استخدام طريقة التقلص في تقدير انحدار كاما باسلوب المحاكاة

الباحث لؤي عادل عبد الجبار   أ.د قتيبة نبيل نايف

كلية الادارة والاقتصاد - جامعة بغداد- قسم الإحصاء

المستخلص

سنوضح في هذا البحث انموذج انحدار كاما على افتراض أن المتغير التابع (Y) هو توزيع كاما وأن متوسطه (μ_i) مرتبط من خلال تركيبة خطية بواسطة □□□□□□□□ □□□□□ □□□□ $identity\ link : g(\mu) = \mu$. ويحتوي أيضًا على معلمة الشكل (α_i) ، والتي تكون غير ثابتة وتعتمد ايضاً على تركيبة خطية بواسطة دالة الربط اللوغارتمية $Log\ link : h(\alpha_i) = log\ log\ (\alpha_i)$، حيث سيتم تقدير معلمات انحدار كاما باستعمال طريقتي للتقدير وهي طريقة الامكان الاعظم (Method Maximum Likelihood) و طريقة التقلص (Shrinkage Method) واجراء المقارنة بين هذه الطرائق بأستعمال معيار المقارنة متوسط مربعات الخطأ التكاملي (IMSE) ، وباستخدام اسلوب المحاكاة باستعمال برنامج (R) وبتجربتين حيث ان كل تجربة فرضت لها قيم افتراضية وبحجوم عينات مختلفة (٥٠،٧٠،١٠٠) ، اظهرت نتائج المحاكاة ان هنالك تناقص في قيم متوسط مربعات الخطأ التكاملي (IMSE) وبازدياد حجم العينة وكانت افضل طريقة للتقدير هي طريقة التقلص (Shrinkage Method) لانها اعطت اقل (IMSE).

<u>المصطلحات الرئيسية بالبحث:</u>انحدار كاما ، طريقة الامكان الاعظم ، طرقة التقلص ، متوسط مربعات الخطأ التكاملي (IMSE).

Abstract:

In this paper, we will illustrate a gamma regression model assuming that the dependent variable (Y) is a gamma distribution and that its mean (μ_i) is related through a linear predictor and with link function which is identity link function g(μ) = μ .It also contains the shape parameter (α_i) which is not constant and also depends on the linear predictor and with link function which is the log link : $h(\alpha_i) = log\ log\ (\alpha_i)$, and we will estimated the parameters of gamma regression by using two estimation methods which are The method of Method Maximum Likelihood and the method of shrinkage and make a comparison between these methods by using the average of the integral error squares (IMSE), and by using the simulation method using (R) program and with two experiments as each experiment has default values and with different sample sizes (50,70,100), the simulation results showed that there is a decrease in the values of (IMSE) and with an increase in the sample size, and the best estimation method was the Shrinkage Method because it gave less (IMSE).

<u>Key word :</u> Gamma Regression, Maximum Likelihood Method, Shrinkage Method**,** Average squares of integrative error(IMSE).

*The research is drawn from master thesis

<u>المقدمة :</u>

يعتبر تحليل الانحدار من أهم الادوات لبناء أنماذج تمثل الظواهر المدروسة وذلك من خلال معرفة العلاقة بين المتغيرات حيث يكون أحد المتغيرات متغير تابع (معتمد) والاخرى تكون متغيرات مستقلة حيث يقوم الأنموذج بربط هذه المتغيرات بمعادلة رياضية ومن ثم نقدر معالم الأنموذج وبعدها يمكننا التنبؤ بالظاهرة المدروسة ولأستخدام هذا الانموذج يتطلب تحقق شروط فروض المربعات الصغرى حيث يم تقدير الانموذج من خلال معالمه مباشرة ، اما في حال المتغير المعتمد (Y_i) يتبع توزيع كاما فنلجأ الى هذا الاسلوب أي بالاعتماد على

*بحث مستل من رسالة ماجستير

دالة الربط بحيث نلجأ الى تقديره من خلال معالمه والتي تمثل كل معلمة معادلة انحدار(تركيبة خطية) حسب دالة الربط التي تم افتراضها ، حيث يكون المتغير الاستجابة (Y_i) يتبع توزيع كاما حيث تكون قيمه ضمن الفتر $(0,\infty)$.

هدف البحث **objective of the study** :

تقدير انموذج انحدار كاما (GRM) بطريقة الامكان الاعظم(Method Maximum Likelihood) و طريقة التقلص (Shrinkage Method) واجراء المقارنة بين الطريقتين بأستعمال المحاكاة من خلال معيار المقارنة (IMSE).

١- الجانب النظري:

١-١ توزيع كاما:

يعد توزيع كاما من احد التوزيعات المستمرة حيث تم تعريفه من قبل الباحثة stacy (١٩٦٢), ويمكن استخدام توزيع كاما بمرونة كبيرة في تحليل المتغيرات العشوائية الايجابية ويكون معتمد دائما في اوسع المجالات الطبية [Bossio , 2015 :p2] [Kalbfleisch, 2011: p42].

