اختيار المتغيرات في التحليل العاملي البيزي مع تطبيق

ذنون يونس ذنون الشكرجي
قسم التقنيات المالية والمصرفية
المعهد التقني – نينوى

الاستلام	القبول
2009 / 12 / 05	2010 / 02 / 16

Summary

Using Bayesian inference in factor analysis is to estimate parameters of the model (Λ, F), Where Λ represents the factor loadings matrix, F represents the factor scores matrix, for that, selected variables of interest relatively through interpretation are the results of Bayesian estimator for factor loadings matrix and application on data taken from Ibn-Alatheer hospital, These data include (150) patients injured with thalassemia, and the results showed the importance all of the variables in the impact on the disease.

الملخص

استخدم أسلوب بيز في التحليل العاملي لتقدير معلمات النموذج (Λ, F) حيث إن Λ تمثل مصفوفة تحميلات العوامل (Factor Loadings Matrix) وان F تمثل مصفوفة قيم العوامل (Factor Scores Matrix) لذلك اختيرت المتغيرات ذات الأهمية نسبياً من خلال تفسير نتائج مقدر بيز لمصفوفة تحميلات العوامل وبالتطبيق على بيانات أخذت من مستشفى ابن الأثير والتي تضم (150) مريضاً مصابين بالثلاسيميا وتبين من النتائج أهمية جميع المتغيرات في تأثيرها على المرض.

المقدمة:

أسلوب بيز في الإحصاء يضع المعلومات الأولية المتوفرة التي لها علاقة بمعلمات النموذج على شكل كميات بصيغة توزيعات أولية , المعلومات الموجودة ضمن البيانات تمثل دالة الإمكان

Likelihood Function ، يتم ربط المعلومات الأولية ودالة الإمكان استناداً إلى قاعدة بيز بحيث تندمج المعلومات من كلا المصدرين مع التداخلات . إن طرق بيز الإحصائية لا تدمج فقط المعلومات الأولية المتوفرة سواءاً كانت مأخوذة من خبراء متمكنين أو من بيانات سابقة وإنما تسمح بتجميع المعلوم ات المتعلقة بقيم المعلمات كبيانات لاحقة (إضافية) تم الحصول عليها (Rowe,2001).

إن نموذج التحليل العاملي يستخدم الارتباطات و التباينات المشتركة بين مجموعة المتغيرات المشاهدة من اجل وصفهم بدلالة مجموعة اصغر مكونة من متغيرات غير مشاهدة (غير ملحوظة) unobserved variables، وتسمى المتغيرات غير المشاهدة عوامل توصف العلاقة الأساسية بين المتغيرات الأصلية.

هناك سببين رئيسيين لاستخدام التحليل العاملي . السبب الأول هو لاستيضاح العلاقة بين مجموعة المتغيرات المشاهدة بدلالة عدد اصغر من متغيرات غير المشاهدة أو عوامل كامنة تؤثر على المشاهدات كما يمكن استعمال هذا العدد الأصغر من المتغيرات لإيجاد تركيب ذو معنى للمتغيرات المشاه دة. وسيساعد هذا التركيب في ترجمة وتفسير العملية التي ولدت المشاهدات. أما السبب الثاني الذي من اجله نستخدم التحليل العاملي فهو لتقليل المشاهدات ، بما أن المتغيرات المشاهدة سيتم تمثيلها بدلالة عدد اقل من متغيرات غير المشاهدة أو متغيرات كامنة فان عدد المتغيرات في التحليل سينخفض وبوجود عدد اقل من العوامل (متجهات بأبعاد اصغر) التي علينا أن نعمل عليها لكي نحصل على حقيقة المتغيرات المشاهدة . وسيكون مطلوب فقط تخزين هذا العدد الأصغر ويمكن أيضاً استعمال هذا العدد الأصغر من العوامل لإجراء تحليلات أخرى.

(Mar and McD)1975 وضعا طريقة لمسالة التحليل العاملي مع النظر إلى حل بيز لحالات هايوود Heywood cases. وقد اعتمدا مصفوفات قطرية للتباين المشترك وقد افترضا العثور على تقديرات للنموذج المشترك اللاحق لمصفوفات تحميلات العوامل والتباين المشترك للمتغيرات و حصلا على حل عددي مبسط كما تم إيجاد تقدير نقطي لمصفوفة تحميلات العوامل . وقد توجه (Won) عام 1980 لمسالة التحليل العاملي من وجهة نظر بيز التجريبية معتمداً مستبقات طبيعية لتحميل العامل . وقد اقترح استعمال خوارزمية تعظيم التوقع (EM Algorithm) لإيجاد نمط لاحق لمصفوفة تحميلات العوامل ولكن لم يتم الحصول على خوارزمية مبسطة . اعتمد Lee عام 1981 أسلوب بيز الهرمي للتحليل العاملي التوكيدي ابتداءاً من فرضية إن المعلمات الحرة في مصفوفة تحميلات العوامل هي معلمات تبادلية وإنها موزعة بشكل طبيعي . وقد تم افتراض إن مصفوفة التباين المشترك قطرية . وقد وجدت تقديرات النموذج المشترك لمصفوفة تحميلات العوامل و تباينات التوزيعات و مصفوفة التباين المشترك . هذا الحل النمطي ضمني

