مجلة التربية والعلم - المجلد (23)، العدد (2)، لسنة 2010

استخدام مصفوفة ارتباط مقترحة في التحليل ألعاملي

مروان عبد العزيز دبدوب

قسم الإحصاء والمعلوماتية / كلية علوم الحاسبات والرياضيات

جامعة الموصل

الاستلام	القبول
2009 / 03 / 15	2009 / 05 / 05

ABSTRACT

some variables may have nominal or ordinal scale, this causes a failure of analyzing the Pearson's correlation matrix by factor analysis. However, it was suggested to replace the elements of the correlation matrix by correlation coefficients which should be calculated according to the scales of the two correlated variables. The suggested correlation matrix might become an indefinite matrix which can not be analyzed by factor analysis, so the suggested methods were established to eduction a positive definite matrix or positive semi-definite matrix.

المستخلص

لا يعول على نتائج التحليل ألعاملي إذا طبق على مصفوفة ارتباط بيرسون وكانت بعض متغيراتها ذات مقاييس أسمية nominal أو رتبيه ordinal، لذا اقترح استبدال عناصرها بمعامل ارتباط ثنائية تحسب وفقاً لنوع مقياس كل متغير من المتغيرين المرتبطين. إن مصفوفة الارتباط المقترحة قد تكون من النوع غير المحدد indefinite matrix والتي لا يمكن تطبيق التحليل ألعاملي عليها، فقد اقترحت طرائق لاستنباط مصفوفة ارتباط من النوع أكيد الايجابية positive definite matrix أو شبه أكيدة الايجابية positive semi-definite matrix.

المقدمة

إن السعي إلى تحديد وقياس العوامل الكامنة وراء الظواهر يعتمد أساسا على ملاحظة طبيعة وعلاقة المتغيرات التي تُكون الأحداث الطبيعية للظاهرة، سواء أكانت بايولوجية أو اقتصادية أو سياسية أو نفسية ... وغيرها. إن العلاقة بين المتغيرات لا تظهر بطريقة عفوية خاضعة للصدفة أو التلقائية، ولكنها تخضع لنظام تحدده طبيعة الظاهرة المدروسة. وبذلك وجب على الباحث التعرف على تلك العلاقات وتوظيفها للتوصل إلى تفسير علمي صحيح لما بين يديه من دراسة مستعيناً بالأساليب الإحصائية.

إن التحليل ألعاملي (factor analysis) هو أحد الأساليب الإحصائية الذي يحلل مصفوفة الارتباط بين المتغيرات ليقدم أكثر ما يمكن من المعلومات عن الظاهرة بأقل عدد من العوامل. ويعد تحليل المكونات الرئيسة (principal component analysis) الأسلوب الأكثر إتباعا في التحليل ألعاملي، إذ يعمل على تحويل المتغيرات المترابطة إلى متغيرات غير مترابطة عن طريق التعامد (orthogonal)، وبذلك يتيح للباحث وضع تراكيب خطية للمتغيرات الأصلية تدعى بالمكونات الرئيسة يفسر كل منها نسبة من التباين الكلي، يتم اختيار بعض تلك التراكيب وفق ما تحتويه من نسبة مقنعه من المعلومات الأصلية (دبدوب، 1996).

تعود فكرة التحليل ألعاملي إلى عام 1904 عندما افترض E. S. Pearson وجود عامل عام يفسر الارتباط بين المتغيرات، وفي عام 1933 قدم H. Hotelling طريقة المكونات الرئيسة، وفي عام 1955 أشار C. R. Rao إلى تقدير كميات الشيوع (communality) للمتغيرات، وبذلك تبلور الأسلوب الإحصائي للتحليل ألعاملي (الشكرجي، 2005).

من المتبع أن التحليل ألعاملي يطبق على مصفوفة ارتباط بيرسون (Pearson's correlation matrix)، عناصرها معامل الارتباط الثنائي لبيرسون والمعروف بمعامل الارتباط الثنائي البسيط (simple correlation coefficient) بين توافيق أزواج المتغيرات، كل معامل ارتباط عبارة عن مقياس لقوة واتجاه العلاقة الخطية بين متغيرين. ذكر عودة والخليلي (1988) و Siegel and Castellan (1988) و Sullivan (2007) أن نوع مقياس المتغيرين يحدد نوع معامل الارتباط الثنائي بينهما، فمعامل ارتباط بيرسون يمثل الارتباط بين متغيرين مقاسين بمقياس نسبي ratio scale أو فتروي interval scale (Agresti and Franklin, 2007).

الهدف:

إن التحليل ألعاملي يتعامل مع أعداد كثيرة من المتغيرات غالباً ما يكون لهـا مقاييس مختلفة. إن وجود متغير واحد ذي مقياس اسمي nominal scale أو رتبي ordinal scale يجعل مصفوفة ارتباط بيرسون تبتعد عن حقيقة تمثيلها للارتباط الثنائي بين بعض أو كل متغيراتها، ومن ثم تبتعد نتائج التحليل ألعاملي عن صحتها في تفسير الظاهرة المدروسة. وللتغلب على هذه المعضلة وضع الهدف الآتي:

إنشاء مصفوفة ارتباط تحدد عناصرها حسب نوع مقياس كل متغير من أزواج المتغيرات الداخلة في عملية الارتباط. وسيتم مناقشة الايجابيات ومعالجة السلبيات نتيجة تطبيق التحليل ألعاملي على مصفوفة الارتباط المقترحة.

تحليل المكونات الرئيسة والتحليل ألعاملي:

إن مصفوفة البيانات $\underline{X}_{n\times m}$ تحتوي في أعمدتها على m من المتغيرات وفي صفوفها على n من الأفراد (i=1,2,…,n; j=1,2,…,m). من مصفوفة البيانات تحسب مصفوفة ارتباط بيرسون بين توافيق أزواج المتغيرات، ومن المتبع أن يطبق التحليل ألعاملي على هذه المصفوفة بإتباع أسلوب المكونات الرئيسة الذي يعد أكثر الأساليب استخداما (الجبوري وعبد، 2000).

