استعمال نماذج انحدار بواسون المختلط للبيانات الطولية لتحليل صفات دم الأغنام

أ. م. إيمان حسن احمد العاني الباحث: حيدر زحام جبر
جامعة بغداد / كلية الإدارة والاقتصاد / قسم الإحصاء

تاريخ استلام البحث: 2014/10/30 تاريخ قبول النشر: 2014/12/4

المستخلص

تم في هذا البحث دراسة استعمال نماذج انحدار بواسون المختلط للبيانات الطولية لتحليل صفات دم الاغنام، اذ تعرف البيانات الطولية على انها بيانات مقطعية مقاسة في فترات زمنية معينة وقد اكتسبت هذه البيانات في العقد الحالي اهتماماً بالغاً ولاسيما في الدراسات الاقتصادية و الطبية.

هذا البحث يتضمن استجابات مختلطة (متقطعة متمثلة بتوزيع بواسون ومستمرة متمثلة بتوزيع كاوس) وتكون هذه الاستجابات مترابطة، وعند حساب تقديرات معلمات الانموذج لكل استجابة ستعطي تقديرات تكون متحيزة وبالتالي لحساب هذه الاستجابات تم استخدام طرائق متعددة المتغيرات لإيجاد تقديرات معلمات الانموذج، ومن هذه الطرائق: طريقة انموذج العوامل (Factorization Model) (FM) وطريقة معادلات التقدير العمومية (Generalized Estimating Equation) (GEE) وقد تم استخدام اسلوب المحاكاة من خلال اجراء مقارنة بين الطريقتين بواسطة البرنامج الاحصائي (R) فضلا عن تطبيق الاسلوبين على بيانات حقيقية لتحليل صفات دم الاغنام.

أظهرت نتائج تجربة المحاكاة أن هناك تناقصا حاصلا في قيم متوسط مربعات الخطأ (Mean Squared Error) كلما ازداد حجم العينة، إذ يظهر ذلك جليا" في كافة الطرائق المستخدمة في الدراسة، وهذا يعكس إحدى الخصائص الجيدة عندما تقترب قيمة المقدر من القيمة الحقيقية للمعلمة بزيادة حجم العينة ، كما نلاحظ أن هناك تقارب بين الطريقتين معادلات التقدير العمومية (GEE) وطريقة أنموذج العوامل (FM) في قيم متوسط مربعات الخطأ، وعلى ضوء نتائج المحاكاة تم تطبيق الطرائق على بيانات حقيقية لتحليل صفات دم الأغنام مأخوذة على عدة فترات فقد تم اختيار نموذج دم من 50 نعجة بواقع 5 فترات لكل نعجة وتم الحصول على 250 مفردة لغرض دراسة تأثر أنزيم (GPT) والدهون الثلاثية (TRI) ببعض العوامل والصفات المدروسة والمتمثلة بـ جنس الحيوان ووزنه ونوع خضاب الدم (HB) لكل حيوان وقد بينت النتائج تقديرات جيدة لمعلمات النماذج المختلطة وان هناك ارتباط بين الاستجابات المختلطة والمتغيرات المستقلة كذلك فأن هناك تقاربا كبيرا بين طريقتي أنموذج العوامل (FM) وطريقة معادلات التقدير العمومية (GEE) .

Using Regression Models of Mixed Poisson for Longitudinal data for Analizing Sheep Blood

Abstract

In this research study using Regression models of mixed Poisson for longitudinal data for analysis of sheep blood, longitudinal data was Known as evidence sectional measured in certain periods of time. These data gained in the current decade great importance, especially in the economic studies and medical.

This research includes mixed responses (discrete response represented by the distribution of Poisson and continuous response represented by the distribution of Gaussian) and these responses are correlated, thus when counting estimations of the parameter for models for every response separately, will give biased estimates, For this reason the joint regression models can be used for mixed responses by using the multivariate methods to estimate the parameter of the model, and these methods: method

Factorization Model (FM) and Generalized Estimating Equation (GEE) has been using the simulation through a comparison between the two methods by statistical program (R) in addition to the application of two methods on real data for analizing of sheep blood.
The results of simulation obtained that they decrease at the MSE whenever the sample size increases, as that a paired in all methods that used in the study, and this reflects one of good properties when the value of estimator is approach from the actual value for the parameter at increase of the sample size, and had been noted that a big converge between the (GEE) and (Factorization) methods in the value of mean squared error.
From the results of simulation was applied at actual data to analize sheep blood that taken at intervals from 50 sheep's select a sample of blood from each sheep at 5 intervals and collected 250 subjects to study the effect of Enzyme (GPT) and (TRI) by some studied properties of animal sex, weight and type of blood (HB) for every animal and the results obtained that good estimations for the parameters of mixed models and there is correlation between the mixed response and independent variables. And they are accounted highly conduce between (FM) and (GEE) methods.

1- المقدمة:

تكمن فلسفة الإحصاء من حيث آلية التطبيق في محاولة نمذجة الظواهر المختلفة بنماذج اقرب ما يمكن إلى الواقع الفعلي، وان هذه النماذج تقاس درجة قوتها بحسب درجة تقاربها مع الخواص الإحصائية وهي على أشكال وأنواع مختلفة فمنها الاحتمالي التي تعتمد في صياغتها على الاحتمالات الصرفة (نماذج السلاسل الزمنية، نماذج سلاسل ماركوف، نماذج المعولية) ومنها النماذج السببية التي تقوم صياغة نماذجها على ما يعرف بالسبب ونتيجة السبب وتأتي في مقدمة هذه النماذج ما تسمى بنماذج الانحدار، اذ تقوم نماذج الانحدار باستكشاف العلاقة ما بين السبب والذي يعرف إحصائياً بالمتغيرات التوضيحية (التفسيرية)، وبين ما هو نتيجة السبب أو ما يعرف بمتغير الاستجابة (المتغير المعتمد).

إن دراسة أي ظاهرة من الظواهر كأن تكون ظواهر اقتصادية أو اجتماعية أو طبية أو تطبيقية تحتاج إلى جمع البيانات لهذه الظاهرة المدروسة وتحليلها باستخدام الطرق والأساليب الإحصائية المتنوعة للحصول على أفضل قرار لهذه الظاهرة ولتسهيل عملية اتخاذ القرار للباحث، وان آلية وطرق جمع البيانات تختلف من ظاهرة إلى أخرى مما يبرز عنها أنواع مختلفة من البيانات اعتماداً على الظاهرة المدروسة فهناك بيانات تجمع على شكل فترات معينة مشكلة سلسلة زمنية تعرف ب بيانات السلاسل الزمنية (Time Series Data)، وكذلك هناك دراسة ظواهر معينة لمجاميع أو قطاعات تعتمد على جمع بيانات مقطعية (Cross Section Data)، ويؤدي الاختلاف في نوعية البيانات إلى بروز اختلافات في النماذج الرياضية كأن تكون نماذج انحدار خطية (Linear Regression Models) أو نماذج السلاسل الزمنية (Time Series Models) أو نماذج تصميم التجارب (Experimental Design Models) وغيرها من النماذج الرياضية، وتبعاً لذلك تظهر اختلافات في الأساليب والطرق الإحصائية المناسبة لها من حيث عملية التقدير (Estimation) لمعلمات الأنموذج.

