# مجلة
# كلية التـراث الجامعة

## مجلة علمية محكمة
## متعددة التخصصات نصف سنوية

**العدد الأربعون**
**30 آب 2024**
**ISSN 2074-5621**

**رقم الايداع في دار الكتب والوثائق 719 لسنة 2011**

**مجلة كلية التراث الجامعة معترف بها من قبل وزارة التعليم العالي والبحث العلمي بكتابها المرقم (ب 3059/4) والمؤرخ في (4/7 /2014)**

# تقدير انموذج انحدار Log-Gamma Normal المتعدد في حالة البيانات غير المتماثلة

**فرح عماد محمد قتيبة نبيل نايف**
قسم الاحصاء ، كلية الادارة والاقتصاد ، جامعة بغداد ، بغداد–العراق

## المستخلص

تم في هذه البحث دراسة نموذج الانحدار الخطي المتعدد باخطاء تتبع توزيع لوغاريتم كاما الطبيعي Log-Gamma- Normal distribution والذي يعد احد وسائل نمذجة البيانات غير المتماثلة ( Asymmetric data) وتقدير معلماته بثلاث طرائق تقدير هي طريقة تعظيم التوزيع اللاحق Maximum a Posteriori (MAP) وطريقة الانتروبي العظمى العامة Generalized Maximum Entropy(GME) وطريقة الامكان الاعظم Maximum Likelihood (ML) بالاضافة الى عرض بعض المفاهيم الاساسية المتعلقة بالرسالة ومنها القيم الشاذة outliers والبيانات المتماثلة symmetric data وغير المتماثلة Asymmetric data ومعاملي الالتواء (skewness) والتفلطح ( Kurtosis ) وخصاص التوزيع (Log-Gamma- Normal distribution).

وقد تم اجراء المحاكاة باحجام عينات (25، 50، 75، 100 ) وبقيم افتراضية لمعلمات نموذج انحدار LGN (β , σ , τ , δ) واظهرت النتائج افضلية طريقة MAP على باقي طرائق التقدير لامتلاكها اقل (RMSE و AIC و BIC ) وامتلاكها صفة الثبات وعدم التذبذب بتغير احجام العينات ومؤكدة حقيقية انها تحسين لتقدير طريقة الامكان الاعظم ML كونها توظف التوزيع الاولي حول المعلمات كاحدى طرائق التقدير البيزية.

### المصطلحات والرموز

| عربي | English | المختصر |
|---|---|---|
| انموذج الانحدار لوغاريتم كاما الطبيعي | Log-Gamma-Normal Regession Model | lGNRM |
| طريقة الامكان الاعظم | Likelihood Method | ML |

| | | |
|---|---|---|
| طريقة تعظيم دالة التوزيع اللاحق | Maximum A Posterior | MAP |
| طريقة التقدير الانتروبي العظمى العامة | Generalized Maximum Entropy | GME |
| دوال الكثافة الطبيعية الملتوية | skew- normal density | SN |
| نموذج القوة العادي | Power Normal | PN |
| دالة الكثافة الاحتمالية | Probability Density Function | pdf |
| الدالة التراكمية | Cumulative distribution function | cdf |
| مشتقة لورغاريتم دالة الكاما | Digamma | $\psi(\tau)$ |
| دالة انتروبي شانون | Shannon’s entropy function | $F(P,m)$ |
| التوزيع الاولي للمعلمة | prior distribution | $p(\pi)$ |
| التوزيع اللاحق | posterior distribution | $p(\pi\|x)$ |
| مقياس اكاكي | Akaike information criterion | *AIC* |
| معيار معلومات بيز | Bayesian information criterion | BIC |
| الوسط الحسابي | Mean Squar Error | $\mu$ |
| الالتواء | Skew | $\tau$ |
| التفرطح | Kurtosis | $\delta$ |

| | | |
|---|---|---|
| دالة الكثافة الاحتمالية للتوزيع الطبيعي القياسي | Probability Density Function For NS | $\varnothing(z)$ |
| الدالة التراكمية للتوزيع الطبيعي القياسي | Cumulative distribution functionFor NS | $\Phi(z)$ |
| مشتقة لورغاريتم دالة الكاما | digamma | $\psi(\tau)$ |
| مشتقة لدالة digamma | trigamm | $\psi_1(\tau)$ |
| متجه المعلمات | Prameter Vactor | $\beta^T$ |

## 1.المقدمة

تعد نماذج الانحدار Regression Model من اهم أساليب الاحصاء الشائعة والمهمه عند الباحثين لنمذجة البيانات وتحليلها وتستعمل لدراسة العلاقة بين اثنين من المتغيرات او مجموعة من المتغيرات احدها يتغير للاستجابة Response variable وبقية المتغيرات سواء كانت متغيرا واحدا او اكثر هي متغيرات توضيحية Explanatory Variables الصفة الشائعة لمثل هكذا علاقة الانحدار والتي تناولته معظم البحوث والدراسات هي امتلاك متغير الاستجابة التوزيع الطبيعي اي ان بياناته متماثلة symmetric data ويمتلك صفة الطبيعية normality الا ان هذه الحالة لاتتحقق دائما عند تحليل البيانات باسلوب الانحدار وعدم التحقق يؤدي الى ظهور درجات التواء في النوزيع الطبيعي Skewness بنوعية الى اليمين او الى اليسار والتفرطح Kurtosis وبالتالي تكون البيانات غير متماثلة Asymmetric data وبوجودها لابد من استعمال طرائق تقدير تقليدية او بيزية تاخذ بعين الاعتبار هذا النوع من البيانات ان امتلاك المتغير المعتمد لهذا النوع من البيانات سببه ان توزيع الاخطاء العشوائية في انموذج الانحدار الخطي لايكون توزيعاً طبيعياً ومن تلك التوزيعات هو توزيع Log-Gamma Normal وفي هذه الحالة تتاثر تقديرات معلمات انموذج الانحدار الخطي فتصبح غير كفوءة لامتلاكها تبايناً كبيرا مما يؤثر على عملية التنبؤ للظاهرة المدروسة.