اذا كان المتغير العشوائي (y) ياخذ شكل توزيع كاما ذو المعلمتين فتكون دالة الكثافة الاحتمالية (pdf) له كالاتي [Adekanmbi ,2017:p4]

$$f(y;\lambda,\alpha)=\frac{\lambda}{\Gamma_{(\alpha)}}(\lambda y)^{\alpha-1}\ exp\ exp\ (-\lambda y) \qquad \dots(1)$$

حيث ان α : تمثل معلمة الشكل (Shape Parameter).

λ : معلمة القياس (Scale Parameter).

$\Gamma_{(.)}$: دالة كاما

وان $\alpha,\lambda>0$

٢-١ خصائص توزيع كاما:

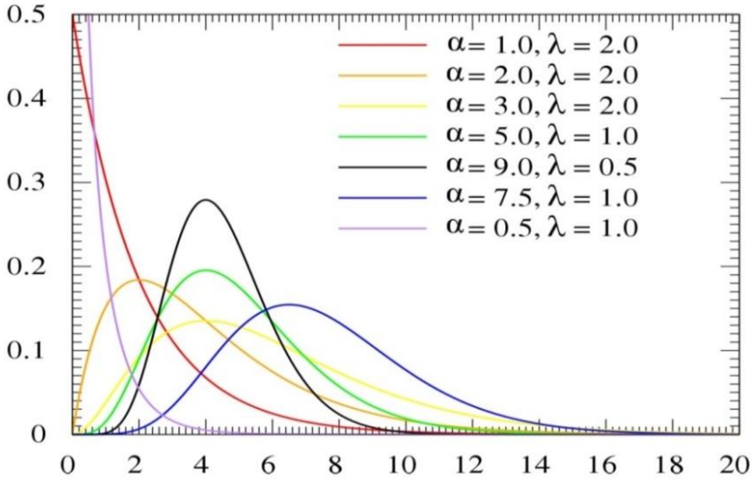
$$Y_i \sim G(\alpha,\lambda) \qquad \dots(3)$$

١- الوسط الحسابي [Cuervo , 2001 : p52]:

$$E(Y_i)=\frac{\alpha}{\lambda} \qquad \dots(4)$$

٢- التباين [Cuervo , 2001 : p52]:

$$V(Y_i)=\frac{\alpha}{\lambda^2}=\mu^2\left(\frac{1}{\alpha}\right)=\sigma^2E(Y_i)^2 \qquad \dots(5)$$



الشكل رقم(١) يبين دالة الكثافة الاحتمالية (pdf) لتوزيع كاما.

ويمكن كتابة الدالة الكثافة الاحتمالية (Pdf) لتوزيع كاما بدلالة معلمة الشكل (α) والوسط الحسابي (μ) وكالاتي Adekanmbi ,2017:p4]:]

$$\mu = \frac{\alpha}{\lambda} \qquad \dots (6)$$

$$\lambda = \frac{\alpha}{\mu} \qquad \dots (7)$$

$$f(\mu,\alpha) = \frac{1}{y\Gamma_{(\alpha)}}\left(\frac{\alpha}{\mu}y\right)^{\alpha}\ exp\ exp\left(-\frac{\alpha}{\mu}y\right) \quad I_{(0,\infty)}y \qquad \dots (8)$$

١-٣ انحدار كاما :

يعتبر انموذج انحدار كاما (GRM) امتداداً لموضوع النماذج الخطية المعممة (Generalized Linear Models) (GLM) اذ ان النماذج الخطية المعممة تختلف عن الانحدار الخطي المعروف كون ان القيم المتوقعة μ_i للمتغير العشوائي Y تستبدل بدالة الربط $g(\mu_i) = \eta$ حيث ان η هي تركيبة خطية من المتغيرات المستقلة والفائدة من دالة الربط هو ان يكون التباين مستقرا في جميع حالاته ، بالاضافة الا انه يمكن اختيار توزيع الخطأ الخاص بالنموذج بصورة يكون مستقل وبعكس الانحدار الخطي الذي يجب دائما ان يكون توزيع الخطأ توزيع طبيعياً ، تعتمد فئة نماذج انحدار كاما على افتراض أن المتغير التابع هو توزيع كاما وأن متوسطه مرتبط بمعادلة انحدار من خلال مؤشر خطي (تركيبة خطية) ووظيفة ارتباط. يمكن أن يكون هذا الارتباط هو الدالة الطبيعية أو الدالة المعكوسة أو دالة اللوغاريتم. يتضمن الانموذج أيضًا معلمة الشكل ، والتي قد تكون ثابتة أو تعتمد بمعادلة انحدار من خلال مؤشر خطي (تركيبة خطية) ووظيفة ارتباط ، كدالة لوغاريتم ،ويتم تطبيق انموذج انحدار كاما في مجموعة واسعة من التطبيقات التجريبية كماهو الحال في عملية تحديد اطار العمل في شركات التأمين ويعتمد في اوسع المجالات الطبية وايضا في حالات الدخول الى المستشفيات بسبب الامراض النادرة وجاءت تسمية انموذج انحدار كاما بهذه التسمية كون متغير الاستجابة (Y_i) يتبع توزيع كاما حيث تكون قيمه ضمن الفترة $(0,\infty)$ [Bossio , 2015:p2] :