وعددي. وقد تم الحصول على تقدير نقطي لمصفوفة تحميلات العوامل . و في العام 1983 كان (Eve and Ver) قد درسا نموذج تحليل بيز العاملي مع مصفوفة التباين المشترك القطرية وأعادا تعيين عدد العوامل . وقد تم وصف روتين عددي على الحاسوب لإيجاد بشكل ضمني نمط التوزيع المشترك اللاحق لتحميلات العوامل وتباين الخطأ. درسMay عام 1985 مشكلة تحليل بيز العاملي مختبراً قيم العوامل بالإضافة إلى تحميلات العامل وتباين الخط أ. وقد تم افتراض إن تحميلات العوامل طبيعية وأولية. وقد استعمل المؤلف خوارزمية تعظيم التوقع EM (Algorithm) لإيجاد التقدير النقطي للمعلمات كأنماط هامشية (حدية) للتوزيعات اللاحقة . لسوء الحظ لا توجد براهين على تقارب خوارزمية تعظيم التوقع EM)) المستخدمة. وقد استعمل Shig عام 1986 إحدى طرق بيز المشتقة اولياً لنموذج التحليل العاملي ووجد حلولاً عددية ضمنية لمصفوفة تحميلات العوامل والمتغيرات المحددة.

اقترح Aka عام 1987 إمكانية استعمال معيار AIC لاختيار العدد المناسب للعوامل التي تستعمل في نموذج بيز (Press and Shigemasu, 1997).

نموذج التحليل العاملي البيزي The Bayesian Factor Analysis Model

في البداية سنقوم بتعريف دالة الإمكان , ثم ندخل التوزيعات الأولية على المعلمات ثم نحسب التوزيع المشترك اللاحق للمعلمات وأخيراً نجد التوزيع اللاحق الحدي للمعلمات.

دالة الامكان The Likelihood Function

نموذج التحليل العاملي البيزي يمكن كتابته بالشكل التالي :(Rowe, 2000c)

$$(x_j \mid \Lambda , f_j) = \Lambda f_j + \varepsilon_j \sim N(\Lambda f_j, \Psi) \quad \ldots(1\text{-}1)$$

(px1) (pxm) (mx1) (px1)

لـ $j = 1,2,\ldots,n$ حيث x_j هي المشاهدة j^{th} Λ هي مصفوفة ثو ابت تسمى مصفوفة تحميلات العوامل. و f_j هي متجهات قيم العوامل للموضوع j و ε_j's وقد تم افتراض انها لا توجد بينها ارتباطات تبادلية وهي موزعة طبيعياً $N(0,\varphi)$.

في النموذج التقليدي تعتبر φ مصفوفة قطرية بحيث يمكن بسهولة تمييز عوام ل عامة وخاصة . في النموذج الحالي تعتبر φ مصفوفة تباين مشترك موجبة التعريف ومتماثلة ذات خاصية أن تكون قطرية عند المتوسط . أي أن $E(\varphi)$ هي مصفوفة قطرية ونفترض أن Λ و f_j's و φ غير مشاهدة ويمكن كتابة توزيع لـ x_j كما يلي:

$$p(x_j \mid \Lambda, f_j, \Psi) = (2\pi)^{-\frac{p}{2}} |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2}(x_j-\Lambda f_j)'\Psi^{-1}(x_j-\Lambda f_j)} \quad \ldots(2\text{-}1)$$

إذا اشرنا لها بالتناسب (α) فان الاحتمال لـ)) Λ, f_j, Ψ هو كما يلي:

$$p(X|\Lambda,F_j,\Psi)\ \alpha\ |\Psi|^{-\frac{N}{2}}\ e^{-\frac{1}{2}(X-F\Lambda')'\Psi^{-1}(X-F\Lambda')} \qquad \ldots(3\text{-}1)$$

حيث أن متجهات مشاهدة المتغير P على N من المواضيع هي $X'=(x_1,\ldots,x_n)$ وقيم العوامل هي $F'=(f_1,\ldots,f_n)$ والأخطاء في المشاهدة هي $E'=(\varepsilon_1,\ldots,\varepsilon_n)$ ويشير الرمز P(.) بشكل عام إلى توزيع يمكن تمييزه ببرهانه . يعتمد ثابت التناسب في (1−3) فقط على (P,N) وليس على (Λ,F,Ψ)

التوزيع الأولي The Prior Distribution

سنستخدم العلاقات المتآلفة الطبيعية المعممة من توزيعات أولية لـ Λ,F,Ψ ((. ونأخذ توزيع أولي لغير المشاهدة (للتعبير عن حالة اللاتاكد). (Press and Shigemasu, 1997)

التوزيع المشترك الأولي لمعلمات النموذج يمكن كتابته بالشكل التالي:−

$$p(\ \Lambda,F,\Psi\ )\ \alpha\ p(\Lambda|\Psi)\,p(\Psi)\,p(F) \qquad \ldots(4\text{-}1)$$

حيث

$$p(\Lambda|\Psi)\ \alpha\ |\Psi|^{-\frac{m}{2}}\ e^{-\frac{1}{2}tr\,\Psi^{-1}\ (\Lambda-\Lambda_0)H(\Lambda-\Lambda_0)'} \qquad \ldots(5\text{-}1)$$

$$p(\Psi\ )\ \alpha\ |\Psi|^{-\frac{v}{2}}\ e^{-\frac{1}{2}tr\,\Psi^{-1}B} \qquad \ldots(6\text{-}1)$$

و B هي مصفوفة قطرية و H>0 و Ψ^{-1} تتبع توزيع Wishart و (v,B) معلمات يجب تقييمها . التوزيع الشرطي لـ Λ على Ψ يمتلك عناصر تتوزع توزيع طبيعي مشترك و (Λ_0 , H) معلمات يجب تقييمها. وفوق ذلك سنأخذ في الغالب $H=n_0I$ لبعض الكميات العددية الأولية التحديد n_0. هذه التفسيرات للمعلمات ستبسط التقييم.