وان من اهم فرضيات انموذج الانحدار الخطي المتعدد هو امتلاك الاخطاء العشوائية فيه للتوزيع الطبيعي والذي يمثل الحالة التقليدية لانموذج الانحدار الذي يسمى انموذج الانحدار الطبيعي Normal Regression Model (NR) التي يتعامل معها اغلب الباحثين الامر الذي يمتد الى ان بيانات المتغير المعتمد تمتلك توزيعاً طبيعياً $Y_i \sim (x_i \beta, \sigma^2)$ لتأثر المتغير المعتمد بالاخطاء العشوائية في الانموذج حيث ان : (i=1,2,..n)، $(\beta = (\beta_0, \beta_1, \dots \beta_k)^T)$ ، $x_i = (1, x_{i1}, \dots x_{ik})$ وان T تعني المبدلة

. ولكن في الواقع التطبيقي قد لاتتحقق هذه الفرضية وعندها تمتلك الاخطاء العشوائية توزيعاً اخراً كتوزيع Log-Gamma Normal موضوع الرسالة وامتلاكها لذلك التوزيع يؤثر على عملية التقدير لمعلماته سواء تمت بطرائق تقليدية classical methods او طرائق بيزية Bayesian methods اذ سيكون لدينا انموذج انحدار لوغاريتم كاما الطبيعي (Log-Gamma- Normal Regression model (LGNRM) هذا التأثير يظهر في كفاءة تقديرات المعلمات وايضاً في عملية التنبؤ للظاهرة المدروسة المعتمدة على المعادلة التقديرية المتضمنة تقديرات المعلمات في الانموذج مما يولد صعوبة لمتخذ القرار بتحديد السبل الكفيلة لمعالجة اسباب الظاهرة. اذ يهدف هذا البحث الى تقديرمعلمات انموذج الانحدار الخطي المتعدد باخطاء عشوائية تتبع توزيع Log-Gamma Normal بطرائق تقديرمختلفة منها تقليدية وهي طريقة الامكان الاعظم Maximum Likelihood method (ML) طرائق بيزية بعدة توزيعات اولية منها التوزيع الاولي غير المعلوماتي Diffuse prior والتوزيع الاولي المرافق الطبيعي Natural Conjugate prior واجراء المقارنة بين التقديرات لتحديد الاكفأ منها من خلال استخدام المحاكاة.

## 2. الدراسات السابقة

وفي عام 2017 قدم الباحث Cordeiro واخرون [6] بحثاً تضمن اقتراحاً حديداً لانموذح الانحدار يكون كفوءاً ومرناً لتمثيل بيانات البقاء وهو انحدار لوغاريتم كاما للبيانات المراقبة وغير المراقبة الذي من خلاله يمكن تعريف نماذج خاصة ودراسة خواصه ومنها العزوم الاعتيادية وغير الكاملة ومقارنتها مع النماذج المتشعبة وغير المتشعبة nested and no-nested models التي وفقاً للتطبيق العملي وقد اظهرت النتائج ان النماذج المقترحة لديها الكفاءة في تمثيل مجموعات من البيانات الحقيقية.

وفي عام 2018 قدم الباحثان Masjkur & Folmer [9] بحثاً لتحليل نماذج الاستجابة للمعلمة العشوائية باعتماد ثلاثة توزيعات اولية هي التوزيع الطبيعي والطبيعي الملتو وتوزيع t الملتو وقد اجريت محاكاة واظهرت نتائجها تفضيل التوزيع الاولي t الملتو مقارنة بباقي التوزيعات الاولية.

وفي عام 2019 قدم الباحثان Bossio & Cuervo [5] بحثاً عرضا فيه انموذج انحدار كاما بتاثيرات مختلطة وتقدير معلماته باسلوب بيز وقد استعمل الباحثان معلمات توزيع كاما للمتوسط والشكل والتي تم نمذجتها على شكل نماذج انحدار ثابتة وعشوائية ولعدة امثلة تم اجراء المحاكاة وايضاً تطبيقاً عملياً عبر اخذ عينات باستعمال معاينة جبس Gibbs Sampling وقد اظهرت النتائج افضلية استعمال انموذج انحدار كاما بتاثيرات مختلطة فضلاً عن مرونته بالتعامل مع البيانات المفقودة من انموذج الانحدار التقليدي باخطاء تتبع التوزيع الطبيعي.

وفي عام 2021 نشر الباحث Halliwell [8] بحثاً قدم فيه بتوزيع لوغاريتم كاما Log-Gamma Distribution وتراكيبه الخطية وخصوصاً التركيب الخطي المعروف على شكل التوزيع اللوجستي العام generalized logistic distribution وبسبب ان الخسائر غالبا ما تكون موجبة نجد ان خسائر التوزيعات تبدأ من الصفر ولها ذيل من اليمين الا ان البواقي او الاخطاء العشوائية المتمركزة حول المتوسط الصفري فانها قد تملك ذيلاً يمنياً او ذيلا يساراً حدا سواء ورغم ذلك نادراً ما تكون تلك الخسائر طبيعية اذا غالباً ماتكون ملتوية موجبة (الى اليمين ) بدلا من كونها متماثلة وهذه الذيول اليمنى

تقاس باستعمال معدلات الفشل المتقاربة. وقد استعمل البحث طريقة الامكان الاعظم ML لتقدير المعلمات التوزيع اللوجستي العام لتحليل البواقي وباجراء التطبيق العملي اظهرت النتائج افضلية التوزيع المقترح لتمثيل البيانات غير المتماثلة. وفي عام 2021 نشر الباحثان عبد الجبار ونايف [1] بحثاً تناولا فيه انموذج انحدار كاما بافتراض ان المتغير المعتمد يتبع توزيع كاما ويمتلك معلمة الشكل والتي تعتمد على تركيبة خطية وقد قدرت معلماته باستعمال طريقتي الامكان الاعظم ML والبيزية Bayesian واجريت المقارنة بينهما من خلال تطبيقهما على بيانات حقيقية تتعلق بمرض يرقان الاطفال (ابو صفار في الدم ) وقد اظهرت النتائج افضلية طريقة الامكان الاعظم على الطريقة البيزية كونها اعطت مقدرات اقل متوسط مربعات الخطأ.

## 3. الجانب النظري

هذا الجانب يبين نمذجة البيانات الغير متماثلة ثم توزيع لورغاريتم كاما الطبيعي وانموذج الانحدار الخاص به وطرائق تقدير معالم هذا الانموذج ومن ثم طرائق المقارنة بين النماذج.