١-٤ انماذج انحدار كاما:

لتكن $Y_i \sim G(\mu_i,\alpha)$ حيث ان : $i = 1,2,\dots\dots,n$ هي متغيرات عشوائية مستقلة و α معلمة الشكل وتكون هنا قيمة ثابتة وبالتالي فان معادلة انحدار كاما هي عبارة عن الوسط الحسابي للمتغير Y وكالاتي [McCullagh,1989 :p287-296] [Cuervo , 2016 :p3]:

$$\eta_i = g(\mu_i) = \acute{x}_i\beta \qquad \dots (9)$$

$\beta = (\beta_0,\beta_1,\dots\dots,\beta_p)$: متجه معلمات الانحدار المجهولة.

$x_i = (x_{i1},\dots\dots\dots,x_{ip})$: متجه لـ p من المتغيرات المستقلة.

η_i : هي تركيبة خطية.

$g(.)$: دالة الربط

حيث هنالك ثلاث ادوات ربط (Link function)لانموذج انحدار كاما وهي :[McCullagh,1989 :p291-295]

١- $log\ link : g(\mu) = log\ log\ (\mu)$ □□□□□□□□□□□□ □□□□□□ □□□□

٢- $identity\ link : g(\mu) = \mu$ □□□□□□□□□ □□□□□□ □□□□

٣- $inverse\ link : g(\mu) = \frac{1}{\mu}$ □□□□□□□□ □□□□□□ □□□□

في حالة معلمة الشكل α غير ثابتة وهذا ماسيتم دراسته في هذا البحث أي يمكن نمذجتها كما في معادلة (٩) حيث ان انحدار كاما يسمح بالنمذجة المشتركة لمعلمات المتوسط والشكل لمتغير توزيع كاما اي ان $Y_i \sim G(\mu_i,\alpha_i)$ حيث ان : $i = 1,2,\dots\dots,n$, وكالاتي: [Cuervo , 2016 :p3]

$$\eta_{1i} = g(\mu_i) = \acute{x}_i\beta \qquad \dots (10)$$

$$\eta_{2i} = h(\alpha_i) = \acute{z}_i\gamma \qquad \dots (11)$$

$\beta = (\beta_0,\beta_1,\dots\dots,\beta_p)$, $\gamma = (\gamma_0,\gamma_1,\dots\dots,\gamma_k)$:هي متجهات معلمات الانحدار المتعلقة بالمتوسط والشكل على التوالي.

$g(\mu)$: دالة الربط للمتوسط.

$h(\alpha)$: دالة الربط للشكل (عادة تكون دالة الربط اللوغارتيمي).

η_{1i} ,η_{2i} : التنبؤات(التركيبات) الخطية.

$x_i = (x_{i1}, \dots \dots \dots, x_{ip})$: متجه لـ p من المتغيرات المستقلة.

$z_i = (z_{i1}, \dots \dots \dots, z_{ik})$: متجه لـ k من المتغيرات المستقلة.

١-٥ طريقة الامكان الاعظم Maximum Likelihood Method)(MLE)):

تعتبر هذه الطريقة من احد الطرائق المهمة لما لها من تطبيقات واسعة لتقدير معلمات النماذج الاحصائية وتتصف هذه الطريقة بعدة خواص استدلالية منها خاصية الأتساق والثبات وعدم التحيز في أغلب الاحيان أذ ان يمكن كتابة دالة الامكان لانحدار كاما بالشكل الاتي [Bossio , 2015:-4]:

$$L = \prod_{i=1}^{n} f(y;\lambda,\alpha) \quad \dots (12)$$

$$L = \prod_{i=1}^{n} \frac{1}{\Gamma_{(\alpha_i)}}\left(\frac{\alpha_i}{\mu_i}\right)^{\alpha_i} {y_i}^{\alpha_i-1} exp\left(-\frac{\alpha_i}{\mu_i}y_i\right) \quad \dots (13)$$

$$log\, log\,(L) = \sum_{i=1}^{n} \left\{-\,log\, log\,(log\, log\, \Gamma_{(\alpha_i)} + \alpha_i\, log\, log\, \left(\frac{\alpha_i y_i}{\mu_i}\right) - log\, log\,(y_i) - \left(\frac{\alpha_i}{\mu_i}\right) y_i \right\} \quad \dots (14)$$

