المجلة العراقية للعلوم الاحصائية (13) 2008
ص ص [10–34]

حول تقدير بيز في النماذج الخطية المختلطة باستخدام معاينة جبس

محمد نذير إسماعيل قاسم* **ايمان طارق فتحي****

الملخص

ان المسألة التي يتناولها هذا البحث هي تقدير معلمات النموذج الخطي المختلط أحادي التقسيم ذي التأثيرات العشوائية وذلك بأستخدام أسلوب بيز الذي يتضمن استخدام معاينة جبس (Gibbs Sampling) ومعاينة جبس حالة خاصة من طريقة مونتي كارلو (Monte Carlo) التي تستخدم سلسلة ماركوف وتسمى MCMC (Markov Chain Monte Carlo Methods) التي تعتمد على تجزئة النماذج الرياضية المركبة إلى عدد من المسائل البسيطة التي يمكن تحليلها ومعالجتها بسهولة ولا سيما في التوزيعات اللاحقة التي ليس بالامكان الحصول على صيغة نهائية لها .

تم في هذا البحث افتراض النموذج الخطي المختلط ذي التأثيرات العشوائية أحادي التقسيم على مجتمع من المعاملات الذي يشمل 15 صنفاً من أصناف القطن إذ أخذت عينة عشوائية بحجم 5 أصناف من هذا المجتمع واستخدمت طريقة تحليل التباين لاختبار فرضية تساوي التباين بين هذه المعاملات والحصول على تقدير معلمات هذا النموذج . كما تم استخدام طريقة معاينة جبس للحصول على تقدير معلمات نفس النموذج .

وتم اختبار فرضية تساوي التباين بين المعاملات في هذه العينة العشوائية فكانت النتائج متقاربة ومشجعة بالمقارنة مع طريقة تحليل التباين .

تمت برمجة جميع الخوارزميات التي وردت في هذا البحث باستخدام البرنامج الجاهز Win BUGS .

---

* استاذ مساعد/ رئيس قسم الرياضيات في كلية التربية/ جامعة الموصل

** مدرس مساعد/ قسم الرياضيات/ كلية التربية/ جامعة الموصل

تاريخ التسلم : 2007/5/22 ـــــــــــــــــــــــــــتاريخ القبول : 2007/8/26

On Bayesian Estimation in Mixed Linear Models Using the Gibbs Sampler

ABSTRACT

This paper tackles the estimation of parameters of linear mixed random effect one–classification model by Bayesian technique which includes Gibbs sampling. Gibbs sampling is a special case of Monte Carlo Method which uses Markov Chain and so called MCMC (Markov Chain Monte Carlo).

This MCMC method depends on partition of difficult and compound models into simple ones which can be manipulated and easily analyzed, specially for the posterior distribution which are not easy to find their final formulae.

In this research the mixed random effect linear one–classification model is proposed on a population of 15 treatments including 15 types of cotton plant. A random sample of 5 types is taken and using the analysis of variance method to test the hypothesis that all the 15 types have equal effect and the estimation of the parameters is obtained Gibbs sampling is also used in order to estimate the parameters and then testing the hypothesis of equal effects of treatments. The results obtained in both ANOVA and Gibbs sampling are nearly the same and encouraging. All algorithms are programmed in this research using WinBUGS program.

1- المقدمة :

لاشك أن لنظرية التقدير أهمية كبيرة في التطبيقات العملية ، اذ أنها تـوفر قواعد يتم بموجبها تقدير معلمات مجهولة لظاهرة معينة في الواقع العملي وبمختلف المجالات .

ان دراسة مكونات التباين وتحليلها تعد من الموضوعات المهمة فـي علـم الإحصاء لما لهذا الموضوع من تطبيقات واسعة . وإن مسألة تقدير مكونات التباين (Variance Component) من النماذج الخطية التي درسها الكثير من الباحثين وبطرائق مختلفة في مجالات البحوث إذ كانت بوادر الإستخدام المبكـر لمكونـات التباين قد طرقت في بحـوث الفلكيـان (1863 – 1861) Airy , Chauvent الذين أستخدما نموذجاً احادي البيانات الذي يتضمن مكونات التباين والبالغ عـددها اثنين بعد ذلك كانت الأعمال الرائدة للعالم (1918) Fisher التي أدت إلى تطوير أسلوب تحليل التباين (ANOVA) في تقدير مكونـات التبـاين وكـذلك تقـدير مكونات التباين بطريقة الإمكان الأعظم ، وكان للباحثين في جامعة الموصـل دور في دراسة مكونات التباين في النماذج الخطية ذات التأثيرات العشوائية منهم الباحث عباس (2000) درس فيها تقدير مكونات التباين في نموذج احادي التقسـيم ذات التأثير العشوائي بواسطة مقدر بيز .

وقدم الباحث سالم (2000) النماذج الهرمية والحركية وتحليل بيز ، وقدم الباحث الطائي (2004) دراسة تناول فيها معايرة بيز وقدم الباحـث إسـماعيل (2005) دراسة تناول فيها الكشف عن القيم الشاذة بأستخدام التوزيـع المخـتلط بالاعتماد على معاينة جبس وأسلوب بيز .

إن العائق الرئيسي في الحصول على التوزيع اللاحق في أسلوب بيز غالبـاً مايتطلب تكامل دوال ذات أبعاد متعددة ، وهذا يشكل صعوبة كبيرة في الحسـابات

للحصول على التوزيع اللاحق ، إلا أن كثيراً من الطرائـق المختصـرة للتكامـل المباشر فـي الحصـول علـى التوزيعـات اللاحقـة اقترحهـا بـاحثون امثـال (Smith 1991 , Evans and Swartz 1995) . ونؤكـد هنـا علـى طريقـة (MCMC) التي تتمثل بعملية محاكاة مباشرة يتم سحبها من التوزيعات المركبـة قيد الدراسة . طرائق (MCMC) سميت بهذا الاسم بسبب ان الباحث يستخدم قيم العينة السابقة لتوليد قيم عينة لاحقـة عشـوائياً بتوليـد سلسـلة مـاركوف اذ ان الاحتمالات الانتقالية بين قيم العينة هي دالة بقيم العينة الأكثـر حداثـة ، إسـتخدم العلماء Gelfand and Smith (1990) طريقة (MCMC) ، اذ ان معاينـة جبس وطريقة (MCMC) طبقت بشكل واسع في حل مسائل بيز (Bayesian Problems) وجاء هذا الإهتمام لزيادة التطبيق في تحليل بيـز لهـذه الطريقـة (MCMC) التي اصل جـذورها خوارزميـة ميتروبـولس Metropolis and Ulam(1949) وكانت هذه الخوارزمية محاولة من قبل علماء الفيزياء للحصول على حسابات التكاملات المعقدة وذلك بالتعبير عن هذه التكاملات بتوقعات لـبعض التوزيعات الاحتمالية ومن ثم تقدير هذا التوقع بواسطة سحب عينـات مـن ذلـك التوزيع ، ان اصل معاينة جبس ظهر في موضوع معالجة الصور، وقد أظهـرت مقدرة عظيمة في التطبيق في البحوث الفيزياوية والهندسية بعيدة عن حقـل علـم الإحصاء.