لذلك اكتسبت البيانات الطولية (Longitudinal Data) في العقد الحالي اهتماماً بالغاً ولاسيما في الدراسات الاقتصادية والطبية (في البيانات التطبيقية) تعرف هذه البيانات على أنها بيانات مقطعية

مقاسة في فترات زمنية معينة وان الفائدة الرئيسة من استعمالها هي زيادة الدقة في التنبؤ من خلال زيادة عدد المشاهدات عن طريق ربط عدد المشاهدات المقطعية بعدد الفترات الزمنية.

هناك بعض البحوث وخاصة البحوث الطبية والبحوث الزراعية تتضمن استجابات مختلطة (متقطعة ومستمرة) وتكون هذه الاستجابات مترابطة وعند تقدير معلمات الأنموذج لكل استجابة على حدة يعطي تقديرات تكون متحيزة، ولأجل ذلك يمكن استعمال نماذج انحدار مشتركة للاستجابات المختلطة باستخدام طرق متعددة المتغيرات لإيجاد تقديرات معلمات الأنموذج.

تم في هذا البحث دراسة أنموذج انحدار بواسون المختلط (Mixed Poisson) من خلال افتراض الاستجابة المتقطعة لبواسون والاستجابة المستمرة لتوزيع الطبيعي وتكوين دالة احتمالية مشتركة للاستجابات المختلطة وتم استعمال أسلوب الإمكان الأعظم في تقدير معلمات الأنموذج فضلا عن طريقة معادلات التقدير العمومية (Generalized Estimating Equations) لمعالجة البيانات الطولية وتقدير معلمات الأنموذج.

2- هدف البحث:

يهدف هذا البحث إلى تقدير معلمات الأنموذج المختلط باستخدام إحدى طرائق متعدد المتغيرات (Multivariate Methods) هو أنموذج العوامل (FM) من خلال الطريقة الاحتمالية (Likelihood Method) وكذلك استخدام طريقة معادلات التقدير العمومية (GEE) لمعالجة البيانات الطولية (Longitudinal Data) للدالة الاحتمالية المشتركة والمتكونة من استجابة بواسون (Poisson Response) كمتغير متقطع واستجابة كاوس (Gaussian Response) كمتغير مستمر، ومن ثم إجراء مقارنة بين الطريقتين باستخدام أسلوب المحاكاة بواسطة معيار متوسط مربعات الخطأ (Mean Squared Error) ومن ثم تطبيقها على بيانات الدراسة الخاصة بتحليل صفات دم الأغنام.

3- نماذج الانحدار المختلطة

في بعض الدراسات وخاصة في الدراسات الطبية والدراسات التجريبية هناك بيانات أو استجابات مختلطة (متقطعة ومستمرة) اذ تكون هذه الاستجابات مترابطة إذ عند حساب تقديرات معلمات الأنموذج لكل استجابة على حدة يعطي تقديرات تكون متحيزة، ولأجل ذلك يمكن استخدام نماذج انحدار مشتركة للاستجابات المختلطة باستخدام طرائق متعددة المتغيرات لإيجاد تقديرات معلمات الأنموذج .

ففي عام 1961 قام الباحثان (Olkin and Tate) [8] باقتراح أسلوب أنموذج الموقع العام (General Location Model) لتحليل النتائج المتقطعة والمستمرة، في هذا الأنموذج اعتماداً على نموذج متعدد الحدود للاستجابة المتقطعة وأنموذج متعدد المتغيرات كاوس (Multivariate Gaussian) للاستجابة المستمرة اذ استخدم الأسلوب الاحتمالي في تقدير معلمات المتوسط الحدي (The Marginal Mean) تبعهم بذلك العديد من العلماء والباحثين في دراسة النماذج المختلطة من ضمنهم الباحثان (Catalano and Ryan) [9] عام 1992 اذ قاما بوصف أنموذج لمتغير ثنائي لنتائج متقطعة ومستمرة بالاعتماد على متغير كامن مستمر (Latent Variable) التي تقوم عليها الاستجابة المتقطعة اذ اقترحا أن التوزيع المشترك يمكن أن يصاغ بشكل منفصل: التوزيع الحدي للاستجابة المستمرة وتوزيع شرطي للاستجابة المتقطعة الثنائية اعتماداً على الاستجابة المستمرة اذ يتم الحصول على التوزيع الحدي بسهولة ويتم الحصول على المكون الثاني باستخدام الافتراضات الطبيعية المشتركة وبنفس الفكرة في عام 2007 قام الباحثان (Ying Yang و Jian Kang) [13] بتطوير أنموذج انحدار لتقييم تأثير التغايرات على الاستجابات المختلطة وكذلك حساب الارتباط بين الاستجابات وذلك باستخدام البيانات الطولية في دراسة التهاب المثانة البولية اذ تم بناء أنموذج مشترك للاستجابات من خلال فرض التوزيع الحدي لاستجابة بواسون والتوزيع الشرطي للاستجابة المستمرة اعتماداً على استجابة بواسون.

4- طرائق تقدير معلمات النماذج المختلطة

توجد هناك العديد من الطرائق المختلفة لتقدير معلمات نماذج الانحدار المختلطة منها ما يعتمد على الأسلوب البيزي ومنها ما يعتمد على طرق متعددة المتغيرات، في هذا البحث سيتم الاعتماد على طريقتي أنموذج العوامل و طريقة معادلات التقدير العمومية وكالآتي :-

1-4 أنموذج العوامل (Factorization Model):

تعد طريقة أنموذج العوامل واسعة الاستخدام في اغلب البحوث والدراسات التي تهتم بدراسة إنشاء نماذج مشتركة من خلال دمج الاستجابات المختلفة كأن تكون استجابات مستمرة واستجابات متقطعة في دالة احتمالية مشتركة تمثل هذه الاستجابات المرتبطة. ومن الباحثين الذين اهتموا بهذه النماذج Olkin and Tate(1961)[8] و Fitzmaurice and Laird (1995) [4] و Y.Yang [13](2007) و Armando Teixeira–Pinto(2009)[10] و F. Razie, E. Bahrami Samani and M.Ganjali (2013)[11] .

فكرة هذه الطريقة تقتصر على بناء الدالة المشتركة الاحتمالية من خلال ضرب التوزيع الحدي للاستجابة المتقطعة والمتمثلة باستجابة بواسون والتوزيع الشرطي للاستجابة المستمرة والمتمثلة باستجابة كاوس اعتماداً على الاستجابة الحدية.