حيث ان

$$\alpha_i = exp\,(\acute{z}\gamma) \text{ و } \mu_i = \acute{x}\beta$$

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{n} -\frac{\alpha_i}{\mu_i}\left(1-\frac{y_i}{\mu_i}\right)x_{ij} \quad ; j = 1, \dots \dots , p \quad \dots (15)$$

$$\frac{\partial L}{\partial \gamma_k} = \sum_{i=1}^{n} -\alpha_i\left[\frac{d}{d\alpha_i} log\, log\, \Gamma_{(\alpha_i)} - log\, log\, \left(\frac{\alpha_i y_i}{\mu_i}\right) - 1 + \frac{y_i}{\mu_i}\right] z_{ik} \quad ; k = 1, \dots \dots , r \quad .(16)$$

من خلال المصفوفة الهيسية (Hessian Matrix), والتي هي مصفوفة الاشتقاق الجزئي من الدرجة الثانية لدالة عددية متعددة المتغيرات وتضم جميع المشتقات الجزئية من الدرجة الثانية الممكنة للدالة [Adekanmbi ,2017:p7] .

$$\frac{\partial^2{}_L}{\partial \beta_k \beta_j} = \sum_{i=1}^{n} \frac{\alpha_i}{{\mu_i}^2}\left(1-\frac{2y_i}{\mu_i}\right)x_{ij}x_{ik} \quad ; j,k = 1, \dots \dots p \quad \dots (17)$$

$$\frac{\partial^2{}_L}{\partial \gamma_k \beta_j} = \sum_{i=1}^{n} -\frac{\alpha_i}{\mu_i}\left(1-\frac{y_i}{\mu_i}\right)x_{ij}z_{ik} \quad ; k = 1, \dots \dots r \quad \dots (18)$$

$$\frac{\partial^2{}_L}{\partial \gamma_k \gamma_j} = \sum_{i=1}^{n} -\alpha_i\left[\frac{d}{d\alpha_i} log\, log\, \Gamma(\alpha_i) - log\, log\, \left(\frac{\alpha_i y_i}{\mu_i}\right) - 1 + \frac{y_i}{\mu_i}\right] z_{ij}z_{ik}$$
$$-\sum_{i=1}^{n} \alpha_i\left[\alpha_i \frac{d^2}{d{\alpha_i}^2} log\, log\, \Gamma(\alpha_i) - 1\right] z_{ij}z_{ik} \quad ; j,k = 1, \dots \dots , r \quad \dots (19)$$

ويتم استخدام مصفوفة فيشر (Fisher information matrix) لحساب مصفوفة التباين المرتبطة بتقديرات الاحتمالات القصوى وكالاتي [Adekanmbi ,2017:p8]:

$$I(\beta) = \left[-E\left(\frac{{\partial_L}^2}{\partial \beta_k \beta_j}\right) - E\left(\frac{{\partial_L}^2}{\partial \gamma_k \beta_j}\right) - E\left(\frac{{\partial_L}^2}{\partial \gamma_k \beta_j}\right) - E\left(\frac{{\partial_L}^2}{\partial \gamma_k \gamma_j}\right)\right] \quad \dots (20)$$

$$-E\left(\frac{\partial_L{}^2}{\partial\beta_k\beta_j}\right)=\sum_{i=1}^{n}\ \frac{\alpha_i}{\mu_i{}^2}x_{ij}x_{ik} \qquad \dots(21)$$

$$-E\left(\frac{\partial_L{}^2}{\partial\gamma_k\beta_j}\right)=0 \quad ,k=1,\dots,r\ ;\ j=1,\dots,p \qquad \dots(22)$$

$$-E\left(\frac{\partial_L{}^2}{\partial\gamma_k\gamma_j}\right)=\sum_{i=1}^{n}\ \alpha_i{}^2\left[\frac{d^2}{d_{\alpha_i}{}^2}log\ log\ \Gamma(\alpha_i)-\frac{1}{\alpha_i}\right]z_{ij}z_{ik}\quad ;\ j,k=1,\dots,r \qquad \dots(23)$$

$$I(\beta)=\left[\sum_{i=1}^{n}\ \frac{\alpha_i}{\mu_i{}^2}x_{ij}x_{ik}\ 0\ 0\ \sum_{i=1}^{n}\ \alpha_i{}^2\left[\frac{d^2}{d_{\alpha_i}{}^2}log\ log\ \Gamma(\alpha_i)-\frac{1}{\alpha_i}\right]z_{ij}z_{ik}\right] \qquad \dots(24)$$

ان مصفوفة المعلومات (Fisher information matrix) هي مصفوفة قطرية حيث تتطابق احدى الكتل مع معلمات الانحدار المتوسط والاخرى مع معلمة انحدار الشكل وبالتالي فأن متجهات المعلمات β,γ تكون بشكل متعامد ومقدرات الاحتمالات القصوى $\hat{\beta},\hat{\gamma}$ مستقلة عن بعضها وغير متناظرة . ومن خلال اعلاه نلاحظ انه لايمكن تقدير معلمات انموذج انحدار كاما (GRM) بالطرائق الاعتيادية وسوف نستعمل خوارزمية درجة فيشر (Fisher score) وهي خوارزمية تكرارية للحصول على الحد الاقصى لتقدير احتمال معلمات انموذج انحدار كاما وهي مشابهة لطريقة نيوتن رافسن او خوارزمية المربعات الصغرى التكرارية الموزونة (Iterative weighted least square) حيث يستخدم فيها القيمة المتوقعة من مشتقات المصفوفة الثانية ومن خلال الخوزارزمية نحصل على تقدير المعلمات $\hat{\gamma}$ □ $\hat{\beta}$ [Cuervo , 2001 :p55-57].

