

مقارنة الأساليب المستخدمة في تحديد عدد المركبات الرئيسية

مع التطبيق العملي

زينة ياوز عبد القادر

أ.م.د. دجلة إبراهيم العزاوي

المستخلص:

استخدم تحليل المركبات الرئيسية في تحليل الكثير من الظواهر الاجتماعية و الاقتصادية، ومن هذه الظواهر اخترنا ظاهرة تخص شريحة واسعة من مجتمعنا ألا وهي ظاهرة تأخر حصول التدريسي الجامعي على اللقب العلمي التالي، ولأن تحديد عدد المركبات الرئيسية الداخلة في تحليل المركبات الرئيسية يعتمد على استخدام عدة أساليب، قمنا بالمقارنة بين ثلاثة من هذه الأساليب والتي هي

(BARTLETT, SCREE DIAGRAM, JOLLIFFE)

واستنتجنا أن أسلوب JOLLIFFE هو الأسلوب الأفضل في تحديد عدد المركبات الرئيسية الداخلة في تحليل بيانات الظاهرة المدروسة من بين الثلاثة أساليب. وبالاعتماد على النتائج المستخلصة لهذا الأسلوب نجد أن أبرز العوامل المؤثرة في الظاهرة هي (نصاب الأستاذ الجامعي، الروتين الإداري، أسعار لوازم القرطاسية، الجوائز المادية، النشاطات الترفيهية، الوضع العائلي، نقص الكادر الأكاديمي، عدد الصفحات، استخدام الانترنت، الاتجاه السياسي والمحسوبة، عدد السنوات) ولذلك يجب أن نبدأ بمعالجة هذه العوامل كخطوة أولى لتحفيز التدريسي الجامعي على تطوير قدراته الفكرية نحو الأفضل.

ABSTRACT:

The principal components analysis is used in analyzing many economic and social phenomena; and one of them is related to a large group in our society who are the university instructors. This phenomenon is the delay occurred in getting university instructor to his next scientific title. And as the determination of the principal components number inside the principal components depends on using many methods, we have compared between three of these methods that are: (BARTLETT, SCREE DIAGRAM, JOLLIFFE).

We concluded that JOLLIFFE method was the best one in analyzing the studying phenomenon data among these three methods, we found the most distinguishing factors effecting on the phenomenon was (the number of the lectures the university instructor has, the administrative routine, the entertainment activities, family situation, shortage in academic staff, pages number, the use of the internet and the political attitude).

So, we should begin with treating these factors as a first step for motivating university instructor to develop his thinking capabilities.

الجانب النظري

المقدمة

يعد تحليل المركبات الرئيسية احد اقدم واوسع الأساليب المستخدمة في تحليل البيانات متعددة المتغيرات^[15]، حيث يستخدم في تحليل البيانات ذات العدد الكبير للمتغيرات التوضيحية نسبة إلى عدد المشاهدات، كما نلجأ إلى تحليل المركبات الرئيسية عندما تكون المتغيرات التوضيحية ذات ارتباط عالي (اي وجود مشكلة التعدد الخطي)^[7]، بالإضافة إلى إمكانية استخدام تحليل المركبات الرئيسية لتحديد الإسقاط الملائم للبيانات، والقيم الشاذة للبيانات^[2].

تحليل المركبات الرئيسية

ان تحليل المركبات الرئيسية يستند إلى ايجاد الجذور المميزة والمتجهات المميزة لمصفوفة التباين والتباين المشترك للمتغيرات التوضيحية، او ايجاد الجذور المميزة والمتجهات المميزة لمصفوفة الارتباطات، وهذا يعتمد على طبيعة البيانات فأذا كانت وحدات القياس متشابهة نستخدم مصفوفة التباينات او مصفوفة الارتباطات بين المتغيرات التوضيحية، اما اذا كانت وحدات القياس مختلفة فيمكننا استخدام مصفوفة الارتباطات^[4]، ولنفرض مصفوفة البيانات هي المصفوفة \times حيث ان:

$$\dots(1) \mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \dots & \mathbf{X}_{1p} \\ \mathbf{X}_{21} & \mathbf{X}_{22} & \dots & \mathbf{X}_{2p} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \mathbf{X}_{n1} & \mathbf{X}_{n2} & \dots & \mathbf{X}_{np} \end{bmatrix}$$

حيث ان:

n تمثل حجم العينة.

p تمثل عدد المتغيرات التوضيحية.