2- **هدف البحث :**

يختص هذا البحث بمسألة تقدير معلمات النمـوذج الخطـي المخـتلط ذي التأثيرات العشوائية أحادي التقسيم الذي يعرف بالصيغة الرياضية الآتية:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \qquad (1)$$

$$i = 1,2,\ldots,a \qquad j = 1,2,\ldots,n$$

إذ إن :

μ : الوسط العام .

τ_i : تأثير المعاملة i.

y_{ij} : المشاهدة j في المعاملة i .

ε_{ij} : تمثل الخطأ العشوائي لكل مشاهدة .

$$\begin{aligned} &\tau_i \sim N(0,\sigma_a^2) \\ &\varepsilon_{ij} \sim N(0,\sigma^2) \qquad (2) \\ &y_{ij} \sim N(\mu,\sigma_a^2 + \sigma^2) \end{aligned}$$

ان هدف البحث هو تقدير المعلمات σ^2 , σ_a^2 , μ بطريقة معاينـة جـبس MCMC ومقارنتها مع تقدير المعلمات بواسطة تحليل التباين .

3- النماذج الخطية Linear Models :

هناك ثلاثة انواع من النماذج الخطية هي :
Montgomery (2001)

1. النموذج I : التأثيرات الثابتة Fixed Effect Model

2. النموذج II : التأثيرات العشوائية Random Effect Model

3. النموذج III : التأثيرات المختلطة Mixed Effect Model

النموذج المختلط هو خليط من التأثيرات الثابتة والعشوائية فـي النمـوذج نفسه مع الأخذ بنظر الاعتبار بعض التأثيرات مثلا μ ثابتة وبعضـها عشـوائية ، ويحدث هذا كذلك في النموذج ثنائي التقسيم وثنائي التقسيم مـع وجـود التفاعـل

وغيرها من النماذج العاملية . إذفي بحثنا هذا لدينا نموذج مختلط يشمل المعلمة μ وتأثير المعاملة τ_i التي لها توزيع احتمالي معلوم ، وفي هـذا البحـث اسـتخدمنا نموذج التاثيرات المختلطة .

4- اسلوب بيز Bayes Approach :

ظهر الاهتمام بأسلوب بيز واستخدامه في الاستدلال الاحصائي بعـد عقـد الستينات من القرن الماضي ، حيث كان ذلك محط اهتمام العديد من العلماء مـنهم على سبيل المثال :

Barnett (1982), Lyle (1985), Mike and Harrison (1989)

هؤلاء وغيرهم بذلوا جهوداً مضنية في هذا الاسلوب وهذا الاهتمام الـذي نجده عند هذه النخبة من الباحثين والعلماء انما يعكس امكانية تطبيق هذا النوع من الاحصاء في مجالات كثيرة في الحياة اليومية ، والاستدلال هو احد فـروع علـم الاحصاء ، إذ يهتم بجمع المعلومات حول المعلمـات باسـتعمال نظريـة التقـدير واختبار الفرضيات ، ويصنف الاستدلال على وفق مدرسة الاحصاء التـي يتبـع اسلوبها في التحليل والاستنتاج الى الاستدلال الكلاسيكي واستدلال بيز .

5- نظرية بيز ودالة الامكان

Bayes Theorem and Likelihood Function

تخبرنا نظرية بيز ان التوزيع اللاحق يتناسب مع حاصل ضرب دالة الامكان والتوزيع الاولي **أي ان :**

Posterior Distribution $\propto$ Likelihood Function $\times$ Prior Distribution

ان دالة الامكان تعرف الى حد ثابت ضربي أي بمعنى اخر ان ضرب دالة الامكان بثابت لا يغير دالة الامكان وبموجب هذا فان ضرب دالة الامكـان بثابـت اعتباطي لن يكون له تأثير في التوزيع اللاحق .

6- الاحتمال الاولي Prior Probability

يعبر الاحتمال الاولي الاساس الذي يعتمد عليه اسلوب بيز في معالجتـه للمسـائل المتناولة وان دوره المهم يبرز في حالة صعوبة الحصول على بيانات كافية وعلى العكس يقل تأثيره اذا كانت لدينا بيانات كثيرة ويقسم الاحتمال الاولي الى :

1- **الاحتمال الاولي ذو المعلومات القليلة** Non–Informative Prior Probability

2- **الاحتمال الاولي ذو المعلومات** Informative Prior Probability

7- توزيع معكوس كاما Inverse Gamma (IG) Distribution:

لتكن x تتوزع توزيع كاما $G_a(\alpha,\beta)$ فان دالة الكثافة الاحتمالية لـ x سوف تكون كالآتي :

$$P(x)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\exp[-\beta x]\ ,\qquad (0<x<\infty)$$

افرض ان $y=\frac{1}{x}$ فان توزيع y سوف يكون معكوس كاما .

$$y\sim IG(\alpha,\beta)$$

$$P(y)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}y^{-\alpha+1}\qquad \exp\left[-\frac{\beta}{y}\right]\ ,\qquad (0<y<\infty)$$

الوسط لتوزيع معكوس كاما هو :

$$E(y)=\frac{\alpha}{\beta-1}\qquad ,\qquad \beta>1$$

التباين لتوزيع معكوس كاما هو

$$Var(y)=\frac{\alpha^2}{(\beta-1)^2(\beta-2)}\qquad ,\qquad \beta>2$$

انظر التفاصيل : الناصر (1988)

8- تكامل مونتي كارلو Monte Carlo Integration :

ان اصل طريقة مونتي كارلو طورها الفيزيائيون لاستخدام توليد الاعـداد العشوائية لحساب التكاملات .