التوزيع اللاحق المشترك The Joint Posterior Distribution

بربط المعادلات (1−3) − (1−6) فان التوزيع اللاحق المشترك للمعلمات سيكون كمايلي:

$$p(\Lambda,F,\Psi|X)\ \alpha\ p(F)\ |\Psi|^{-\frac{(N+m+v)}{2}}\ e^{-\frac{1}{2}tr\,\Psi^{-1}G} \qquad \ldots(7\text{-}1)$$

حيث

$$G=\ (X-F\Lambda')'(X-F\Lambda')+(\Lambda-\Lambda_0)H(\Lambda-\Lambda_0)'+B$$

التوزيع اللاحق الحدي Distribution Marginal Posterior

بتكامل العلاقة (1-7) بالنسبة لـ Ψ و استعمال خصائص توزيع ويشارت المعكوس تعطينا التوزيع اللاحق الحدي لـ (Λ , F).

$$p(\Lambda,F|X)\ \alpha\ p(F)\,|G|^{-\frac{(N+m+v-p-1)}{2}} \qquad \ldots(8\text{-}1)$$

ثم بعد ذلك نعمل تكامل لـ (1−8) بالنسبة لـ Λ لنحصل على التوزيع الحدي لـ F. نحقق ذلك بجعل G عامل بالصيغة التي تجعله بدلالة Λ. وهكذا بإكمال المربع Λ في الدالة G المحددة في (1−7) و (1−8) وباستخدام بعض العمليات الجبرية يمكن إعادة كتابتها كما يلي:−

$$\begin{aligned}
G &= (X-F\Lambda')'(X-F\Lambda')+(\Lambda-\Lambda_0)H\,(\Lambda-\Lambda_0)'+B\\
&=(X'-\Lambda F')(X-F\Lambda')+(\Lambda H-\Lambda_0 H)\,(\Lambda'-\Lambda_0')+B\\
&=X'(X-F\Lambda')-\Lambda F'(X-F\Lambda')+\Lambda H(\Lambda'-\Lambda_0')-\Lambda_0 H(\Lambda'-\Lambda_0')+B\\
&= X'X-X'F\Lambda'-\Lambda F'X+\Lambda F'F\Lambda'+\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B\\
&= X'X+\Lambda_0 H\Lambda_0'+B+\Lambda H\Lambda'+\Lambda F'F\Lambda'-\Lambda F'X-\Lambda H\Lambda_0'-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+\Lambda[(H+F'F)\Lambda'-F'X-H\Lambda_0']-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+\Lambda Q_F[\Lambda'-Q_F^{-1}(F'X+H\Lambda_0')]-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+\Lambda Q_F[\Lambda-(X'F+\Lambda_0 H)Q_F^{-1}]'-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+\Lambda Q_F(\Lambda-\Lambda_F)'-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'+\Lambda_F Q_F(\Lambda-\Lambda_F)'-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'+(X'F+\Lambda_0 H)Q_F^{-1}Q_F\Lambda'-\Lambda_F Q_F\Lambda_F'-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'+X'F\Lambda'-\Lambda_0 H\Lambda'-\Lambda_F Q_F\Lambda_F'-X'F\Lambda'-\Lambda_0 H\Lambda'\\
&=W+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'-\Lambda_F Q_F\Lambda_F'\\
&=W+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'-(X'F+\Lambda_0 H)Q_F^{-1}Q_F Q_F^{-1}(F'X+H\Lambda_0')\\
&=W+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'-(X'F+\Lambda_0 H)Q_F^{-1}(F'X+H\Lambda_0')\\
&=W-(X'F+\Lambda_0 H)Q_F^{-1}(F'X+H\Lambda_0')+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'\\
&=R_F+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'
\end{aligned}$$

وبالتعويض عن قيمة G في المعادلة (1-8) يمكن إعادة كتابتها كما يلي:−

$$p(\Lambda,F|X)\ \alpha\ \frac{p(F)}{|R_F+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'|^{\frac{\gamma}{2}}} \qquad \ldots(9\text{-}1)$$

حيث

$$Q_F=H+F'F$$

$$R_F=X'X+B+\Lambda_0 H\Lambda_0'-(X'F+\Lambda_0 H)Q_F^{-1}(X'F+\Lambda_0 H)'$$

$$\Lambda_F=(X'F+\Lambda_0)(H+F'F)^{-1}$$

$$\gamma=N+m+v-p-1$$

يمكن تكامل (1−9) بسهولة بالنسبة لـ Λ (باستعمال ثابت يجعلها طبيعية لمصفوفة توزيع T) لتعطينا التوزيع اللاحق الحدي لـ F.

$$p(F|X)\ \alpha\ \frac{p(F)}{|R_F|^{\frac{\gamma-m}{2}}|Q_F|^{\frac{p}{2}}} \qquad \ldots(10\text{-}1)$$

بتطبيق بعض العمليات الحسابية يمكن إعادة كتابة التوزيع اللاحق الحدي لـ F في (1−10) بالصيغة التالية:

$$p(F|X)\alpha\frac{p(F)\left|H+F'F\right|^{\frac{\gamma-m-p}{2}}}{\left|A+(F-\hat{F})'(I_N-XW^{-1}X')(F-\hat{F})\right|^{\frac{\gamma-m}{2}}} \qquad \ldots(11\text{-}1)$$

