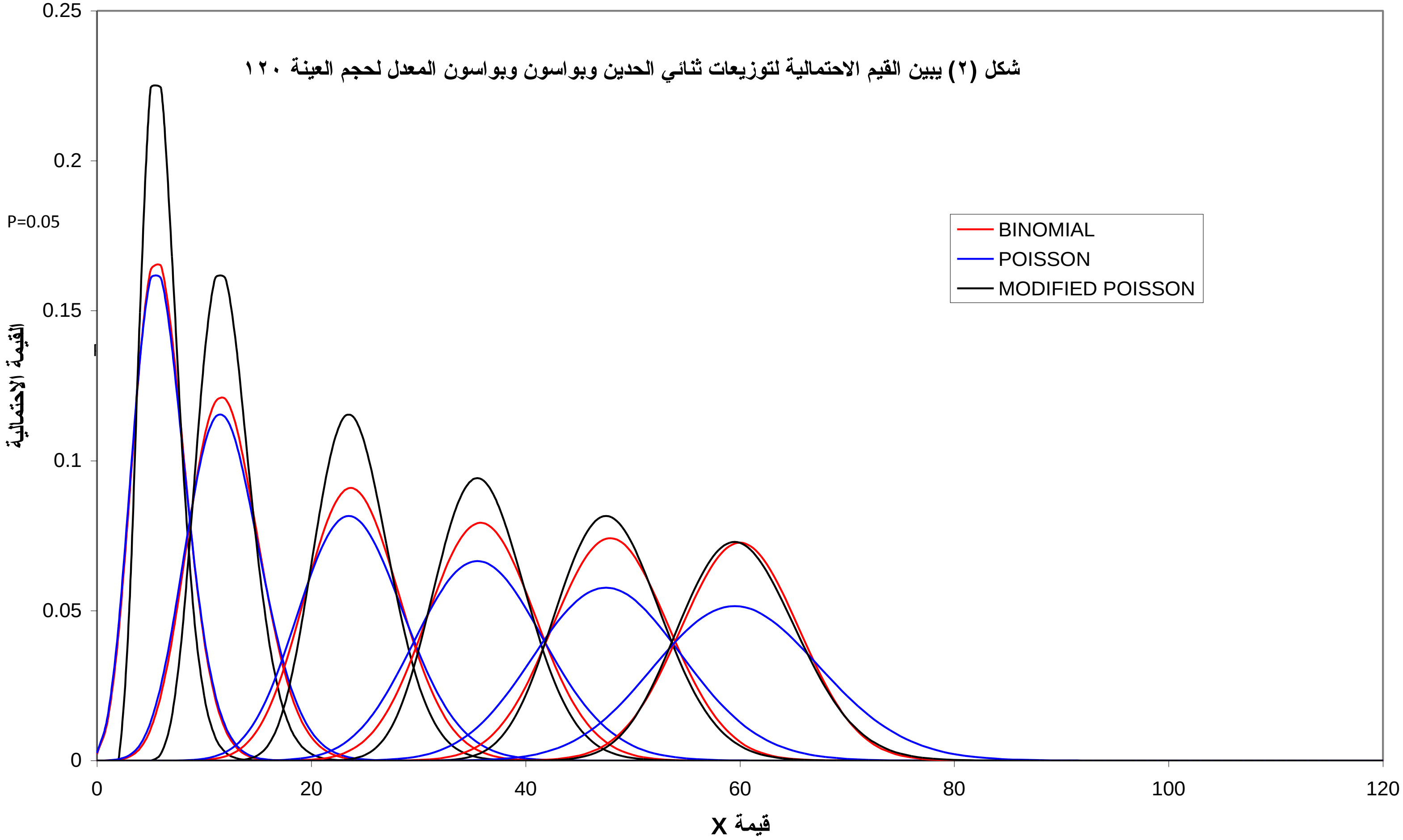
٦. تقارب توزيعات Poisson و Poisson المعدل من توزيع Binomial