افرض اننا نرغب في حساب التكامل المعقد :

$$\int_a^b h(x)\ dx \qquad (3)$$

إذا استطعنا تجزئة الدالة $h(x)$ إلى حاصل ضرب دالتـين همـا $f(x)$ ودالة كثافة احتمالية $P(x)$ معرفة في الفترة (a,b) عندئذ :

$$\int_a^b h(x)\ dx = \int_a^b f(x)P(x)\ dx$$

$$= E_{P(x)}\left[f(x)\right] \qquad (4)$$

أي ان التكامل (3) يمكن التعبير عنه كتوقع للدالة $f(x)$ على دالة الكثافة الاحتماليـة $P(x)$. وعليــه إذا ســحبنا عــدداً كبيــراً مــن الأعــداد العشوائية $x_1, x_2, \ldots, x_n$ من دالة الكثافة $P(x)$ عندئذ :

$$\int_a^b h(x)\ dx = E_{P(x)}\left[f(x)\right] \simeq \frac{1}{n}\sum_{i=1}^{n} f(x_i) \qquad (5)$$

وهذه العملية لغاية المعادلة (5) تمثل تكامل مونتي كارلو .

يمكن استخدام تكامل مـونتي كـارلو لتقريـب التوزيعـات اللاحقـة (أو التوزيعات الحدية اللاحقة) التي نحتاج اليها في تحليل بيز .

اعتبر التكامل

$$I(y) = \int f(y/x)P(x)\ dx$$

ممكن تقريبه بواسطة

$$\hat{I}(y)=\frac{1}{n}\sum_{i=1}^{n} f\left(y/x_i\right) \qquad (6)$$

إذ ان : x_i تحسب من دالة الكثافة $p(x)$.

الخطأ المعياري المقدر لطريقة مونتي كارلو يعطى بالقانون .

$$SE^2\left[\hat{I}(y)\right]=\frac{1}{n}\left(\frac{1}{n-1}\sum_{i=1}^{n}\left(f\left(y/x_i\right)-\hat{I}(y)\right)^2\right) \qquad (7)$$

انظر التفاصيل : Walsh (2002) .

9- معاينة جبس Gibbs Sampler

ظهرت معاينة جبس في بحوث معالجة الصور والتي تعد حالة خاصة من خوارزمية ميتروبولس هاستنك عندما تكون القيمة العشوائية التي تكون مقبولة دائماً متمثلة $(\alpha=1)$ ويبقى من الخوارزمية تحديد كيفية تكوين سلسلة ماركوف التي سوف تتقارب إلى حالة الاستقرارية .

ان المفتاح الاساسي في طريقة معاينة جبس هو في معرفة التوزيعات الشرطية أحادية المتغير فقط ومثل هذه التوزيعات الشرطية يكون أسهل في تطبيق اسلوب المحاكاة من التوزيعات المشتركة المركبة وللتوزيعات الشرطية صيغ بسيطة وغالباً ما يكون لها التوزيع الطبيعي أو توزيع مربع كاي أو توزيعات أولية معروفة . وعلى هذا النحو نقوم بمحاكاة n من المتغيرات العشوائية بشكل تتابعي من التوزيعات الشرطية أحادية المتغير .

للدخول في معاينة جبس نعد المتغير العشوائي ثنائي المتغيرين (x,y) ونفترض اننا راغبين في حساب التوزيعات الحدية $p(y), p(x)$ وفكرة المعاينة هو ان اعتبار التوزيعات الشرطية $p(y/x), p(x/y)$ يكون ابسط من الحصول على التوزيعات الحدية بواسطة تكامل دالة الكثافة الاحتمالية المشتركة $p(x,y)$ ، مثلا :

$$p(x) = \int p(x, y)\ \ dy$$

المعاينة تبدأ بقيمة ابتدائية $y_\circ$ للمتغير y والحصول على $x_\circ$ من توليـد متغير عشوائي من التوزيع الشرطي $p(x/y = y_\circ)$. وبعد ذلك تستخدم المعاينة $x_\circ$ لتوليد قيمة جديدة للمتغير y هي y_1 الذي يسحب مـن التوزيـع الشـرطي $p(y/x = x_\circ)$ المعتمد على القيمة $x_\circ$.

المعاينة تجري كالآتي :

$$p(x/y = y_{i-1}) \qquad \leftarrow \text{من } x_i \qquad (8)$$

$$p(y/x = x_i) \qquad \leftarrow \text{من } y_i \qquad (9)$$

وبتكرار هذه العملية k من المرات ستتولد متتابعة جبس بطـول k إذ إن المجموعة الجزئية من النقاط $(x_j, y_j), 1 \le j \le m < k$ تعد سحبات المحاكـاة من التوزيع المشترك الكامل . كل تكرار من جميع التوزيعات احادية المتغير غالباً ما يطلق عليه فحص المعاينة Scan of the Sampler . للحصول على عينة مـن النقاط حجمها m (إذ إن كل نقطة من المعاينة عبارة عن متجـه مـن قيمتـين) . بحيث ان المعاينة الفعلية تبدأ بعد إزالة تأثير قيم المعاينة الابتدائيـة التـي تسـمى الحرق في المعاينة Burning in the Sampler .

وبعد مجموعة من النقاط (مثلاً بعد كل n من العينات) تتبع الحـرق فـي المعاينة فان متتابعة جبس ستؤول إلى التوزيع المستقر الذي لا يعتمد علـى القـيم الابتدائية وبهذا نكون قد ولدنا العينة المرغوب فيها من التوزيع قيد الدراسة . انظر التفاصيل :

Tierney (1994) , Walsh (2002)

10- مفهوم معاينة جبس Concept of Gibbs Sampling

تعود معاينة جبس إلى عمليات جبس Gibbs Processes المـأخوذة مـن علم الفيزياء الاحصائية التي تمثل طريقة لبناء عمليات جديدة من العمليات القديمة ،

إذ اننا نبدأ بعملية ابتدائية $y^{(\circ)}$ ثم نعرف عملية جديدة y وذلـك بأخـذ الدالـة الاحتمالية $p(y)$ لـ y بالنسبة إلى $y^{(\circ)}$. وللتعرف على كيفية عمل طريقـة معاينة جبس افرض ان لدينا k من المتغيرات العشوائية التي نرمز لهـا بـالرمز $y=(y_1, y_2, \ldots, y_k)$ ، ونفرض لدينا التوزيعات الشرطية التامة بالشكل :

$$p(y_i/y_j)\,,\ i\neq j \quad , \quad i=1,2,\ldots,k$$

وفي هذه الخوارزمية لا تحتاج الى الشكل التام للتوزيعات الشرطية :

$$\lfloor \mathbf{y_i}/\mathbf{y_j} \quad , \quad \mathbf{i}\neq \mathbf{j} \rfloor \ , \ \mathbf{i}=\mathbf{1,2,\ldots,k}$$

لكن نحتاج الى كتابتها بشكل ثابت متناسب ، وتحت شروط معتدلة .

