الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة الشهيد حمة لخضر-الوادي-

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

مطبوعة دروس على الخط:

تحليل المعطيات

موجهة لطلبة السنة الثانية ماستر اقتصاد كمي

إعداد: د.جديدي موسى

2022/2021

فهرس المحتويات

المحور الاول: طريقة تحليل المركبات الاساسية ACP

1. التقديم النظري لطريقة المركبات الاساسية

استعملت هذه الطريقة لأول مرة من طرف pearson karl سنة 1901 ، و أول من ضمها إلى الإحصاء الرياضي هو Hotelling Harold سنة 1933 ، غير أنها لم تصبح واسعة الإستعمال إلا في التسعينات، ذلك لظهور الحاسوب و البرامج الإحصائية التي سهلت العمل بهذه التقنية.

تعتبر طريقة تحليل المركبات الرئيسية أو المكونات الأساسية Analyse en composantes principales

إحدى أهم طرق التحليل العاملي حيث تهتم بتقليص أكبر قدر ممكن من المتغيرات أو الأفراد إلى عوامل أو مركبات

رئيسية تحتوي على أكبر قدر من المعلومات أو التباينات الموجودة في المتغيرات أو الأفراد باستخدام التحليل الوصفي، وأيضا تعمل على اسقاطات النقاط في المستوي باستخدام العديد من أساليب تدوير المحاور لأجل الحصول على أحسن تمثيل.

كما يعتبر التحليل في المركبات ألاساسية أحد تقنيات تحليل التي تختص في اختزال البيانات ذات الأبعاد الكبير ، أي اختزال عدد كبير من المتغيرات الخام إلى عدد أقل من المتغيرات الجديدة في شكل مركبات وتكون غالبا أهدادها أقل بكثير من المتغيرات الخام

وعلى خلاف تحليل الإنحدار أو تحليل التباين الذان يعتبران متغير واحد تابع و البقية مستقلة فإن التحليل في المكونات الأساسية كل متغير يقارن ببقية المتغيرات، وإيجاد الجذور المميزة و المتجهات المميزة لمصفوفة التباين و التباين

المشـــترك للمتغيــرات التوضـــيحية، أو إيجـــاد الجـــذور المميـــزة لمصـــفوفة الارتبـــاط، و هذا يعتمد على طبيعة البيانات.

2. التحليل في فضاء المتغيرات

تشـــكل مصـــفوفة البيانـــات x_{np} ســـحابة مـــن n نقطـــة فـــي فضـــاء بعـــده p ، أو ســـحابة مـــن نقـــاط p فـــي فضـــاء ذو بعـــد n. يتكـــون الرســـم التخطيطـــي الثنـــائي مـــن إســـقاط هـــذه النقـــاط علـــى اثنـــين أو أكثـــر مـــن الأبعـــاد المختـــارة بشـــكل عشـــوائي. يتمثـــل PCA مـــن إســـقاط النقـــاط علـــى خـــط مســـتقيم ، مســـتوى ... مســـاحة فرعيـــة مـــن أبعـــاد s (مـــع p s) تـــم اختيارهـــا لتحســـين معيـــار معـــين. بشـــكل حدســـي ، ســـوف نبحـــث عـــن الفضـــاء الجزئـــي الـــذي يعطـــي أفضـــل تصـــور ممكـــن لســـحابة النقـــاط الخاصـــة بنـــا. الاختيـــار الجيـــد هـــو البحـــث عـــن أكبـــر تشـــتت ممكـــن أكبـــر انتشـــار فـــي الفضـــاء الجزئـــي المختـــار. وبالتـــالي ، فإننـــا قـــادرين علـــى البحـــث عـــن دوران لنظامنـــا الأولـــي مـــن المحـــاور (المتغيـــرات) ممـــا يســـمح لنـــا برؤيـــة الســـحابة بشـــكل أفضـــل. دعونـــا نحـــدد u_1 الشـــعاع المطلـــوب ($u_1'u_1=1$)؛ والـــذي يعطـــي أكبـــر تشـــتت للإسقاطات.

لـــتكن المصـــفوفة X_{nxp} ؛ كـــل ســـطر يمثـــل ملاحظـــة. يمثـــل كـــل عمـــود متغيـــرًا. ســـنفترض أن كـــل متغيـــر يـــتم توســـيطه ، أي أننـــا طرحنـــا متوســـط كـــل متغيـــر مسبقًا. يتم ذلك بطريقة تجعل مركز ثقل نقطة السحابة متزامنًا مع المركز.

تُعطى إسقاطات الملاحظات n على الشعاع u_1 من خلال:

$$C = Xu_1$$

مجموع مربعات هذه السقاطات هو:

$$C'C = u_1'X'Xu_1$$

سنختار u_1 لتعظيم هذه الكمية الأخيرة. لذا فإن المشكلة هي:

تعظيم $u_1'X'Xu_1$ وفقًا لـ $u_1'u_1 = 1$

هذه مشكلة يمكن حلها بطريقة لاغرانج.