المركبة الرئيسية الأولى M_1 يمكن التعبير عنها بالصيغة

$$\mathbf{M}_1 = \mathbf{a}_{11}\mathbf{X}_1 + \mathbf{a}_{12}\mathbf{X}_2 + \dots + \mathbf{a}_{1p}\mathbf{X}_p \quad \dots(2)$$

تمتلك هذه المركبة اكبر نسبة تباين بالنسبة إلى المركبات الرئيسية، وان $\mathbf{a}_{11}, \mathbf{a}_{12}, \dots, \mathbf{a}_{1p}$ تمثل معاملات المركبة الرئيسية الأولى والتي يمكن التعبير عنها بالمتجه \mathbf{a}_1 والذي يمثل المتجه الذاتي الأول المقابل للجذر المميز الأول λ_1 يحقق الشرط التالي

$$\mathbf{a}'_1 \mathbf{a}_1 = 1 \quad \dots(3)$$

اما الجذر المميز الأول فيمكن الحصول عليه بحل إحدى المعادلتين

$$|\mathbf{S} - \lambda_1 \mathbf{I}| = 0 \quad \dots(4)$$

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdot & \cdot & \cdot & S_{1p} \\ S_{21} & S_{22} & \cdot & \cdot & \cdot & S_{2p} \\ & & \cdot & & & \cdot \\ & & & \cdot & & \cdot \\ & & & & \cdot & \cdot \\ & & & & & S_{pp} \end{bmatrix}$$

S: تمثل مصفوفة التباين و التباين المشترك.

$$|R - \lambda_1 I| = 0 \quad \dots(6)$$

R: تمثل مصفوفة الارتباطات .

$$R = \begin{bmatrix} 1 & r_{12} & \cdot & \cdot & \cdot & r_{1p} \\ r_{21} & 1 & \cdot & \cdot & \cdot & r_{2p} \\ & & \cdot & & & \cdot \\ & & & \cdot & & \cdot \\ & & & & \cdot & \cdot \\ & & & & & 1 \end{bmatrix}$$

اما المركبة الرئيسية الثانية

$$M_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

تمتلك ثاني اكبر نسبة تباين وتحقق الشرطين:

$$a'_2 a_2 = 1 \quad \dots(6)$$

$$a'_2 a_i = 0 \quad \dots(7)$$

وهكذا يمكن ايجاد P من المركبات الرئيسية حيث ان عدد المركبات الرئيسية مساوي إلى عدد المتغيرات التوضيحية الداخلة في الظاهرة قيد البحث، وبصورة عامة فأن

$$a'_j a_j = 1 \quad \dots(8)$$

$$a'_j a_i = 0 \quad (i < j) \quad \dots(9)$$

اما اهم القوانين التي استخدمت في هذا البحث فهي:

١- تباين المركبة الرئيسية j

$$\text{var}(M_j) = \lambda_j \quad \dots(10)$$

٢- نسبة التباين التي تفسرها المركبة الرئيسية إلى التباين الكلي لجميع المركبات الرئيسية

$$P. \text{var}(M_j) = \frac{\lambda_j}{\text{trace}(S)} \times 100\% \quad \dots(11)$$

٣- التباين المشترك بين المتغيرات التوضيحية والمركبة الرئيسية

$$\text{Cov}(X, M_j) = \lambda_j a_j \quad \dots(12)$$

٤- الارتباط بين المتغيرات التوضيحية والمركبات الرئيسية

$$r_{X_i, M_j} = \frac{\sqrt{\lambda_j} a_{ji}}{S_{ii}^{1/2}} \quad \dots(13)$$

٥- يمكن تحديد محددة وأثر مصفوفة التباين والتباين المشترك من خلال [7]

$$\sum_{j=1}^p V(M_j) = \text{tr}(S) \quad \dots(14)$$

$$\prod_{j=1}^p V(M_j) = |S| \quad \dots(15)$$

٦- يمكن تحديد الإسقاط الملائم للبيانات V والذي يكون مساوي إلى المتجه الذاتي الأول a'_1 [15]

اخيار عدد المركبات الرئيسية،

CHOOSING THE NUMBER OF PRINCIPLE COMPONENTS

هناك عدة أساليب تستخدم في تحديد عدد المركبات الرئيسية الداخلة في تحليل المركبات الرئيسية، تناولنا في هذا البحث ثلاثة من هذه الأساليب وقمنا بالمقارنة بينها وهذه الأساليب هي

١. أسلوب JOLLIFFE

اقترح هذا الأسلوب JOLLIFFE عام (١٩٧٢) أسلوب يعتمد على ايجاد الجذور المميزة لمصفوفة الارتباطات بين المتغيرات التوضيحية، ومن ثم نلاحظ الجذور المميزة التي تزيد قيمتها عن ٠.٧، المركبات الرئيسية المقابلة لتلك الجذور تبقى في التحليل، أما المركبات الرئيسية المقابلة للجذور المميزة الأخرى فيتم إهمالها من التحليل وبذلك يكون عدد المركبات الرئيسية مساوي إلى عدد الجذور المميزة الداخلة في التحليل [7].

٢. أسلوب SCREE Diagram

اقترح هذا الأسلوب من قبل CATTELL عام (١٩٦٥) يعتمد هذا الأسلوب على الرسم البياني من خلال تعيين الجذور المميزة لمصفوفة الارتباطات على المحور العمودي مقابل رقم الجذر المميز على المحور الأفقي، الشكل الذي ينتج يشبه شكل الصخرة الجذور المميزة التي تقع على السطح العلوي تؤخذ مركباتها بنظر الاعتبار أما الجذور المميزة التي تقع أسفل الصخرة فيتم إهمال مركباتها وقد تم تطورة من قبل JOLLIFFE عام (١٩٨٦) حيث استخدم لوغارتم قيمة الجذر المميز مقابل رقم الجذر المميز (I) فيكون عدد المركبات الرئيسية مساوي إلى رقم الجذر المميز المناظر إلى الانحناء المرفقي للمنحنى (Corresponding to an elbow in curve) فالنقطة التي ينتهي عندها هذا الانحناء تتوقف عندها الجذور المميزة الكبيرة وتبدأ الجذور المميزة الصغيرة، والتي يتم استبعادها من التحليل [10].