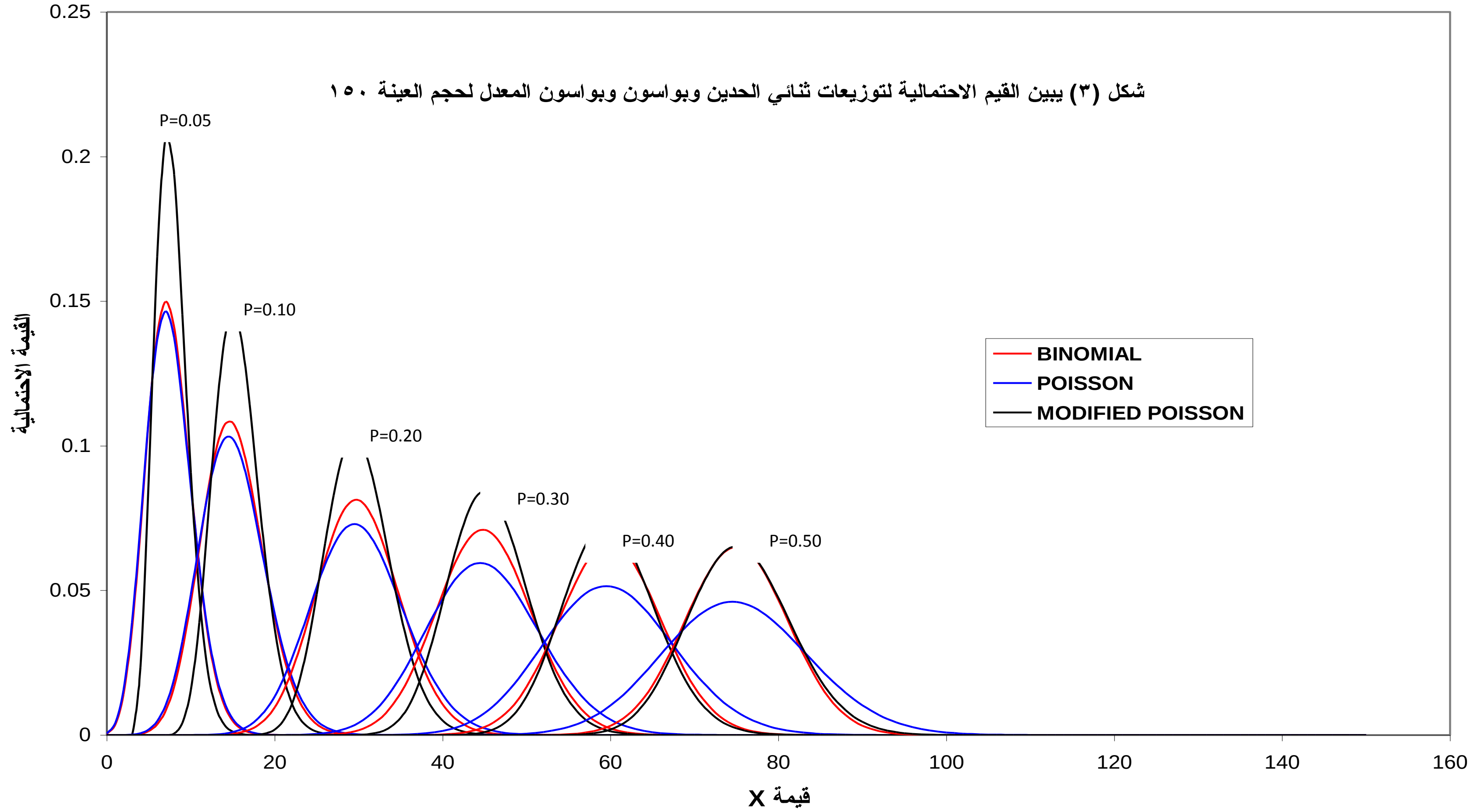
تم إجراء ثلاث تجارب بأحجام العينات ١٠٠ و١٢٠ و١٥٠ لدراسة مدى تقارب توزيع Poisson وتوزيع Poisson المعدل من توزيع Binomial حيث أخذت لكل عينة احتمالات النجاح

($p = 0.05$ و $p = 0.1$ و $p = 0.2$ و $p = 0.3$ و $p = 0.4$ و $p = 0.5$)

وتم رسم دالة الكثافة الاحتمالية للتوزيعات الثلاث والأشكال التالية تمثل مدى تقارب التوزيعات من بعضها البعض.