$$\hat{\beta}^{(k+1)}=(XW_1{}^{(k)}X)^{-1}XW_1{}^{(k)}Y \qquad \dots(25)$$

حيث ان :

$W_1{}^k$: مصفوفة قطرية تحتوي عناصر قطرها على

$$W_i{}^k=\frac{(\mu_i{}^2)^{(k)}}{\alpha_i{}^{(k)}} \qquad \dots(26)$$

$$\hat{\gamma}^{(k+1)}=\left(ZW_2{}^{(k)}Z\right)^{-1}XW_2{}^{(k)}Y \qquad \dots(27)$$

$W_2{}^k$: مصفوفة قطرية تحتوي عناصر قطرها على

$$W_i{}^k=\frac{1}{d_i{}^{(k)}} \qquad \dots(28)$$

$$di=\alpha_i{}^{-2}\left[\frac{d^2}{d_{\alpha_i{}^2}}log\ log\ \Gamma_{(\alpha_i)}-\frac{1}{\alpha_i}\right]^{-1} \qquad \dots(29)$$

١-٦ طريقة التقلص :(Shrinkage Method)

تعتمد هذه الطريقة على المعلومات السابقة حيث تعتمد فكرة مقدر التقلص على استعمال ماهو متوفر من المعلومات السابقة حول المعلمة المراد تقديرها على شكل قيمة اولية ولتكن β_0 مع قيمة تقديرية اخرى $\hat{\beta}$ والتي تحسب باحدى الطرائق التقليدية وصيغة مقدر التقلص تحسب كالاتي [احمد، ٢٠١٣ :٩- ١٠].

$$\hat{\beta}_{sh}=K\hat{\beta}+(1-K)\beta_0 \qquad \dots(30)$$

حيث ان :

β_0: مقدر اولي لمتجه المعلمات β.

$\hat{\beta}$: مقدر لمتجه المعلمات β.

K : مقدار التقلص وينتمي للفترة (٠،١)

وبما ان K كمية ثابتة تقع بين الصفر والواحد لذلك من الممكن اشتقاقها لجعل قيمة متوسط مربعات الخطأ ((MSE اقل مايمكن للمقدر $\hat{\beta}_{sh}$ بالاعتماد على المقدر $\hat{\beta}$] احمد، ٢٠١٣ :١٠-١١:]
]من قبل الباحث]

$$MSE(\hat{\beta}_{sh}) = E(\hat{\beta}_{sh} - \beta)^2 \qquad \dots(31)$$

$$MSE(\hat{\beta}_{sh}) = E\left[K\hat{\beta}_{MLE} + (1-K)\hat{\beta}_0 - \beta\right]^2$$

بأضافة وطرح $K\beta$ نحصل على :

$$MSE(\hat{\beta}_{sh}) = E\left[K\hat{\beta}_{MLE} - K\beta + K\beta + (1-K)\hat{\beta}_0 - \beta\right]^2$$

$$MSE(\hat{\beta}_{sh}) = E\left[K(\hat{\beta}_0 - \beta) - (\hat{\beta}_0 - \beta)K + \hat{\beta}_0 - \beta\right]^2$$

$$MSE(\hat{\beta}_{sh}) = E\left[K^2(\hat{\beta}_{MLE} - \beta)^2 + 2K(1-K)(\hat{\beta}_{MLE} - \beta)(\hat{\beta}_0 - \beta) + (1-K)^2(\hat{\beta}_0 - \beta)^2\right]$$

$$MSE(\hat{\beta}_{sh}) = K^2E(\hat{\beta}_{MLE} - \beta)^2 (1-K)^2 + (\hat{\beta}_0 - \beta)^2$$

وبأشتقاق المعادلة اعلاه بالنسبة الى K:

$$\frac{\partial MSE(\hat{\beta}_{sh})}{\partial K} = 2KE(\hat{\beta}_{MLE} - \beta)^2 - 2(1-K)(\hat{\beta}_0 - \beta)^2 \qquad \dots(32)$$

$$\frac{\partial MSE(\hat{\beta}_{sh})}{\partial K} = 2KMSE(\hat{\beta}_{MLE}) - 2(1-K)(\hat{\beta}_0 - \beta)^2 \qquad \dots(33)$$