حيث

$$\hat{F}=(I_N-XW^{-1}X')^{-1}XW^{-1}\Lambda_0 H$$
$$=(I_N-X(X'X-W)^{-1}X')XW^{-1}\Lambda_0 H \quad \ldots(12\text{-}1)$$
$$W=X'X+B+\Lambda_0 H\Lambda_0'$$
$$A=H-(\Lambda_0 H)'W^{-1}\Lambda_0 H-(XW^{-1}\Lambda_0 H)(I_N-X(X'X-W)^{-1}X')(XW^{-1}\Lambda_0 H)$$

لذلك فان المعادلة (1−12) تمثل مصفوفة قيم العوامل.

تقدير مصفوفة تحميلات العوامل

نعود الآن إلى التوزيع اللاحق لـ (Λ,F) والمعطى في (1−8) للحصول على التوزيع اللاحق الحدي لـ Λ. ثم يمكن استخدام نوع من مقياس المكان للتوزيع كمقدر نقطي لـ Λ. لسوء الحظ بينما يمكن عمل التكامل فان التوزيع الحدي الناتج معقد جداً، ولا يبدو انه من الممكن الحصول على وسط أو المنوال لأي من التوزيعات الأولية الحقيقية لـ F.

$$p(\Lambda,F|X)\,\alpha\,p(F)\,|G|^{-\frac{(N+m+v-p-1)}{2}}$$

$$p(\Lambda,F|X)\,\alpha\,\frac{p(F)}{|G|^{\frac{\gamma}{2}}} \qquad \ldots(13-1)$$

وبتطبيق بعض العمليات الجبرية على G ينتج ما يلي:−

$$G=(X-F\Lambda')'(X-F\Lambda')+(\Lambda-\Lambda_0)H(\Lambda-\Lambda_0)'+B$$
$$=(X'-\Lambda F')(X-F\Lambda')+(\Lambda H-\Lambda_0 H)(\Lambda'-\Lambda_0')+B$$
$$=X'(X-F\Lambda')-\Lambda F'(X-F\Lambda')+\Lambda H(\Lambda'-\Lambda_0')-\Lambda_0 H(\Lambda'-\Lambda_0')+B$$
$$=X'X-X'F\Lambda'-\Lambda F'X+\Lambda F'F\Lambda'+\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda')'(F\Lambda')-(F\Lambda')'X-X(F\Lambda')+X'X+\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda')'[F\Lambda'-X]-X'(F\Lambda')+X'X+\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda'-X)'(F\Lambda'-X)+X'[(F\Lambda')-X]-X'(F\Lambda')+X'X+$$
$$\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda'-X)'(F\Lambda'-X)+X'F\Lambda'-X'X-X'(F\Lambda')+X'X+$$
$$\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda'-X)'(F\Lambda'-X)+\Lambda H\Lambda'-\Lambda H\Lambda_0'-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda'-X)'(F\Lambda'-X)+\Lambda H(\Lambda'-\Lambda_0')-\Lambda_0 H\Lambda'+\Lambda_0 H\Lambda_0'+B$$
$$=(F\Lambda'-X)'(F\Lambda'-X)+(\Lambda-\Lambda_0)H(\Lambda-\Lambda_0)'+B$$
$$=(F\Lambda'-X)'(F\Lambda'-X)+P_\Lambda$$

وبالتعويض عن قيمة G بما يساويها في المعادلة (1-13).

$$p(\Lambda,F|X)\ \alpha \frac{p(F)}{|P_\Lambda+(F\Lambda'-X)'(F\Lambda'-X)|^{\frac{\gamma}{2}}} \qquad ...(14-1)$$

نستطيع إعادة كتابة المعادلة (1−14) بما يشابه دالة توزيع T وكمايلي:−

$$p(\Lambda,F|X)\ \alpha \frac{1}{|P_\Lambda|^{\frac{(\gamma-N)}{2}}|I_N|^{\frac{m}{2}}} \frac{|P_\Lambda|^{\frac{(\gamma-N)}{2}}|I_N|^{\frac{m}{2}}\ p(F)}{|P_\Lambda+(F\Lambda'-X)'\ I_N\ (F\Lambda'-X)|^{\frac{\gamma}{2}}}$$

$$p(\Lambda,F|X)\ \alpha \frac{1}{|P_\Lambda|^{\frac{(\gamma-N)}{2}}|I_N|^{\frac{m}{2}}} \frac{1}{|P_\Lambda|^{\frac{N}{2}}|I_N|^{\frac{(\gamma-m)}{2}}} \frac{|P_\Lambda|^{\frac{(\gamma-N)}{2}}|I_N|^{\frac{m}{2}}\ p(F)}{|I_N+(F\Lambda'-X)'\ P_\Lambda^{-1}(F\Lambda'-X)|^{\frac{\gamma}{2}}}$$

$$p(\Lambda,F|X)\ \alpha \frac{p(F)\ |P_\Lambda|^{\frac{-\gamma}{2}}}{|I_N+(F\Lambda'-X)'\ P_\Lambda^{-1}(F\Lambda'-X)|^{\frac{\gamma}{2}}} \qquad ...(15\text{-}1)$$