٣. أسلوب BARTLETT

وضع BARTLETT عام (١٩٥١) صيغة اختبار بموجبها تساوي $P - K$ من الجذور المميزة المستخرجة من مصفوفة التباين والتباين المشترك وقد تم تعديل هذا الأسلوب من قبل K.V.MARDIA عام (١٩٧٩) و اقترح الصيغة التالية

$$\chi^2 = n'(p - k) \log \left(\frac{a_0}{g_0} \right)$$

$$n' = n - \left(\frac{2p + 11}{6} \right)$$

حيث ان :-

n : تمثل حجم العينة.

p : تمثل عدد المتغيرات التوضيحية .

K : رقم آخر جذر مميز يدخل إلى التحليل.

a_0 : الوسط الحسابي لجميع الجذور المميزة المستخرجة من مصفوفة التباين والتباين المشترك والمستبعدة من التحليل.

g_0 : الوسط الهندسي لجميع الجذور المميزة المستخرجة من مصفوفة التباين والتباين المشترك والمستبعدة من التحليل.

لاختبار الفرضية التالية

$$H_0 : \lambda_p = \lambda_{p-1} = \dots = \lambda_{k+1}$$

فإذا كانت قيم χ^2 المحسوبة اقل من χ^2 الجدولة ودرجة

حرية $\frac{1}{2}(p - k + 2)(p - k - 1)$ نقبل الفرضية H_0 أي إن المركبات الرئيسية المقابلة لتلك الجذور متساوية فيما تفسره من نسبة التباين الكلي لجميع المركبات الرئيسية ولذلك يتم إهمالها^[10].

الجانب التطبيقي

إن المجتمع الذي اعتمدهنا في هذا البحث هو مجتمع التدريسيين في الجامعة المستنصرية والعينه التي تم سحبها من هذا المجتمع هي عينه طبقية عشوائية ، قسم فيها التدريسيون على أربع طبقات (مدرس مساعد، مدرس، أستاذ مساعد، أستاذ)، اما عدد التدريسيين في جميع الكليات فقد بلغ 1497 وحجم العينه الطبقية الذي بلغ 150 تدريسياً من مختلف الاختصاصات في الجامعة المستنصرية.

واعتمدهنا في هذا البحث على تصميم استمارة استبيان يتم من خلالها حصر اغلب العوامل المؤثرة في الظاهرة قيد البحث باستخدام مجموعة من الأسئلة، كل سؤال من هذه الأسئلة تم وضعه وصياغته بعد التحاور مع مجموعة من التدريسيين لكي تظهر الاستمارة بشكل واضح من حيث المعنى والغرض.

تضمنت الاستمارة واحداً وعشرين متغيراً هي:-

Y	اللقب العلمي	.١
X ₁	المصادر	.٢
X ₂	عدد الصفحات	.٣
X ₃	الزمالات البحثية	.٤
X ₄	الواجبات الإدارية	.٥
X ₅	الوضع العائلي	.٦
X ₆	الوضع الصحي	.٧
X ₇	الأجور	.٨
X ₈	الانترنت	.٩
X ₉	نصاب الاستاذ	.١٠
X ₁₀	الاتجاه السياسي والمحسوبية	.١١
X ₁₁	قلة التخصص الدقيق	.١٢
X ₁₂	البنى التحتية	.١٣
X ₁₃	زيادة عدد الجامعات	.١٤
X ₁₄	نقص الكادر الأكاديمي	.١٥
X ₁₅	القيود على الطبع والنشر والترجمة	.١٦
X ₁₆	الجوائز المادية	.١٧
X ₁₇	عدد السنوات	.١٨
X ₁₈	توفير نشاطات ترفيهية	.١٩
X ₁₉	الروتين الإداري	.٢٠
X ₂₀	أسعار لوازيم القرطاسية	.٢١

وقد تم استخدام النظام الإحصائي (SPSS VERSION 10) في تحليل النتائج، وبتطبيق الاختبار χ^2 وجد أن توزيع العينة مطابق لتوزيع المجتمع، كما اختبرنا وجود مشكلة التعدد الخطي باستخدام اختبار دليل الحالة CONDITION INDEX وجدنا ان البيانات تعاني من مشكلة التعدد الخطي.

تطبيق أسلوب JOLLIFFE: الجذور المميزة لمصفوفة الارتباطات R, موضحة في

الجدول رقم (١)، الجذور المميزة التي أخذت بنظر الاعتبار هي تلك الجذور التي تزيد قيمتها على 0.7 وعددها 14 جذرا مميزا والمركبات الرئيسية المقابلة لتلك الجذور تفسر لنا 83.623 % من مجموع التباين الكلي لجميع المركبات الرئيسية، أي أن مجموع الجذور المميزة التي تم استبعادها من التحليل 6 جذور مميزة والتي تقل قيمتها عن 0.7.