### *نمذجة البيانات غير المتماثلة*

هناك العديد من محاولات مختلفة لنمذجة البيانات غير المتماثلة ومن ضمنها محاولة الباحث Amini واخرون [4] فقد قدموا عائلة جديدة من التوزيعات الاحتمالية المستمرة وهي مفيدة لنمذجة البيانات غير المتماثلة وتم توليدها من توزيع F ومعلمتين موجبتين حقيقيتين احدهما للسيطرة على الالتواء والثانية للسيطرة على التفرطح (وزن الذيل ) لذلك التوزيع ، وتعرف صيغة دالة الكثافة الاحتمالية لعائلة التوزيعات المستمرة المشار اليها كالاتي [10] :

$$g(z;\tau,\delta)=\frac{\delta^{\tau}}{\Gamma(\tau)}\,[-\ln F(z)]^{\tau-1}\,[F(z)]^{\delta-1}\,f(z) \qquad )1\ldots($$

حيث ان $\tau,\delta$ معلمتان للسيطرة على الالتواء $(\tau)$ والتفرطح $(\delta)$ وحدودهما تنتمي الى مجموعة الاعداد الحقيقية الموجبة $R^{+}$.و$\Gamma(\tau)$ تشير دالة كاما الكاملة complete gamma function .و $F(z)$ تشير الى الدالة التراكمية لتوزيع المتغير z .و $f(z)$ تشير الى دالة الكثافة الاحتمالية للمتغيرz .

### *توزيع لورغاريتم كاما الطبيعي*

اعتمد الباحث Tover-Falon واخرون على المعادلة (1) في تعريف توزيع لورغاريتم كاما الطبيعي Log-Gamma- Normal distribution (LGN) من خلال تعويض الدالتين التراكمية والاحتمالية للتوزيع الطبيعي القياسي standard normal distribution $z\sim N(0\,,1)$في المعادلة(1) ليكون $X\sim LGN(\tau\,,\delta)$ وقاموا بدراسة بعض الخواص له بعد عرض دالته كالاتي : [10]

$$f(z;\tau,\delta)=\frac{\delta^{\tau}}{\Gamma(\tau)}\,[-\ln \Phi(z)]^{\tau-1}\,[\Phi(z)]^{\delta-1}\,\emptyset(z) \quad ,z\in R\ \ldots(2)$$

حيث ان $\tau, \delta$ معلمتان للسيطرة على الالتواء ($\tau$) والتفرطح ($\delta$) وحدودهما تنتمي الى $R^+$ . و $\Phi(z)$ تشير الى الدالة التراكمية للتوزيع الطبيعي القياسي . و $\emptyset(z)$ تشير الى دالة الكثافة الاحتمالية للتوزيع الطبيعي القياسي .

### ***انموذج انحدار لورغاريتم كاما – الطبيعي***

ان كتابة انموذج انحدار LGN تستوجب توضيح طريقة الوصول اليه وهذه الطريقة تبدأ من تعريف متغيرجديد w بدالة متغير z المعرفة دالته في المعادلة (2) باستعمال التحويل الخطي وفق توسيع لمعلمتي الموقع والقياس Location-Scale Extension للمتغير العشوائي z المعرف بالمعادلة (2) وكالاتي [10] :

$$w = \mu + \sigma z \quad \rightarrow \quad z = \frac{w-\mu}{\sigma} \ldots (3)$$

وعلى ضوء المعادلة (3) يمكن كتابة التوزيع الاحتمالي للمتغير w كالاتي:

$$f(w; \mu, \sigma, \tau, \delta) = \frac{\delta^{\tau}}{\sigma\,\Gamma(\tau)} \left[-\ln \Phi\left(\frac{w-\mu}{\sigma}\right)\right]^{\tau-1} \left[\Phi\left(\frac{w-\mu}{\sigma}\right)\right]^{\delta-1} \emptyset\left(\frac{w-\mu}{\sigma}\right) \quad , w \in R \quad \ldots (4)$$

ويمكن الاشارة الى توزيع المتغير w المشار اليه في المعادلة (4) بالشكل الاتي :

$$w \sim LGN(\mu, \sigma, \tau, \delta) \quad \ldots (5)$$

حيث ان $\mu$ تمثل معلمة الموقع location parameter و $\sigma$ تمثل معلمة القياس scale parameter .

وبافتراض توفر انموذج الانحدار الخطي المتعدد التقليدي الاتي[2] :

$$Y_i = x_i^T \beta + \varepsilon_i \quad , i = 1,2, \ldots n \quad \ldots (6)$$

حيث ان $Y_i$ المشاهدة i من مشاهدات متغير الاستجابة (المعتمد ) . $x_i^T$ موجه صف من مرتبة (1 × p) للمشاهدة i من مشاهدات كل المتغيرات التوضيحية بما فيها الحد الثابت (p=k+1). $\beta$ موجه من مرتبة (p × 1) من معلمات انموذج الانحدار الخطي التقليدي المجهولة . $\varepsilon_i$ الخطأ العشوائي i والذي يملك التوزيع الطبيعي اذ ان $\varepsilon_i \sim N(0, \sigma^2)$ وبذلك فان فرض الطبيعية لمتغير الاستجابة ( المعتمد ) $Y_i$ اي ان $Y_i \sim (x_i\beta, \sigma^2)$ وبالتالي فان مقدرت طريقة المربعات الصغرى الاعتيادية OLS تكون مناسبة لتقدير معلمات الانموذج التقليدي لتحقق فرضيات اجراءها.

ان ماتقدم كان بافتراض ان الخطأ العشوائي يتملك توزيعاً طبيعياً ولكن بافتراض ان الخطأ العشوائي في المعادلة (6) يمتلك توزيع لورغاريتم كاما الطبيعي LGN اي ان $\varepsilon_i \sim \mathrm{LGN}(0\,,\sigma\,,\tau\,,\delta)$ وهذا يعني تاثر متغير الاستجابة وتغير توزيعه الى ان يكون $Y_i \sim \mathrm{LGN}(x_i^T\beta\,,\sigma\,,\tau\,,\delta)$ بعد الاعتماد على المعادلتين (4) و(5) من خلال الاتي [10]:

$$Y_i = w_i \;\; , x_i^T\beta = \mu \;\; ,$$
$$z = x_i = (1, x_{i1}, \dots x_{ik}) \;\; , (\beta = (\beta_0, \beta_1, \dots \beta_k)^T) \quad \dots (7)$$