وبالتعويض عن مقام المعادلة (1-15) بـ K وتطبيق بعض العمليات الجبرية نحصل عل ى ما يلي:−

$$\begin{aligned}
K &= I_N+(F\Lambda'-X)P_\Lambda^{-1}\ (F\Lambda'-X)' \\
&= I_N+(F\Lambda'P_\Lambda^{-1}-X\,P_\Lambda^{-1})\ (\Lambda F'-X') \\
&= I_N+F\Lambda'P_\Lambda^{-1}(\Lambda F'-X')\ -X\,P_\Lambda^{-1}\ (\Lambda F'-X') \\
&= I_N+F\,(\Lambda'P_\Lambda^{-1}\Lambda)F'-F\Lambda'P_\Lambda^{-1}X'\ -X\,P_\Lambda^{-1}\,\Lambda F'+X\,P_\Lambda^{-1}X' \\
&= I_N+X\,P_\Lambda^{-1}X'+F[(\Lambda'P_\Lambda^{-1}\Lambda)F'-\Lambda'P_\Lambda^{-1}X']-\ X\,P_\Lambda^{-1}\,\Lambda F' \\
&= I_N+X\,P_\Lambda^{-1}X'+F\,(\Lambda'P_\Lambda^{-1}\Lambda)[F'-(\Lambda'P_\Lambda^{-1}\Lambda)^{-1}(\Lambda'P_\Lambda^{-1}X')]-\ X\,P_\Lambda^{-1}\,\Lambda F' \\
&= I_N+X\,P_\Lambda^{-1}X'+F\,(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'-\ X\,P_\Lambda^{-1}\,\Lambda F' \\
&= I_N+X\,P_\Lambda^{-1}X'+(F-F_0)(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'-\ X\,P_\Lambda^{-1}\,\Lambda F'+ \\
&\quad F_0\ (\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)' \\
&= I_N+X\,P_\Lambda^{-1}X'+(F-F_0)(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'-\ X\,P_\Lambda^{-1}\,\Lambda F'+ \\
&\quad F_0\ (\Lambda'P_\Lambda^{-1}\Lambda)\ F'-F_0\,(\Lambda'P_\Lambda^{-1}\Lambda)\ F_0' \\
&= I_N+X\,P_\Lambda^{-1}X'+(F-F_0)(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'-\ X\,P_\Lambda^{-1}\,\Lambda F'+ \\
&\quad X\,P_\Lambda^{-1}\,\Lambda(\Lambda'P_\Lambda^{-1}\Lambda)^{-1}(\Lambda'P_\Lambda^{-1}\Lambda)F'-\ X\,P_\Lambda^{-1}\,\Lambda(\Lambda'P_\Lambda^{-1}\Lambda)^{-1}(\Lambda'P_\Lambda^{-1}\Lambda)(\Lambda'P_\Lambda^{-1}\Lambda)^{-1} \\
&\quad \Lambda'P_\Lambda^{-1}X' \\
&= I_N+X\,P_\Lambda^{-1}X'+(F-F_0)(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'-\ X\,P_\Lambda^{-1}\,\Lambda F'+X\,P_\Lambda^{-1}\,\Lambda F' \\
&\quad -\ X\,P_\Lambda^{-1}\,\Lambda(\Lambda'P_\Lambda^{-1}\Lambda)^{-1}\ \Lambda'P_\Lambda^{-1}X' \\
&= I_N+X\,P_\Lambda^{-1}X'+(F-F_0)(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'-(X\,P_\Lambda^{-1}\,\Lambda)\ (\Lambda'P_\Lambda^{-1}\Lambda)^{-1}(X\,P_\Lambda^{-1}\,\Lambda)' \\
&= Z+\ (F-F_0)(\Lambda'P_\Lambda^{-1}\Lambda)\ (F-F_0)'
\end{aligned}$$

$$p(\Lambda,F|X)\ \alpha\ \frac{p(F)\,|P_\Lambda|^{\frac{-\gamma}{2}}}{\left|Z+(F-F_0)(\Lambda' P_\Lambda^{-1}\Lambda)\,(F-F_0)'\right|^{\frac{\gamma}{2}}} \qquad \ldots(16\text{-}1)$$

$$Z = I_N + X\,P_\Lambda^{-1}X'(X\,P_\Lambda^{-1}\,\Lambda)\,(\Lambda' P_\Lambda^{-1}\Lambda)^{-1}(X\,P_\Lambda^{-1}\,\Lambda)'$$

$$F_0 = (X\,P_\Lambda^{-1}\,\Lambda)\,(\Lambda' P_\Lambda^{-1}\Lambda)^{-1}$$

بالامكان الحصول على التوزيع اللاحق الحدي لـ $p(\Lambda,|X)$ بتكامل الدالة $p(\Lambda,F|X)$ بالنسبة لـ F وبلستخدام خصائص توزيع T نحصل على ما يلي:-

$$p(\Lambda|X)\ \alpha\ |P_\Lambda|^{-\frac{\gamma}{2}}\left|\ \Lambda' P_\Lambda^{-1}\Lambda\right|^{-\frac{N}{2}}|Z|^{-\frac{(\gamma-m)}{2}}$$

حيث

$$P_\Lambda = B+(\Lambda-\Lambda_0)\,H\,(\Lambda-\Lambda_0)'$$

و

$$Z = I_N + XP_\Lambda^{-1}\,X' - (XP_\Lambda^{-1}\,\Lambda)(\Lambda P_\Lambda^{-1}\,\Lambda)^{-1}(XP_\Lambda^{-1}\,\Lambda)'$$