التوزيعات الشرطية احادية البعد تحدد بشكل وحيد

$$[y_1, y_2, \ldots, y_k]$$

ومن خلاله يمكن ايجاد التوزيعات الهامشية :

$$[y_i] \quad , \qquad i=1,2,\ldots,k$$

وفيما يأتي توضيح لاجراء الخوارزمية :

1. افرض لدينا مجموعة من القيم الابتدائية :

$$\left\lfloor y_1^{(0)}, y_2^{(0)}, \ldots, y_k^{(0)} \right\rfloor$$

2. اسحب $y_1^{(1)}$ من توزيع :

$$\left\lfloor y_1 / y_2^{(0)}, \ldots, y_k^{(0)} \right\rfloor$$

3. اسحب $y_2^{(1)}$ من توزيع :

$$\left\lfloor y_2 / y_1^{(1)}, y_3^{(0)}, \ldots, y_k^{(0)} \right\rfloor$$

4. اسحب $y_k^{(1)}$ من توزيع :

$$\left\lfloor y_k / y_1^{(1)}, y_2^{(1)}, \ldots, y_{k-1}^{(1)} \right\rfloor$$

وهذا يكمل تكراراً واحداً لمعاينة جبس وعليه بعد تكرار واحد نكـون قـد حصلنا على :

$$\left[y_1^{(1)}, y_2^{(1)}, \ldots, y_k^{(1)}\right]$$

وبعد t من هذه التكرارات نكون قد حصلنا على :

$$\left[y_1^{(t)}, y_2^{(t)}, \ldots, y_k^{(t)}\right]$$

وبتكرار هذه العملية **m** من المرات فان معاينة جـبس يمكـن ان تولـد مشاهدات مستقلة عددها **m** للمتغير مثلاً y_1 الذي يعد عينة عشوائية من التوزيع $p(y_1)$ بشرط ان تكون التكرارات مستقلة وباستخدام قيم اولية (ابتدائية)

$$\left[y_1^{(0)}, y_2^{(0)}, \ldots, y_k^{(0)}\right]$$

يمكن الحصول على الخصائص الاحصائية للمتغيـر y_1 مـثلا الوسـط Mean ، التباين Variance أو أية خواص احصائية للمتغير y_1 .

أسماعيل (2005)

11- معاينة جبس لنموذج بيز الخطي

Gibbs Sampling for Bayesian Linear Model :

افرض لدينا نموذج بيز الخطي المذكور في المعادلتين (1) و (2) .

$$P\left(y/\mu, \sigma_a^2, \sigma^2\right) = \frac{1}{\sqrt{2\pi\left(\sigma_a^2 + \sigma^2\right)}} \exp\left[\frac{-(y-\mu)^2}{2\left(\sigma_a^2 + \sigma^2\right)}\right] \qquad (10)$$

وان دالة الامكان الاعظم تحسب كالآتي :

$$L\left(y/\mu, \sigma_a^2, \sigma^2\right) = \prod_{i=1}^{a}\prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\left(\sigma_a^2 + \sigma^2\right)}} \exp\left[\frac{-\left(y_{ij}-\mu\right)^2}{2\left(\sigma_a^2 + \sigma^2\right)}\right] \qquad (11)$$

$$L\left(y/\mu,\sigma_a^2,\sigma^2\right)=\left(\frac{1}{\sqrt{2\pi\left(\sigma_a^2+\sigma^2\right)}}\right)^{an}\exp\left[\frac{-\sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij}-\mu\right)^2}{2\left(\sigma_a^2+\sigma^2\right)}\right]$$

إذ إن :

$$\sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij}-\mu\right)^2=\sum_{i=1}^{a}\sum_{j=1}^{n}\left[y_{ij}-\bar{y}_{i\bullet}+\bar{y}_{i\bullet}-\bar{y}+\bar{y}-\mu\right]^2$$

$$=\sum_{i=1}^{a}\sum_{j=1}^{n}\left[\left(y_{ij}-\bar{y}_{i\bullet}\right)+\left(\bar{y}_{i\bullet}-\bar{y}\right)+\left(\bar{y}-\mu\right)\right]^2$$

$$\begin{aligned}=\sum_{i=1}^{a}\sum_{j=1}^{n}\Big[&\left(y_{ij}-\bar{y}_{i\bullet}\right)^2+\left(\bar{y}_{i\bullet}-\bar{y}\right)^2+\left(\bar{y}-\mu\right)^2+\\ &2\left(y_{ij}-\bar{y}_{i\bullet}\right)\left(\bar{y}_{i\bullet}-\bar{y}\right)+\left(y_{ij}-\bar{y}_{i\bullet}\right)\left(\bar{y}-\mu\right)+\\ &\left(\bar{y}_{i\bullet}-\bar{y}\right)\left(\bar{y}-\mu\right)\Big]\end{aligned}$$

$$\therefore L\left(y/\mu,\sigma_a^2,\sigma^2\right)\propto\left(\sigma_a^2+\sigma^2\right)^{-\frac{an}{2}}\exp\left[\frac{-\left(SS_t+SS_E+na\left(\bar{y}-\mu\right)^2\right)}{2\left(\sigma_a^2+\sigma^2\right)}\right]\quad(12)$$

إذ إن : $y=\left(y_{11},y_{22},\ldots,y_{an}\right)$

وان SS_t تمثل مجموع المربعات لبين المعاملات و SS_E يمثل مجموع مربعات الخطأ .

وللبدء في تحليل بيز نفرض التوزيعات الاولية للمعلمات μ , σ_a^2 , σ^2 وللسهولة نختار التوزيعات الاولية المتآلفة Conjugate prior أي ان التوزيعين

الاولي واللاحق ينتميان إلى عائلة التوزيع الاحتمالي نفسها لكن بمعلمـات مختلفـة ونفرض ان المعلمات مستقلة ولها التوزيعات الاولية الآتية :

$$p(\mu) \sim N\left(A_m, \sigma_m^2\right) = \left(2\pi\sigma_m^2\right)^{-\frac{1}{2}} \exp\left[\frac{-(\mu - A_m)^2}{2\sigma_m^2}\right] \qquad (13)$$

$$p(\sigma_a^2) \sim IG(\alpha_1, \beta_1) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)}\left(\sigma_a^2\right)^{-\alpha_1+1} \exp\left[\frac{-\beta_1}{\sigma_a^2}\right] \qquad (14)$$

$$p\left(\sigma^2\right) \sim IG(\alpha_2, \beta_2) = \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)}\left(\sigma^2\right)^{-\alpha_2+1} \exp\left[\frac{-\beta_2}{\sigma^2}\right] \qquad (15)$$

إذ إن IG تمثل توزيع كاما المعكوس انظر الفقرة (7) .