جدول رقم (١)
قيم الجذور المميزة ونسب التباين لكل مركبه لأسلوب jolliffe

المركبات	الجذور المميزة	نسبة تباين لمركبه %	تجميع نسب التباين %	الجذور المميزة لأسلوب jolliffe	نسبة تباين المركبة %	تجميع نسب التباين %
1	1.953	9.763	9.763	1.953	9.763	9.763
2	1.792	8.958	18.721	1.792	8.958	18.721
3	1.659	8.294	27.015	1.659	8.294	27.015
4	1.471	7.356	34.371	1.471	7.356	34.371
5	1.266	6.332	40.703	1.266	6.332	40.703
6	1.201	6.006	46.709	1.201	6.006	46.709
7	1.127	5.637	52.346	1.127	5.637	52.346
8	1.053	5.267	57.613	1.053	5.267	57.613
9	.997	4.983	62.596	.997	4.983	62.596
10	.970	4.852	67.448	.970	4.852	67.448
11	.942	4.711	72.159	.942	4.711	72.159
12	.794	3.970	76.129	.794	3.970	76.129
13	.767	3.837	79.966	.767	3.837	79.966
14	.731	3.657	83.623	.731	3.657	83.623
15	.694	3.469	87.092			
16	.635	3.174	90.266			
17	.560	2.797	93.063			
18	.555	2.776	95.839			
19	.443	2.216	98.055			
20	.389	1.945	100.00			

أما مصفوفة التشبعات التي تتضمن التشبعات للمركبات الرئيسية الباقية في التحليل (المركبات 14 الأولى) حيث ان التشبع هو عبارة عن معامل الارتباط البسيط بين المركبة الرئيسية والمتغير التوضيحي، فقد تم تمثيلها في الجدول رقم (٢)، الأعمدة تمثل المركبات والصفوف تمثل المتغيرات.

جدول رقم (٢)

مصفوفة التشيعات

المركبات

	2	3	4	5	6	7	8	9	10	11	12	13	14
43	-.372	-.010	.244	-.315	.461	-.108	-.172	.283	-.103	-.020	.193	.215	.038
55	.096	-.206	-.030	-.014	.582	.398	-.111	.033	.238	.053	.223	-.243	-.030
78	.151	.412	.463	.287	-.047	-.256	-.182	.021	-.020	-.070	-.212	.158	.291
71	.308	-.171	-.043	.021	.096	-.345	-.099	-.108	.416	.126	.000	.347	.184
72	.432	.154	.586	-.205	.049	.087	-.125	.115	.111	.212	-.080	-.142	.002
07	.472	-.020	.456	-.194	-.127	.261	.266	-.011	.016	.335	.050	.022	.154
38	.152	-.497	.178	.282	-.213	.227	-.233	-.097	-.382	.057	.276	.341	-.090
24	-.119	-.149	-.253	-.047	-.506	-.103	.035	.506	.328	.261	.251	-.058	.144
63	-.372	.068	.081	-.096	.048	.070	.374	-.230	.009	.119	-.116	-.020	-.111
08	-.315	-.040	.053	-.118	-.120	.573	-.076	.175	.412	-.389	-.219	.254	.069
24	-.104	.005	-.406	.422	.174	.262	-.007	.118	-.174	.342	-.080	-.015	.370
65	.465	.054	-.230	.087	.274	-.258	.352	.057	.108	-.226	.028	-.041	.117
45	.422	.378	-.340	.023	.064	.219	.078	-.190	.149	.093	.000	.444	-.185
18	.211	-.201	.279	.605	.153	.098	.254	.240	-.010	-.286	-.050	-.037	.099
173	-.269	.233	.207	.471	-.179	.115	.069	-.183	.304	.195	.020	-.154	-.187
18	.141	.543	-.130	-.230	-.217	.286	.024	-.158	-.215	-.238	.224	-.100	.407
08	-.430	-.268	.201	-.111	.050	-.034	.612	-.176	.009	.035	.200	.194	.175
27	-.121	.632	-.004	.091	.165	.040	.186	.466	-.173	.192	.089	.197	-.249
24	-.313	.212	.005	.132	.221	-.027	-.268	-.349	.190	.199	.198	-.015	.178
32	-.050	-.254	-.162	-.187	.101	.101	.023	.029	-.159	.308	-.553	.037	.128