يمكن الحصول على التوزيع اللاحق الحدي لـ $p(\Lambda,F|X)$ باستخدام العلاقة التالية:-

$$p(\Lambda|F,X)\ \alpha\ \frac{p(\Lambda,F|X)}{p(F|X)} \qquad \ldots(17\text{-}1)$$

بما أن هذا التوزيع معقد جداً سنقوم بتقدير Λ بشكل متناوب لـ $F=\hat{F}$ لاحظ أولاً من (1-9) أن:

$$p(\Lambda|F,X)\ \alpha\ \left|R_F+(\Lambda-\Lambda_F)Q_F(\Lambda-\Lambda_F)'\right|^{\frac{-Y}{2}} \quad \ldots(18\text{-}1)$$

حيث Λ_F هي كما موضح سابقاً وكما يلي:-

$$\Lambda_F = (X'F+\Lambda_0)(H+F'F)^{-1}$$

وذلك هو التوزيع الشرطي لـ Λ بدلالة F محددة مسبقاً وهي مصفوفة T.وبالتالي $\hat{\Lambda}$ هو النقدي النقطي لـ Λ

$$\hat{\Lambda} = \Lambda_{\hat{F}} = (X'\hat{F}+\Lambda_0 H)(H+\hat{F}'\hat{F})^{-1} \quad \ldots(19\text{-}1)$$

وهكذا فان المعادلة (1-19) تمثل مصفوفة تحميلات العوامل.

الجانب التطبيقي

جمع البيانات:- أخذت البيانات من مستشفى ابن الأثير التعليمي للولادة والأطفال في الموصل من مرضى مصابين بفقر دم البحر الأبيض المتوسط من نوع بيتا أو ما يسم ى بالثلاسيميا، وكان عدد المشاهدات (150) مشاهدة، كما واختير عدد من المتغيرات التي يعتقد بأنها تؤثر فيه بعد مراجعة أطباء اختصاصيين وممارسين في مرض الثلاسيميا وهذه المتغيرات هي :- (النعيمي، 2005)

X_1 العمر الحقيقي مقاساً بالشهر	Real age (Month)
X_2 عمر المريض عند المرض مقاساً بالشهر	Age of patient about disease (Month)
X_3 تضخم الكبد مقاساً (بالسنتمتر)	Enlargement of liven (c.m)

X_4 هيموكلوبين الدم	Hemoglobin
X_5 مكداس الدم	Packed Cells Volume(P.C.V)
X_6 الخلايا الشبكية	Reticulacyte
X_7 أرومة حمراء	Normooblast
X_8 الهيموكلوبين الجيني	Fetal Haemoglobin
X_9عدد وحدات الدم	Numbers of Blood units
X_{10} بداية نقل الدم حسب العمر مقاساً (بالشهر)	Onset of Blood transfusion to according

استخدم البرنامج الإحصائي (V13.20) Minitab لإيجاد مقدر بيز لقيم العوامل وتحميلات العوامل من خلال المعادلات (1−12) و (1−19) على التوالي.

الآن سوف نختار نموذج بأربعة متغيرات، هذا الاختيار يعتمد على تنفيذ تحليل المكونات الرئيسة وبالاعتماد على نسبة التباين المفسرة نجد بان هذه العوامل فسرت ما يقارب (76.2%) من التباين الكلي (الراوي، 1987) وكما هو موضح بالجدول (1)

جدول (1): يبين الجذور المميزة المختارة من مصفوفة الارتباط البسيط

المكونات الرئيسة	الجذور المميزة	نسبة تبلين المكون الرئيس	التباين المتجمع
1	2.566	0.257 0.222	0.257 0.479
2	2.221	0.181 0.103	0.660 0.762
3	1.808		
4	1.028		

نفترض بان $B=0.2I_{10}$ (مصفوفة قطرية)، $H=10I_4$ ، Λ_0 كما هو موضح في جدول (2) وبالتطبيق على المعادلات (1−12) و (1−19) على التوالي نجد بان مقدر بيز لقيم العوامل كما هو موضح في الجدول (4) (الملحق 1) كما و جد مقدر بيز لتحميلات العوامل في الجدول (3).

جدول (2): التوزيع الأولي Λ_0

0.0	0.7	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.7	0.0
0.7	0.0	0.0	0.0
0.0	0.0	0.0	0.7
0.0	0.0	0.0	0.7
0.7	0.0	0.0	0.0
0.0	0.0	0.0	0.7
0.0	0.7	0.0	0.0
0.0	0.0	0.0	0.7

جدول (3): يبين مقدر بيز لمصفوفة تحميلات العوامل

communality	Factor 4	Factor 3	Factor 2	Factor 1
0.494	0.067	0.698	0.001	0.040
0.324	−0.013	−0.006	0.085	0.563
0.529	−0.000	0.000	0.727	−0.000
0.488	0.673	−0.007	−0.114	0.146
0.405	0.524	0.018	−0.120	0.339
0.812	−0.210	0.064	−0.036	0.873
0.601	0.752	0.008	0.114	−0.145
0.693	−0.171	−0.098	0.007	0.808
0.536	−0.067	0.728	−0.000	−0.040
0.687	−0.141	0.015	0.149	0.803