لنفترض ان Y_{pi} تمثل الاستجابة المتقطعة وهي استجابة بواسون وان Y_{ci} تمثل الاستجابة المستمرة ولتكن التوزيع الطبيعي وان X_{pi} و X_{ci} تمثلان متجه المتغيرات المستقلة ذات البعد $(p \times 1)$ و $(q \times 1)$ على التوالي وان (i= 1,…,n).

إن الدالة الاحتمالية لتوزيع بواسون و المتمثلة بالتوزيع الحدي يمكن كتابتها كما في المعادلة الآتية -: [13],[12]

$$f(Y_{pi} \backslash X_{pi}) = \frac{\lambda_i^{Y_{pi}}}{Y_{pi}!} e^{-\lambda_i} \quad , \qquad 1 \le i \le n \quad \dots\dots\dots\dots\dots (1)$$

حيث أن

$log\ \lambda_i = X_{pi}^T \beta_p$ تمثل نموذج انحدار بواسون

λ_i: تمثل معلمة توزيع بواسون

β_p: تمثل متجه المعالم ذات بعد $p \times 1$

T : يرمز إلى مبدلة المصفوفة أو المتجه

من خلال استجابة بواسون Y_{pi} المعطاة أعلاه فإننا سنفرض أن التوزيع الشرطي للاستجابة المستمرة هو توزيع كاوس والتي يرمز لها بالرمز Y_{ci} أي أن [3],[4],[12],[13] :-

$$f_{Y_{ci}|Y_{pi}}\left(Y_{ci} \middle| Y_{pi}\right) = \frac{1}{\sqrt{2\pi\sigma_c^2}}\ exp\left\{-\frac{\left[Y_{ci} - \mu_i - \gamma\left(Y_{pi} - \lambda_i\right)\right]^2}{2\sigma_c^2}\right\} \dots\dots (2)$$

حيث ان

$\mu_i = X_{ci}^T\ \beta_c$ تمثل أنموذج الانحدار الطبيعي

β_c : متجه معلمات الانحدار الطبيعي ذات بعد $q \times 1$

γ : معلمة الانحدار التي تصف الارتباط بين استجابة بواسون Y_{pi} واستجابة كاوس Y_{ci}

لتبسيط الصيغة أعلاه واختصاراً لأنموذج الانحدار للدالة الشرطية أعلاه نفرض أن

$\alpha = (\beta_c^T, \gamma)^T$

وان

$$W_i = (X_{ci}^T, Y_{pi} - \lambda_i)^T$$

وبالتالي فان التوقع الشرطي للمعادلة أعلاه هو[4],[13]

$$E\left(Y_{ci} \middle| Y_{pi}\right) = W_i^T \alpha$$

لذلك يمكن الحصول على دالة التوزيع الاحتمالي المشترك للاستجابات المرتبطة لـ (Y_{ci} ، Y_{pi}) كما يلي [4],[11],[13]:-

$$f_{Y_{pi}, Y_{ci}}(Y_{pi}, Y_{ci}) = f_{Y_{pi}}(Y_{pi}) f_{Y_{ci}|Y_{pi}}(Y_{ci}|Y_{pi})$$

ومن خلال تعويض المعادلة (1) والمعادلة (2) يمكن الحصول على دالة التوزيع المشترك للاستجابات المختلطة كالاتي :

$$f_{Y_{pi}, Y_{ci}}(Y_{pi}, Y_{ci}) = \frac{\lambda_i^{Y_{pi}}}{Y_{pi}!} e^{-\lambda_i} \times \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{\left[Y_{ci} - W_i^T\alpha\right]^2}{2\sigma^2}\right\} \quad ....(3)$$

مع العلم إن أنموذج انحدار بواسون له دالة ربط تعرف بدالة الربط اللوغارتمية وكذلك أنموذج انحدار كاوس يمتلك دالة ربط تعرف بدالة الربط الطبيعية، لذا فان أنموذج الانحدار الخطي لكلا الاستجابتين يمكن تعريفها كما يلي :

$$\log \lambda_i = X_{pi}^T \beta_p \qquad (4)$$

$$Y_{ci} \backslash Y_{pi}, X_{ci}, X_{pi} = \mu_i + \gamma(Y_{pi} - \lambda_i) + \varepsilon_i \qquad (5)$$

حيث ان :

$$\mu_i = X_{ci}^T \beta_c$$

$$Y_{pi} | X_{pi} \sim poi(\lambda_i)$$

$$\varepsilon_i \sim N(0, \sigma_c^2)$$

حيث أن ε_i تمثل قيمة الأخطاء العشوائية لأنموذج الانحدار الطبيعي.

ولأجل تقدير معلمات الأنموذج المشترك للاستجابات المختلطة من خلال تطبيق طريقة الإمكان الأعظم على الدالة الاحتمالية المشتركة لـ (Y_{ci} ، Y_{pi}) حيث يتم تعظيم مشاهدات الاستجابات المختلطة [11]، أي أن :

$$l(Y_p, Y_c) = \prod_{i=1}^{n} f(Y_{pi}, Y_{ci} \backslash X_{pi}, X_{ci}) \qquad (6)$$

$$= \prod_{i=1}^{n} f(Y_{pi} \backslash X_{pi}) f(Y_{ci} \backslash Y_{pi}, X_{pi}, X_{ci})$$

ومن ثم اخذ اللوغارتم لدالة الإمكان الأعظم أي أن :

$$l(Y_p, Y_c) = log \prod_{i=1}^{n} f(Y_{pi}, Y_{ci} \backslash X_{pi}, X_{ci})$$

$$= \sum_{i=1}^{n} \left\{-\frac{1}{2}\log(2\pi\sigma_c^2) - \frac{\left(Y_{ci} - \mu_i - \gamma(Y_{pi} - \lambda_i)\right)^2}{2\sigma_c^2}\right\} + \sum_{i=1}^{n} \{Y_{pi} log\lambda_i - \lambda_i - \prod_{i=1}^{n} Y_{pi}!\} \qquad (7)$$

يتم تعويض نماذج الانحدار الخطية (انحدار بواسون والمتمثل بالمعلمة $\log\lambda_i = X_{pi}^T\beta_{pi}$ وانحدار كاوس المتمثل بالمعلمة $\mu_i = X_{ci}^T\beta_c$ في المعادلة (7) ولإيجاد تقديرات معلمات الانموذج $(\beta_p, \beta_c, \gamma, \sigma^2)$ نقوم باشتقاق الدالة الاحتمالية المشتركة ووفقاً للاتي : [13],[4]

باشتقاق المعادلة (7) بالنسبة لمتجه معلمات الانحدار الخطي لبواسون β_p نحصل على:

$$\sum_{i=1}^{n}\frac{\partial l_i(y_{pi},y_{ci})}{\partial\beta_p}=\sum_{i=1}^{n}\{\frac{\partial l_i(y_{pi})}{\partial\beta_p}+\frac{\partial l_i(y_{ci}\setminus y_{pi})}{\partial\beta_p}\}$$

$$=\sum_{i=1}^{n}\{\frac{\partial l_i(y_{pi})}{\partial\beta_p}+\frac{\partial\lambda_i}{\partial\beta_p}\frac{\partial l_i(y_{ci}\setminus y_{pi})}{\partial\lambda_i}\}$$

$$=\sum_{i=1}^{n}\{x_i(y_{pi}-\lambda_i)-x_i\Delta_i(y_{ci}-W_i^T\alpha)\gamma\sigma_c^{-2}\}\ldots\ldots\ldots\ldots\ldots(8)$$