12- التوزيع اللاحق Posterior Distribution

في ضوء مبرهنة بيز التوزيع اللاحق يتناسب مع حاصـل ضـرب دالـة الامكان الاعظم والتوزيعات الاولية أي ان

$$p\left(\mu, \sigma_a^2, \sigma^2 / y\right) \propto L\left(y / \mu, \sigma_a^2, \sigma^2\right) p(\mu)\, p(\sigma_a^2)\, p(\sigma^2) \qquad (16)$$

13- التوزيعات الشرطية التامة

Complete Conditional Distributions :

لتطبيق معاينة جبس يجب ان تكون لدينا التوزيعات الشرطية الكاملة لكـل معلمـة مـن المعلمـات $\mu, \sigma_a^2, \sigma^2$ وهـذه التوزيعـات هـي $p\left(\mu / \sigma_a^2, \sigma^2, y\right)$ ، $p\left(\sigma_a^2 / \mu, \sigma^2, y\right)$ ، $p\left(\sigma^2 / \mu, \sigma_a^2, y\right)$ وعادة يشار الـى هذه التوزيعات بشكل مختصر بالشكل $p(\mu/R)$ ، $p(\sigma_a^2/R)$ ، $p(\sigma^2/R)$ إذ إن R تعني بقية المتغيرات التي تكون معلومة عند أخذ التوزيع الشرطي .

$$p(\mu / R) \propto L\left(y / \mu, \sigma_a^2, \sigma^2\right) p(\mu)$$

$$p(\mu/R)\propto\left(\sigma_a^2+\sigma^2\right)^{-\frac{na}{2}}exp\left[\frac{-\left(SS_t+SS_E+na(\bar{y}-\mu)^2\right)}{2\left(\sigma_a^2+\sigma^2\right)}\right]\times$$
$$\left(2\pi\sigma_m^2\right)^{-\frac{1}{2}}exp\left[\frac{-(\mu-A_m)^2}{2\sigma_m^2}\right]$$

$$p(\mu/R)\propto exp\left[\frac{-SS_t}{2\left(\sigma_a^2+\sigma^2\right)}-\frac{SS_E}{2\left(\sigma_a^2+\sigma^2\right)}-\frac{na\bar{y}^2}{2\left(\sigma_a^2+\sigma^2\right)}-\right.$$
$$\left.\frac{2na\bar{y}\mu}{2\left(\sigma_a^2+\sigma^2\right)}-\frac{na\mu^2}{2\left(\sigma_a^2+\sigma^2\right)}-\frac{\mu^2}{2\sigma_m^2}+\frac{2\mu A_m}{2\sigma_m^2}-\frac{A_m^2}{2\sigma_m^2}\right]$$

$$\therefore\ p(\mu/R)\propto exp\left[\mu^2\left(\frac{-na}{2\left(\sigma_a^2+\sigma^2\right)}-\frac{1}{2\sigma_m^2}\right)+\mu\left(\frac{na\bar{y}}{\left(\sigma_a^2+\sigma^2\right)}+\frac{A_m}{\sigma_m^2}\right)\right]$$

وحسب مبرهنة أي متغير له شكل ثنائي في الاس هو توزيع طبيعي . انظر فتحي (2006)

$$p(\mu/R)=N\left[\frac{-\left(\frac{na\bar{y}}{\sigma_a^2+\sigma^2}+\frac{A_m}{\sigma_m^2}\right)}{-2\left(\frac{na}{2\left(\sigma_a^2+\sigma^2\right)}+\frac{1}{2\sigma_m^2}\right)},\frac{-1}{-2\left(\frac{na}{2\left(\sigma_a^2+\sigma^2\right)}+\frac{1}{2\sigma_m^2}\right)}\right]$$

$$p(\mu/R)\sim N\left[\frac{\frac{na\bar{y}}{\sigma_a^2+\sigma^2}+\frac{A_m}{\sigma_m^2}}{\frac{na}{\left(\sigma_a^2+\sigma^2\right)}+\frac{1}{\sigma_m^2}},\frac{1}{\frac{na}{\left(\sigma_a^2+\sigma^2\right)}+\frac{1}{\sigma_m^2}}\right]\qquad(17)$$

ولايجاد التوزيع الشرطي الكامل لـ σ_a^2 نعد كلا من μ,σ^2 ثوابت

$$p(\sigma_a^2/R)\propto\left(\sigma_a^2+\sigma^2\right)^{-\frac{na}{2}}exp\left[\frac{-\left(SS_t+SS_E+na(\bar{y}-\mu)^2\right)}{2\left(\sigma_a^2+\sigma^2\right)}\right]\times\frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)}\left(\sigma_a^2\right)^{-\alpha_1+1}exp\left[\frac{-\beta_1}{\sigma_a^2}\right]$$

نفرض ان $RSS_1 = \left(SS_t + SS_E + na(\bar{y}-\mu)^2\right)$ إذ إن RSS تمثل مجموع مربعات البواقي

$$p\left(\sigma_a^2/R\right) \propto \left(\sigma_a^2\right)^{-\frac{na}{2}} exp\left[\frac{-RSS_1}{2\sigma_a^2}\right]\left(\sigma_a^2\right)^{-\alpha_1+1} exp\left[\frac{-\beta_1}{\sigma_a^2}\right]$$

$$p\left(\sigma_a^2/R\right) \sim IG\left[\alpha_1 + \frac{na}{2}\,,\ \beta_1 + \frac{RSS_1}{2}\right] \qquad (18)$$

وبالطريقة نفسها نجد التوزيع الشرطي الكامل لـ σ^2

$$p\left(\sigma^2/R\right) \propto \left(\sigma_a^2+\sigma^2\right)^{-\frac{na}{2}} exp\left[\frac{-\left(SS_t + SS_E + na(\bar{y}-\mu)^2\right)}{2\left(\sigma_a^2+\sigma^2\right)}\right] \times \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)}\left(\sigma^2\right)^{-\alpha_2+1} exp\left[\frac{-\beta_2}{\sigma^2}\right]$$

نفرض ان $RSS_2 = \left(SS_t + SS_E + na(\bar{y}-\mu)^2\right)$

$$p\left(\sigma^2/R\right) \propto \left(\sigma^2\right)^{-\frac{na}{2}} exp\left[\frac{-RSS_2}{2\sigma^2}\right]\left(\sigma^2\right)^{-\alpha_2+1} exp\left[\frac{-\beta_2}{\sigma^2}\right]$$

$$p\left(\sigma^2/R\right) \propto IG\left[\alpha_2 + \frac{na}{2}\,,\ \beta_2 + \frac{RSS_2}{2}\right] \qquad (19)$$

ومن الجدير بالذكر أن البرنـامج الجـاهز WinBUGS يقـوم بايجـاد التوزيعات الشرطية التامة كافة التي حصلنا عليها من الصيغ (17) ، (18) ، (19) إذ إن المطلوب هو افتراض التوزيع الاولي فقط للمتغيرات $\sigma^2, \sigma_a^2, \mu, y_{ij}$.