تفسير نتائج مقدر بيز لمصفوفة تحميلات العوامل:-

استخدم البرنامج الإحصائي Minitab v(13.20) لإيجاد مقدر بيز لمصفوفة تحميلات العوامل ومن خلال النتائج الموضحة في الجدول (3) تبين أهمية جميع المتغيرات في تشخيص مرض الثلاسيميا ولكن حسب الأهمية فمن تسلسل العامل ونسبة ما يفسرهكل عامل من التباين الكلي وكميات الشيوع الخاصة بكل متغير تبين أن المتغيرات الثاني (عمر المريض عند المرض) والسادس (الخلايا الشبكية) والثامن (الهيموكلوبين الجيني) والعاشر (بداية نقل الدم حسب العمر) تأتي بالمرتبة الأولى بليها المتغير الثالث (العمر الحقيقي) بالمرتبة الثانية ، فضلاً عن ذلك نجد أن المتغيرات الأول (العمر الحقيقي) والتاسع (عدد وحدات الدم) بالمرتبة الثالثة , وأخيراً ظهرت المتغيرات الرابع (هيموكلوبين الدم) والخامس (مكداس الدم) والسابع (أرومة حمراء) بالمرتبة الأخيرة.

مصفوفة تحميلات العو امل الموضحة في الجدول (3) لا تحتا ج إلى عملية تدوير لكون قيم التحميلات (التشبعات) متباعدة فيما بينها في كل عامل وهذا ما اكده (Press and Shigemasu, 1997).

الاستنتاجات:

يمكن توضيح أهم الاستنتاجات بما يلي:-

1) المعلومات الأولية في نموذج تحليل بيز العاملي تأتي من خبراء ماهرين أو من تجارب سابقة.

2) ظهور جميع المتغيرات المختارة بالنسبة لمرض الثلاسيميا ذات تأثير معنوي ولكن حسب الأهمية حيث تم اختيار أهمية كل متغير بالاعتماد على عدة أسباب أولاً عدد الجذور المميزة يجب أن يكون اكبر من الواحد ثانياً نسبة ما يفسرهكل عامل من التباين الكلي ثالثاً كميات الشيوع المستخلصة من كل متغير لذلك فان المتغيرات الثاني (عمر المريض عند المرض) والسادس (الخلايا الشبكية) والثامن (الهيموكلوبين الجيني) والعاشر (بداية نقل الدم حسب العمر) تأتي بالمرتبة الأولى بليها المتغير الثالث (العمر الحقيقي) بالمرتبة الثانية , فضلاً عن ذلك نجد أن المتغيرات الأول (العمر الحقيقي) والتاسع (عدد وحدات الدم) بالمرتبة الثالثة , وأخيراً ظهرت المتغيرات الرابع (هيموكلوبين الدم) والخامس (مكداس الدم) والسابع (أرومة حمراء) بالمرتبة الأخيرة.

3) نموذج تحليل بيز العاملي لا يحتا ج إلى عملية تدوير لمصفوفة تحميلات العوامل لان التدوير موجود اوتوماتيكياً.

المصادر

أ- المصادر العربية

1) الراوي، خاشع محمود (1987): "**المدخل إلى تحليل الانحدار** ". مديرية دار الكتب للطباعة والنشر –جامعة الموصل –الطبعة الأولى.

2) النعيمي، أسوان محمد طيب رشيد (2005): "**اختيار المتغيرات في انحدار الحرف** "، رسالة ماجستير، كلية علوم الحاسبات والرياضيات – جامعة الموصل.

ب- المصادر الإنكليزية

3) Daniel B. Rowe (2003): "**Multivariate Bayesian Statistics**, Models for source separation and signal unmixing", USA.

4) Press S. James & Shigemasu K. (1997): "**Bayesian inference in factor analysis- Revised**". Technical Report No. 243, Department of Statistics, University of California, USA.

5) Daniel B. Rowe (2000a): "**On Using the Sample Mean in Bayesian Factor Analysis**", social science working paper 1096, USA.

6) Daniel B. Rowe (2000b): "**Incorporating prior knowledge regarding the mean in Bayesian factor analysis**", social science working paper 1097, USA.

7) Daniel B. Rowe (2000c): "**A Bayesian factor analysis model with generalized prior information**" ,social science working paper 1099, USA.

الملحق (1)