كذلك بالنسبة لمتجه المعلمات $\alpha=(\beta_c^T,\gamma)^T$ يمكن الحصول عليه من المعادلة الآتية :

$$\sum_{i=1}^{n}\frac{\partial l_i(y_{pi},y_{ci})}{\partial\alpha}=\sum_{i=1}^{n}\{\frac{\partial l_i(y_{pi})}{\partial\alpha}+\frac{\partial l_i(y_{ci}\setminus y_{pi})}{\partial\alpha}$$

$$=\sum_{i=1}^{n}\frac{\partial l_i(y_{ci}\setminus y_{pi})}{\partial\alpha}$$

$$=\sum_{i=1}^{n}\{W_i(y_{ci}-W_i^T\alpha)\sigma_c^{-2}\}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

أما المعادلة الاحتمالية لحساب قيمة المعلمة σ^2 فيمكن الحصول عليها كما يلي :-

$$\sum_{i=1}^{n}\frac{\partial l_i(y_{pi},y_{ci})}{\partial\sigma_c^2}=\sum_{i=1}^{n}\{\frac{\partial l_i(y_{pi})}{\partial\sigma_c^2}+\frac{\partial l_i(y_{ci}\setminus y_{pi})}{\partial\sigma_c^2}\}$$

$$=\sum_{i=1}^{n}\frac{\partial l_i(y_{ci}\setminus y_{pi})}{\partial\sigma_c^2}$$

$$=\phi^{-1}\sum_{i=1}^{n}(C_i-\sigma_c^2)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

حيث أن

$$\Delta_i=\mathrm{var}(Y_{pi})=\lambda_i$$

$$C_i=(Y_{ci}-W_i^T\alpha)^2$$

$$\emptyset=var(C_i)$$

كذلك يمكن كتابة المعادلات الاحتمالية عن طريق المصفوفات كالآتي:-

$$\sum_{i=1}^{n}\begin{pmatrix}\frac{\partial l_i}{\partial \beta_p}\\ \frac{\partial l_i}{\partial \alpha}\\ \frac{\partial l_i}{\partial \sigma^2}\end{pmatrix}=\sum_{i=1}^{n}\begin{pmatrix}\Delta_i X_i & -\gamma\Delta_i X_i & 0\\ 0 & W_i & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\Delta_i^{-1} & 0 & 0\\ 0 & \sigma^{-2} & 0\\ 0 & 0 & \Phi^{-1}\end{pmatrix}\begin{pmatrix}y_{pi}-\lambda_i\\ y_{ci}-W_i^T\alpha\\ C_i-\sigma^2\end{pmatrix}$$

$$=\begin{pmatrix}\frac{\partial E(y_{pi})}{\partial \beta_p} & \frac{\partial E(y_{pi})}{\partial \alpha} & \frac{\partial E(y_{pi})}{\partial \sigma^2}\\ \frac{\partial E(y_{ci}|y_{pi})}{\partial \beta_p} & \frac{\partial E(y_{ci}|y_{pi})}{\partial \alpha} & \frac{\partial E(y_{ci}|y_{pi})}{\partial \sigma^2}\\ \frac{\partial E(C_i|y_{pi})}{\partial \beta_p} & \frac{\partial E(C_i|y_{pi})}{\partial \alpha} & \frac{\partial E(C_i|y_{pi})}{\partial \sigma^2}\end{pmatrix}^T \times cov^{-1}\begin{pmatrix}y_{pi}\\ y_{ci}\\ C_i\end{pmatrix}|y_{pi}\begin{pmatrix}y_{pi}-E(y_{pi})\\ y_{ci}-E(y_{ci}|y_{pi})\\ c_i-E(C_i|y_{pi})\end{pmatrix}$$

وبحل المعادلة (9) يتم الحصول على تقديرات المعلمات $\alpha=(\beta_c^T,\gamma)^T$ وكما يلي :

$$\sum_{i=1}^{n}\frac{\partial l_i(y_{pi},y_{ci})}{\partial \alpha}=0$$

$$\sum_{i=1}^{n}W_iW_i^T\hat{\alpha}=\sum_{i=1}^{n}W_iy_{ci}$$

لذا فأن الدالة التقديرية لـ $\hat{\alpha}$ ستكون كالاتي :-

$$\hat{\alpha}=(\sum_{i=1}^{n}W_iW_i^T)^{-1}\sum_{i=1}^{n}W_iy_{ci} \dots\dots\dots\dots\dots\dots\dots\dots (11)$$

كذلك يمكن حساب تقدير المعلمة σ^2 من خلال مساواة الاشتقاق بالصفر للمعادلة (10) كما يلي :

$$\sum_{i=1}^{n}\frac{\partial l_i(y_{pi},y_{ci})}{\partial \sigma_c^2}=0$$

$$n\hat{\sigma}_c^2=\sum_{i=1}^{n}(y_{ci}-W_i^T\hat{\alpha})^2$$

$$\hat{\sigma}_c^2=\frac{\sum_{i=1}^{n}(y_{ci}-W_i^T\hat{\alpha})^2}{n} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (12)$$

أما بالنسبة للمعادلة (8) الخاصة بتقدير معلمة الانحدار β_p فيتم الحصول عليها من خلال تعويض قيمة المعلمات $\hat{\boldsymbol{\alpha}}$ و $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ في المعادلة الاحتمالية (8) لـ β_p وبما إن هذه المعادلة تكون غير خطية يتم استخدام إحدى الطرائق الرقمية التكرارية كخوارزمية نيوتن رافسون (Newton Raphson algorithm) أو خوارزمية ناتج فشر (Fisher Scoring algorithm) ومن ثم استخدام طريقة

المربعات الصغرى التكرارية الموزونة (iteratively reweighed least squares method) وإيجاد قيمة $\hat{\beta}_p$ حيث يتم حساب قيم أولية لـ β_p بصورة تكرارية لحين الوصول إلى قيمة تقديرية جيدة لـ $\hat{\beta}_p$.

2-4 معادلات التقدير العمومية Generalized Estimating Equations(GEE)

تعتمد طريقة معادلات التقدير العمومية (GEE) على تقديرات (quasi-likelihood) و تستخدم بصورة واسعة في تقدير النماذج الحدية للبيانات الطولية [5] . تتميز هذه الطريقة بأنها لا تتطلب جميع خواص التوزيعات متعددة المتغيرات للاستجابات المكررة إلا أنها تحتاج إلى تحديد العزميين الأوليين لمتجه الاستجابات (outcomes vector) .