نشكل لاغرانج:

$$L = u_1'X'Xu_1 - \lambda\,(u_1'u_1-1)$$

نشتق المعادلة بالنسبة لكل مكونات p للشعاع u_1 وكذلك بالنسبة إلى مضاعف لاغرانج (λ) ونساوي المشتقات الجزئية لصفر.

$$2(X'Xu_1- \lambda u_1) = 0$$

$$u_1'u_1 = 1$$

بعد التبسيط نجد:

$$X'X\,u_1 = \lambda u_1$$

$$u_1'u_1 = 1$$

الشكل السابق يمثل معادلة الأشعة الذاتية والقيم الذاتية للمصفوفة X'X. وبالتالي ، فإن الشعاع الذي يعطي الإسقاطات ذات التشتت الأكبر هو الشعاع الذاتي الأول لمصفوفة التباين - التغاير لـ X.

ملاحظات:

- مصفوفة العددية (X'X) هي مصفوفة متماثلة وموجبة. هذا يعني أن القيم الذاتية و الأشعة الذاتية ستكون عددية. بالاضافة الى ذلك ، ستكون قيم القيم الذاتية موجبة أو صفرية.
- الأشعة الذاتية لمصفوفة متماثلة تكون دائمًا متعامدة مع بعضها البعض ، أي $u_1'u_2=0$.
- تحدد الإسقاطات الموجودة على الشعاع u_1 متغيرًا جديدًا وهو عبارة عن مجموعة خطية من المتغيرات الأصلية. يتم إعطاء تباين هذا المتغير الجديد من خلال:

$$1/n\ (u_1'X'Xu_1) = 1/n\ (u_1'\lambda 1 u_1) = 1/n\ \lambda_1$$

وبالتالي فإن تباين الإسقاطات يساوي (للعامل 1/n) للقيمة الذاتية. وهكذا يتم الوصول إلى الحد الأقصى مع القيمة الذاتية الأولى. هذا هو السبب في الاحتفاظ بالشعاع الذاتي الأول.

- نعلم أن مجموع القيم الذاتية للمصفوفة يساوي أثر المصفوفة الأصلية. لذلك لا يتغير الحجم الإجمالي للتباين.

- تشير النسبة $\lambda_1/(\Sigma\lambda_i)$ إلى نسبة التباين الكلي الذي يدعمه الشعاع الذاتي الأول.

السؤال 1: ما هي العلاقة بين الأشعة الذاتية والقيم الذاتية لـ $X'X$ وتلك الخاصة بـ $X'X / n$؟

إيجاد الشعاع الثاني u_2.

نبحث عن الشعاع u_2 الذي يحقق:

تعظيم $u_2'X'Xu_2$

يخضع لـ $u_2'u_1 = 0$

و $u_2'u_2 = 1$

الحل يكمن في استخدام طريقة لاغرانج كما يلي:

$$L = u_2'X'Xu_2 - w\,(u_2'u_1) - \lambda_2\,(u_2'u_2 - 1)$$

نشتق لاغرانج فيما يتعلق بكل مكون من مكونات الشعاع u_2 وفيما يتعلق بـ u_1.

بعد التبسيط ، نجد:

$$X'X\,u_2 = \lambda_2 u_2$$

$$u_2'u_2 = 1$$

$$u_1'u_2 = 0$$

الشعاع الذي تم العثور عليه هو الشعاع الذاتي المرتبط بثاني أكبر قيمة ذاتية لمصفوفة التباين والتغاير. يمكن تعميم هذه النتائج بسهولة على عدة أبعاد ونجد ما يلي:

الفضاء ذو البعد s بحيث s≤p الذي يعطي أفضل تفسير يتم تعريفه بواسطة الأشعة الذاتية لـ X'X المرتبطة بأكبر قيم للقيم الذاتية.

3. التحليل في فضاء العينة للعينات

لقد درسنا كيف يمكن تطبيق ACP لتحليل سحابة n من المشاهدات في فضاء المتغيرات p. يمكن أيضًا أن نقوم بالتحليل سحابة المتغيرات p في فضاء n من المشاهدات. نحن نبحث عن فضاء فرعي للبعد s (s ≤ p) حيث يكون مجموع مربعات الإسقاطات هو الحد الأقصى.

نطبق نفس التقنية السابقة ونجد أن حل الشعاع الأول يتم الحصول عليه من خلال:

$$XX'\nu_1 = \beta_1 \nu_1$$

$$\nu_1'\nu_1 = 1$$

الشعاع الذاتي الأول لـ XX' هو الذي يزيد من تباين الإسقاطات. ، نشكل معادلة تعظيم تباين الإسقاطات من الشعاعين الذاتيين الأول والثاني ، إلى آخره.