وعلى ضوء ملاحظات المعادلة (7) وتوزيع الخطأ العشوائي $\varepsilon_i \sim \mathrm{LGN}(0\,,\sigma\,,\tau\,,\delta)$ يمكن الحصول على انموذج انحدار لورغاريتم كاما الطبيعي Log-Gamma-Normal Regression Model (LGNRM) وتكون صيغة انموذج الانحدارLGN كما معرفة في الصيغة (7) باستثناء ان الخطأ العشوائي فيه له توزيع LGN اي $\varepsilon_i \sim \mathrm{LGN}(0\,,\sigma\,,\tau\,,\delta)$ والذي يمكن كتابته كالاتي :

$$Y = X\beta + \varepsilon \quad \dots (8)$$

وعلى ضوء ما تقدم يمكن كتابة الدالة الاحتمالية لمتغير الاستجابة ( المعتمد ) $Y_i$ في انموذج الانحدار LGN كالاتي:

$$f(Y_i;\theta) = \frac{\delta^{\tau}}{\sigma\,\Gamma(\tau)} \left[-\ln\Phi\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)\right]^{\tau-1} \left[\Phi\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)\right]^{\delta-1} \emptyset\left(\frac{Y_i - x_i^T\beta}{\sigma}\right) \;\; , Y_i \in R \quad \dots (9)$$

حيث ان : $\theta = \beta^T\,,\sigma\,,\tau\,,\delta \quad , \quad i = 1,2,\dots,n$

### ***طرائق التقدير لانموذج انحدار*** LGN

سيتم في هذه البحث تقدير معلمات انموذج انحدار LGN باستعمال بطريقتين الاولى والتي تفترض ان المعلمات المراد تقديرها هي كميات ثابتة متمثلة بطريقة الامكان الاعظم ML وطرائق تقدير تندرج ضمن طرائق التقدير البيزية المعتمدة على بعض التوزيعات الاولية التي تفترض بان المعلمات المراد تقديرها هي كميات عشوائية [2] متمثلة بطريقة تعظيم دالة التوزيع اللاحق Maximum a Posteriori (MAP) بتقدير بيز المعتمد على التوزيع الاولي غير معلوماتي Bayes Diffuse prior (BD) وكالاتي :

### ***طريقة الامكان الاعظم***

يمكن الحصول على تقديرات طريقة الامكان الاعظم ML بالاعتماد على المعادلة(9) اذ يمكن كتابة دالة الامكان لحجم عينة n من مشاهدات المتغير الاستجابة (المعتمد ) $Y = (Y_i\,, Y_2, \dots , Y_n)$ وكالاتي :

$$L(\,\theta; Y) = \left[\frac{\delta^{\tau}}{\sigma\,\Gamma(\tau)}\right]^{n} \times \prod_{i=1}^{n}[-\ln\Phi(\tfrac{Y_i - x_i^T\beta}{\sigma})]^{\tau-1} \times \prod_{i=1}^{n}\left[\Phi\left(\tfrac{Y_i - x_i^T\beta}{\sigma}\right)\right]^{\delta-1} \times$$

$$(2\pi)^{-n0.5}\exp\left(\left(-\tfrac{1}{2\sigma^2}(Y - X\beta)^T\,(\,Y - X\beta)\right)\right) \quad \dots (10)$$

ولغرض ايجاد مصفوفة معلومات فشر Fisher information matrix لموجه المعلمات $(\theta = \beta^T\,, \sigma\,, \tau\,, \delta)$ يتم البدء بإيجاد المشتقات التفاضلية الجزئية الثانية التي يمكن الحصول عليها كالاتي [10]:

$$\frac{\partial^2 \ln L}{\partial\,\beta\beta^T} = \frac{1}{\sigma^2}\,X^TX + \frac{\tau-1}{\sigma^2}\,X^T\,A_2X + \frac{\delta-1}{\sigma^2}X^T\,A_3X \qquad \dots (11)$$

$$\frac{\partial^2 \ln L}{\partial\,\beta\sigma} = \frac{2}{\sigma^3}\,X^T(\,Y - X\beta) + \frac{\tau-1}{\sigma^2}\,X^T\,A_4 + \frac{\delta-1}{\sigma^2}X^T\,A_5 \quad \dots (12)$$

$$\begin{aligned}\frac{\partial^2 \ln L}{\partial\,\sigma^2} = {} & -\frac{n}{\sigma^2} + \frac{3}{\sigma^2}\sum_{i=1}^{n}\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)^2 + \frac{\tau-1}{\sigma^2}\sum_{i=1}^{n}\{\,-2\frac{a_i}{r_i}(\frac{Y_i - x_i^T\beta}{\sigma}) \\ & + \frac{a_i^2}{r_i^2}\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)^2 + \frac{a_i\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)^3 + a_i^2\,\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)^2}{r_i}\} \\ & + \frac{\delta-1}{\sigma^2}\sum_{i=1}^{n}[\,-2\,a_i(\frac{Y_i - x_i^T\beta}{\sigma}) + a_i\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)^3 \\ & + a_i^2\,\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)^2\,] \quad \dots (13)\end{aligned}$$

$$\frac{\partial^2 \ln L}{\partial\,\beta\delta} = \frac{1}{\sigma}\,X^TA_1\,, \frac{\partial^2 \ln L}{\partial\,\beta\tau} = \frac{1}{\sigma}\,X^TA_6\,, \frac{\partial^2 \ln L}{\partial\,\sigma\tau} = \frac{1}{\sigma}\,Z^TA_6\,, \frac{\partial^2 \ln L}{\partial\,\sigma\delta} = \frac{1}{\sigma}\,Z^TA_1$$

$$\frac{\partial^2 \ln L}{\partial\,\tau^2} = n\,\psi_1(\tau)\,, \frac{\partial^2 \ln L}{\partial\,\tau\delta} = -\frac{n}{\delta'}, \quad \frac{\partial^2 \ln L}{\partial\,\delta^2} = \frac{n\tau}{\delta^{2'}}$$

حيث ان:

$$A_1 = (a_1, a_2, \dots, a_n)^T \quad a_i = \frac{\emptyset\,(\frac{Y_i - x_i^T\beta}{\sigma})}{\Phi\left(\frac{Y_i - x_i^T\beta}{\sigma}\right)} \quad \dots (16)$$

$$R = \operatorname{diag}\left[\frac{1}{r_1} \quad \dots \quad \frac{1}{r_n}\right] \quad , r_i = \ln\Phi\left(\frac{Y_i - x_i^T\beta}{\sigma}\right), i = 1,2,\dots,n \dots\dots (17)$$

حيث ان $I_n$ مصفوفة الوحدة من مرتبة n. و$\psi(\tau)$ دالة digamma والتي تعرف بانها مشتقة لورغاريتم دالة الكاما ويتم كتابتها كالاتي :

$$\psi(\tau) = \frac{d}{d(\tau)}\ln\Gamma(\tau) = \frac{\acute{\Gamma}(\tau)}{\Gamma(\tau)} \dots (18)$$