نفرض ان N من المفردات المستقلة تمتلك مشاهدات n_i على متجه الاستجابات المختلطة (Y_{pi}, Y_{ci}) حيث ان :

$$Y_{pi} = (Y_{pi1}, Y_{pi2}, Y_{pi3}, ..., Y_{pini})^T$$

حيث ان Y_{pi} تمثل استجابة بواسون

$$Y_{ci} = (Y_{ci1}, Y_{ci2}, Y_{ci3}, ..., Y_{cini})^T$$

حيث ان Y_{ci} تمثل استجابة توزيع كاوس

$$X_i = (X_{i1}, X_{i2}, X_{i3}, ..., X_{ini})$$

حيث ان X_i تمثل مصفوفة المتغيرات المستقلة (التغايرات covariates) ذات بعد $p \times n_i$

وان X_{ik} متجه ذات بعد $p \times 1$ يمثل المتغيرات المستقلة داخل كل وحدة.

تستخدم هذه الطريقة في معالجة البيانات الطولية للبيانات المختلطة المتقطعة و المستمر .

في البداية يتم بناء نموذج الانحدار للاستجابات المختلطة [4],[13] كما يلي :

$$\log[E(Y_{pi})] = X_i^T \beta_p \qquad (13)$$

$$E(Y_{cik} \backslash Y_p) = X_{ik}^T \beta_c + \gamma_1 (Y_{pik} - \lambda_{ik}) + \gamma_2 S_i \qquad (14)$$

حيث ان :

$$S_i = \sum_{j=1}^{ni} (Y_{pij} - \lambda_{ij})$$

وان :

γ_1 و γ_2 يمثلان الارتباط بين Y_{pi} و Y_{ci}

γ_1 تمثل تأثير الـ Y_{pi} على Y_{ci} داخل المشاهدات

γ_2 تمثل تأثير بقية المشاهدات داخل هذه الوحدة

لتبسيط النموذج (14) نفرض ان :

$$W_{ik} = (X_{ik}^T, Y_{pik} - \lambda_{ik}, S_i)^T$$

$$\alpha = (\beta_2^T, \gamma_1, \gamma_2)^T$$

لذلك فأن أنموذج التوقع الشرطي يصبح كالآتي :

$$E(Y_{cik} \backslash Y_{pi}) = W_{ik}^T \alpha \qquad (15)$$

لذا فأن الدالة المشتركة للاستجابات المختلطة و المتمثلة باستجابة بواسون و استجابة كاوس يمكن كتابتها كما في المعادلة التالية :-

$$f_{Y_{pi},\,Y_{ci}}\left(Y_{pi},\,Y_{ci}\right)=\frac{\lambda_i^{\,Y_{pi}}}{Y_{pi}!}e^{-\lambda_i}\times\frac{1}{\sqrt{2\pi\sigma^2}}\;exp\left\{-\frac{\left[Y_{ci}-W_{ik}^{T}\alpha\right]^2}{2\sigma^2}\right\}$$

اعتماداً على نفس فكرة الباحثان (Fitzmaurice and Laird)[4] و(Y.Yang)[13] حيث افترضوا أن بين الارتباط (ρ_{Y_c} وρ_{Y_p}) يمثلان لكل من استجابة بواسون و استجابة كاوس على التوالي وان المشاهدات داخل الوحدة تكون مرتبطة ايجابيا" .

يتم الحصول على مصفوفة الارتباط العاملة (Working Correlation Matrix) كما يأتي :[13],[6],[4]

$$cov(Y_{pi)}=\mathrm{V_{1i}}\approx\Delta_{\mathrm{i}}^{\frac{1}{2}}\left[\left(1-\rho_{Y_p}\right)I_i+\rho_{Y_p}J_i\right]\Delta_i^{\frac{1}{2}}\qquad(16)$$

$$cov\left(Y_{ci}\backslash Y_{pi}\right)=V_{2i}\approx\sigma^2\left[\left(1-\rho_{Y_c}\right)I_i+\rho_{Y_{ci}}J_i\right]\qquad(17)$$

حيث ان

Δ_i مصفوفة قطرية تحتوي على عناصر التباين لتوزيع بواسون $var\left(Y_{pi}\right)=\lambda_i$

I_{i} مصفوفة الوحدة ذات بعد $n_i\times n_i$

J_{i} مصفوفة احادية ذات بعد $n_i\times n_i$

ولإيجاد قيم كل من $\hat{\rho}_{Yc}$, $\hat{\rho}_{Y_p}$, $\hat{\sigma}_c^2$ من خلال تقدير العزوم وكالاتي [4],[6],[7],[13]:-

$$r_{Y_{pij}}=\frac{Y_{pij}-\lambda_{ij}(\hat{\beta_p})}{\sqrt{\lambda_{ij}(\hat{\beta_p})}}$$

$$r_{Y_{cij}}=Y_{cij}-W_{ij}^{T}\,\hat{\alpha}$$

$$\phi^{-1}=(\sum_{i=1}^{N}\sum_{j=1}^{ni}r_{Y_{pij}}^2)/(\sum_{i=1}^{N}n_i-p)$$

$$\hat{\rho}_{Y_p}=\hat{\phi}(\sum_{i=1}^{N}\sum_{j\rangle j'}r_{Y_{pij}}r_{Y_{pij'}})/(\sum_{i=1}^{N}\frac{1}{2}n_i(n_i-1)-p)\qquad\ldots\ldots\ldots\ldots\ (18)$$

$$\hat{\sigma}_c^2=(\sum_{i=1}^{N}\sum_{j=1}^{ni}r_{Y_{cij}}{}^2)/(\sum_{i=1}^{N}n_i-p-1)\qquad\ldots\ldots\ldots\ldots\ldots\ (19)$$

$$\hat{\rho}_{Y_c}=\hat{\sigma}_c^{-2}(\sum_{i=1}^{N}\sum_{j\rangle j'}r_{Y_{cij}}r_{Ycij'})/(\sum_{i=1}^{N}\frac{1}{2}n_i(n_i-1)-p-1)\qquad\ldots\ldots\ldots\ldots\ (20)$$

بعد إيجاد المقدرات في المعادلات (18) و (19) و (20) يتم تعويضها في مصفوفة الارتباط العاملة (16) و (17) لذلك فان نظام معادلات التقدير لحساب قيم α و β_p يتم من خلال تطبيق المعادلة الآتية :

$$\sum_{i=1}^{n}\begin{pmatrix}\frac{\partial l_i}{\partial \beta_p}\\ \frac{\partial l_i}{\partial \alpha}\end{pmatrix}=\begin{pmatrix}\frac{\partial E(y_{pi})}{\partial \beta_p} & \frac{\partial E(y_{pi})}{\partial \alpha}\\ \frac{\partial E(y_{ci}|y_{pi})}{\partial \beta_p} & \frac{\partial E(y_{ci}|y_{pi})}{\partial \alpha}\end{pmatrix}^{T}$$
$$\times cov^{-1}\begin{pmatrix}y_{pi}\\ y_{ci}|y_{pi}\end{pmatrix}\begin{pmatrix}y_{pi}-E(y_{pi})\\ y_{ci}-E(y_{ci}|y_{pi})\end{pmatrix}$$