14- وصف البيانات :

تم تطبيق النتائج النظرية في هذا البحث على بيانـات حقيقيـة مصـدرها تجارب نفذت من قبل البرنامج الوطني لتطوير زراعة القطن فـي العـراق لسـنة (2002) ، وتمثل البيانات متوسط وزن جوزة القطن مقاس بوحدة الغرام لـ 15 صنفاً من اصناف نبات القطن .

جدول (1) :بيانات متوسط وزن جوزة القطن بالغرام

الصنف	المكرر الاول	المكرر الثاني	المكرر الثالث	المكرر الرابع	المكرر الخامس
نازلي 5.3	4.68	5.03	4.89	4.33	4.99
ربيع Verd	4.45	4.55	4.55	4.25	4.60
نازلي 87	4.27	4.54	4.23	4.20	4.34
نامنعان	4.69	4.73	4.47	4.49	4.80
تامكرت انفكس	3.77	3.89	5.89	3.57	3.97
طاقة 1	3.95	4.06	4.96	3.92	3.98
طاقة 2	3.49	3.81	4.02	3.40	3.58
فيرقانا 2	4.77	4.48	4.27	4.72	4.81
كريمة 11	3.88	4.49	4.76	3.90	3.82
مرسومي 1	4.43	5.19	5.87	4.42	5.13
كوندور	4.85	4.98	4.43	4.80	4.90
1k347	4.58	4.09	4.49	4.50	4.66
1k359	5.28	4.49	4.82	5.20	4.82
1k327	4.48	4.25	4.03	4.42	4.54
كوكر 310	4.19	5.82	4.12	4.10	4.28

15- تقدير المعلمات بطريقة تحليل التباين :

جدول (2): نتائج تقدير المعلمات μ ، σ_a^2 ، σ^2 بواسطة تحليل التباين

$\hat{\mu}$	$\hat{\sigma}_a^2$	$\hat{\sigma}^2$
4.178	0.09321	0.0887

16- تقدير بيز للمعلمات بطريقة معاينة جبس :

نفرض ان التوزيعات الاولية للمعلمات $\mu, \sigma_a^2, \sigma^2$ كالاتي:

$$\mu \sim N\left(A_m, \sigma_m^2\right)$$

$$p(\mu) \sim N\left(A_m, \sigma_m^2\right) = \left(2\pi\sigma_m^2\right)^{-\frac{1}{2}} \exp\left[\frac{-(\mu - A_m)^2}{2\sigma_m^2}\right]$$

$$\sigma_a^2 \sim IG(\alpha_1, \beta_1)$$

$$p\left(\sigma_a^2\right) \sim IG(\alpha_1, \beta_1) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)}\left(\sigma_a^2\right)^{-\alpha_1+1} \exp\left[\frac{-\beta_1}{\sigma_a^2}\right]$$

$$\sigma^2 \sim IG(\alpha_2, \beta_2)$$

$$p\left(\sigma^2\right) \sim IG(\alpha_2, \beta_2) = \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)}\left(\sigma^2\right)^{-\alpha_2+1} \exp\left[\frac{-\beta_2}{\sigma^2}\right]$$

ولعدم توافر قـيم ابتدائيـة للمعلمـات نعـد القـيم الابتدائيـة للمعلمـات $\sigma^2, \sigma_a^2, \mu$ هي التقديرات التي حصلنا عليها بطريقة تحليـل التبـاين أي أن :

$$\mu_\circ = \mathbf{4.178}$$

$$\sigma_\circ^2 = \mathbf{0.0887}$$

$$\sigma_{a\circ}^2 = \mathbf{0.09321}$$

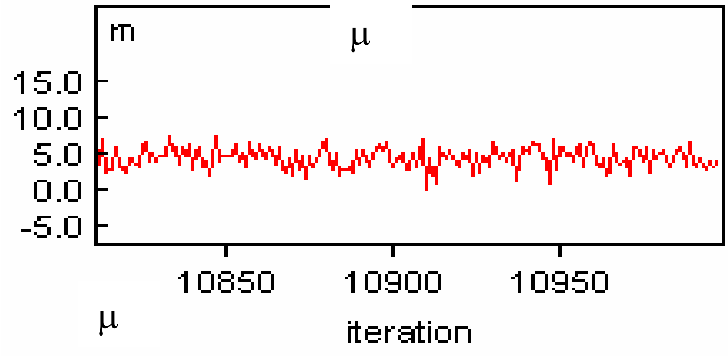
وللحصول على تقـديرات هـذه المعلمـات نسـتخدم البرنـامج الجـاهز WinBugs، انظر التفاصيل فتحي،(2006) الذي يستخدم التوزيعـات الشـرطية التامة التي تم الحصول عليها وهـي المعـادلات (17) ,(18) ,(19) والتقـديرات موضحة في الجدول (3) .

وتمت اعادة تكرار البرنامج 10000 مرة الى حين الوصول إلى تقدير جيد للمعلمات ، وكما هو موضح في النماذج (1) ,(2) ,(3) إذ يمكن اسـتنتاج جميـع خواص المعلمة μ الاحصائية (طبعا بعد عدها متغيراً عشـوائياً) وهـي الوسـط والتباين والانحراف المعياري والوسيط وغيرها من الخواص الاحصائية للمعلمـة μ ، كما ان الرسم البياني الموضح في النموذج (1) يبين التوزيع الطبيعي للمعلمة μ .

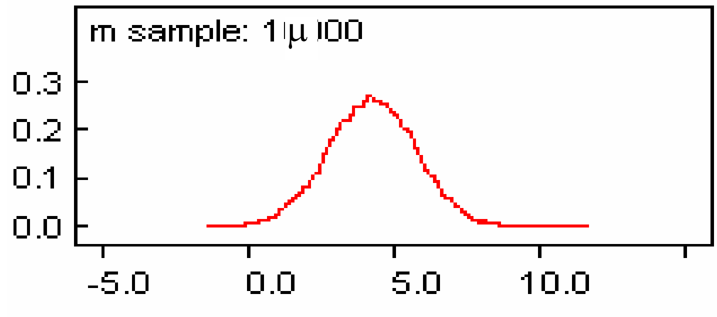
كذلك بالنسبة إلى المعلمة σ_a^2 فالتقدير موضح في النموذج (2) مع الرسم البياني الذي يظهر توزيع معكوس كاما مع جميع الخواص الاحصائية للمعلمة σ_a^2 (بعد عدها متغيراً عشوائياً) كالوسط والتبـاين والانحـراف المعيـاري والوسـيط وغيرها من الخواص الاحصائية .