بتطبيق تحليل المكونات الرئيسة على مصفوفة الارتباط نحصل على m من الجذور المميزة (characteristic roots) الموجبة (λ_j) مرتبة تنازلياً $\lambda_1>\lambda_2>...>\lambda_m$، ترافقها متجهات مميزة ($\underline{a}_j$) (characteristic vectors) (Kuo and Golnaraghi, 2003)، توضع المتجهات المميزة بوصفها عوامـل في تركيبة خطية للمتغيرات الأصلية ذات الوحدات القياسية ($z_{ij} = (x_{ij} - \bar{x}_{ij})/s_j$) لتعطي ($pc_{ij}$) القيمة i للمكون الرئيس j :

$$pc_{ij} = a_{1j}z_{i1}+a_{2j}z_{i2}+\ldots+a_{mj}z_{im} \quad \ldots \quad (1)$$

ويمكن الحصول على متجه $\underline{pc}_j$ يحتوي على n من قيم المكون الرئيس j:

$$\underline{pc}_j = \underline{Z}\,\underline{a}_j \quad \ldots \quad (2)$$

حيث أن $\underline{Z}_{n\times m}$ هي المصفوفة القياسية لمصفوفة البيانات $\underline{X}_{n\times m}$.

إن اختيار k من المكونات الرئيسة هو بعينه اختيار k من عوامل التحليل ألعاملي، وهنالك عدة طرائق لتعيين عدد العوامل المختارة ولكل منها ميزاتها وعيوبها (الجبوري وعبد، 2000). وقد أشار Kline (1994) إلى رأي الباحث في ذلك مع الأخذ بنظر الاعتبار ارتفاع نسبة التفسير من التباين الكلي والتي تحسب كالآتي:

$$\% \text{ of total variance explained} = \sum_{j=1}^{k} \frac{\lambda_j}{\sum_{j=1}^{m} \lambda_j} \quad \dots \quad (3)$$

وذكر دبدوب والشكرجي (2007) أن عدد العوامل المختارة يعتمد بالدرجة الأولى على رأي المختص في دراسة الظاهرة ومدى قناعته بنسبة التباين المفسر، وعلى هذا الأساس استخلص الشكرجي (2005) تلك العوامل التي فسرت فقط 68% من التباين الكلي.

يعمل التحليل ألعاملي على فصل تأثير العوامل عن تأثير الأخطاء، ويساعد ذلك على اختيار عدد k من العوامل التي تفسر الاختلافات في عدد m من المتغيرات (حيث أن: m>k)، وتكون العلاقة بين المتغيرات في العامل الواحد أقوى من العلاقة مع المتغيرات في العوامل الأخرى. ويعبر عن كل متغير بنموذج يضم العوامل المختارة (Johnson, 1998)، وهو:

$$X_j = p_{1j} F_1 + p_{2j} F_2 + \dots + p_{gj} F_g + \dots + p_{kj} F_k + U_j \quad \dots \quad (4)$$

F_g = عدد k من العوامل المشتركة common factors المختارة من بين عدد m من العوامل، حيث أن: g=1,2,…,k.

P_{gj} = تحميل العامل المشترك g (factor loading) للمتغير، حيث أن: j=1,2,…,m .

U_j = الجزء العائد للمتغير j الذي لا تفسره العوامل المشتركة، ويتكون من مصدرين هما: خطأ القياس error of measurement والعامل الخاص specific factor.

يتم تعيين أهمية المتغير في العامل من خلال محورين: المحور الأول: دراسة كل عامل مختار بصورة منفردة، إذ تعين أهمية المتغير من خلال قيمتين: الأولى: نسبة ما يفسره العامل من التباين الكلي، فالعامل الأول له أكبر نسبة تفسير يليه العامل الثاني وهكذا البقية، القيمة الثانية: هي قيمة تحميل العامل g على المتغير j والتي تدل على نسبة ما يفسره المتغير من تباين العامل المختار، وأشار Afifi and Clark (1884) إلى الدلالة الإحصائية لقيم تحميل العامل من خلال قيمة حرجة (critical value) تحسب لكل عامل مختار كالآتي:

$$\text{critical value} = (0.5)(\text{variance of } F_g)^{-1/2} \quad \dots \quad (5)$$

بانخفاض القيمة الحرجة العائدة للعامل g (g=1,2,…,k) عن قيمة تحميل العامل على المتغير j (j=1,2,…,m) يعد المتغيرات ذات دلالة إحصائية في العامل g. المحور الثاني: دراسة المتغير بصورة عامة من خلال وجودة في العوامل المختارة. تعتمد هذه الطريقة على قيمة تسمى كميات الشيوع (communality) وهي نسبة التباين المشاع الذي يفسره التداخل بين المتغير والعوامل المشتركة، وتقع قيمته بين الصفر والواحد، فإذا اقتربت من الواحد فالمتغير يزداد تداخله مع العوامل المشتركة وبذلك تزداد أهميته في تفسير الظاهرة لمدروسة، وإذا اقتربت القيمة من الصفر

فإن العوامل المشتركة لا تستطيع تفسير الشيء المهم من تباين المتغير فتتلاشى أهمية المتغير في تأثيره على الظاهرة المدروسة.