وبمساواة المشتقة بالصفر نحصل على :

$$KMSE(\hat{\beta}_{MLE}) - (1-K)(\hat{\beta}_0 - \beta)^2 = 0 \qquad \dots(34)$$

$$KMSE(\hat{\beta}_{MLE}) - (\hat{\beta}_0 - \beta)^2 + K(\hat{\beta}_0 - \beta)^2 = 0$$

$$K = \frac{(\hat{\beta}_0 - \beta)^2}{MSE(\hat{\beta}_{MLE}) + (\hat{\beta}_0 - \beta)^2} \qquad \dots(35)$$

وبنفس الاسلوب اعلاه سيكون تقدير المعلمة γ بطريقة التقلص وكالاتي:

$$\gamma_{sh} = S\hat{\gamma}_{MLE} + (1-S)\hat{\gamma}_0 \qquad \dots(36)$$

حيث ان :

γ_0: مقدر اولي لمتجه المعلمات γ.

$\hat{\gamma}_{MLE}$: مقدر طريقة الامكان الاعظم (MLE) لمتجه المعلمات γ.

S : مقدار التقلص وينتمي للفترة (٠،١)

وبما ان S كمية ثابتة تقع بين الصفر والواحد لذلك من الممكن اشتقاقها لجعل قيمة متوسط مربعات الخطأ ((MSE اقل مايمكن للمقدر $\hat{\gamma}_{sh}$ بالاعتماد على المقدر $\hat{\gamma}$] احمد، ٢٠١٣ :١٠-١١.]

وبنفس اسلوب الاشتقاق اعلاه نحصل على:

$$S = \frac{(\hat{\gamma}_0 - \gamma)^2}{MSE(\hat{\gamma}_{MLE}) + (\hat{\gamma}_0 - \gamma)^2} \qquad \dots(37)$$

٢-الجانب التجريبي:

يعتبر اسلوب المحاكاة من الاساليب التي لها اهمية كبيرة في حل المشاكل التي تواجه الباحثين بأساليب رياضية ، حيث ان المحاكاة تعطي معلومات وافية ومفيدة على الواقع العملي الذي نريد دراسته ، وتبرز اهمية استخدام اسلوب المحاكاة بصعوبة ايجاد عينات من المجتمع المراد دراسته لما يتطلب من وقت وكلفة وجهد لذلك يلجأ الباحثون الى استخدام هذا الاسلوب ، خصوصاً بعد استحداث علوم جديدة في مجال البرمجة والحاسبات.

في هذا المبحث سنستعرض كيفية كتابة البرنامج باستعمال (R) وكما يلي:

١- اختيار احجام عينات مختلفة (٥٠,٧٠,١٠٠) وذلك لمعرفة تأثير تقديرات معلمات الانموذج المستخدم بتغيير احجام العينة.

٢- اختيار قيم افتراضية لمعلمات انحدار كاما ، اذ قمنا بفرض اربعة قيم لمعلمة المتوسط (mean parameters) وكذلك اربعة قيم لمعلمة الشكل (shape parameters) وسبب اختيار هذه القيم بالنسبة للمعالم الى انه في حال تغيير قيم المعالم وحجم العينة ستكون لدينا فكرة عن نمط وسلوك المقدرات في حال تغيير هذه القيم وكما موضح في الجدول الاتي :

models	Mean parameters				Shape parameters			
	β_0	β_1	β_2	β_3	γ_0	γ_1	γ_2	γ_3
Model 1	**1.8**	**-0.1**	**-0.01**	**16.2**	**7.1**	**-0.2**	**-0.1**	**-1.9**
Model 2	**2.9**	**-0.75**	**-0.24**	**18**	**8.9**	**-0.35**	**-0.27**	**-2.5**

جدول رقم (١) يبين القيم الافتراضية للمعلمات وعدد الانماذج المفترضة.

٣- توليد المتغيرات العشوائية (المستقلة) وكمايلي:

$$X1\sim U(1,5)$$
$$X2\sim U(2,9)$$
$$X3\sim U(0.30,0.90)$$

٤- توليد المتغير المعتمد (Y) وكما يلي :[Bossio , 2015:p7]

$$Y_i\sim G(\mu_i,\alpha_i)$$

حيث ان:

$$\mu_i=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3 \qquad \dots(37)$$

وهنا تم استعمال □□□□□□□□□ □□□□□ □□□□ $.identity\ link : g(\mu_i)=\mu_i$

$$\alpha_i = exp\ exp\ (\gamma_0+\gamma_1Z_1+\gamma_2Z_2+\gamma_3Z_3) \qquad \dots(38)$$

هنا تم استعمال دالة □□□□□□□□□□□□ □□□□□□ $log\ link : h(\alpha_i)=lo(\alpha_i)$

حيث ان مصفوفة Z تحتوي على نفس المتغيرات المستقلة الموجودة في مصفوفة X [DeJong,2008: p143]

٥- تقدير معلمات انموذج انحدار كاما وذلك حسب طرائق التقدير وباحجام عينات مختلفة (٥٠,٧٠,١٠٠) وبتكرار (١٠٠) .