<u>نظرية: قيم الذاتية λ_i و β_i متطابقة.</u>

لنضرب المعادلة بـ X':

نجد:

$$X'XX'v_i = \beta_i X'v_i$$

نلاحظ من خلال هذه المعادلة أن $X'v_i$ هو الشعاع ذاتي لـ X'X مرتبط بقيمة β_i. ولدينا القيم الذاتية لـ X'X تعطى بواسطة λ_i. ومنه نستنتج أن $\beta_i = \lambda_i$

ملاحظة: من خلال مما سبق فإنه ليس من الضروري البحث بشكل مباشر عن القيم الذاتية والأشعة الذاتية لـ XX' ، ويتم ايجاد الاشعة الذاتية u_i بواسطة $X'v_i$ ويمكن توضيح ذلك كما يلي:

$$v_i'XX'v_i = \lambda_i$$

ومنه:

$$u_i = \frac{1}{\sqrt{\lambda_i}} X'v_i$$

وبالضرورة يكون

$$v_i = \frac{1}{\sqrt{\lambda_i}} Xu_i$$

تسمى هذه الصيغة بصيغة الانتقال.

يمكن كتابتها في شكل مصفوفة كالتالي :

$$V = XU\Lambda^{-1/2} \quad U = X'V\Lambda^{-1/2}$$

مثال:

x	3	4	6	6	6	7	7	8	9	9	9	01	11	21	21	31	31	31	31	41	51	71	71	81	02
y	2	01	5	8	01	2	31	9	5	8	41	7	21	01	11	6	41	51	71	7	31	31	71	91	02

نطبق التحليل بالمركبات الرئيسة لمصفوفة التغاير

$$\Sigma = \begin{pmatrix} 20.28 & 15.59 \\ 15.59 & 24.06 \end{pmatrix}$$

4. طرق التدوير:

أ. التدوير المتعامد:

هناك أكثر من طريقة تحقق التعامد بين العوامل، أشهرها طريقة تعظيم التباين Varimax، ويدل التعامد على استقلالية المحاور العاملية، وتسمح هذه الطريقة بالاحتفاظ بالتعامد القائم بين العوامل الأصلية. مما يعني أن المحور العاملي الأول و المحور العاملي الثاني متعامدين لا يرتبطان مهما تغير موضع المحورين.

ب. التدوير المائل:

يختلف هذا النوع من التدوير عن السابق في كونه يفترض ارتباط العوامل. هذا النوع من التدوير صعب ومعقد. ومع هذا بعض الباحثين يفضلون استخدامه بدلا عن التدوير المتعامد لسببين:

- يتصف بالفاعلية في التعامل مع البيانات كبيرة الحجم

- يسمح بتفسير أبسط تكوين ممكن للبنية العاملية والتوصل إلى حلول يمكن من خلالها تفسير وتعليل العلاقات البينية.

ت. طريقة تعظيم التباين:

كطريقة التعظيم الرباعي Quartimax وطريقة ايكويماكس Equimax،

تهدف هذه الطريقة الى تقليل عدد المتغيرات التي تتشبع بشكل قوي بكل عامل، عن طريق تدوير جميع العوامل الممكنة مثنى مثنى على حدة إلى أن يتم تعظيم مجموع تباينات مربعات التشبعات في أعمدة المصفوفة العاملية، والوصول في النهاية إلى أن يكون لكل متغير تشبع واحد عال على أحد العوامل ومنخفض على العوامل الأخرى. وينتج من هذه العمليات إحداث تكافؤ في أهمية المتغيرات مما يؤدي إلى التقليل من عدد المتغيرات لكل عامل، وبالتالي يكون العامل أكثر وضوحا ويسهل قراءته وتفسير مكوناته.

5. مثال حسابي

لتكن المصفوفة X بحيث

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 4 \\ 3 & 1 \end{pmatrix}$$

1. احسب المتوسط لكل متغير
2. احسب التباين لكل متغير σ
3. استنتج المصفوفة المركز والمعممة
4. احسب مصفوفة الارتباط

5. احسب القيم الذاتية لمصفوفة الارتباط

6. احسب الأشعة الذاتية لمصفوفة الارتباط

7. قارن بين $\sum\lambda i$ و tr(r)

الحل:

.1

$$\bar{X}_1 = \frac{1}{\mathrm{n}}\sum X_{1i} = \frac{1}{3}(1+1+3) = \frac{5}{3}$$

$$\bar{X}_2 = \frac{1}{\mathrm{n}}\sum X_{2i} = \frac{1}{3}(1+4+1) = 2$$

.2

$$\sigma^2_{\mathrm{x}_1} = \frac{1}{\mathrm{n}}\sum(X_{1i} - \bar{X}_1)^2$$

$$= \frac{1}{3}\left[(1-\tfrac{5}{3})^2 + (1-\tfrac{5}{3})^2 + (3-\tfrac{5}{3})^2\right] = \frac{8}{9}$$