إن مقدار كمية الشيوع h^2_j للمتغير j هو مجموع مربعات تحميلات العوامل المختارة للمتغير، فمن المعادلة (4) تحسب القيمة كالآتي:

$$h_j^2 = \sum_{j'=1}^{k} p_{jj'}^2 = p_{j1}^2 + p_{j2}^2 + ... + p_{jk}^2 \quad \ldots \quad (6)$$

أنواع مقاييس المتغيرات:

المتغير عبارة عن أية ظاهرة تتغير كمياً أو نوعياً من فرد إلى آخر في مجتمع معين. ولكل متغير أحد أنواع المقاييس الآتية (عودة والخليلي، 1988)، (Agresti and Franklin, 2007):

1) المقياس الاسمي nominal scale: يحدد قيم عددية للخصائص، وللمتغير عدد محدد من الأرقام الرمزية التي ليس لها وزن أو معنى حقيقي. ويتكون من مقياس أسمي طبيعي كالجنس، أو مقياس أسمي محول كالحكم الشخصي على تصنيف ظاهرة ما.
2) المقياس ألرتبي ordinal scale: وهو ذو مدلول رتبي تصاعدي أو تنازلي حسب مستوى الصفة المقاسة، ويحتوي على عدد محدد من الرتب، ولا يمكن تحديد قيمة الفرق بين أي رتبتين. كما لا يمكن الحكم على تساوي الفرق بين الرتب المتتالية، فمثلاً الفرق بين رتبة امتياز ورتبة جيد جداً لا يمكن تحديدها، كما أنها لا تساوي الفرق بين رتبة جيد جداً ورتبة جيد.
3) المقياس الفتروي interval scale: يمكن لهذا المقياس من تحديد الفرق بين أي مستويين، كما أن الفرق بين المستويات المتتالية تكون متساوية، وإن الصفر لا يعني غياب الصفة، فمثلاً الصفر المئوي هي درجة حرارة تجمد الماء، في حين يتجمد الماء أيضاً عند درجة حرارة 32 فهرنهايت. إن المتغيرات الفترية هي متغيرات كمية يمكن إجراء العمليات الحسابية عليها دون أن تتأثر المسافة النسبية بين قيمها.
4) المقياس النسبي ratio scale: يختلف عن المقياس الفتروي في أن نقطة الصفر فيه تعني غياب الصفة، وأن النسبة بين أي نقطتين لها تفسيراً منطقياً مستقلاً عن وحدة القياس.

يتم التعامل إحصائياً مع المقياسين الفتروي والنسبي بذات الطريقة، ويطلق على متغيراتهما بالمتغيرات الكمية quantitative variables.

الارتباط الثنائي بين المتغيرات:

إن حساب معامل الارتباط الثنائي يعتمد على نوع مقياس كل متغير داخل في عملية الارتباط. هنالك أربعة مجاميع من أنواع مقاييس المتغيرات وهي: الأولى تتمثل بالمقياسين النسبي والفتروي، والثانية المقياس ألرتبي، والثالثة المقياس الاسمي الطبيعي، والمجموعة الرابعة المقياس الاسمي المحول. إن هذه المجاميع تحدد عشرة احتمالات لتكوين أزواج المقاييس، فمثلاً متغيران يعود مقياساهما إلى المجموعة الأولى ليكون معامل ارتباط بيرسون (Person's correlation coefficient) هو الممثل لقوة واتجاه العلاقة بينهما، في حين انتماء مقياسيهما إلى المجموعة الثانية يكون معامل ارتباط سبيرمان (Spearman's correlation coefficient) الممثل الصحيح للعلاقة بين المتغيرين، وعندما يكون أحد المتغيرات ذو مقياس أسمي محول والمتغير الثاني ذو مقياس نسبي فإن العلاقة بينهما تمثل بمعامل ارتباط ثنائي التسلسل (Biserial correlation coefficient). وهكذا بقية معاملات الارتباط المختلفة (عودة والخليلي، 1988).

ليست كافة توافيق أزواج المقاييس لها معامل ارتباط معرف يمكن حسابه، ففي بعض الحالات يجب تحويل مقياس متغير إلى نوع آخر يسمح بإيجاد معامل ارتباط بين المتغيرين، مثلاً لا يعرف معامل ارتباط بين متغيرين أحدهما ذي مقياس نسبي أو فتروي والآخر ذي مقياس رتبي، في هذه الحالة يحول المتغير ذو المقياس النسبي أو الفتروي إلى متغير ذو مقياس رتبي، وبذلك يمكن حساب معامل ارتباط سبيرمان ليمثل العلاقة بين المتغيرين.

عند عدم خضوع أحد المتغيرين أو كلاهما إلى مقياس نسبي أو فتروي، فان ذلك يؤدي إلى عدم إمكانية الاعتماد على معامل ارتباط بيرسون في تمثيل العلاقة بين المتغيرين، ومن ثم لا يعول على نتائج التحليل ألعاملي الناتجة من مصفوفة الارتباط لبيرسون. لذا أقترح إنشاء مصفوفة ارتباط عناصرها معاملات الارتباط الثنائي وفقاً لنوع مقياس كل متغير من أزواج المتغيرات المترابطة، ليطبق التحليل ألعاملي على مصفوفة الارتباط المقترحة.

تطبيق المقترح ومعالجة سلبياته:

لتوضيح الايجابيات ومعالجة السلبيات فقد طبق المقترح على بيانات جمعت كالآتي:

قبل انتهاء العام الدراسي 2006-2007 بستة أسابيع، أخذت عينة عشوائية حجمها 55 طالباً من طلاب المرحلة الرابعة البالغ عددهم 84 طالباً في قسم الإحصاء والمعلوماتية-كلية علوم الحاسبات والرياضيات–جامعة الموصل-العراق. أُعطي لكل فرد استمارة استبيان (الملحق 1) احتوت على سبعة أسئلة كلٌ منها احتوى على متغير بمقياس معين (الجدول 1).

الجدول 1: المتغيرات التي طبق عليها المقترح: تقسيماتها ورموزها وأنواع مقاييسها.

المتغيرات	الفئات	نوع المقياس
X_1= الجنس	أنثى:0 ، ذكر:1	أسمي طبيعي
X_2= القناعة بالانتساب إلى الكلية	بدون قناعة:0 ، بقناعة:1	أسمي محول
X_3= القناعة بالانتساب إلى القسم	بدون قناعة:0 ، بقناعة:1	أسمي محول
X_4= التحصيل العلمي للأب	أمي:1 ، ابتدائية:2 ، متوسطة:3 ، إعدادية:4 ، معهد أو كلية:5 ، عليا:6	رتبي
X_5= التحصيل العلمي للام	أمي:1 ، ابتدائية:2 ، متوسطة:3 ، إعدادية:4 ، معهد أو كلية:5 ، عليا:6	رتبي
X_6= مجموع درجات الامتحان العام للسادس الإعدادي العلمي		فتروي
X_7= الدخل السنوي للأسرة (بالدينار العراقي)		نسبي

يبين الملحق (2) البيانات التي أعطتها العينة. وقد استخدم في تحليل البيانات البرنامجان الحاسوبيان الجاهزان SPSS (Brace et al., 2006) و Minitab (Carver, 1999).