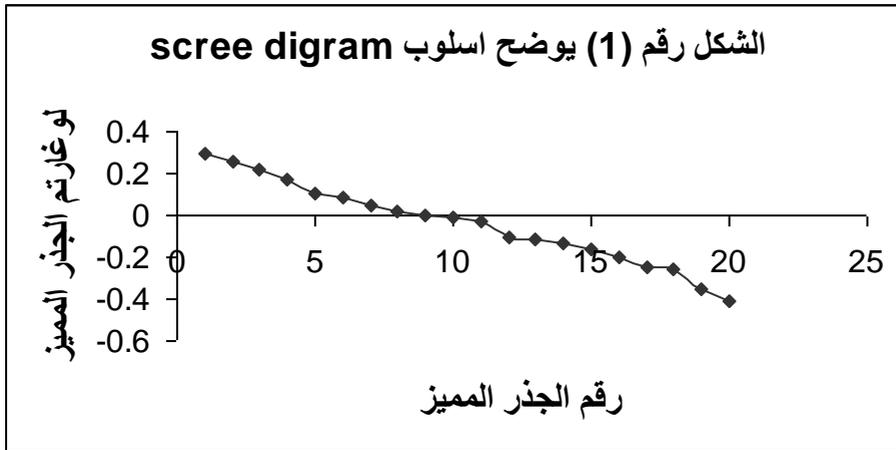
نلاحظ أن أقوى المتغيرات ارتباطا بالمركبة الرئيسية الأولى هي (نصاب الأستاذ الجامعي X_9 ، الروتين الإداري X_{19} ، أسعار لوازم القرطاسية X_{20})، أما المركبة الرئيسية الثانية فتظهر ارتباطات ضعيفة لجميع المتغيرات التوضيحية (معامل الارتباط اقل من 0.5 بغض النظر عن الإشارة).

في حين أن المتغيرين التوضيحين (الجوائز المادية X_{16} ، النشاطات الترفيهية X_{18}) يرتبطان ارتباطاً قوياً بالمركبة الرئيسية الثالثة، والمتغير التوضيحي (الوضع العائلي X_5) يرتبط ارتباطاً قوياً بالمركبة الرئيسية الرابعة، والمتغير التوضيحي (نقص الكادر الأكاديمي X_{14}) يرتبط ارتباطاً مؤثراً بالمركبة الرئيسية الخامسة، أما المركبة الرئيسية السادسة فترتبط بالمتغيرين التوضيحين (عدد الصفحات X_2 ، الانترنت X_8)، والمركبة الرئيسية السابعة ترتبط بالمتغير التوضيحي (الاتجاه السياسي والمحسوبية X_{10})، أما المركبة الرئيسية الثامنة فترتبط

ارتباطاً قوياً بالمتغير التوضيحي (عدد السنوات X_{17})، والمركبة الرئيسية التاسعة تظهر ارتباطاً مؤثراً مع المتغير التوضيحي (الانترنت X_8)، أما المركبتان الرئيستان العاشرة والحادية عشر فلهما ارتباطات ضعيفة مع جميع المتغيرات التوضيحية في حين أن المركبة الرئيسية الثانية عشر لها ارتباطاً مؤثراً مع المتغير (أسعار لوازم القرطاسية X_{20}).

وأخيراً فإن المركبتين الرئيسيتين الثالثة عشر والرابعة عشرة تمتلكان ارتباطات ضعيفة مع جميع المتغيرات التوضيحية.

تطبيق أسلوب **SCREE DIAGRAM**: من خلال تطبيق الأسلوب السابق تم الحصول على الجذور المميزة لمصفوفة الارتباطات والموضحة في الجدول رقم (١) نستخدم لوغاريتم الجذر المميز (المحور العمودي) في تمثيل الرسم البياني مقابل رقم الجذر المميز (المحور الأفقي)، كما في الشكل التالي.



نلاحظ أن المركبة الرئيسية رقم 12 تمثل نقطة انتهاء الإنحناء المرفقي للمنحنى، أي أن الجذر المميز رقم 12 يمثل نقطة توقف الجذور الكبيرة وبدء الجذور الصغيرة، ولذلك فإن عدد الجذور المميزة المستخرجة لهذا الأسلوب هو اثنا عشر جذراً مميزاً يتم إدخالها في تحليل المركبات الرئيسية في حين أن الثمانية جذور الأخيرة يتم إهمالها من التحليل. كما يمكن ملاحظة الجذور المميزة لأثنى عشر جذراً مميزاً والتي تم اختيارها وفق أسلوب SCREE DIAGRAM ونسب التباين المقابلة لمركباتها في الجدول رقم (1). بعد حذف الجذرين المميزين الثالث عشر والرابع عشر ومركبتهما. المركبات الرئيسية الأثنا عشرة تفسر لنا نسبة تباين مقدارها 76.129% من نسبة التباين الكلي لجميع المركبات الرئيسية. أما مصفوفة التشبعات التي تتضمن التشبعات للمركبات الرئيسية الباقية في التحليل (المركبات 12 الأولى) فهي نفس التشبعات التي تم الحصول عليها من أسلوب jolliffe بعد حذف العمودين الأخيرين من الجدول رقم (٢).