وهي تمثل الجزء الاول من دوال متعددة كاما polygamma functions وتسمى ايضا دالة كاما الثنائي double –gamma ، ولكونها هي دالة متزايدة بقوة strictly increasing ومقعرة بقوة strictly concaveعلى الفترة $(0,\infty)$ يمكن ان يكون سلوكها التقاريبي كالاتي :

$$\psi(\tau) \sim \ln\tau - \frac{1}{2\tau} \dots (19)$$

اذ ان :

$$A_2 = \begin{bmatrix} \left(\frac{a_1^2}{r_1^2} + \frac{\left( a_1 \left(\frac{Y_1 - x_1^T\beta}{\sigma}\right),\ + a_1^2 \right)}{r_1}\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(\frac{a_n^2}{r_n^2} + \frac{\left( a_n \left(\frac{Y_n - x_n^T\beta}{\sigma}\right),\ + a_n^2 \right)}{r_n}\right) \end{bmatrix} \ldots.(20)$$

$$A_3 = \begin{bmatrix} \left(a_1 \left(\frac{Y_1 - x_1^T\beta}{\sigma}\right) + r_1^2\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(a_n \left(\frac{Y_n - x_n^T\beta}{\sigma}\right) + r_n^2\right) \end{bmatrix} \ldots.(21)$$

$$A_4 = \begin{bmatrix} -\frac{a_1}{r_1} + \frac{a_1^2\left(\frac{Y_1 - x_1^T\beta}{\sigma}\right)}{r_1^2} + \frac{\left( a_1 \left(\frac{Y_1 - x_1^T\beta}{\sigma}\right)^2 + a_1^2\left(\frac{Y_1 - x_1^T\beta}{\sigma}\right)\right)}{r_1} \\ \vdots \\ -\frac{a_n}{r_n} + \frac{a_n^2\left(\frac{Y_n - x_n^T\beta}{\sigma}\right)}{r_n^2} + \frac{\left( a_n \left(\frac{Y_n - x_n^T\beta}{\sigma}\right)^2 + a_n^2\left(\frac{Y_n - x_n^T\beta}{\sigma}\right)\right)}{r_n} \end{bmatrix} \ldots(22)$$

$$A_5 = \begin{bmatrix} a_1\left(\frac{Y_1 - x_1^T\beta}{\sigma}\right)^2 + \ a_1^2\left(\frac{Y_1 - x_1^T\beta}{\sigma}\right) - \ a_1 \\ \vdots \\ a_n\left(\frac{Y_n - x_n^T\beta}{\sigma}\right)^2 + a_n^2\left(\frac{Y_n - x_n^T\beta}{\sigma}\right) - \ a_n \end{bmatrix} \ldots .(23)$$

$$A_6 = \begin{bmatrix} \frac{a_1}{r_1} \\ \vdots \\ \frac{a_1}{r_1} \end{bmatrix} \ldots ..(24)$$

وان $\left(\psi_1(\tau)\right)$ هي دالة trigamma والتي تعرف بانها مشتقة لدالة digamma وكالاتي :

| | |
|---|---|
| )25( | $\psi_1(\tau) = \frac{d^2}{d(\tau^2)} \ln \Gamma(\tau) = \frac{d}{d(\tau)}\ \psi(\tau)$ |

وهي تمثل الجزء الثاني من دوال متعددة كاما polygamma functions يمكن تعريفها على شكل مجموع لسلسلة لانهائية وكالاتي :

| | |
|---|---|
| )26( | $\psi_1(\tau) = \sum_{q=1}^{\infty} \frac{1}{(\tau-q)^2}$ |

واعتماداً على المشتقات الجزئية الثانية للمعلمات والمعرفة بالصيغ من (11) الى (15) نحصل على مصفوفة معلومات فيشر Fisher information matrix والتي يرمز لها بـ $I(\theta)$ لموجه المعلمات $(\theta = \beta^T , \sigma , \tau , \delta)$ عددياً من خلال حساب $n^{-1}$ من المرات لقيمة التوقع الرياضي لمصفوفة المعلومات ومنها يمكن الحصول على معكوس مصفوفة معلومات فيشر $I^{-1}(\theta)$ اذن ان محددة المصفوفة $I(\theta)$ تكون مساوية الى الاتي :

| | |
|---|---|
| )27( | $\lvert I(\theta)\rvert = \lvert \acute{X}X\rvert\left[-0.3137\ \lvert\acute{X}X\rvert + 0.3093 \sum_{j=0}^{k} \bar{x}_j^2\right]$ |

اذ ان :

| | |
|---|---|
| )28( | $\bar{x}_j = \frac{\sum_{1=1}^{n} x_{ij}}{n} \quad , \ \ i = 1,2, . . , n \quad , j = 0,1, \ldots , k$ |

ويلاحظ من الصيغة (27) ان محددة مصفوفة المعلومات فيشر لايمكن ان تكون قيمتها صفراًوبالتالي فهي مصفوفة غير مفردة non-singular وهذا يعني ان معكوسها سيكون معرفاً دائماً.

ان الغاية من استخراج معكوس مصفوفة المعلومات فيشر $I^{-1}(\theta)$ هي الحصول على مصفوفة التباين والتباين المشترك covariance's matrix لموجه المعلمات $(\theta = \beta^T , \sigma , \tau , \delta)$ من خلال تطبيق الشروط اللازمة للحصول على التقريب المقارب للتوزيع الطبيعي لموجه مقدرات المعلمات $(\theta = \beta^T , \sigma , \tau , \delta)$ بطريقة الامكان الاعظم ML اذ تعرف مصفوفة التباين والتباين المشترك covariance's matrix لموجه مقدرات المعلمات $(\theta = \beta^T , \sigma , \tau , \delta)$ كالاتي :

| | |
|---|---|
| )29( | $\text{covariance's matrix}\,(\hat{\theta}\,) = I^{-1}(\theta)$ |

### اسلوب بيز في التقدير

تستند نظرية بيز في عملية التقدير على اعتبار ان المعلمات متغيرات عشوائية يمكن صياغة المعلومات المسبقة عنها على شكل دوال اولية يمكن دمجها مع بيانات العينة الجديدة التي تمثلها دالة الامكان للحصول على التوزيع اللاحق وبوجود دالة خسارة يمكن الوصول الى المقدرات المطلوبة بعد تقليل التوقع الرياضي للخسارة وهذا هو مفهوم مبسط لنظرية بيز في التقدير [2] وهناك انواعاً متعددة من طرائق بيز في التقدير سيتم في الرسالة طريقة تعظيم دالة التوزيع اللاحق:

### طريقة تعظيم دالة التوزيع اللاحق

تعد طريقة تعظيم التوزيع اللاحق MAP اسلوباً احصائياً لتقدير معلمات التوزيع الاحتمالي لمجموعة من البيانات من خلال توظيف المعلومات الاولية ( المسبقة ) prior information المتوفرة حول المعلمات المراد تقديرها مع دالة الامكان Likelihood function مما يبدو انه توسيعاً لمقدرات طريقة الامكان الاعظم ML التي يمكن الحصول عليها من تعظيم دالة الامكان لوحدها دون وجود التوزيع الاولي [3][7].