سيطبق التحليل ألعاملي على نوعين من مصفوفة الارتباط بين ستة متغيرات ($X_1,\ldots,X_6$) باستخدام البيانات في الملحق (2). النوع الأول: مصفوفة ارتباط عناصرها قيم معامل بيرسون بين أزواج المتغيرات، ويتم الحصول على المصفوفة مباشرة بالاستعانة بالبرنامج ألحاسوبي الجاهز SPSS. والنوع الثاني: مصفوفة الارتباط المقترحة، وقد أشار عودة والخليلي (1988) و(Siegel and Castellan (1988 و(Sullivan (2007 إلى طرائق حساب عناصر هذه المصفوفة واختبار دلالتها الإحصائية وفق نوع مقياس المتغيرين المترابطين.

استخدام مصفوفة معامل ارتباط بيرسون:

وضعت قيم معامل ارتباط بيرسون بين أزواج المتغيرات في الجدول (2)، وأشير إلى قيمة مستوى المعنوية (α) التي ترفض عندها فرضية العدم: $H_O : \rho = 0$ وتقبل الفرضية البديلة: $H_1 : \rho \neq 0$ (Berenson et al., 2006).

الجدول 2: مصفوفة الارتباط عناصرها معامل بيرسون بين أزواج المتغيرات.

	X_1	X_2	X_3	X_4	X_5
X_2	0.1180				
X_3	0.0790	0.6710**			
X_4	-0.0690	-0.0140	0.1330		
X_5	-0.1100	-0.1000	-0.0830	0.4580**	
X_6	-0.1580	0.1360	0.2020	-0.0180	0.0760
** ذو دلالة إحصائية عند $\alpha = 0.01$.					

يشير الجدول (2) إلى معاملين ارتباط لهما دلالة إحصائية: القناعة بدخول الكلية (X_2) مع القناعة بالانتساب إلى القسم (X_3) وكذلك التحصيل الدراسي للأب (X_4) مع التحصيل الدراسي للام (X_5). بتطبيق التحليل ألعاملي على مصفوفة ارتباط بيرسون تم الحصول على ستة عوامل اختير منها أول ثلاثة عوامل (الجدول 3) فسرت 73.40% من التباين الكلي (المعادلة 3).

الجدول 3: نتائج التحليل ألعاملي لمصفوفة ارتباط بيرسون.

Variable	Factor Loading			Communality
	Factor-1	Factor-2	Factor-3	
X_1	0.1910	-0.3380	0.6980	0.6380
X_2	0.8870	0.0200	0.0850	0.7950
X_3	0.8940	0.1580	0.0610	0.8270
X_4	-0.0130	0.8040	0.3520	0.7700
X_5	-0.2120	0.8070	0.1540	0.7200
X_6	0.3290	0.2660	-0.6870	0.6510
Variance (λ_j)	1.7755	1.5082	1.1180	4.4016
% of Var.	0.2960	0.2510	0.1860	0.7340

تشير تحميلات العوامل (factor loading) إلى أهمية المتغير من ناحيتين: الأولى: كمية ما يفسره العامل المختار من التباين الكلي (قيمة λ_j)، الناحية الثانية: قيمة تحميل العامل g على المتغير j والتي تدل على نسبة ما يفسره المتغير من تباين العامل المختار. وفق المعادلة (5) تم الحصول على القيم الحرجة للعوامل المختارة ومقدارها على التوالي: 0.3752 و 0.4071 و0.4729. عند انخفاض القيمة الحرجة g (g=1,2,3) عن قيمة تحميل العامل g على المتغير j (j=1,2,…,m) يعد ذلك المتغير ذو دلالة إحصائية في ذلك العامل، على هذا الأساس تبين أن القناعة بدخول الكلية (X_2) والقناعة بالانتساب إلى القسم (X_3) لهما التأثير الرئيس على الدراسة الجامعية، ليأتي بالمرتبة الثانية التحصيل الدراسي للأب (X_4) والأم (X_5)، وأخيراً: الجنس (X_1) ومجموع درجات الامتحان العام (X_6) الذي تدل إشارته السالبة أن له تأثيراً عكسياً.

إن كميات الشيوع (communality) تشير إلى مدى التداخل بين المتغير والعوامل المشتركة المختارة لتفسر بشكل عام تأثير المتغير من خلال العوامل المختارة، وبذلك يمكن تحديد أهمية المتغيرات في تأثيرها على الظاهرة المدروسة. من النتائج التحليلية في الجدول (3) يمكن إعطاء تسلسل لأهمية المتغيرات بصورة عامة كالآتي: X_3 يليه X_2، X_4، X_5، X_6، وأخيراً X_1.

استخدام مصفوفة الارتباط المقترحة:

لاختلاف المتغيرات في أنواع مقاييسها (الجدول 1) يقترح استحداث مصفوفة ارتباط تحسب عناصرها وفقاً لنوع تلك المقاييس وكذلك حال اختبار فرضياتها، ويشير الجدول (4) إلى أنواع معاملات الارتباط الثنائية بين الستة متغيرات. حول مقياس المتغير (X_6) من الفتروي إلى الرتبي (الملحق 2) لنتمكن من حساب ارتباطه مع المتغيرين الرتبيين: التحصل الدراسي للأب (X_4) والتحصيل الدراسي للام (X_5). ويشير الجدول (5) إلى قيم عناصر المصفوفة المقترحة.

الجدول 4: أنواع معاملات الارتباط الثنائية لاستحداث مصفوفة الارتباط المقترحة.