تطبيق أسلوب BARTLETT:

يعتمد أسلوب BARTLETT على إيجاد الجذور المميزة لمصفوفة التباين والتباين المشترك للمتغيرات التوضيحية، ومن ثم نقوم باختبار الفرضية

$$H_0 : \lambda_p = \lambda_{p-1} = \dots = \lambda_{k+1}$$

ولجميع قيم k وجدنا أن اقل قيمة لـ k تحقق الفرضية السابقة هي $k=10$ (أي أن عدد الجذور المميزة الباقية في التحليل عشرة جذور مميزة)، قيمة χ^2 المحسوبة مساوية إلى 66.222 بينما قيمة χ^2 الجدولة ودرجة حرية 54 وبمستوى معنوية 0.05 و0.01 هي على التوالي 73.293، 82.266 أي أن قيمة χ^2 المحسوبة اقل من قيمة χ^2 الجدولة ولذلك نقبل الفرضية $H_0 : \lambda_{20} = \dots = \lambda_{11}$ كل قيم k الأكبر من 10 تحقق الفرضية الصفرية، بينما قيم k الأقل من 10 تعطينا قيم χ^2 المحسوبة اكبر من χ^2 الجدولة مما يعني رفض الفرضية H_0 ،

أي أن هناك جذوراً مميزة تم إهمالها ويجب إعادتها إلى التحليل، والجدول رقم (٣) يوضح أسلوب BARTLETT والذي تضمن استخدام عشرة جذور مميزة فقط في التحليل من اصل 20 جذراً مميزاً ونسب التباين للمركبات الرئيسية المقابلة لتلك الجذور. فمن هذا الجدول يمكن ملاحظة أن نسبة التباين التي تفسرها المركبات الرئيسية العشرة الداخلة في التحليل هي 77.153 من مجموع التباين الكلي لجميع المركبات الرئيسية.

جدول رقم (٣)

قيم الجذور المميزة ونسب التباين لكل مركبة لأسلوب Bartlett

المركبات	الجذور المميزة	نسبة تباين المركبة %	تجميع نسبة تباين المركبة %	الجذور المميزة لأسلوب Bartlett	نسبة تباين المركبة %	تجميع نسبة التباين %
----------	----------------	----------------------	----------------------------	--------------------------------	----------------------	----------------------

1	.941	18.645	18.645	.941	18.645	18.645
2	.672	13.304	31.949	.672	13.304	31.949
3	.409	8.095	40.044	.409	8.095	40.044
4	.356	7.057	47.101	.356	7.057	47.101
5	.301	5.966	53.067	.301	5.966	53.067
6	.294	5.817	58.884	.294	5.817	58.884
7	.256	5.074	63.958	.256	5.074	63.958
8	.237	4.695	68.653	.237	4.695	68.653
9	.230	4.550	73.203	.230	4.550	73.203
10	.199	3.950	77.153	.199	3.950	77.153
11	.192	3.806	80.959			
12	.162	3.212	84.171			
13	.159	3.143	87.314			
14	.125	2.476	89.790			
15	.117	2.318	92.108			
16	.107	2.112	94.220			
17	.105	2.079	96.299			
18	.092	1.807	98.106			
19	.066	1.326	99.432			
20	.028	.568	100.00			

مصفوفة التشبعات التي تتضمن التشبعات للمركبات الرئيسية الباقية في التحليل (المركبات العشر الأولى)، فقد تم تمثيلها في الجدول رقم (4). نلاحظ أن أقوى متغيرين توضيحيين مرتبطين بالمركبتين الرئيسيتين الأولى والثانية هما (زيادة عدد الجامعات X_{13} ، عدد السنوات X_{17}) والمركبة الرئيسية الثالثة أظهرت ارتباطاً مع المتغير التوضيحي (القيود على الطبع والنشر والترجمة X_{15}).

في حين أن باقي المركبات الرئيسية (الرابعة، الخامسة، السادسة، السابعة، الثامنة، التاسعة، العاشرة) أظهرت ارتباطات ضعيفة لجميع المتغيرات التوضيحية

جدول رقم (٤)

مصفوفة التشتيعات لأسلوب bartlett

١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
-0.085	.029	.042	-0.034	-0.060	-0.019	.159	-0.095	-0.020	.120
.008	-0.085	-.186	-0.006	.273	-0.084	.130	-.100	.131	.111
-.008	-0.030	.017	.039	.028	.038	-.013	.002	-.023	.004
.026	-0.059	-.132	-0.035	.086	-0.017	-.108	.051	-.041	.059
.055	-0.024	.014	-0.063	.075	.186	.172	.202	-.090	.166
.030	.025	-.042	-0.051	.049	.120	.087	.206	-.036	.024
-.027	-0.026	.008	-0.083	.062	-0.041	.027	.129	-.023	-.206
-.029	-0.002	.091	-0.003	-.191	-.188	-.190	.221	.149	.142
-.103	-0.016	-.103	.158	.081	-0.021	-.050	-.051	-.038	.012
-.054	.019	.046	.154	.037	-.149	.187	.080	.282	.036
.007	-0.035	-.041	.045	.061	-0.081	-.075	-.094	.046	-.081
.147	.129	.012	-.319	-.075	.184	-.105	-.140	.178	.114
.673	.513	-.101	.049	.067	-.112	-.010	.040	-.041	-.009
-.010	.039	.102	-.215	.180	.079	-.006	.044	.182	-.169
-.031	.154	.500	.108	.197	.043	-.067	-.014	-.022	.035
.157	.051	-.021	.282	-.193	.282	.084	-.019	.148	-.101
-.641	.571	-.142	.015	.008	.041	-.019	.024	.001	.000
.054	.019	.013	.107	.017	.050	-.043	-.062	-.006	.086
.010	.097	.162	.011	-.013	-.031	.062	.134	-.048	.073
.002	.108	.111	-.186	-.189	-.172	.237	-.052	-.059	-.070