٦- في هذه المرحلة سيتم استعمال متوسط مربعات الخطأ التكاملي (IMSE) للمقارنة بين الطرائق المستخدمة.

$$MSE=\frac{1}{K}\sum_{j=1}^{k}\left\{\frac{1}{n}\sum_{i=1}^{n}\left[Y_i-\hat{Y}_i\right]^2\right\} \qquad \dots(39)$$

n تمثل حجم العينة.

K : يمثل مرات تكرار التجربة

٢- تحليل نتائج المحاكاة:

١- الانموذج الاول : ((Model 1

سيتم هنا تقدير معلمات الانموذج وبالطريقتين المستخدمتين وبحجم عينة (٥٠,٧٠,١٠٠) ، نلاحظ من الجدول رقم (٢) ان هناك تقارب بين القيم التقديرية والقيم المفترضة وبكافة احجام العينات المستخدمة وبالطريقتين التقديريتين ، وكذلك نلاحظ من الجدول (٣) ان نتائج المحاكاة التي اظهرت الينا بان هنالك تفوق لطريقة التقلص (SH) عند حجم عينة (٥٠) اذا بلغت قيمة (IMSE) له (1.385816) وهي اقل قيمة من قيمة الـ (IMSE) بالنسبة لطريقة الامكان الاعظم (MLE) والتي بلغت (1.410543) ،وهذه يدل على عدم وجود قيم

n		Mean parameters				Shape parameters			
		β_0	β_1	β_2	β_3	γ_0	γ_1	γ_2	γ_3
	True values	1.8	-0.1	-0.01	16.2	7.1	-0.2	-0.1	-1.9
50	est. MLE	1.8229187	0.1346999-	0.0551813-	16.652305	6.1817213	0.1067694-	0.1284851-	0.6715530-
	est. SH	1.8664337	0.1242760-	0.0270167-	16.336584	6.8154434	0.1558849-	0.1370809-	1.4102760-
70	est. MLE	1.5325638	0.1603792-	0.0073327-	16.871042	5.5387489	0.0481402-	0.0033553-	1.1605984-
	est. SH	1.7691978	0.1526447-	0.0116958-	16.491136	6.2218009	0.1099942-	0.0011194-	1.9427132-
100	est. MLE	2.1494128	0.0934597-	0.1256411-	16.687232	6.0637366	0.1435337-	0.1032861-	0.6509311-
	est. SH	1.9495299	0.0994446-	0.0790561-	16.606735	6.4790191	0.1672249-	0.1106000-	1.1427399-

شاذة في البيانات وتفشل طريقة التقلص (SH) عندما يكون هنالك تلوث في البيانات ، اما باقي حجوم العينات فنلاحظ ان هنالك تناقص في قيم مقياس معيار المقارنة (IMSE) عندما تزداد حجم العينة وان افضل طريقة هي طريقة التقلص (SH) في كل حجم عينة لانها تمتلك اقل (IMSE).

جدول رقم (٣) يبين متوسط مربعات الخطا التكاملي (IMSE) للانموذج الاول وباحجام عينات (٥٠,٧٠,١٠٠).

Method n	MLE	SH	Best
50	1.410543	1.385816	SH
70	1.324726	1.288537	SH
100	1.273056	1.239834	SH

جدول رقم (٢) يبين تقدير معلمات انحدار كاما وباستخدام الطريقتين للانموذج الاول وباحجام عينات (٥٠,٧٠,١٠٠).

٢- الانموذج الثاني : ((Model 2

سيتم تقدير معلمات الانموذج وبالطريقتين المستخدمتين وبحجم عينة (٥٠,٧٠,١٠٠) ، نلاحظ من الجدول رقم (٤) ان هناك تقارب بين القيم التقديرية والقيم المفترضة وبكافة احجام العينات المستخدمة ، وكذلك نلاحظ من الجدول (٥) ان نتائج المحاكاة التي اظهرت الينا بان هنالك تفوق لطريقة التقلص (SH) عند حجم عينة (٥٠) اذا بلغت قيمة (IMSE) له (1.495651) و هي اقل قيمة من قيمة الـ (IMSE) بالنسبة لطريقة الامكان الاعظم (MLE) والتي بلغت (1.567416) على ، اما باقي حجوم العينات فنلاحظ ان هنالك تناقص في قيم مقياس معيار المقارنة (IMSE) عندما تزداد حجم العينة وان افضل طريقة هي طريقة التقلص (SH) في كل حجم عينة لانها تمتلك اقل (IMSE).