$$\sigma^2_{\mathrm{x}_2} = \frac{1}{\mathrm{n}}\sum(X_{2i} - \bar{X}_2)^2$$

$$= \frac{1}{3}[(1-2)^2 + (4-2)^2 + (1-2)^2] = 2$$

.3

نحسب المصفوفة الممركزة والمعممة من خلال العلاقة التالية $Z = \frac{X_i - \bar{X}_i}{\sigma_i}$

$$Z = x_i - \bar{x}_i = \begin{pmatrix} 1-\frac{5}{3} & 1-2 \\ 1-\frac{5}{3} & 4-2 \\ 3-\frac{5}{3} & 1-2 \end{pmatrix} = \begin{pmatrix} \frac{-2}{3} & -1 \\ \frac{-2}{3} & 2 \\ \frac{4}{3} & -1 \end{pmatrix}$$

$$Z = \frac{x_i - \bar{x}_i}{\sigma_i} = \begin{pmatrix} \frac{\frac{-2}{3}}{\frac{2}{3}\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{\frac{-2}{3}}{\frac{2}{3}\sqrt{2}} & \frac{2}{\sqrt{2}} \\ \frac{\frac{4}{3}}{\frac{2}{3}\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

$$Z = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 \\ -1 & 2 \\ 2 & -1 \end{pmatrix}$$

.4

نحسب مصفوفة الارتباط من خلال العلاقة التالية $\mathrm{r} = \frac{1}{n} Z'Z$

$$\mathrm{r} = \frac{1}{3} \times \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \end{pmatrix} \times \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 \\ -1 & 2 \\ 2 & -1 \end{pmatrix}$$

$$\mathrm{r} = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$$

.5

نقوم بحساب القيم الذاتية بتطبيق العلاقة التالية $|r - \lambda \mathrm{I}| = 0$

$$\left|\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| = 0$$

$$\left|\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right| = 0$$

$$\left|\begin{pmatrix} 1-\lambda & -0.5 \\ -0.5 & 1-\lambda \end{pmatrix}\right| = 0$$

$$(1-\lambda)^2 \times \left(-\frac{1}{2}\right)^2 = 0$$

$$\left(1-\lambda-\frac{1}{2}\right) \times \left(1-\lambda+\frac{1}{2}\right) = 0$$

$$\begin{cases} \lambda_1 = \frac{2}{3} = 1.5 \\ \lambda_2 = \frac{1}{2} = 0.5 \end{cases}$$

.6

<u>الشعاع الذاتي الاول u_1</u>

نحسب الشعاع u_1 من خلال العلاقة التالية:

$$(r-\lambda_1 \mathrm{I})\mathrm{u}_1 = 0$$

$$\left(\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - 1.5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = 0$$

$$\begin{cases} -0.5x - 0.5y = 0 \\ -0.5x - 0.5y = 0 \end{cases}$$

حلول المتراجحة السابقة من الشكل $y=-x$

ومنه $\mathrm{u}_1 = \begin{pmatrix} x \\ -x \end{pmatrix}$

$$u_1 = x\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

لدينا $u_1' u_1 = 1$

$$u_1' u_1 = x(1 \quad -1) \times x\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1$$

بعد حل المعادلة نختار $x = \sqrt{\frac{1}{2}}$

اذا الشعاع الذاتي الأول هو $u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

<u>الشعاع الذاتي الثاني u_2</u>

نحسب الشعاع u_2 من خلال العلاقة التالية:

$$(r - \lambda_2 I)u_2 = 0$$

$$\left(\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} - 0.5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{cases} 0.5x - 0.5y = 0 \\ -0.5x + 0.5y = 0 \end{cases}$$

حلول المتراجحة السابقة من الشكل $y=x$

ومنه $u_1 = \begin{pmatrix} x \\ x \end{pmatrix}$

$$u_1 = x\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

لدينا $u_1' u_1 = 1$

$$u_1' u_1 = x(1 \quad 1) \times x\begin{pmatrix}1\\1\end{pmatrix} = 1$$

بعد حل المعادلة نختار $x = \sqrt{\frac{1}{2}}$

اذا الشعاع الذاتي الثاني هو $u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$

.7

$tr(r) = \sum a_{ii} = 1+1=2$

$\sum \lambda_i = 1.5+0.5=2$

نلاحظ أن $tr(r) = \sum \lambda_i$

6. مثال تدريبي:

مصفوفة البيانات

$$\begin{pmatrix} 21 & 38 & 52 \\ 4 & 51 & 67 \\ 67 & 83 & 0 \\ 67 & 3 & 38 \\ 93 & 5 & 6 \end{pmatrix}$$

المصفوفة الممركز

$$\begin{pmatrix} -29.4 & 2 & 19.4 \\ -46.4 & 15 & 34.4 \\ 16.6 & 47 & 32.6 \\ 16.6 & -33 & 5.4 \\ 42.6 & -31 & -26.6 \end{pmatrix}$$

مصفوفة التباين والتباين المشترك

$$\begin{pmatrix} 1076.6 & -368.6 & -750.24 \\ -368.6 & 897.6 & -66.2 \\ -750.24 & -66.2 & 671.84 \end{pmatrix}$$