مقارنة نسب الاشتراكات للمتغيرات التوضيحية مع المركبات الرئيسية

ان القيمة المستخلصة لأشترافية اي متغير مع المركبات الرئيسية هي عبارة عن قيمة مربع الارتباط المتعدد بين المتغير التوضيحي وجميع المركبات الرئيسية الداخلة في التحليل وبتطبيق هذه القاعدة على جميع الأساليب السابقة وبملاحظة النتائج في الجدول رقم (٥) التالي

جدول رقم (٥) اشتراكيات المتغيرات التوضيحية في المركبات الرئيسية للأساليب الثلاثة			
المتغيرات	JOLLIFFE	SCREE DIAGRAM	BARTLETT
X1	.850	.797	.431
X2	.858	.798	.734
X3	.826	.717	.148
X4	.841	.687	.337
X5	.725	.705	.796
X6	.763	.739	.492
X7	.893	.769	.482
X8	.905	.881	.860
X9	.716	.703	.382
X10	.936	.866	.749
X11	.862	.724	.275
X12	.761	.745	.905
X13	.853	.622	.997
X14	.827	.816	.744
X15	.803	.744	.940
X16	.869	.693	.898
X17	.840	.772	.997
X18	.895	.794	.228
X19	.826	.794	.436
X20	.877	.859	.769

نجد ان اسلوب JOLLIFFE اعطى نسب عالية القيمة لجميع المتغيرات التوضيحية في حين ان الأسلوبين الآخرين اعطت نسباً اقل مما يدل على انه الأسلوب الأفضل لتحليل بيانات الظاهرة المدروسة.

تحليل الانحدار الخطي المعدل للمركبات الرئيسية:

بعد أن تم تحديد المركبات الرئيسية الداخلة في التحليل لكل اسلوب من الأساليب الموضحة سابقاً ، ندخل هذه المركبات كمتغيرات مستقلة في نموذج الانحدار العام ونستخرج جدول تحليل التباين لكل اسلوب من الأساليب الثلاثة. والجدول رقم (6) يشير إلى تحليل التباين للأساليب الثلاثة، وعند النظر إلى هذا الجدول نجد أن قيمة متوسط مربعات الخطأ لأسلوب JOLLIFFE هي 0.806 ولأسلوب SCREE DIAGRAM هي 0.812 أما قيمة متوسط مربعات الخطأ لأسلوب BARTLETT في 0.815 ، وهذه القيم على الرغم من إنها متقاربة فإنها تؤكد حقيقة ان زيادة عدد المركبات الرئيسية تؤدي إلى تقليل متوسط مربعات الخطأ . وبما أن MSE لأسلوب JOLLIFFE هو اقل من MSE للأسلوبين الآخرين، ولذلك يعد اسلوب JOLLIFFE هو الأسلوب الأفضل

جدول رقم (6)
جدول تحليل الانحدار للأساليب الثلاثة

الأسلوب	مصدر التباين S.O.V	d.f	s.s	m.s	F ₀	sig.
JOLLIFFE	Reg.(m1,..., m14)	14	32.766	2.336	2.899	.001
	Residual	135	108.794	.806		
	Total	149	141.500			
SCREE DIAGRAM	Reg.(m1,..., m12)	12	30.190	2.516	3.096	.001
	Residual	137	111.390	.812		
	Total	149	141.500			
BARTLET T	Reg.(M1,..., M10)		28.188	2.819	3.458	.004
	Residual		113.312	.815		
	Total		141.500			

الاستنتاجات

عند مقارنة نسب التباين التي تفسرها المركبات الرئيسية المختارة وفق كل اسلوب نجد أن اكبر نسبة للتباين (التباين الذي تفسره المركبات الرئيسية المختارة من مجموع التباين الكلي لجميع المركبات الرئيسية) قد أعطيت عند تطبيق اسلوب JOLLIFFE وقد بلغت %83.623، في حين أن اقل نسبة للتباين كانت لأسلوب SCREE DIAGRAM وقد بلغت %76.129 كما ان عدد المركبات الرئيسية المستخرجة بأسلوب BARTLETT بلغ 10 مركبات رئيسية وهو اقل من العدد المستخرج للأسلوب JOLLIFFE، SCREE DIAGRAM، ولكن يعد اسلوب JOLLIFFE هو الأسلوب الأفضل لأن مركباته تعبر عن متغيرات الظاهرة المدروسة بشكل أفضل وهذا يمكن ملاحظته من خلال نسب الاشتراكات (مربعات معاملات الارتباطات الخطية المتعددة) المرتفعة لمتغيرات الظاهرة المدروسة في مركباته الرئيسية حيث ان اقل قيمة للاشتراكات كانت 0.716 بين المتغير التوضيحي نصاب الأستاذ الجامعي و جميع المركبات الرئيسية المختارة.