واخيرا بالنسبة للمعلمة σ^2 فالتقدير موضح بـالنموذج (3) مـع الرسـم البياني الذي يظهر توزيع معكوس كاما مع جميع الخواص الاحصائية للمعلمة σ^2 (بعد عدها متغيراً عشوائياً) كالوسط والتبـاين والانحـراف المعيـاري والوسـيط وغيرها من الخواص الاحصائية .

الشكل (A)



الشكل (B)



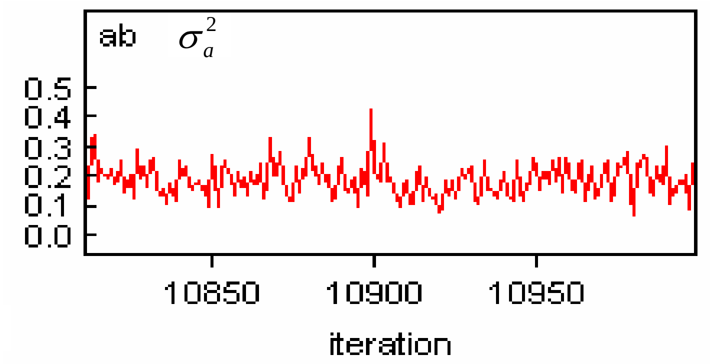
node	mean μ	sd	MC error	2.5%	median	97.5%	start	sample
m	4.223	1.507	0.01477	1.282	4.222	7.165	1001	10000

النموذج (1) التوزيع الطبيعي للمعلمة μ

1- الشكل (A) يمثل رسماً لقيمة المعلمة غير المعلومة μ مع عدد التكرارات وان المخطط أعلاه حركي .

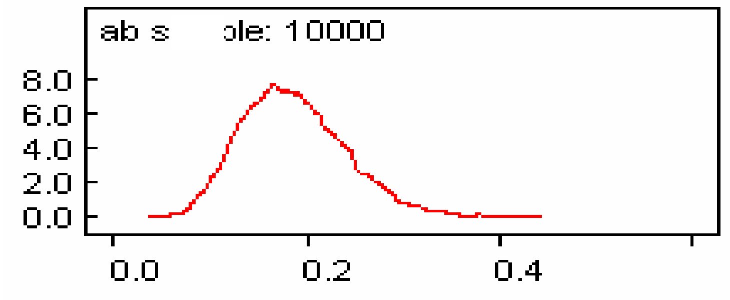
2- الشكل (B) يمثل رسم مقدر نواة (Kernel) الكثافة الهامشية للمعلـمة μ اذ يلاحظ من الرسم ان هذا التوزيع اللاحق مقترب من التوزيـع الطبيعـي ، ان الكمية المؤشرة في عمود MC error تعطي تقدير لـ $\frac{\sigma}{N^{\frac{1}{2}}}$ الـذي يسـمى بالخطأ المعياري لمونتي كارلو .

الشكل (E)



σ_a^2

الشكل (F)



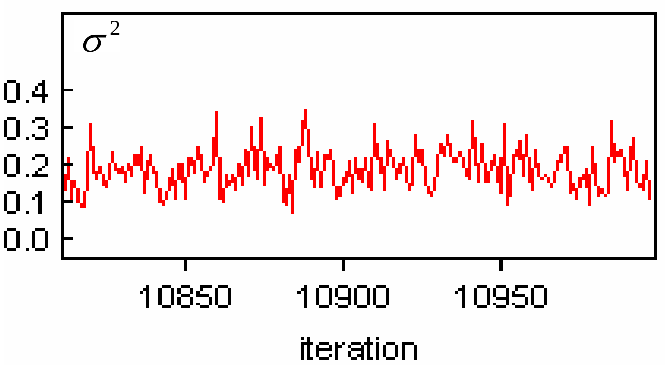
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
ab	0.1868	0.05398	5.367E-4	0.09529	0.1821	0.3046	1001	10000

النموذج (2) توزيع معكوس كاما للمعلمة σ_a^2

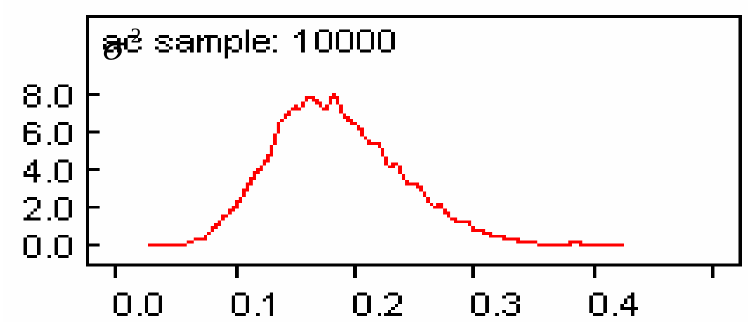
1- الشكل (E) يمثل رسماً لقيمة المعلمة غير المعلومة σ_a^2 مع عدد التكرارات وان المخطط أعلاه حركي .

2- الشكل (F) يمثل رسم مقدر نواة (Kernel) الكثافة الهامشية للمعلمة σ_a^2 اذ يلاحظ من الرسم ان هذا التوزيع اللاحق مقترب من توزيع معكوس كاما ، ان الكمية المؤشرة في عمود MC error تعطي تقديراً لـ $\frac{\sigma}{N^{\frac{1}{2}}}$ والـذي يسمى بالخطأ المعياري لمونتي كارلو .

الشكل (I)



الشكل (J)



node	mean	sd	MC error	2.5%	median	97.5%	start	sample
ac	0.1853	0.05357	6.181E-4	0.09326	0.1804	0.3009	1001	10000

النموذج (3) توزيع معكوس كاما للمعلمة σ^2

1- الشكل (I) يمثل رسماً لقيمة المعلمة غير المعلومة σ^2 مع عدد التكرارات وان المخطط أعلاه حركي .

2- الشكل (J) يمثل رسم مقدر نواة (Kernel) الكثافة الهامشية للمعلمة σ^2 اذ يلاحظ من الرسم ان هذا التوزيع اللاحق مقترب الى توزيع معكوس كاما ، ان الكمية المؤشرة في عمود MC error تعطي تقديراً لـ $\frac{\sigma}{N^{\frac{1}{2}}}$ والذي يسمى بالخطأ المعياري لمونتي كارلو .