القيم الذاتية:

$\lambda_1 = 1732.20$

$\lambda_2 = 906.65$

$\lambda_3 = 7.28$

الاشعة الذاتية:

	u_1	u_2	u_3
x_1	0.79	-0.06	0.62
x_2	-0.30	-0.90	0.31
x_3	-0.54	0.43	0.73

احداثيات المشاهدات

	u_1	u_2	u_3
1	-34.15	8.16	-3.40
2	-59.54	3.84	1.03
3	16.25	-57.29	1.07
4	20.21	31.14	3.94
5	57.24	14.14	-2.65

إحداثيات المتغيرات

	u_1	u_2	u_3
x_1	-32.73	-1.66	1.66

x_2	12.68	-27.13	0.83
x_3	22.37	12.95	1.96

جودة تمثيل المشاهدات

	u_1	u_2	u_3
1	0.94	0.05	0.01
2	1.00	0.00	0.00
3	0.07	0.93	0.00
4	0.29	0.70	0.01
5	0.94	0.06	0.00

جودة تمثيل المتغيرات

	u_1	u_2	u_3
x_1	0.99	0.00	0.00
x_2	0.18	0.82	0.00
x_3	0.74	0.25	0.01

مساهمة المشاهدات

	u_1	u_2	u_3
1	0.13	0.01	0.32
2	0.41	0.00	0.03
3	0.03	0.72	0.03
4	0.05	0.21	0.43
5	0.38	0.04	0.19

مساهمة المتغيرات

	u_1	u_2	u_3
x_1	0.62	0.00	0.38
x_2	0.09	0.81	0.10
x_3	0.29	0.19	0.53

المحور الثاني: التحليل العاملي المتناظر AFC

يعد التحليل العاملي للمشاهدات من مجموعة التحليل العاملي التي طورت للتعامل مع معطيات تختلف عن طريقة المكونات الرئيسية، غير أن الاختلاف الجوهري يكمن في نوع تغيرات لا في الطريقة، وبالتالي فإنهما لا يختلفان من حيث الهدف في تقديم اختزال للمعلومات في فضاء متعدد الأبعاد بحيث تكون القراءة واضحة ولا تختلف مع المعطيات الاولية عن طريق الحفاظ على أكبر قدر من المعطيات الخام .

يعود أصول الطريقة إلى jean paul benzcri حيث قام في سنة 1961 بدراسة العلاقة بين متغيرين كيفيين باستعمال طريقة المركبات الرئيسية ليتم تطويرها و تعديلها لاحقا من قبله في 1990 عموما تستعمل طريقة أو تقنية AFC في الجداول المزدوجة (الاقتران أو التوافق) ذات المتغيرات الكيفية حيث تسمح بتمثيل نقاط الأسطر و الأعمدة لجدول فئات أو صفات (modalité) المتغيرين انوعيين.

من الواضح أن جدول البيانات أكثر قابلية للإدارة ، من الناحية العملية ، من الجداول المنفصلة الكاملة ، خاصة إذا كان عدد الأفراد كبيرًا.

من الضروري أن نلاحظ الآن أن هناك اختلافين رئيسيين بين ACP و AFC.

- المقياس المستخدم في AFC لتحديد القرب بين صفين أو عمودين هو مقياس Chi-square ، بينما تُستخدم المسافة الإقليدية في PCA.

- يسمح AFC بتمثيل متداخل لصفوف وأعمدة الجدول، بدلاً من تقديم رسمين بيانيين مستقلين ، أحدهما للمتغيرات والآخر للأفراد.

1. مجالات التطبيق

تــم اســتخدام هــذا التحليــل عمليًــا لأنــه مصــمم لجــداول التقابليــة وبالتــالي يســمح بدراســة الــروابط الموجــودة بــين متغيــرين اســميين. وبالتــالي فــإن مجــالات تطبيــق AFC تختلــف عــن تلــك الخاصــة بــ PCA ، وهــي مناســبة لجــداول القياســات المتجانسة أو غير المتجانسة.

بالنســبة لهــذا التحليــل ، يمكننــا أيضًــا تقــديم قائمــة طويلــة مــن التخصصــات التــي وجــدت إجابــة لمشــكلتها مــن خــلال AFC. وهكــذا ، فــإن علــم البيئــة، وعلــم الــنفس ، والاقتصــاد ، وغيــر ذلــك مــن العلــوم التــي قــد يكــون مــن المثيــر للاهتمــام فيهــا بدراســة الــروابط بــين متغيــرين اســميين ، قــد قــدمت قــدرًا كبيــرًا مــن البيانــات. المصــــمم للجــــداول التقابليــــة (أي التــــرددات) ، يمكــــن تطبيقــــه علــــى جــــداول المقــايس المتجانســـة (أي نفــس نظــام الوحــدات) ، وجــداول الــدرجات ، والرتــب ، والتفضــيلات ، والجــداول ذات القــيم المنطقيــة (0 أو 1) ، وايظــا علــى الجــداول المأخوذة من استبيانات .