	X_1	X_2	X_3	X_4	X_5
X_2	phi				
X_3	Phi	Tetrachoric			
X_4	rank biserial	rank biserial	rank biserial		
X_5	rank biserial	rank biserial	rank biserial	Spearman	
X_6	point biserial	biserial	biserial	Spearman	Spearman

الجدول 5: مصفوفة الارتباط المقترحة بين أزواج ستة متغيرات حسب أنواع مقاييسها.

	X_1	X_2	X_3	X_4	X_5
X_2	0.1181				
X_3	0.0795	0.9165**			
X_4	0.0063	-0.0063	0.0154		
X_5	-0.0119	0.0005	-0.0113	0.4470**	
X_6	-0.1567	0.2199	0.5328**	-0.0280	0.1520

** ذو دلالة إحصائية عند $\alpha = 0.01$.

نتج عن تطبيق التحليل ألعاملي على مصفوفة الارتباط المقترحة ستة عوامل اختير منها ثلاثة عوامل فسرت 97.20% من التباين الكلي (الجدول 6).

الجدول 6: نتائج التحليل ألعاملي لمصفوفة الارتباط المقترحة.

Variable	Factor Loading			Communality
	Factor-1	Factor-2	Factor-3	
X_1	0.0740	-0.1200	-0.8710	0.7790
X_2	0.8960	-0.1090	-0.2280	0.8660
X_3	0.9870	-0.0770	-0.0420	0.9810
X_4	0.0340	0.8240	-0.2090	0.7240
X_5	0.0830	0.8540	-0.0340	0.7370
X_6	0.6170	0.1330	0.5170	0.6660
Variance (λ_j)	2.1710	1.4575	1.1257	4.7542
% of Var.	0.3620	0.2430	0.1880	0.7920

بتطبيق المعادلة (5) على تباينات العوامل في الجدول (6) تم حساب القيم الحرجة للعوامل الثلاثة المختارة ومقدارها على التوالي: 0.3393 و0.4142 و0.4713، وبمقارنتها مع قيم تحميلات العامل المنسوبة إليه تم تعين المتغيرات ذات الدلالة الإحصائية في العامل المعني، فكانت النتائج الآتية: ثلاثة متغيرات لها التأثير الرئيس على الدراسة الجامعية وهي: القناعة

بدخول الكلية (X_2) والقناعة بالانتساب إلى القسم (X_3) ومجموع درجات الامتحان العام (X_6)، من ثم يأتي تأثير التحصيل الدراسي للأب (X_4) والأم (X_5)، وأخيراً تأثير الجنس (X_1).

إن النتائج في الجدول (6) اختلفت عما في الجدول (3) وهذا يوصل إلى اختلاف في تفسير الظاهرة واتخاذ القرارات، وقد درجت الاختلافات في النقاط الآتية:

1- زادت نسبة ما تفسره العوامل المختارة من التباين الكلي، فقد أصبحت 79.2% بعد أن كانت 73.40% في الجدول (3)، فضلاً عن اختلاف نسبة ما يفسره كل عامل.

2- إن الاختلاف بين قيم تحميل العوامل المتناظرة أدى إلى اختلاف في الدلالة الإحصائية للمتغيرات، فقد برزت أهمية مجموع درجات الامتحان العام (X_6) في العامل الأول لم يكن للمتغير في هذا العامل دلاله إحصائية في الجدول (3). إن ظهور تأثير مجموع درجات الامتحان العام (X_6) في العاملين الأول والثالث أعطى أهمية ايجابية لهذا المتغير في تأثيره على الدراسة الجامعية، وهذه النتيجة تعزز من ملائمة تطبيق التحليل ألعاملي على مصفوفة الارتباط المقترحة، في حين ظهر المتغير X_6 في العامل الثالث وبتأثير سلبي عند استخدام مصفوفة ارتباط بيرسون.

3- اختلفت أهمية المتغيرات بصورة عامة من خلال وجودها في العوامل المختارة، فحسـب قيم كميات الشـيوع (communality) أصبحت الأهمية التنازلية للمتغيرات (الجدول 6) كالآتي: X_3 يليه X_4, X_5, X_1, X_2 وأخيراً X_6. وقد كانت عند استخدام مصفوفة ارتباط بيرسون (الجدول 3) كالآتي: X_3 يليه X_2، X_4، X_5، X_6، وأخيراً X_1.

سلبيات المصفوفة المقترحة، ومقترح لمعالجتها:

لاختلاف طرائق حساب قيم عناصر مصفوفة الارتباط المقترحة فإن ذلك قـد يؤدي إلى أن تكون المصفوفة من النوع غير المحدد indefinite matrix، إن هذا النوع من المصفوفات يكون له جذور مميزة موجبة وسالبة وقد يشارك الصفر في ذلك، وهذا لا ينطبق مع المفهوم الإحصائي للمكونات الرئيسة ومن ثم التحليل ألعاملي، إذ تعد الجذور المميزة فيهما تباينات، فهذا النوع من المصفوفات لا يمكن تطبيق التحليل ألعاملي عليها (Johnson, 1998) (الجبوري وعبد، 2000).

لتوضيح هذه الحالة نضيف إلى المتغيرات الستة المدروسة متغير آخر هو الدخل السنوي للأسرة (X_7). حول مقياس المتغير المضاف (X_7) من النسبي إلى الرتبي (الملحق 2) حتى يكون بالمستطاع حساب ارتباطه مع المتغيرين الرتبيين: التحصل الدراسي للأب (X_4) والتحصيل الدراسي للام (X_5). ويبين الجدول (7) أنواع الارتباطات الثنائية بين المتغير المضاف مع المتغيرات الستة السابقة.

الجدول 7: أنواع الارتباطات الثنائية بين المتغير المضاف مع المتغيرات الستة.

	X_1	X_2	X_3	X_4	X_5	X_6
X_7	Point biserial	Biserial	Biserial	Spearman	Spearman	Pearson

تضاف قيم معامل الارتباط في الجدول (7) إلى مصفوفة الارتباط المقترحة في الجدول (5) لتكون الصف السابع والعمود السابع، وبذلك تنشأ مصفوفة ارتباط مقترحة (الجدول 8).