ولذا فان طريقة MAP للتقدير تقوم بتركيب التوزيع الاولي prior distribution مع دالنة الامكان Likelihood function للحصول على التوزيع اللاحق posterior distribution بتطبيق نظرية بيز Bayes theorem وكالاتي [2] :

| | |
|---|---|
| )30( | $p(\pi\|x) = \frac{p(X\|\pi)*p(\pi)}{f(x)}$ |

حيث ان $p(\pi|x)$ يمثل التوزيع اللاحق للمعلمة $\pi$ بثبوت البيانات x . و$p(x|\pi)$ يمثل دالة الامكان لبيانات المتغيرx بثبوت البيانات المعلمة $\pi$. و$p(\pi)$ يمثل التوزيع الاولي للمعلمة $\pi$. و$p(\pi)$ يمثل التوزيع االحدي للمتغير x .

وتهدف طريقة MAP الى ايجاد قيمة للمعلمة $\pi$ بحيث تجعل التوزيع اللاحق $p(\pi|x)$ في نهايته العظمى وبذلك يكون هذا الهدف مكافئ للتعظيم العددي لنظرية بيز المعرف في المعادلة(20) ، ولان

التوزيع الحدي $f(x)$ هو دالة ثابتة لكل قيم المعلمة $\pi$ لذا فان مقدر طريقة MAP يمكن تعريفه كالاتي [37]:

| | |
|---|---|
| )31( | $\hat{\pi}_{MAP} = \arg\max_{\pi} p(x|\pi) * p(\pi)$ |

وهذا التقدير يكون لكمية غير معلومة (غير مشاهدة ضمن العينة ) بحيث تكون مساوية للمنوال.

واعتماداً على ماتقدم فانه يمكن ان تبدء طريقة MAP في التقدير بالحصول على التوزيع اللاحق posterior distribution لموجه معلمات $\beta$ انموذج انحدار LGN من خلال الاعتماد على البيانات data والتوزيع الاولي والحصول على مقدرات المعلمات لغرض تعظيمه وبذلك يمكن عدها شكلا من اشكال التقدير البيزي باختلاف ان الاخير يعتمد على المقدرات بحسب نوع دالة الخسارة loss function المستعملة وخطوات طريقة MAP كالاتي :

**الخطوة الاولى** : تعريف التوزيع الاولي لموجه المعلمات $\beta$ prior distribution على شرط ان تكون باقي المعلمات $\sigma , \tau , \delta$ معلومة $f(\beta|\sigma , \tau , \delta)$.

**الخطوة الثانية** : تعريف دالة الامكان لمتغير الاستجابة ( المتغير المعتمد ) Likelihood function $f(Y|\beta, \sigma , \tau , \delta) = L$ وهي معرفة في المعدلة (10) .

**الخطوة الثالثة** : تعريف دالة التوزيع الاحتمالي لبيانات متغير الاستجابة $f(Y)$ المستقل عن جميع المعلمات ويتم الحصول عليها كالاتي :

| | |
|---|---|
| )32( | $f(Y) = \int f(Y|\beta, \sigma , \tau , \delta) f(\beta|\sigma , \tau , \delta) d\beta$ |

وبتوظيف الخطوات الثلاث يمكن الحصول على التوزيع اللاحق posterior distribution لموجه المعلمات $\beta$ بمعلومية باقي المعلمات مع التوزيع الاحتمالي لمتغير الاستجابة (المعتمد ) وكالاتي :

| | |
|---|---|
| )33( | $f(\beta|Y, \sigma , \tau , \delta) = \frac{f(Y|\beta, \sigma , \tau , \delta) f(\beta|\sigma , \tau , \delta)}{f(Y)}$ |

واشارة الى المعادلة (31) يمكن وضع المعادلة (33) كالاتي :

| | |
|---|---|
| )34( | $f(\beta|Y, \sigma , \tau , \delta) = f(Y|\beta, \sigma , \tau , \delta) f(\beta|\sigma , \tau , \delta)$ |

وللحصول على مقدر طريقة MAP يتم تبسيط المعادلة (34) وكالاتي [37] :

)35(

$$\begin{aligned}\hat{\beta}_{MAP} &= \text{org } \max_{\beta} f(\beta|Y,\sigma,\tau,\delta)\\ &= \text{org } \max_{\beta}[\log f(Y|\beta,\sigma,\tau,\delta)+\log f(\beta|\sigma,\tau,\delta)]\end{aligned}$$

)36(

$$\hat{\beta}_{MAP} = \text{org } \max_{\beta}\Big[\ln f(\beta|\sigma,\tau,\delta) + n[\tau\ln(\delta) - \ln(\Gamma(\tau) - \ln(\sigma) - 0.5\,(2\pi]\,\frac{1}{2\sigma^2}(Y-X\beta)^T\,(Y-X\beta) + (\tau-1)\sum_{i=1}^{n}\ln[-\ln\Phi(\frac{Y_i-x_i^T\beta}{\sigma})] + (\delta-1)\sum_{i=1}^{n}\ln\left[\Phi\left(\frac{Y_i-x_i^T\beta}{\sigma}\right)\right]\Big]$$

وعندما تكون هناك اشارات سالبة للمقادير في المعادلة (36) يمكن كتابتها كالاتي [37] :

)37(

$$\begin{aligned}\hat{\beta}_{MAP} = \text{org } \min_{\beta}\Bigg[&\ln f(\beta|\sigma,\tau,\delta)\\ &- n[\tau\ln(\delta) - \ln(\Gamma(\tau) - \ln(\sigma) - 0.5(2\pi]\\ &+ \frac{1}{2\sigma^2}(Y-X\beta)^T\,(Y-X\beta)\\ &- (\tau-1)\sum_{i=1}^{n}\ln[-\ln\Phi(\frac{Y_i-x_i^T\beta}{\sigma})] - (\delta\\ &- 1)\sum_{i=1}^{n}\ln\left[\Phi\left(\frac{Y_i-x_i^T\beta}{\sigma}\right)\right]\Bigg]\end{aligned}$$