تصبح المعادلة التقديرية كالاتي [6] :

$$\sum_{i=1}^{n}\begin{pmatrix}\Delta_i X_i & -(\gamma_1+\gamma_2)\Delta_i X_i\\ 0 & W_i\end{pmatrix}\begin{pmatrix}V_{1i}^{-1} & 0\\ 0 & V_{2i}^{-1}\end{pmatrix}\begin{pmatrix}y_{pi}-\lambda_i\\ y_{ci}-W_i^{T}\alpha\end{pmatrix}=0 \dots (21)$$

يتم تعويض قيم V_{1i}^{-1} و V_{2i}^{-1} في المعادلة التقديرية (21) حيث تحل المعادلة بصورة تكرارية باستخدام احدى الطرق الرقمية كطريقة نيوتن رافسون (Newton Raphson)أو طريقة ناتج فشر (Fisher Scoring algorithm) وذلك بإعطاء قيم أولية لكل من α و β_p ومن ثم استخدام هذه الطرائق لحين الحصول على قيم تقديرية جيدة لكل من $\hat{\alpha}$ و $\hat{\beta}_p$.

$$G^{(t+1)}=G^{(t)}-\{\sum_{i=1}^{N}DVD^{T}\}\times\{\sum_{i=1}^{N}DVC\}$$

حيث أن :

$$G=[\beta_p,\alpha]^{T}$$
$$D=\begin{pmatrix}\Delta_i X_i & -(\gamma_1+\gamma_2)\Delta_i X_i\\ 0 & W_i\end{pmatrix}$$
$$V=\begin{pmatrix}V_{1i}^{-1} & 0\\ 0 & V_{2i}^{-1}\end{pmatrix}$$
$$C=\begin{pmatrix}y_{pi}-\lambda_i\\ y_{ci}-W_i^{T}\alpha\end{pmatrix}$$

5- المحاكاة simulations

يعتبر أسلوب المحاكاة واحداً من الأساليب المهمة التي تساعد الباحثين في دراساتهم، إذ أن المحاكاة تشبه مختبر الباحثين حيث يقوم الباحث بتوليد بيانات أو مشاهدات من النموذج المستخدم لدراسة ظاهرة معينة.

فأسلوب المحاكاة مفيد جدا اذ هو البديل في حالة عدم توفر البيانات المطلوبة الخاصة بموضوع الدراسة لأسباب قد تكون متعلقة بكلفة الحصول على البيانات أو استحالة الحصول على نوع البيانات الخاصة بالدراسة.

وبعد توليد هذه البيانات من الممكن أن نستخدم فيها النتائج والعلاقات الإحصائية التي تم التوصل إليها من خلال نموذج الدراسة المستخدم، وفي بعض الأحيان تفيد في بعض الحقائق التي من الصعب التوصل إليها رياضيا .

في هذه الدراسة تم توليد المتغيرات العشوائية باستخدام البرنامج الإحصائي R من خلال توليد المتغيرات المستقلة X_{p1}, X_{p2}, X_{p3} و X_{c1}, X_{c2}, X_{c3} كالاتي:-

$$X_{p1}, X_{c1} \sim Normal(n,5)$$

$X_{p2}, X_{c2} \sim binomil(n, 8, 0.5)$

$X_{p3}, X_{c3} \sim Gama(n, 2, 1/2)$

$\varepsilon_i \sim N(0,1)$

وقد تم تكرار التجربة 500 تكرار لنماذج الانحدار الخطي المختلطة وكانت نتائج المحاكاة كما يلي:

1-5 نتائج طريقة أنموذج العوامل Factorization Model :

تم تقدير معلمات الأنموذج من خلال حل الصيغة (8) و(11) و(12)حيث يتم حسابها من خلال حجوم عينات مختلفة (40,80,260) وتم تكرار التجربة 500 تكرار وكانت نتائج المحاكاة لطريقة (FM) كما يلي:

جدول (1): تقدير معالم الأنموذج المختلط باستخدام طريقة أنموذج العوامل FM ولأحجام مختلفة من العينات (صغيرة، متوسطة، كبيرة)

parameter	N=40		N=80		N=260	
	Estimate	S.E	Estimate	S.E	Estimate	S.E
Poisson						
β_0	0.119	0.013	0.117	0.025	0.090	1.376329e-03
β_1	0.098	0.007	0.098	0.002	0.100	9.564817e-05
β_2	0.099	0.011	0.098	0.001	0.100	9.648013e-05
β_3	0.099	0.001	0.099	0.001	0.101	1.297452e-04
Gaussian						
β_0	0.850	0.263	0.870	0.642	0.663	0.572
β_1	0.098	0.007	0.075	0.042	0.073	0.020
β_2	0.074	0.030	0.074	0.026	0.072	0.019
β_3	0.074	0.012	0.076	0.015	0.070	0.012
σ_c^2	1.359		2.635		6.642	
γ	0.768		0.838		0.796	

***S.E : Standard Error**

2-5 نتائج طريقة معادلات التقدير العمومية GEE

تم تقدير معلمات الانموذج من خلال حل الصيغة (21) حيث تم حسابها من خلال حجوم عينات مختلفة (40,80,260) و تم تكرار التجربة 500 تكرار وكانت نتائج المحاكاة لطريقة GEE كما يلي:

جدول (2) : تقدير معالم الأنموذج المختلط باستخدام طريقة معادلات التقدير العمومية (GEE) و لأحجام مختلفة من العينات (صغيرة ، متوسطة ، كبيرة)

parameter	N=40		N=80		N=260	
	Estimate	S.E	Estimate	S.E	Estimate	S.E
Poisson						
β_0	0.120	0.019	0.103	0.025	0.103	0.010
β_1	0.098	0.004	0.100	0.006	0.099	0.002
β_2	0.099	0.003	0.100	0.004	0.099	0.003
β_3	0.100	0.0009	0.100	0.0006	0.099	0.001
Gaussian						
β_0	0.766	0.132	0.752	0.093	0.750	0.093
β_1	0.068	0.035	0.061	0.009	0.061	0.016
β_2	0.064	0.012	0.060	0.017	0.060	0.006
β_3	0.070	0.003	0.070	0.010	0.070	0.005
σ_c^2	0.058		0.061		0.062	
γ_1	0.028		0.016		0.004	
γ_2	0.025		0.015		0.004	

يظهر الجدولان (1) و (2) نتائج تجربة المحاكاة والتي تم الحصول عليها من خلال تطبيق البرنامج الإحصائي R لإيجاد تقدير معلمات النماذج المستخدمة في الدراسة وكذلك إيجاد متوسط مربعات الخطأ (MSE) من خلال تطبيقها على الأساليب أو الطرائق المستخدمة في البحث (أسلوب نموذج العوامل (FM) و أسلوب معادلات التقدير العمومية (GEE))، إذ تعكس النتائج قيم متوسط مربعات الخطأ (MSE) لإحجام مختلفة من العينات (صغيرة، متوسطة، كبيرة) وللطريقتين سابقة الذكر وكما في الجدول (3) الآتي:

جدول (3): نتائج MSE للأساليب الإحصائية المستخدمة في الدراسة

FM	GEE	Sample Size
10.75481	10.67137	**N=40**
10.73167	10.61265	**N=80**
10.70219	10.60226	**N=260**

إن الملاحظة الدقيقة للجدول (3) تبين التراجع الحاصل في قيم متوسط مربعات الخطأ (MSE) كلما ازداد حجم العينة، إذ يظهر ذلك جلياً في كافة الطرائق المستخدمة في الدراسة، وهذا يعكس إحدى الخصائص الجيدة عندما تقترب قيمة المقدر من القيمة الحقيقية للمعلمة بزيادة حجم العينة، كما يمكن ملاحظة أن هناك تقاربا كبيرا بين طريقة (GEE) و(FM) اذ تشير قيم متوسط مربعات الخطأ هذا التقارب كلما ازدادت عملية التكرار وكذلك بزيادة حجم العينة، كذلك فان استخدام النماذج الخطية للاستجابات المختلط للبيانات الطولية أعطت نتائج جيدة وقريبة من القيم الأولية لكل من معلمات الانحدار لكل من طريقة العوامل (FM) وطريقة معادلات التقدير العمومية (GEE).

6- الجانب التطبيقي للبيانات:

يتناول هذا الجانب تطبيقاً عملياً لما تم دراسته نظرياً وتجريبياً في النقاط (4) و (5) السابقة بالاعتماد على بيانات حقيقية للبرهنة على إمكانية تطبيق الأسلوب التجريبي على بيانات واقعية لعينة من البيانات مأخوذة من محطة أبحاث الأغنام و الماعز التابعة للهيئة العامة للبحوث الزراعية/ وزارة الزراعة.

تم جمع البيانات من دراسة نفذت في محطة أبحاث الأغنام و الماعز التابعة للهيئة العامة للبحوث الزراعية / وزارة الزراعة، تقع المحطة في أبي غريب (20 كم غرب بغداد) وكان الهدف دراسة تأثير أنزيم GPT في الكبد و الدهون الثلاثية TRI في الدم على بعض الصفات المدروسة الجنس و الوزن ونوع خضاب الدم لعدد من الأغنام العواسية التركية و لـ 250 مفردة مقسمة على 5 فترات زمنية.

اذ يعرف أنزيم (Glutamic Pyruvic Transaminase) والمعروف اختصارا (GPT) على انه احد الأنزيمات الموجودة في الكبد اذ تعمل على نقل مجموعة الامين من الأحماض الامينية إلى الأحماض الكيتونية وبالعكس، وهي تنتشر في الأنسجة و تنتقل إلى المصل بعد تحطم الأنسجة، وان من وظائفها الرئيسة تكون داخل الخلية، وهي توجد بتراكيز عالية في الكبد والعضلات القلبية والهيكلية والكليتين وتوجد في أعضاء أخرى كنواتج تحفيز لنقل مجموعة الامين علماً أنها تنتقل إلى المصل بعد تحطيم الخلايا وهو يدل على زيادة نشاطها وتركيزها في ذلك الوسط [1].

كما تعرف الدهون الثلاثية (triacylglyceride) (TRI) بأنها تركيب عضوي يتألف من الكليسيريد مع ثلاث مجموعات من الاسيد الدهني وهو مصدر كبير لتخزين الطاقة. تتحرك الدهون الثلاثية (TRI) عبر الدم وتخزن في الخلايا الدهنية.

فقد تم اختيار نموذج دم من 50 نعجة بواقع 5 فترات لكل نعجة وتم الحصول على 250 مفردة لغرض دراسة تأثر أنزيم GPT والدهون الثلاثية (TRI) ببعض العوامل والصفات المدروسة والمتمثلة بـ جنس الحيوان ووزنه و نوع خضاب الدم (HB) لكل حيوان.

إذ يطلق على المعلومات التي يتم جمعها من خلال دراسة طولية بالبيانات الطولية التي تعرف بأنها القياسات المأخوذة بشكل متكرر على كل قطاع (subject) خلال مدة من الزمن (Time) كأن تكون أشهر أو سنوات، كذلك تعرف على أنها قياسات متكررة (Repeated Measurement)

لكل قطاع من القطاعات المشاهدة، إذ المشاهدات من القطاع نفسه تكون مرتبطة وان الارتباط هو الصفة المميزة لهذه البيانات [2].

تم تلخيص المتغيرات المعتمدة و المستقلة المستخدمة في الدراسة كما يلي:

Y_{pi} : يمثل معدل انزيم GPT في الكبد

Y_{ci} : يمثل نسبة الدهون الثلاثية TRI في الدم

SEX : يمثل جنس الحيوان

WIEGHT : يمثل وزن الحيوان

HB_TYPE : يمثل نوع خضاب الدم

ومن خلال استخدام الأساليب الإحصائية و تطبيقها بواسطة البرنامج الإحصائي R تم الحصول على النتائج الآتية:

أ- طريقة أنموذج العوامل Factorization Model

إن الأنموذج الخطي لهذه الأسلوب يكون كالآتي:-

$$\log[E(Y_p)] = \beta_{p0} + \beta_{p1}SEX + \beta_{p2}WIEGHT + \beta_{p3}HB_TYPE$$

$$E(Y_c \backslash Y_p) = \beta_0 + \beta_{c1}SEX + \beta_{c2}WIEGHT + \beta_{c3}HB_{TYPE} + \gamma \ (Y_p - E[Y_p])$$

اذ تم التطبيق من خلال الصيغ (8) و (11) و (12) وتم إيجاد معلمات الانحدار لكل من بواسون وكاوس وكما في الجدول رقم (4):

جدول (4): نتائج التطبيق العملي لطريقة Factorization

Parameter	Estimate	S.E
Poisson		
intercept	2.3148	0.0666
SEX	0.0039	0.0431
WIEGHT	0.0144	0.0216
HB_TYPE	0.0393	0.0198
Gaussian		
intercept	0.0672	0.2339
SEX	0.1316	0.1589
WIEGHT	0.1621	0.0758
HB_TYPE	0.0890	0.0703
σ_c^2	0.2600	
γ	- 0.0085	

ب- طريقة معادلات التقدير العمومية Generalized Estimating Equations

إن الأنموذج الخطي للاستجابات المختلطة لهذا الأسلوب هو :

$$\log[E(Y_p)] = \beta_{p0} + \beta_{p1}SEX + \beta_{p2}WIEGHT + \beta_{p3}HB_TYPE$$

$$E(Y_c \backslash Y_p) = \beta_0 + \beta_{c1}SEX + \beta_{c2}WIEGHT + \beta_{c3}HB_{TYPE} + \gamma_1(Y_p - E[Y_p]) + \gamma_2 \sum_{j=1}^{ni} (Y_{pi} - E[Y_{pi}])$$