الجدول 8: مصفوفة الارتباط المقترحة بين السبعة متغيرات حسب أنواع مقاييسها.

	X_1	X_2	X_3	X_4	X_5	X_6
X_2	0.1181					
X_3	0.0795	0.9165**				
X_4	0.0063	-0.0063	0.0154			
X_5	-0.0119	0.0005	-0.0113	0.4470**		
X_6	-0.1567	0.2199	0.5328**	-0.0280	0.1520	
X_7	0.0867	0.0947	0.5709**	0.3560**	0.3790**	0.0980

** ذو دلالة إحصائية عند $\alpha = 0.01$.

فشلت محاولة تطبيق التحليل ألعاملي على مصفوفة الارتباط المقترحة (الجدول 8)، فقد تبين أن المصفوفة هي من النوع غير المحدد فقد كان أحد جذورها المميزة سالباً (الجدول 9).

الجدول 9: الجذور المميزة للمصفوفة المقترحة ذات السبعة متغيرات.

2.4093	1.6621	1.1475	0.7275	0.6708	0.5001	- 0.117380

الكشف عن المتغيرات المولدة للمصفوفة غير المحددة:

يتم الكشف عن المتغيرات المولدة للمصفوفة غير المحددة بإتباع الخطوات المقترحة الآتية:

أ- نشرك المتغيرات كافة في تكوين مصفوفة الارتباط المقترحة، كما في الجدول (8).

ب- تنشأ مصفوفات جزئية من مصفوفة الارتباط المقترحة ويتم بحذف صف وعمـود متناظرين في كل مرة وبالتعاقب، وبذلك يستبعد أحد المتغيرات من كل مصفوفة جزئية.

ت- عن طريق البرنامج الحاسوبي الجاهز Minitab بأي من إصداراته، تحليل كل مصفوفة جزئية ناتجة من الفقرة (ب) بأتباع المسار:

Calc > Matrices > Eigen Analysis > Analyze matrix ☐ ;

Column of eigen values ☐

يعطي التحليل قيم الجذور المميزة، قد تكون قيم بعضها مساوية للصفر، إذا كانت علامات الجذور المميزة موجبة دل على أن المتغير المستبعد هو المولد للمصفوفة غير المحددة، واستبعاده جعل مصفوفة الارتباط المقترحة أن تكون من النوع أكيد الايجابية أو شبة أكيدة الايجابية. أما إذ ظهرت علامة سالبة لنفس الجذر المميز و/أو عند غيره فإن ذلك يدل على أمرين هما: الأول: وجود متغير غير المتغير المستبعد هو المسبب للجذر المميز السالب، في هذه الحالة يمكن التعرف على المتغير المقصود بالاستمرار في تطبيق الفقرتين (ب) و(ت) على التوالي. الأمر الثاني: وجود متغير آخر أو أكثر مسببين للجذر المميز السالب، في هذه الحالة يستبعد توافيق أزواج المتغيرات. في كلا الأمرين يجب الانتهاء من الفقرتين (ب) و(ت).

طبقت الفقرتين (ب) و(ت) على المصفوفة المقترحة المشار إليها في الجدول (8)، ويشير الجدول (10) إلى نتائج التحليل الذي يؤكد أن وجود الثلاثة متغيرات X_2 وX_3 وX_7 سوية في المصفوفة المقترحة تجعل المصفوفة أن تكون من النوع غير المحدد، وأن استبعاد إحداها يعمل على جعل المصفوفة من النوع أكيد الايجابية positive definite matrix.

الجدول 10: الجذور المميزة لستة متغيرات متبقية بعد استبعاد متغير منفرد.

المتغير المستبعد						
X_1	X_2	X_3	X_4	X_5	X_6	X_7
2.4043	2.0497	1.8213	2.3756	2.3629	2.2194	2.1710
1.6618	1.4345	1.2176	1.2784	1.2935	1.6434	1.4575
0.8682	1.1091	1.1287	1.1433	1.1158	0.9803	1.1257
0.6717	0.7126	0.7066	0.7108	0.7016	0.6709	0.7260
0.5114	0.5301	0.6265	0.6074	0.6248	0.5521	0.5006
- 0.1174	0.1640	0.4995	- 0.1154	- 0.0986	- 0.0661	0.0193

طرائق مقترحة لاستنباط مصفوفة ارتباط مقترحة أكيدة الايجابية أو شبة أكيدة الايجابية:

بعد تحديد المتغيرات المسببة للمصفوفة غير المحددة، اقترحت طرائق لاستنباط مصفوفة ارتباط أكيدة الايجابية أو شبة أكيدة الايجابية، وهي:

1- يستبعد متغير حسب رأي متخصص علمي بالظاهرة المدروسة بعد أن يعين إحصائي المتغيرات التي يمكن أن يستبعد منها. فقد لاحظنا في الجدول (10) يمكن استبعاد أحد المتغيرات الآتية: X_2 وX_3 وX_7.

2- بما أن التحليل ألعاملي هو الأسلوب المستخدم في تفسير الظاهرة، عليه يستبعد المتغير الذي تعطي المتغيرات الأخرى بدونه أعلى نسبة تفسير من التباين الكلي. يتبين من الجدول (11) أن المتغيرات X_1 إلى X_6 أعطت أعلى نسبة تفسير مقدارها 79.2%، عليه يفضل استبعاد المتغير X_7 من الدراسة إذا لم يكن للباحث رأي علمي يختار بموجبة متغير آخر كأن يكون X_2 أو X_3. وبذلك نحصل على مصفوفة ارتباط مقترحة يمكن تطبيق التحليل ألعاملي عليها كما في الجدول (5) والنتائج في الجدول (6).

الجدول 11: نسبة التباين الذي تفسره ثلاثة عوامل مختارة ناتجة من تحليل ستة متغيرات.