٢- أفضل اسقاط ملائم (PROJECTION PURSUIT) لبيانات الظاهرة المدروسة يمكن الحصول عليه باستخدام أسلوب (JOLLIFFE , SCREE DIAGRAM)، حيث إن المتجه الذاتي الأول الذي يعطي اكبر نسبة للتباين هو متجه الأسقاط الملائم (PROJECTION PURSUIT).

٣- عند مقارنة متوسطات مربعات الخطأ للأساليب الثلاثة نجد أن أسلوب JOLLIFFE أعطى اقل قيمة لمتوسط مربعات الخطأ وقد بلغت 0.806.

٤- من خلال دراستنا لأبرز العوامل المسببة لتأخر حصول التدريسي الجامعي على اللقب العلمي، حيث اعتمدنا على استخدام اسلوب JOLLIFFE في تحديد عدد المركبات الرئيسية الداخلة في تحليل المركبات الرئيسية نستطيع ان نحدد ابرز العوامل المؤثرة في الظاهرة المدروسة وهذه العوامل هي (نصاب الأستاذ الجامعي، الروتين الإداري، أسعار لوازم القرطاسية، عدد الصفحات، عدد السنوات، النشاطات الترفيهية، الجوائز المادية، الوضع العائلي، الاتجاه السياسي والمحسوبية، نقص الكادر الأكاديمي)، أما بقية العوامل فأن تأثيرها كان ضعيفا.

المصادرالمصادر العربية:

- ١- الألوسي، احمد صالح، (1987) "مقدمة تحليل متعدد المتغيرات"، الجزء الثاني، المكتب المركزي للطباعة، جامعة بغداد.
- ٢- بشير، سعد زغول، (2003) "دليلك الى البرنامج الاحصائي SPSS"، الإصدار العاشر، المعهد العربي للتدريب والبحوث.
- ٣- الجنابي، ضوية سلمان، (١٩٩٠) "دراسة تطبيقية لاستخدام التحليل العامل- المركبات الرئيسية على مرضى السكر" مجلة كلية الإدارة والاقتصاد، جامعة بغداد، العدد الأول، ص ٥٣٥-٥٦٤
- ٤- كاظم، اموري هادي وباسم شليبه مسلم، (٢٠٠٢)، "القياس الاقتصادي المتقدم النظرية والتطبيق" مطبعة الطيف، العراق .
- ٥- المشكوري، سعاد خلف، (2001) "مقارنة أسلوبى المركبات الأساسية والتحليل العاملى لمعالجة مشكلة التعدد الخطي"، رسالة دكتوراه في الإحصاء، جامعة بغداد.
- ٦- المشهداني، إيمان محمد عبد الله، (1994) "استخدام المركبات الرئيسية في تشخيص ومعالجة مشكلة التعدد الخطي مع تطبيق عملي لبعض الظواهر الاقتصادية" رسالة ماجستير في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد.
- ٧- عبد القادر، زينة ياوز، (٢٠٠٤) "مقارنة الاساليب المستخدمة في تحديد عدد المركبات الرئيسية مع الجانب التطبيقي" رسالة ماجستير علوم في الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد.

المصادر الأجنبية:

- 7- Brian S. Everitt and Graham Dunn, (2001), "applied multivariate data analysis", Oxford university, p- 48-69.
- 8- Guoying and Zhonguan chen, (1985)," projection pursuit approach to Robust dispersion matrices and principal components", JASA, vol. 80, no.391.
- 9- Henry S. Lynn and Charies E. Mcculloch, (2000)," using principal component analysis and correspondence analysis for estimation in latent variable model", JASA, vol. 95, no. 450.
- 10- K. V. Mardia, J. T. Kent, J. M. Bibby, (1979)," multivariate analysis", Academic press, inc. London.
- 11- Kristin M. Branson and Sameer Agarwal, (2002)" structured principal component analysis" University of California.
- 12- Rafael A. Caivo, Matthew Partidg, marwan a. Jabri, (1998) "acomparative study of principal component analysis techniques"citeseer. its. psu/135968. html.
- 13 Richard Hill,(1996),"elementary linear algebra with applications", third edition, Michigan state university.
- 14- Richard H. Mccuen, (1985), "statistical methodas for engineers", Englewood Cliffs, Newjersey.

- 15-Richard J. Bolton and Ojtek J. Krzanowskik, (1999), "acharacterization of principal component for projection pursuit", JASA, vol. 53, no.2, p- 108-109 .
- 16-Richard J. Harris, (1975), " A primer of multivariate statistics", Academic press.
- 17-Robb J. Muirhead, (1995), " aspects of multivariate statistical theory" John Wiley.
- 18- Thaddeus Tarpey, (1999), " self consistency and principal component analysis", JASA, vol. 94, no. 446, p-456-467.
- 19- Y. L. Tong, (1990), " the multivariate normal distribution", springer series in st.