## مقاييس المقارنة

ان استعمال اكثر من طريقة تقدير يفرض على الباحث المقارنة بينها لغرض تحديد الطريقة الافضل في عملية التقدير وهي كالاتي [10]:

1- مقياس التحيز bias ومقياس التحيز النسبي Relative bias (RB) والذي يعرف كالاتي:

$$RB = \frac{|\text{التحيز}|}{\text{القيمة الافتراضية للمعلمة}} \quad .....(38)$$

2- مقياس الجذر التربيعي لمتوسط مربعات الخطأ RMSE والذي يعرف كالاتي:

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} \quad ....(39)$$

3- مقاييس المقارنة لاختيار النموذج الافضل لتطبيق البيانات بحسب حجم العينة هي كالاتي :

اولاً : مقياس اكاكي *(AIC)* Akaike information criterion ويعرف كالاتي :

$$\mathrm{AIC} = -2\,\widehat{\log \mathrm{L}} + 2(k+1) \quad ....(40)$$

ثانياً: معيار معلومات بيز Bayesian information criterion (BIC) ويعرف كالاتي :

$$\mathrm{BIC} = -2\,\widehat{\mathrm{LOG}\,\mathrm{L}} + (\mathrm{k}+1)\log(\mathrm{n}) \quad ....(41)$$

**4.الجانب التجريبي**

من اجل المقارنة بين مقدرات معلمات نموذج انحدار LGN التي وردت في الجانب النظري في هذه الرسالة سيتم الاعتماد على اسلوب المحاكاة (Simulation) بطريقة (Monte Carlo) ولمختلف من احجام العينات والقيم الاولية للمعلمات، وقد تم الاعتماد على المعايير الاحصائية (التحيز ، التحيز النسبي ،الجذر التربيعي لمتوسط مربعات الخطأ RMSE) ، ومعياري AIC وBIC. اما نتائج المحاكاة فقد تم الحصول عليها بالاعتماد على برامج كتبت بلغة (Matlab). وسوف نستعمل حالتين من القيم الافتراضية للمحاكاة وكما يلي:

**خطوات اجراء المحاكاة**

من اهم خطوات اجراء تجربة المحاكاة هي توليد البيانات data لمتغير الاستجابة (المعتمد Y ) والمتغيرات التوضيحية $X^{'S}$ وكل ذلك يعتمد على نوع انموذج الانحدار الخطي المراد تطبيقه في تجربة المحاكاة سواء كان بسيطاً ام متعدداً ، واعتماداً على ماورد في المصدر [10] فقد اعتمدت الرسالة نفس انموذج الانحدار بالقيم الاولية المفترضة وهو يمثل وكالاتي :

$$\mathrm{Y_i} = \beta_0 + \beta_1 \mathrm{X_i} + \mathrm{u_i} \quad , \mathrm{i} = 1,2,\ldots,\mathrm{n} \quad ....(42)$$

للحصول على بيانات المتغيرين Y و X نتبع الاتي :

1- توليد قيم عشوائية للمتغير z .
2- يتم حساب قيم دالة pdf لقيم المتغيرz باعتبار انها تتوزع توزيعاً طبيعياً بوسط حسابي 0 وتباين 1 .

3- يتم حساب قيم cdf لقيم المتغير z باعتبار انها تتوزع توزيعاً طبيعياً بوسط حسابي 0 وتباين 1 .

4- بالاعتماد على الخطوتين (1) و(2) نستعمل طريقة الامكان الاعظم لتقدير قيم معلمتي دالة توزيع لوغاريتم كاما الطبيعي Log Gamma Normal ، اذ تم تكرار عملية التوليد لحجم توليد محدد ومن ثم حساب الوسط الحسابي لقيم المعلمتين المقدرة ومن ثم حساب قيم الاخطاء العشوائي التي لها توزيع Log Gamma Normal علماً بانه تم استعمال القيم الافتراضية الاولية للمعلمات $\tau$ و $\gamma$ للبدء بعملية توليد الاخطاء العشوائية والهدف من ذلك اعطاء الحصانة للاخطاء لتوزيعها توزيع LGN.

5- يتم توليد قيم عشوائية من جديد للمتغير التوضيحي X.

6- بالاعتماد على القيم الافتراضية الاولية لمعلمات الانحدار $\beta_0$ , $\beta_1$ الموضحة في الجدول (1) وقيم الاخطاء العشوائية المستحصل عليها من الخطوة (4) وقيم المتغير التوضيحي X المستحصل عليها في الخطوة (5) يتم الحصول على قيم متغير الاستجابة (المعتمد) Y تبعاً لانموذج الانحدار الخطي المعرف في اعلاه والذي سيملك توزيع لوغاريتم كاما الطبيعي Log Gamma Normal.

اما احجام العينات التي ساعتمدها في هذه الدراسه فهي (25، 50 ، 75 ، 100 ) ، وسيتم المقارنة تبعاً لمقاييس المقارنة التي تم توضيحا في الجانب النظري لتحدي الطريقة الافضل في عملية التقدير لمعلمات نموذج الانحدار الخطي باخطاء تتبع توزيع Log Gamma Normal.

والقيم الافتراضية الاولية تكون بحسب الحالات التي اعتمدتها الباحثة موضحة بالجدول الاتي :

الجدول (1) : القيم الافتراضية الاولية للمحاكاة

| القيم الافتراضية الاولية | | | | | حالات المحاكاة |
|---|---|---|---|---|---|
| $\Delta$ | $\tau$ | $\sigma$ | $\beta_1$ | $\beta_0$ | |
| 0.75 | 0.75 | 0.5 | 1 | 2 | الاولى |
| 0.75 | 1.5 | 0.5 | 1 | 2 | الثانية |

الجدول (2) : نتائج قيم المحاكاة للحالة الاولى والحالة الثانية

| الحالة الثانية | الحالة الاولى | حالة المحاكاة |
|---|---|---|

| حجم العينة | طريقة التقدير | MLE | MAP | MLE | MAP |
|---|---|---|---|---|---|
| 25 | RMSE | 1.26235 | 0.55743 | 1.38109 | 0.32192 |
| | AIC | 11.5251 | -27.8489 | 17.6427 | -59.5152 |
| | BIC | 13.9629 | -25.4112 | 20.0805 | -57.0774 |
| 50 | RMSE | 1.30584 | 0.44151 | 1.28567 | 0.41238 |
| | AIC | 25.5402 | -79.8679 | 23.4125 | -89.2862 |
| | BIC | 29.3643 | -76.0439 | 27.2366 | -85.4621 |
| 75 | RMSE | 1.33114 | 0.44928 | 1.28567 | 0.41238 |
| | AIC | 40.1062 | -118.046 | 23.4125 | -89.2862 |
| | BIC | 44.7412 | -113.411 | 27.2366 | -85.4621 |
| 100 | RMSE | 1.27826 | 0.47011 | 1.30204 | 0.36071 |
| | AIC | 46.3931 | -149.184 | 49.602 | -206.719 |
| | BIC | 51.6034 | -143.973 | 54.8124 | -201.509 |