المتغير المستبعد	الجذور المميزة			نسبة التباين المفسر
X_2	2.0497	1.4345	1.1091	76.5550
X_3	1.7937	1.2385	1.1353	69.4583
X_7	2.1710	1.4575	1.1257	79.2367

3- هنالك مقترحات خارج نطاق التحليل ألعاملي يمكن بموجبها تحقيق مصفوفة أكيدة الايجابية أو شبة أكيدة الايجابية بالاعتماد على نتائج تحليل الانحدار وعلى أسلوب تحويل مقاييس المتغيرات، وهي:

أ- اختزال متغيرات: تختار المتغيرات المراد اختزالها وفق رأي اختصاصي في الظاهرة المدروسة، ويعين إحداها كمتغير معتمد (dependent variable) في أنموذج انحدار، والبقية تعامل كمتغيرات توضيحية في الأنموذج (explanatory variables)، ومن ثم توضع قيم متغير الاستجابة (response) الناتجة من الأنموذج بدلاً من المتغيرات المختزلة (كاظم ومسلم، 2002).

ب- حذف متغير أو عدد من المتغيرات ذات الأهمية الضعيفة (إبراهيم وآخرون، 2002)، ويتم ذلك بتطبيق اختبار الفرضيات الجزئية في أنموذج الانحدار للتعرف على أهمية استبعاد كل متغير من الأنموذج. يمكن إجراء اختبار الفرضية الجزئية على مجموعة من المتغيرات في حاله الحاجة إلى حذف أكثر من متغير واحد. وتتخذ القرارات المناسبة وفق نتائج اختبار الفرضيات وملاحظة قيمة معامل التحديد (R^2) ومتوسط مربعات الخطأ (MSe). يمكن الاستعانة بإحدى طرائق تحديد متغيرات الأنموذج كطريقة الانحدار التدريجية (stepwise procedure) (كاظم ومسلم، 2002).

ت- تحويل مقياس بعض المتغيرات إلى نوع آخر من المقاييس، كتحويل المقياس النسبي إلى رتبي والمقياس الاسمي الطبيعي إلى مقياس أسمي محول، وبذلك تتغير قيم عناصر مصفوفة الارتباط المقترحة (عودة والخليلي، 1988). إن هذا المقترح يحتاج إلى عمليات تحليلية طويلة وحسابات إحصائية معقدة خاصة وأن البرامج الحاسوبية الإحصائية الجاهزة تفتقد إلى بعض هذه التحاليل. وهنالك احتمالية من عدم جدوى الطريقة لتحقيق الغاية من التحويل.

الاستنتاجات والتوصيات:

1) نتيجة لتطبيق التحليل ألعاملي أعطت مصفوفة الارتباط المقترحة نتائج اختلفت عما أعطته مصفوفة ارتباط بيرسون. مهما ما أعطت مصفوفة ارتباط بيرسون من نتائج منطقية ومقبولة للباحث، فمن المفروض أن يستغني الباحث عنها لعدم تمثيل المصفوفة للارتباط الحقيقي بين المتغيرات. بينما مصفوفة الارتباط المقترحة تمثل الارتباط الثنائي الحقيقي بين المتغيرات، وعند إعطاء التحليل ألعاملي المطبق عليها نتائج غير منطقية وغير مرضية للباحث فإن الخلل ليس في استخدامها بل يكمن عند مسبب آخر.

2) اختلاف طرائق حساب عناصر مصفوفة الارتباط المقترحة قد يؤدي إلى أن تكون المصفوفة من النوع غير المحدد indefinite matrix والتي لا يمكن تطبيق التحليل ألعاملي عليها. يمكن تعيين المتغيرات المسببة لهذا النوع من المصفوفات بتجزئة مصفوفة الارتباط المقترحة وملاحظة علامات الجذور المميزة.

3) يمكن استنباط مصفوفة من النوع أكيد الايجابية positive definite matrix أو شبه أكيد الايجابية Positive semi-definite matrix باستبعاد متغير أو أكثر، وقد يستبعد المتغير الذي تعطي المتغيرات الأخرى بدونه أعلى نسبة تفسير من التباين الكلي، ويعتمد في ذلك على رأي متخصص في الظاهرة المدروسة.

المصادر العربية

1) إبراهيم، بسام يونس وحاجي، أنمار أمين ويونس، عـادل موسـى، 2002، "الاقتـصاد القياسي"، دار عزه للنشر والتوزيع، الخرطوم، جمهورية السودان.

2) البلداوي، عبد الحميد عبد المجيد، 2004، "الأساليب الإحصائية التطبيقيـة"، دار الـشرق الأوسط للنشر والتوزيع، المملكة الأردنية الهاشمية.

3) الجبوري، شلال حبيب وعبد، صلاح حمزة، 2000، "تحليل متعدد المتغيرات"، دار الكتب للطباعة والنشر، جامعة بغداد، العراق.

4) دبدوب، مروان عبد العزيز، 1996، "استخدام مصفوفتي الأعمدة والصفوف في تحليـل المكونات الرئيسية"، مجلة تنمية الرافدين [18] [49]، (185-192)، جامعـة الموصـل، العراق.

5) دبدوب، مروان عبد العزيز والشكرجي، ذنون يونس ذنون، 2007، "الانتقال المختـصر بين نتائج R-mode و Q-mode فـي التحليـل ألعـاملي"، المجلـة العراقيـة للعلـوم الإحصائية، [7] [11]، (153-174)، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق.

6) الشكرجي، ذنون يونس، 2005، "استخدام مصفوفتي R-mode و Q-mode في التحليل ألعاملي"، رسالة ماجستير، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق.

7) عودة، أحمد سليمان والخليلي، خليل يوسف، 1988، "الإحصاء للباحث والعلوم الإنسانية"، دار الفكر للنشر والتوزيع، عمان، المملكة الأردنية الهاشمية.

8) كاظم، أموري هادي و مسـلم، باسـم شـليبة، 2002، "القياسي الاقتـصادي المتقـدم: النظرية والتطبيق"، مطبعة الطيف، بغداد، العراق.

المصادر الأجنبية

9) Afifi, A. and Clark, V.; 1984; "Computer Aided Multivariate Analysis"; Lifetime Learning Publications; U.S.A..