2. البيانات المستخدمة

علــى عكــس ACP فإنــه عنــد اســتعمال AFC يجــب تنظــيم البيانــات، فــي جــداول تقابلية (تسمى أيضًا جداول التبعية أو الجداول المتقاطعة).

الجــدول التقابليــة هــي جــدول للأرقــام تــم الحصــول عليهــا مــن خــلال تقــاطع متغيرين نوعيين محددين على نفس المجموعة المكونة من n من الأفراد.

يمكــن أيضًــا توســيع CFA ليشــمل المتغيــرات الكميــة المتجانســة عــن طريــق تحديــد بعــض الخصــائص لهــذه المتغيــرات. بالامتــداد ، و ينطبــق أيضًــا علــى جــداول المتغيــرات الفرديــة للمتغيــرات الكميــة المتجانســة ، وفــي هــذه الحالــة يــتم اعتبــار الأفراد كمتغيرات.

وبالتــالي فــإن البيانــات الأوليــة تكــون منظمــة بالطريقــة التــي ســبق شــرحها كمــا هو موضح في الجدول التالي:

	فئات المتغير الثاني
فئات المتغير الأول	$\begin{array}{cccccc} & 1 & \cdots\cdots & j & \cdots\cdots & J \\ 1 & & & & & \\ \vdots & & & \vdots & & \\ \vdots & & & \vdots & & \\ i & & \cdots\cdots & k_{ij} & \cdots\cdots & \\ \vdots & & & \vdots & & \\ \vdots & & & \vdots & & \\ I & & & & & \end{array}$

- يمثل i عدد الصفوف وا مجموعة الصفوف
- يمثل j عدد الأعمدة و J مجموعة الأعمدة
- kij هــو عــدد الأفــراد الــذين يمتلكــون كــل مــن الصــفة i للمتغيــر الأول والصفة j للمتغير الثاني.

وبالتالي لدينا:

$$\sum_{i\in I}\sum_{j\in J} k_{ij} = n$$

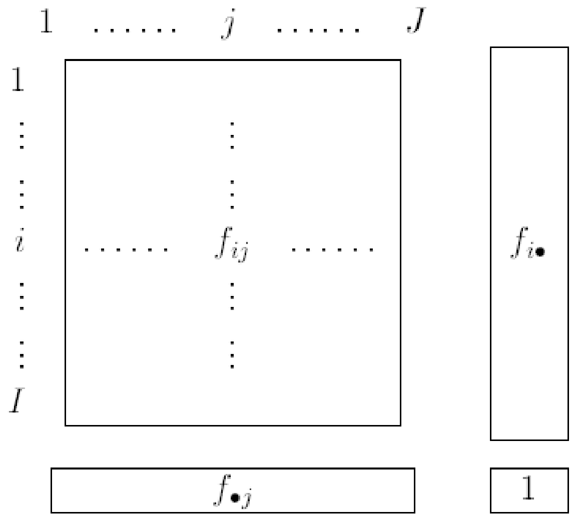
مع n العدد الإجمالي للأفراد في الجدولالأولي. نلاحظ أن الصفوف والأعمدة في هذا النوع من الجداول تلعب دورًا متماثلًا. يمكن تعريف جدول الترددات النسبية الذي يؤخذ في الاعتبار ترددات fij تعطى بواسطة:

$$f_{ij} = \frac{k_{ij}}{n}$$

والهوامش بالعلاقة التالية:

$$f_{i\bullet} = \sum_{j \in J} f_{ij}$$

كما هو موضح في جدول التكرار النسبي



بحيث:

$$f_{\bullet j} = \sum_{i \in I} f_{ij}$$

و

$$\sum_{i\in I} f_{i\bullet} = \sum_{j\in J} f_{\bullet j} = \sum_{i\in I}\sum_{j\in J} f_{ij} = 1.$$

3. أهدف التحليل العاملي المتناظر AFC

أهـــداف AFC هـــي نفـــس أهـــداف ACP بمعنـــى أن AFC يســـعى إلـــى الحصـــول علـــى تصنيف الصفوف وتصنيف الأعمدة ، ثم ربط هذين النموذجين.