يلاحظ من الحالة الاولى للمحاكاة افضلية طريقة تعظيم التوزيع اللاحق MAP في عملية التقدير لامتلاكها اقل قيمة لمعيار RMSE وAIC وBIC ولجميع احجام العينات وتفوقها على طريقة طريقة الامكان الاعظم ML مما يدل على افضلية الطريقة في عملية تقدير معلمات نموذج انحدار لوغاريتم كاما الطبيعي LGN. وفي الحالة الثانية للمحاكاة افضلية طريقة تعظيم التوزيع اللاحق MAP في عملية التقدير لامتلاكها اقل قيمة لمعيار RMSE وAIC وBIC ولجميع احجام العينات

وتفوقها على طريقة طريقة الامكان الاعظم ML مما يدل على افضلية الطريقة في عملية تقدير معلمات نموذج انحدار لوغاريتم كاما الطبيعي LGN.

## 5. الاستنتاجات

1- تفوق طريقة التقدير تعظيم التوزيع اللاحق MAP على بقية طرائق التقدير وعدم تاثرها بتغير حجم العينة واكتسابها صفة ثبات التقدير واحتلالها المرتبة الاولى عند عملية التقدير وقد اظهر معيار RMSE ومعياري AIC و BIC ذلك بوضح.

2- تأثر مقدرات طريقة الامكان الاعظم ML لبعض معلمات انموذج الانحدار LGN اي ان هذه الطريقة لا تتصف بصفة الثبات وأنها تتذبذب بشكل عام عند حالة المحاكاة الثانية وهي بذلك تشابه ما تم التوصل اليه في تجربة المحاكاة الأولى.

## 6.المصادر

1- عبد الجبار، لؤي عادل، نايف ، قتيبة نبيل (2021) " مقارنة بين طريقة الامكان الاعظم والطريقة البيزية في تقدير انحدار كاما مع تطبيق عملي " Journal of Economics and Administrative Sciences ، Vol.27 (NO. 125), pp. 477-492.

2- كاظم ، اموري هادي ، مسلم ، باسم شليبه (2002) " القياس الاقتصادي المتقدم – النظرية والتطبيق " مكتبة الامل ، بغداد شارع الصناعة مجاور الجامعة التكنولوجية .

3- مسلم ، باسم شليبه ، قمر ، سيف الدين هاشم (2019) " دراسة مقارنة بين طريقة بيز وطريقة تعظيم دالة التوزيع اللاحق لتقدير معلمات انموذج نمو ويبل ذو الاربع معلمات" Journal of Wasit for Human and Social Sciences ،Volume15, Issue 43 ، p.p 73-88 .

4- Amini, M., MirMostafaee, S.M.T.K. and Ahmadic, J. (2012)" Log-gamma-generated families of distributions" Statistics, iFirst, 1–20.

5- Masjkur, M. & Folmer. H. (2018)" BAYESIAN ESTIMATION OF RANDOM PARAMETER MODELS OF RESPONSES WITH NORMAL AND SKEW-t DISTRIBUTIONS EVIDENCE FROM MONTE CARLO SIMULATION" J. Indones. Math. Soc. Vol. 24, No. 1, pp. 27–50.

6- Cordeiro, G.M., Marcelo, Bourguignon, E., Ortega M. M., Ramires T. G.(2017) " General Mathematical Properties, Regression and Applications of the Log-Gamma-Generated Family " Communications in Statistics - Theory and Methods.

7- Cousineau, D., Helie, S. (2013) "Improving maximum likelihood estimation using prior probabilities: A tutorial on maximum a posteriori estimation and an examination of the weibull distribution" Tutorials in Quantitative Methods for Psychology, Vol. 9(2), p. 61-71.

8- Halliwell, L. J. (2021) ”The Log-Gamma Distribution and Non-Normal Error" CASUALTY ACTUARIAL SOCIETY, vol. 13/Issue 2.

9- Masjkur, M. & Folmer. H. (2018)" BAYESIAN ESTIMATION OF RANDOM PARAMETER MODELS OF RESPONSES WITH NORMAL AND SKEW-t DISTRIBUTIONS EVIDENCE FROM MONTE CARLO SIMULATION" J. Indones. Math. Soc. Vol. 24, No. 1, pp. 27–50.

10- Tovar-Falón, R., Martínez-Flórez,G. ; Bolfarine, H. (2022)" Modelling Asymmetric Data by Using the Log-Gamma-Normal Regression Model "Mathematics , 10, 1199.

# Estimation of a Multiple Log-Gamma Normal Regression Model for Asymmetric Data

**Farah Emad Mohammed**

(Statistics,Administration and Economics,University of Baghdad,Baghdad-Iraq)

**Qutaiba Nabeel Nayef**

(Statistics,Administration and Economics,University of Baghdad,Baghdad-Iraq)

KeyWords:
Asymmetric Data
Multiple Regression Model
Log-GammaNormal Distribution
Bayesian Estimation

## Abstract

In this research, the multiple linear regression model with errors following the Log-Gamma-Normal distribution was studied, which is one of the methods for modeling asymmetric data and estimating its parameters using three estimation methods: Maximum a Posteriori (MAP), Generalized Maximum Entropy (GME), and Maximum Likelihood (ML). In addition, some basic concepts related to the thesis were presented, including outliers, symmetric data, asymmetric data, skewness and kurtosis coefficients, and Log-Gamma-Normal distribution characteristics. The simulation was conducted with sample sizes (25, 50, 75, 100) and default values for the LGN regression model parameters (β, σ, τ, δ). The results showed the superiority of the MAP method over the rest of the estimation methods because it has the lowest RMSE, AIC, and BIC, and it has the property of stability and non-fluctuation with changing sample sizes. It is confirmed that it is an improvement in estimation of the maximum likelihood ML method because it employs the initial distribution around the parameters as one of the Bayesian estimation meth