10) Agresti, A. and Franklin, C.; 2007; "Statistics The Art And Science Of Learning From Data"; Pearson Prentice-Hall; U.S.A..

11) Berenson, M. L.; Levine, D. M. and Krehbiel, T. C.; 2006; "Basic Business Statistics Concepts and Applications"; 10th Edition; Pearson Prentice Hall; U.SA..

12) Brace, N.; Kemp, R. and Snelgar, R.; 2006; "SPSS For Psychologists"; 3rd Edition; Palgrave Macmillan; U.S.A..

13) Carver, R.; 1999; "Doing Data Analysis With Minitab"; Duxbury Press, U.S.A..

14) Johnson, D. E.; 1998; "Applied Multivariate Methods for Data Analysis"; Duxbury Press; U.S.A..

15) Kline, P.; 1994; "An Easy Guide To Factor Analysis"; TJ Press (Padstow) Ltd, Cornwall; G. B..

16) Kuo, B. and Golnaraghi F.; 2003; "Automatic Control Systems"; 8th Edition; John Wiley and Sons; U.S.A..

17) Siegel; S. and Castellan, N.; 1988; "Nonparametric Statistics for the Behavioral Sciences"; 2nd Edition; McGraw-Hill Book Company; U.S.A..

18) Sullivan, M.; 2007; "Statistics Informed Decisions Using Data"; 2nd Edition; Pearson Prentice-Hall; U.S.A..

الملحق 1: استمارة استبيان للحصول على متغيرات مختلفة المقاييس.

استمارة استبيان مقدمة إلى طلبة المرحلة الرابعة

قسم الإحصاء والمعلوماتية/ كلية علوم الحاسبات والرياضيات/ جامعة الموصل

أبنائي الطلبة الأعزاء، السلام عليكم ورحمة الله وبركاته.

إن استمارة الاستبيان هذه وسيلة تمكننا من الحصول على بيانات سيتم استخدامها في بحـث إحصائي يهتم في أسلوب تحليلي مقترح، ليس لهذه الاستمارة أية استخدامات أو أهداف أخرى.

للمحافظة على سرية انتماء المعلومات يرجى عدم إدراج أية معلومات غير مطلوبة.

نشكر لكم تعاونكم الصادق معنا.

1- الجنس:	أنثى	☐	ذكر	☐
2- القناعة بالانتساب إلى الكلية:	غير مقتنع	☐	مقتنع	☐
3- القناعة بالانتساب إلى القسم:	غير مقتنع	☐	مقتنع	☐
4- التحصيـل العلمـي للأب:	أمـي	☐	ابتدائية	☐
	متوسطة	☐	إعدادية	☐
	معهد أو كلية	☐	عليـا	☐
5- التحصيـل العلمـي للأم:	أمـي	☐	ابتدائية	☐
	متوسطة	☐	إعدادية	☐
	معهد أو كلية	☐	عليـا	☐
6- مجموع درجات الامتحان العام للسادس الإعدادي العلمي:				
7- الدخل السنوي للأسـرة (بالدينار العراقي):				

الملحق 2: بيانات سبعة متغيرات تم الحصول عليها من 55 طالباً، آخر عمودين عبارة عن القيم الرتبيه الناتجة من تحويل المتغيرين X6 وX7 إلى متغيرين رتبيين.

التسلسل	X_7	X_6	X_1	X_2	X_3	X_5	X_4	X_6	X_7
1	2	2	1	1	1	3	1	550	350
2	2	4	1	1	1	2	1	548	500
3	3	2	0	1	1	5	3	568	300
4	3	3	0	1	1	5	3	579	400
5	3	2	0	1	1	3	4	584	300
6	3	3	1	1	0	2	1	576	400
7	3	2	1	1	1	1	1	577	350
8	3	5	1	1	1	6	5	578	600
9	3	2	1	1	1	5	5	581	300
10	3	4	1	1	1	5	3	582	550
11	6	4	0	1	1	5	5	664	500
12	6	4	0	1	1	3	1	687	500
13	3	4	0	1	1	5	2	584	500
14	3	5	0	1	1	5	3	575	600
15	3	5	1	1	1	5	5	588	600
16	3	1	0	1	1	4	4	584	200
17	3	5	0	1	1	5	5	582	600
18	3	4	0	1	1	5	5	576	500
19	3	4	0	1	1	5	1	576	500
20	3	3	0	1	1	5	5	589	400
21	3	6	0	1	1	2	2	575	800
22	3	4	0	1	1	2	3	574	500
23	3	2	0	1	0	2	3	575	300
24	3	4	0	1	0	4	4	577	500
25	3	5	0	0	0	6	6	588	675
26	1	2	0	0	0	5	3	508	350
27	3	3	0	0	0	3	3	581	400

4	5	0	1	1	4	5	605	600	28
3	5	0	1	1	5	5	595	650	29
3	3	0	1	1	5	2	588	400	30
3	2	0	0	0	3	2	583	350	31
3	1	1	1	1	2	5	573	250	32
3	6	1	1	1	5	5	579	750	33
3	4	1	1	0	5	4	579	500	34
3	6	1	1	1	6	4	575	800	35
3	1	1	1	1	5	2	573	160	36
3	3	1	1	1	5	3	573	400	37
3	4	1	1	1	5	2	572	500	38
3	5	1	1	1	4	2	572	600	39
3	6	1	0	0	5	5	586	700	40
3	6	1	1	1	5	3	576	700	41
3	4	1	1	1	5	3	583	550	42
3	5	1	1	1	5	4	570	650	43
3	5	1	1	1	3	2	574	600	44
3	4	0	1	1	5	5	574	550	45
3	3	0	1	1	5	2	574	400	46
2	2	0	1	1	5	3	567	300	47
2	2	0	1	1	5	3	564	300	48
2	5	0	1	0	5	5	539	600	49
3	4	0	1	1	5	2	588	500	50
3	6	0	1	1	5	5	574	750	51
3	2	0	1	1	5	1	576	350	52
3	6	0	1	1	6	5	574	800	53
3	6	0	1	1	5	5	581	750	54
3	3	0	1	1	3	3	584	400	55