جدول (4): مقدر بيز لقيم العوامل

Factor 4	Factor3	Factor2	Factor1	Factor 4	Factor3	Factor2	Factor1
−0.194	0.537	−0.246	−0.125	0.513	2.623	1.977	0.217
−1.584	−0.888	3.089	0.670	−0.553	2.908	0.865	−0.813
0.222	−1.373	0.309	2.394	0.008	1.381	−0.246	0.387
0.095	−0.906	0.587	−0.015	−0.610	0.341	1.421	0.788
2.067	2.786	−0.524	−0.041	−0.309	0.688	0.865	−0.412
2.244	3.419	0.587	−0.439	0.599	0.082	−0.246	−0.359
−1.378	−0.140	−0.246	−0.366	−0.486	−0.211	−0.802	−0.661
−1.211	−0.978	−0.246	−0.532	−0.375	−1.798	−2.470	1.322
1.230	1.039	1.977	0.015	−0.722	−0.235	0.031	−0.949
1.553	−0.924	1.977	1.468	−0.759	0.352	3.367	−0.927
0.136	2.504	−0.802	1.518	−0.705	−0.945	0.865	−0.652
4.220	2.073	−0.524	1.383	0.089	2.513	1.977	−0.375
−0.152	−1.245	0.031	−0.208	0.797	0.400	−0.246	−0.095
−0.192	−1.257	0.865	1.533	0.206	0.581	1.143	0.119
−0.710	0.590	2.533	−0.497	−0.575	−0.131	0.309	−0.883
0.020	−0.383	−2.470	0.477	−1.507	−0.862	1.977	−1.168
2.353	−0.122	1.421	0.011	−0.392	−1.233	−1.358	−0.122
−0.682	2.985	−2.470	−0.012	−0.404	0.002	0.309	−0.347
−1.118	0.091	3.645	−0.763	0.102	−0.870	−2.470	−0.223
0.390	−0.237	−0.246	2.634	−0.405	−1.798	−2.470	0.992
1.026	−1.607	−0.802	1.334	1.980	0.382	1.421	0.539
0.126	−1.658	−2.470	0.039	0.715	0.721	1.421	−0.578
−0.357	−1.732	−0.246	0.564	−0.732	−1.869	−0.802	0.939
2.169	−0.419	0.865	0.858	−0.515	1.232	0.865	−0.291
1.195	3.704	−0.246	0.787	−0.771	−1.111	0.309	−0.905
−0.247	−1.441	−1.914	0.714	−0.905	−1.905	0.865	−0.848
−0.746	−1.426	1.421	0.999	0.932	1.232	0.309	−0.440
0.008	−1.010	−1.358	−0.910	−0.486	0.127	−1.358	−0.742
0.974	−1.566	−0.802	1.516	2.353	−0.814	1.421	−0.287
0.601	1.812	0.865	0.962	−0.682	−1.049	0.865	−0.123
3.377	0.635	0.309	1.028	−1.119	0.073	0.309	−0.621
0.527	0.786	0.865	2.003	1.099	0.977	−0.524	0.361
−0.204	0.192	3.367	0.073	0.390	−1.067	−0.802	0.475
0.598	0.240	1.977	−0.684	0.721	0.038	−1.080	0.128
0.781	−0.364	−0.246	−0.105	0.438	−0.039	−0.802	−0.334
−1.085	1.719	2.255	−1.062	−0.151	−1.816	−1.358	1.362
0.102	0.949	0.031	0.950	−0.339	1.108	1.977	−1.052

0.397	3.321	0.031	−0.276	−2.037	−0.826	−1.358	−1.561
−1.242	0.252	−1.080	−0.502	−0.619	1.375	1.421	−1.447
1.908	−0.052	0.309	−0.233	1.553	−0.389	0.309	1.373
−0.488	0.064	1.699	−0.692	−1.902	−0.818	0.865	−0.818
0.079	−1.506	−1.080	1.304	−1.241	−1.067	−0.802	1.510
1.116	−0.948	−2.470	0.086	0.303	−1.260	−2.470	0.694
−0.746	−0.131	−0.524	−0.552	−1.206	−1.762	0.309	2.221
0.425	0.773	1.977	0.644	−1.518	−1.893	0.587	−1.288
0.303	0.595	−0.802	−0.049	−0.728	−1.822	−0.802	−0.679
−1.064	−0.294	−0.246	−0.872	2.405	2.284	−0.246	−0.614
0.444	2.178	−1.636	0.195	−0.964	−1.168	−0.802	−0.079
−0.020	−0.192	−0.246	−0.167	−0.286	−1.414	−0.246	−0.376
2.358	2.264	−1.358	0.142	−0.681	−0.134	−0.246	−0.853
−0.346	−1.471	0.031	1.196	0.368	0.073	0.309	−0.513
−0.571	−0.732	−1.636	0.029	−0.062	−1.676	0.309	−0.057
−0.078	−0.163	−0.246	−0.043	−1.043	0.037	−2.470	−0.937
0.333	2.117	0.031	0.273	−0.711	0.079	0.865	−0.657
−1.524	−1.723	−1.080	−0.833	−0.864	−0.265	0.865	−0.898
1.430	3.657	1.977	0.774	0.644	−0.838	−0.246	0.625
1.059	1.477	0.309	−0.434	−0.455	−0.178	0.309	0.240
0.931	0.701	1.421	−0.124	−0.463	−0.493	−0.802	−0.203
−1.161	−1.295	−0.802	0.206	−0.517	−0.098	1.977	−0.683
−0.752	−1.127	−2.470	−0.707	−0.965	−1.506	−1.080	1.256
−1.848	0.105	0.865	−1.462	−0.109	1.357	−1.914	−0.600
0.152	−0.933	0.031	0.011	−0.026	1.072	−1.358	−0.546
−0.728	0.233	0.309	−0.872	−1.133	0.934	1.977	−1.214
0.279	−0.260	0.309	−0.441	0.114	−1.364	−2.470	−0.525
0.356	−0.675	−1.636	−0.348	−1.270	2.961	0.865	−0.506
1.651	−0.911	0.309	2.026	−0.551	−1.718	−2.470	0.512
0.137	−0.909	−0.524	0.189	−0.551	−1.561	−0.246	−0.643
−0.380	−1.291	−0.524	−0.330	−0.464	0.041	−0.802	−0.441
3.319	0.516	−1.080	1.445	−0.469	1.552	1.977	0.299
−0.160	−0.980	0.865	0.821	−1.407	−1.962	−0.246	−0.379
0.941	4.043	0.865	−0.196	−1.472	−1.106	0.865	−1.135
−0.964	1.004	−1.080	−0.985	0.309	−0.742	−0.802	−0.296
0.790	−0.519	1.421	0.318	−0.256	−1.320	−2.470	−0.325
				−1.037	−0.028	−0.246	−1.270
				−0.440	−0.478	−2.470	0.150
				−0.576	0.447	−1.358	−0.176