تم التطبيق من خلال حل الصيغة (21) وإيجاد معلمات الانحدار المختلط لكل من بواسون و كاوس وكما في الجدول رقم (5):-

جدول (5): نتائج التطبيق العملي للبيانات لطريقة GEE

parameter	Estimate	S.E
Poisson		
intercept	0.3147	0.0677
SEX	0.0039	0.0359
WIEGHT	0.0143	0.0192
HB_TYPE	0.0393	0.0325

Gaussian		
intercept	0.1118	0.1705
SEX	0.1329	0.1121
WIEGHT	0.1645	0.0452
HB_TYPE	0.0902	0.0456
σ_c^2	0.6049	
γ_1	0.0132	
γ_2	0.0004	

بعد أن تم تطبيق البرنامج الإحصائي R على بيانات الدراسة، أظهرت النتائج لكل من الأساليب الإحصائية المستخدمة في هذه الدراسة إن هناك تقديرات و نتائج جيدة بالنسبة لمعلمات الأنموذج المختلط و المتمثل بأنموذج انحدار بواسون و أنموذج الانحدار الطبيعي (كاوس) اذ يمكن تفسير نتائج التطبيق بأن فعالية أنزيم GPT و مستوى الدهون الثلاثية بالدم TRI تتأثر بالصفات المدروسة (الجنس SEX والوزن WIEGHT ونوع خضاب الدم HB_TYPE) ولكن بنسب متفاوتة.

كانت نتائج طريقتي أنموذج العوامل و معادلات التقدير العمومية GEE متقاربة جداً من ناحية تقدير معلمات النماذج (SEX,WIEGHT,HB_TYPE) فضلا عن التقارب الكبير في قيم MSE لهذه الطريقتين وهي نفس النتيجة التي تم التوصل إليها في الجانب التجريبي لهذه الدراسة.

7- الاستنتاجات :

1- يلاحظ من الجدولان (1) و(2) ان القيم التقديرية لمعلمات الأنموذج قريبة من القيم الأولية ولكافة حجوم العينات (40، 80، 260) ولطريقة أنموذج العوامل (FM) وطريقة معادلات التقدير العمومية (GEE).

2- اظهرت نتائج المحاكاة ان قيم المؤشرات (MSE) تتناقص بزيادة حجم العينة ولكافة الطرائق المستعملة في البحث وهذا يتطابق مع النظرية الاحصائية.

3- من خلال دراسة المحاكاة تبين أن هناك تقارب كبير بين طريقتي أنموذج العوامل (FM) وطريقة معادلات التقدير العمومية (GEE) اذ بينت قيم متوسط مربعات الخطأ MSE هذا التقارب.

4- أظهرت النتائج التطبيق العملي أن لكل من الأساليب الإحصائية المستخدمة في هذه الدراسة ان هناك تقديرات ونتائج جيدة بالنسبة لمعلمات الأنموذج المختلط والمتمثل بأنموذج انحدار بواسون وأنموذج الانحدار الطبيعي (كاوس) اذ بينت أن فعالية أنزيم GPT ومستوى الدهون الثلاثية بالدم TRI تتأثر بالصفات المدروسة (الجنس SEX والوزن WIEGHT ونوع خضاب الدم HB_TYPE) ولكن بنسب متفاوتة.

8- التوصيات :

1- اعتماد طريقة أنموذج العوامل (FM) وطريقة معادلات التقدير العمومية (GEE) في تقدير معلمات نماذج انحدار بواسون المختلط للبيانات الطولية.

2- استخدام طرائق متعددة المتغيرات مثل طريقة المتغير الكامن (Latent Variable) والاسلوب البيزي في الدراسات المستقبلية في تقدير معالم النماذج المختلطة والتي تتميز بها الاستجابات بخاصية الارتباط.

3- إشاعة استخدام الأنظمة البرمجية الحديثة في تبويب وأرشفة البيانات في المؤسسات الحكومية كافة.

4- القيام بدراسة مماثلة باستخدام أساليب أخرى في التقدير لم يتم التطرق إليها في هذا البحث مثل (Copulas Model, Graphical Model) .

5- نوصي باستخدام الأساليب الإحصائية الحديثة وتطبيقها في الدراسات التجريبية وخاصة الدراسات الزراعية.

9- المصادر:

1- المشـهداني، محمـد احمـد شـويل، (2004)، "تـأثير إضـافة الفطـر المحـاري Pleurotus ostreatus ومخلفات زراعته الى العليقة في بعض الصفات الانتاجية والفسلجية لذكور فروج اللحم"، رسالة ماجستير علوم في الزراعة، كلية الزراعة، جامعة بغداد.

2- عبد الحافظ، علي سيف الدين، (2012)، "تقدير أنموذج لا معلمي للبيانات المزدوجة في بعض الأنشطة الاقتصادية في العراق"، رسالة دكتوراه في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد.

3- Fitzmaurice G. M. and Laird N. M,(1997). "Regression models for mixed discrete and continuous responses with potentially missing values". Biometrics, 53: 110-122.

4- Fitzmaurice, G. M. and Laird, N. M. (1995)." Regression models for Bivariate discrete and continuous outcome with clustering", Journal of the American Statistical Association, 845-852.

5- Jang, M.J. (2011)." Working Correlation Selection in Generalized Estimating Equations". Thesis, Biostatistics, College of The University of Iowa.

6- Liang K, Zeger S.(1986). "Longitudinal data analysis using generalized linear models". Biometrika. 73(1):13–22.

7- Liang KY, Zeger SL, Qaqish BF.(1992)." Multivariate regression analyses for categorical data". Journal of the Royal Statistical Society, Series B 54:3–40.

8- Olkin L. and Tate R. F. (1961). "Multivariate correlation models with mixed discrete and continuous variables", Annals of Mathematical Statistics, 448-456.

9- P. J., and Ryan, L. M. (1992), "Bivariate Latent Variable Models for Clustered Discrete and Continuous Outcomes," Journal of the Amer-ican S tatistical Association, 87, 651-658.

10- Pinto, A. T. and Normand, S. L. T. (2009)." Correlated bivariate continuous and bi- nary outcomes: Issues and applications". Statistics in Medicine, 1753-1773.

11- Razi , F. , Bahrami Samani , E. and Ganjali , M.(2013) " Analysis of Mixed Correlated Bivariate Negative Binomial and Continuous Responses". Applications and Applied Mathematics: An International Journal (AAM) Vol. 8, Issue 2, pp. 404 – 415.

12- Yang, Y. and Kang, J. (2011). "Joint analysis of mixed Poisson and continuous longitudinal data with non-ignorable missing values", Computational Statistics Data Analysis, 193-207.

13- Yang, Y., Kang, J., Mao, K. and Zhang, J. (2007), Regression models for mixed Poisson and continuous longitudinal data, Statistics in Medicine; 3782-3800.