n		Mean parameters				Shape parameters			
		β_0	β_1	β_2	β_3	γ_0	γ_1	γ_2	γ_3
	True values	2.9	0.75-	0.24-	18	8.9	0.35-	0.27-	2.5-
50	est. MLE	3.1290137	0.7499242-	0.2304758-	17.4081010	6.6234698	0.2308431-	0.1811346-	0.3944994-
	est. SH	2.8937763	0.7217299-	0.2397336-	17.7770044	7.4525152	0.2739319-	0.2245078-	1.1545810-
70	est. MLE	3.4105349	0.7258762-	0.2673302-	17.2013649	7.3264927	0.3297559-	0.1901048-	0.9391617-
	est. SH	3.0169948	0.7357795-	0.2436353-	17.7402808	8.1641454	0.3367406-	0.2198327-	1.8658303-
100	est. MLE	3.4523816	0.7023083-	0.2829363-	17.2084015	7.1357031	0.2854625-	0.1944454-	0.8615503-
	est. SH	3.1882709	0.7338650-	0.2685536-	17.6898696	7.7617758	0.2794178-	0.2237926-	1.6002210-

جدول رقم (٤) يبين تقدير معلمات انحدار كاما وباستخدام الطريقتين للانموذج الثاني وباحجام عينات (٥٠,٧٠,١٠٠).

Method n	MLE	SH	Best
50	1.567416	1.495651	SH
70	1.369032	1.278184	SH
100	1.256861	1.117652	SH

جدول رقم (٥) يبين متوسط مربعات الخطا التكاملي (IMSE) للانموذج الثاني وباحجام عينات (٥٠,٧٠,١٠٠).

الاستنتاجات:

١- من الجداول (٢,٤) نلاحظ ان جميع تقديرات المعلمات قريبة من المعلمات المفترضة لقيم الانموذج وبكافة احجام العينات (٥٠,٧٠,١٠٠) وباستعمال طريقتي التقدير.

٢- باستعمال طريقتي التقدير المستخدمة في البحث اظهرت لدينا نتائج المحاكاة بان قيم مقياس معيار المقارنة متوسط مربعات الخطأ التكاملي (IMSE) يتناقص بازدياد حجم العينة و هذا مايدل على صحة النظرية الاحصائية.

٣- من خلال تجارب المحاكاة اظهرت النتائج وباستعمال متوسط مربعات الخطأ التكاملي (IMSE) بان افضل طريقة لتقدير انموذج انحدار كاما هو طريقة التقلص (SH).

التوصيات :

١- استعمال طريقة التقلص (SH) في تقدير معلمات انحدار كاما كونها اعطت تقديرات مرنه وكفوءة في التقدير.

٢- استعمال طرائق تقدير اخرى في تقدير معلمات انحدار كاما مثل طريقة مقدر (Liy-Type) او طريقة العزوم او طرائق حصينة.

٣- في الدراسات المستقبلية يمكن استعمال دالة الربط اللوغاريتمية لمعلمة المتوسط لتوزيع كاما ($log\ link: g(\mu) = \mu_i$) ، ($log\ log\ (\mu)$ او □□□□□□□ □□□□□ □□□□ $inverse\ link: g(\mu) = \frac{1}{\mu}$ في تقدير معلمات انحدار كاما.

المصادر:

١- احمد ، احمد ذياب ومهدي ، دجلة ابراهيم (٢٠١٣), "اختيار القيمة الامثل لمعلمة التقلص **(K)** لطريقة **Shrinkage** لبيانات مراقبة هجينة من النوع الثاني" ، بحث منشور , مجلة العراقية للعلوم الادارية , المجلد ٩ , ص١-٢٠.

2-Adekanmbi, D. B. (2017), "**Generalized Gamma Regression Models with Application to CD4 Cell Counts Data of Aids Patients**" , *International Journal of Applied Mathematics & Statistical Sciences (IJAMSS)*, *6*(4), 19-36.

3-Algamal, Z. Y. (2018), "**Shrinkage estimators for gamma regression model**" , *Electronic Journal of Applied Statistical Analysis*, *11*(1), 253-268.

4-Algamal,Z.Y.,& Rasheed, K. B. (2011), "**Paired bootstrapping procedure in gamma regression model using R**" ,*Journal of Basrah Researches*, *37*(4), 201-211.

5- Bossio, M. C., & Cuervo, E. C. (2015), "**Gamma regression models with the Gammareg R package**", *Comunicaciones en EstadÃstica*, *8*(2), 211- 223.

6- Cuervo, E. C, Corrales, M., Cifuentes, M. V., & Zarate, H. (2016), "**On gamma regression residuals**".

7- Cuervo, E. C. (2001), ***"Modelagem da variabilidade em modelos lineares generalizados***" (Doctoral dissertation, Tese de D. Sc., IM– UFRJ, Rio de Janeiro, RJ, Brasil).

8- De Jong, P., & Heller, G. Z. (2008),"**Generalized linear models for insurance data**",Cambridge Books.

9- Kalbfleisch, J. D., & Prentice, R. L. (2011), "**The statistical analysis of failure time data**" , (Vol. 360). John Wiley & Sons.

10- McCullagh, P., & Nelder, J. A. (1989), "**Generalized Linear Models**" 2nd Edition Chapman and Hall. London, UK.

Using the Shrinkage Method For Estimation Gamma Regression By Using Simulation