ومن ملاحظة نتائج معاينة جبس من البرنامج الجاهز WinBUGS نحصل على التقديرات الآتية للمعلمات :

$$\hat{\mu}_{GS} = 4.223$$
$$\hat{\sigma}^2_{aGS} = 0.1868$$
$$\hat{\sigma}^2_{GS} = 0.1853$$

حيث انGS تشير الى معاينة جبس

الجدول (3):نتائج تقدير المعلمات μ ، σ_a^2 ، σ^2 بواسطة معاينة جبس

$\hat{\mu}_{GS}$	$\hat{\sigma}^2_{aGS}$	$\hat{\sigma}^2_{GS}$
4.223	0.1868	0.1853

ومن مقارنة هذه النتائج مع النتائج التي حصلنا عليها بطريقة تحليل التباين في الجدول (2) . نلاحظ ان هذه التقديرات قريبة من بعضها وخاصة بالنسبة إلـى تقدير μ , σ^2 ، أما بالنسبة إلى σ_a^2 فهناك اختلاف طفيف في التقدير وهنا نعتقد بأنه يعزى إلى القيمة الابتدائية $\sigma^2_{a\circ}$ إذ يمكن الوصول إلى قيمة σ^2_{aGS} أقرب إلى $\hat{\sigma}_a^2$ إذا توفرت قيم ابتدائية جيدة .

لكن لو اردنا اختبار الفرضية

$$H_\circ : \ \sigma_a^2 = 0$$
$$H_1 : \ \sigma_a^2 \neq 0$$

على اساس التقديرات $\sigma^2_{aGS} = 0.1868$ و $\sigma^2_{GS} = 0.1853$ فنلاحظ ان :

$$MSS_E = \hat{\sigma}^2 = \sigma^2_{GS} = 0.1853$$
$$MSS_t = MSS_E + n\hat{\sigma}^2 = MSS_E + n\sigma^2_{aGS}$$
$$= 1.853 + 5(0.1868)$$
$$= 2.787$$

$$\therefore \mathbf{F} = \frac{\mathbf{MSS_t}}{\mathbf{MSS_E}} = \frac{\mathbf{2.787}}{\mathbf{0.1853}} = \mathbf{15.040474}^{**}$$

وفي ضوء نتيجة F المحسوبة ومقارنتها مع القيمة الجدولية فان الفرضية سترفض وهذه النتيجة نفسها التي حصلنا عليها بواسطة طريقة تحليل التباين .

17- الاستنتاجات والتوصيات

1. مقدرة طريقة معاينة جبس بالحصول على تقديرات جيدة مقارنة مـع الطرائـق الكلاسيكية مثل تحليل التباين .
2. تم الحصول على نتائج التقدير بعد افتراض التوزيعات الاولية للمعلمـات مـن العائلة المتآلفة وذلك لعدم توافر معلومات اولية او سابقة عن موضوع الدراسة.
3. عندما تكون القيم الابتدائية للمعلمات قريبة من القيم الابتدائية الواقعيـة تكـون النتائج أفضل وهذا ما تم الأخذ به في هذا البحث .
4. تم رفض الفرضية الصفرية حول تساوي التباين بالطريقتين طريقة تحليل التباين وطريقة معاينة جبس .
5. امكانية استخدام طريقة معاينة جبس على نماذج احصائية اخرى مثل التوزيـع العشوائي ثنائي التقسيم مع وجود التفاعل ونماذج عاملية اخرى .
6. إمكانية افتراض توزيعات أولية سواء كانت قليلة المعلومات أم ذات المعلومات أم أية توزيعات احتمالية معروفة أخرى .
7. يمكن تطبيق معاينة جبس على نماذج خطية أو غير خطية معرفة على تصاميم خاصة .
8. يمكن تطبيق النموذج الخطي المختلط على بيانات ناتجة من الدراسات الزراعية او الهندسية او الطبية او بيانات البيئة او في العلوم التربوية والنفسية ... الخ .

18- المصادر

1. إسماعيل ، يونس حازم ، (2005) ، "الكشف عن القيم الشـاذة باسـتخدام التوزيـع المختلط بالاعتماد على معاينة جبس واسلوب بيز" ، (رسالة ماجستير غير منشورة) ، كلية التربية ، جامعة الموصل .
2. الطائي ، وائل عبد اللطيف جاسم ، (2004) ، "معايرة بيز" ، (رسالة ماجستير غير منشورة) ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل .
3. عباس ، سمير عبد الجبار ، (2000) ، "تقدير بيز لمكونات التباين لنمـوذج التـأثير العشوائي الاحادي التقسيم" ، (رسالة ماجستير غير منشورة) ، كلية التربية ، جامعـة الموصل ، العراق .
4. فتحي ، ايمان طارق ، (2006) ، "تقدير بيز في النماذج الخطية المختلطة باسـتخدام معاينة جبس" ،(رسالة ماجستير غير منشورة) ، كلية التربية ، جامعة الموصل .
5. الناصر ، عبد المجيد حمزة وظافر حسين رشيد ، (1988) ، "الاستدلال الاحصائي" ، وزارة التعليم العالي والبحث العلمي ، جامعة بغداد ، العراق .
6. وزارة الزراعة ، مركز إباء للأبحاث الزراعية ، (2002) ، التقرير السنوي للبرنامج الوطني لتطوير زراعة القطن في العراق .
7. Evans, M. and Swartz, T., (1995), “Methods for Approximating Integrals in Statistics with Special Emphasis on Bayesian Integration Problems”, Statistical Science 10: 254–272.
8. Gelfand, A.E. and Smith, A.F.M. (1990), “Sampling–Based Approaches to Calculating Marginal Densities”, J. Am. Stat. Asso. 85: 398–409.
9. Metropolis, N. and Ulam, S., (1949), “The Monte Carlo Method”, J. Amer. Statist. Assoc. 44: 335–341.
10. Montgomery, D.C., (2001), “Design and Analysis of Experiment”, 5th edition, John Wiley, New York.
11. Rao, C.R. and Kleffe, J., (1988), “Estimation of Variance Component and Application”, North Holland, Amsterdam.
12. Smith, A.F.M., (1991), “Bayesian Computational Methods”, Phil. Trans. R. Soc. Lond. A337: 369–386.
13. Tierney, L., (1994), “Markov Chains for Exploring Posterior Distributions”, (with discussion), Ann. Statist. 22: 1701–1762.
14. Walsh, B., (2002), “Markov Chain Monte Carlo and Gibbs Sampling”, Lecture Notes for EEB 596Z.