شكل (١) يبين القيم الاحتمالية لتوزيعات ثنائي الحدين وبواسون وبواسون المعدل لحجم العينة ١٠٠
P=0.05
القيمة الاحتمالية
قيمة X
BINOMIAL
POISSON
MODIFIED POISSON
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
60
70
80
90
100


شكل (٢) يبين القيم الاحتمالية لتوزيعات ثنائي الحدين وبواسون وبواسون المعدل لحجم العينة ١٢٠
P=0.05
القيمة الاحتمالية
قيمة X
BINOMIAL
POISSON
MODIFIED POISSON
0.25
0.2
0.15
0.1
0.05
0
0
20
40
60
80
100
120


شكل (٣) يبين القيم الاحتمالية لتوزيعات ثنائي الحدين وبواسون وبواسون المعدل لحجم العينة ١٥٠
P=0.05
P=0.10
P=0.20
P=0.30
P=0.40
P=0.50
BINOMIAL
POISSON
MODIFIED POISSON
القيمة الاحتمالية
0.25
0.2
0.15
0.1
0.05
0
0
20
40
60
80
100
120
140
160
قيمة X

٧. النتائج

يلاحظ من الأشكال الثلاثة (شكل (١) وشكل (٢) وشكل (٣)

- إن توزيع Poisson أكثر تفلطحا من توزيع Poisson المعدل المقترح وذلك لكون تباينه (٨) يساوي ضعف تباين Poisson المعدل المقترح $\frac{\lambda}{2}$.

- عندما تكون قيمة $p = 0.05$ ينطبق مخطط دالة Poisson الاعتيادية مع دالة Binomial بينما يبتعد مخطط دالة Poisson المعدل عن كلا المخططين ثم يبدأ مخطط دالة Poisson الاعتيادية بالابتعاد عن دالة Binomial كلما ازدادت قيمة P وازداد حجم العينة وهكذا يكون الابتعاد اكبر مايمكن عندما تكون قيمة $P=0.50$ وحجم العينة ١٥٠ بينما اقترب مخطط دالة Poisson المعدلة عن مخطط دالة Binomial كلما ازداد حجم العينة وازدادت قيمة P ويكون الاقتراب في أفضل حالاته عندما تكون قيمة $P=0.50$ وحجم العينة ١٥٠ حيث يكاد ينطبق مخطط دالة Poisson المعدلة مع مخطط دالة Binomial وبذلك فأن دالة Poisson المعدلة تكون أفضل من دالة Poisson الاعتيادية كلما اقتربت قيمة P من النصف وازداد حجم العينة وبذلك فإننا نوصي باستخدام دالة Poisson المعدلة في جميع التجارب والتطبيقات التي تكون نتائجها الفشل او النجاح والتي يكون فيها احتمال حدوث الحدث قريب من النصف وحجم العينة كبير.

٩. المصادر

1. Hogg Robert V. & Allen T. Graig (1978). Introduction to mathematical statistics . forth edition Macmillan publishing co. inc,newyork .

2. pebesma,Edzer (2010).introduction to geostatistics
http://ifgi.uni-muenstrede/epeb.ol/

3. Teerapabolarn,kanint (2008).Poisson approximation to Beta- Negative Binomial distribution .Int.j. contempt . math . sciences , vol 3 , no 10 ,457-461 .

4. Teerapabolarn,kanint (2009).Poisson approximation to Beta- Binomial distribution .international mathematical forum ,no25,1253-1256 .

5. Vu,H.T.V & Maller ,R.A. (1996) . The likelihood ratio test for poisson versus Binomial distributions ,Journal of American Statistical Association .

1. Walis ,Sean (2012).Binomial distribution probability and confidence interval .