لـــذلك مـــن الضـــروري إجـــراء تقيـــيم لأوجـــه التشـــابه بـــين الســـطور (علـــى التـــوالي الأعمدة) الإجابة على أسئلة من كالتالي:

- ما هي الصفوف المتشابهة (الأعمدة)؟

- أيهما مختلف؟

- هل هناك مجموعات متجانسة من الصفوف (أعمدة على التوالي)؟

- هل من الممكن إبراز تصنيف السطور (على التوالي أعمدة)؟

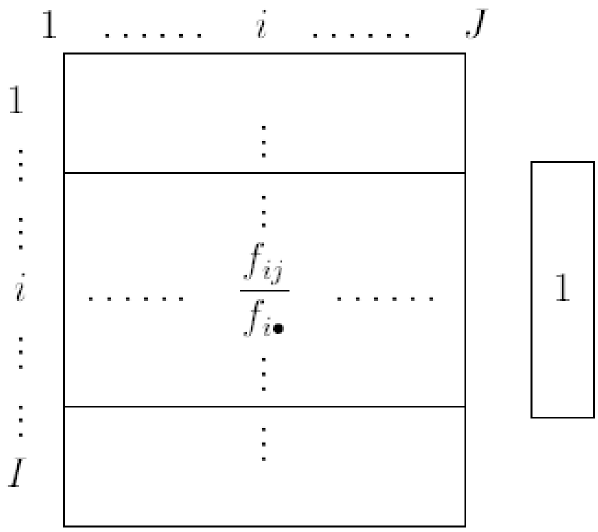
ومـــع ذلـــك ، فـــإن فكـــرة التشـــابه بـــين صـــفين أو عمـــودين تختلـــف عـــن ACP. فـــي الواقـــع يكـــون صـــفان (علـــى التـــوالي عمـــودين) قـــريبين إذا كانـــا مـــرتبطين بـــنفس الطريقة مع جميع الأعمدة.

لـــذلك مـــن الضـــروري البحـــث عـــن الصـــفوف (علـــى التـــوالي الأعمـــدة) التـــي يكـــون توزيعهـــا أكثـــر انحرافًـــا عـــن البقيـــة ، وتلـــك التـــي تتشـــابه مـــع بعضـــها الـــبعض وتلـــك التـــي تتعـــارض . مـــن أجـــل ربـــط تصـــنيف الصـــفوف بمجموعـــة الأعمـــدة ، تتميـــز كـــل مجموعة من الصفوف بالأعمدة التي ترتبط بها هذه المجموعة.

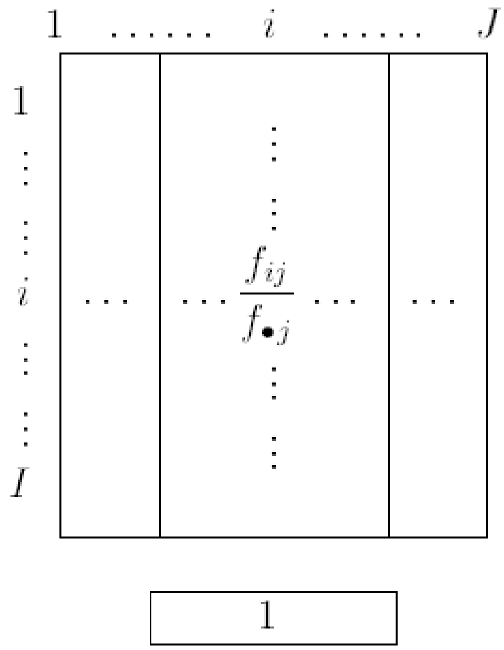
ومــن خــلال التنــاظر ، تتميــز كــل مجموعــة مــن الأعمــدة بالصــفوف التـي تــرتبط بهــا هــذه المجموعــة. وبالتــالي يمكننــا تفكيــك الــرابط بــين متغيــرين إلــى مجمــوع ميــول بسيطة وقابلة للتفسير وقياس أهمية كل منهما.

ســتعامل مــع الجــدول مــن ناحيــة كسلســلة مــن الصــفوف ، ثــم كسلســلة مــن الأعمــدة. عنــدما يــتم التعامــل مــع الجــدول فـي الصــف ، تــتم تســوية البيانــات عــن طريق القسمة على fi_0.

والغــرض مــن هــذا التســوية هــو النظــر فـي الــروابط بــين المتغيــرين مــن خــلال الفــرق بين النسب المئوية في الصفوف.



ونفس الشيء بالنسبة للصفوف



4. التشابه والجمع بين الفئات

يتم تحديد التشابه بين صفين أو بين عمودين بالمسافة بين الفئات. المسافة المستخدمة هي x^2 ويتم تحديدها بشكل متماثل للصفوف والأعمدة. وبالتالي بين صفين i و 'i يتم الحصول عليها بواسطة:

$$d_{\chi^2}(i,i') = \sum_{j \in J} \frac{1}{f_{\bullet j}} \left(\frac{f_{ij}}{f_{i\bullet}} - \frac{f_{i'j}}{f_{i'.}} \right)^2$$

وبين عمودين j و 'jبواسطة:

$$d_{\chi^2}(j,j') = \sum_{i \in I} \frac{1}{f_{i\bullet}} \left(\frac{f_{ij}}{f_{\bullet j}} - \frac{f_{ij'}}{f_{\bullet j'}} \right)^2 .$$

5. مثال حسابي:

	D	E	F	المجموع
A	15	12	3	30
B	10	18	4	32
C	15	5	8	28
المجموع	40	35	15	90

$$V_{11} = \begin{bmatrix} \frac{30}{90} & 0 & 0 \\ 0 & \frac{32}{90} & 0 \\ 0 & 0 & \frac{28}{90} \end{bmatrix}$$

$$V_{22} = \begin{bmatrix} \frac{40}{90} & 0 & 0 \\ 0 & \frac{35}{90} & 0 \\ 0 & 0 & \frac{15}{90} \end{bmatrix}$$

$$V_{12} = \begin{bmatrix} \frac{15}{90} & \frac{12}{90} & \frac{3}{90} \\ \frac{10}{90} & \frac{18}{90} & \frac{4}{90} \\ \frac{15}{90} & \frac{5}{90} & \frac{8}{90} \end{bmatrix}$$

$$(V_{11})^{-1} * V_{12} = \begin{bmatrix} \frac{15}{30} & \frac{12}{30} & \frac{3}{30} \\ \frac{10}{32} & \frac{18}{32} & \frac{4}{32} \\ \frac{15}{28} & \frac{5}{28} & \frac{8}{28} \end{bmatrix}$$

$$V_{12} * (V_{22})^{-1} = \begin{bmatrix} \frac{15}{40} & \frac{12}{35} & \frac{3}{15} \\ \frac{10}{40} & \frac{18}{35} & \frac{4}{15} \\ \frac{15}{40} & \frac{5}{35} & \frac{8}{15} \end{bmatrix}$$

$$(V_{11})^{-1} * V_{12} * (V_{22})^{-1} * V_{21} = (V_{11})^{-1} * V_{12} * (V_{22})^{-1} * (V_{12})' = A$$

$$A = \begin{bmatrix} 0.345 & 0.357 & 0.298 \\ 0.335 & 0.400 & 0.265 \\ 0.320 & 0.301 & 0.379 \end{bmatrix}$$

$$(V_{22})^{-1} * V_{21} * (V_{11})^{-1} * V_{12} = (V_{22})^{-1} * (V_{12})' * (V_{11})^{-1} * V_{12} = B$$

$$B = \begin{bmatrix} 0.466 & 0.358 & 0.176 \\ 0.409 & 0.452 & 0.139 \\ 0.469 & 0.325 & 0.206 \end{bmatrix}$$

المحور الثالث: التحليل العاملي التمييزي AFD

يعد التحليل التمييزي أحد المناهج الإحصائية المهمة في التحليل الإحصائي متعددة المتغيرات والذي يعني في كيفية التمييز بين مجموعتين أو أكثر من المجتمعات وأن الفكرة الأساسية للتمييز هي التفرقة بين مجتمعات متداخلة أو متشابهة لها الخصائص أو الصفات بمعنى آخر أن التحليل التمييزي هو أسلوب إحصائي يتم بموجبه استعمال مجموعه من المتغيرات للتمييز بين مجموعتين أو أكثر عن طريق دالة تمييزية محددة وهي توليفة خطية للمتغيرات التوضيحية و طريقة إيجاد هذه الدالة بإيجاد المعاملات للدالة وفقا للقياسات أو المعايير التي يتم الحصول عليها من المشاهدات .

إن عملية التصنيف وهي العملية اللاحقة لعملية تكوين الدالة التمييزية إذ يتم الاعتماد على هذه الدالة بالتنبؤ وتصنيف المفردة الجديدة لإحدى المجموعات قيد الدراسة بأقل خطأ تصنيف ممكن والتحليل التمييزي أيضا يمكن أستخدمه في مختلف المجالات وفي حالة المجتمعات المتجانسة وغير المتجانسة.

أما الهدف الرئيسي من التحليل التميزي تصنيف المشاهدة او مجموعة من المشاهدات إلى مجاميعها التصنيفية وبأقل خطا تصنيف ممكن.

إن دالـــة التمييـــز الخطيـــة والمســـتندة إلـى تركيـــب خطـــي للمتغيـرات لكـــي تكـــون مثاليـــة يجـــب أن تنـــتج اصـــغر احتمـــال لخطـــا التصـــنيف علمـــا بـــان هنـــاك افتراضـــات يجـــب توافرهـــا عنـــد البيانـــات المســـتخدمة فـي التحليـل ولكـن غالبـا مـا نواجـه اختراقـا لـبعض الفرضـيات ومـن ثـم فـان تقـديرات دالـة التمييـز الخطيـة تفقـد خواصـها فمـثلا: عنـد عـدم تسـاوي مصـفوفة التبـاين والتبـاين المشـترك فان استعمال دالة التمييز التربيعية يكون ضروريا.

1. الدالة التمييزية:

هـي دلـة يمكـن مـن خلالهـا التمييـز بـين المجموعـات (الفصـل بـين المشـاهدات) ووضـع كـل مشـاهدة في المجموعـة التي تتبـع لهـا بمعنى آخـر هي الدالـة التي يمكـن بواســطتها تمييـــز (تصـــنيف) المشـــاهدات الجديـــدة (مجهولـــة الانتمـــاء) الـــى المجموعـــة الصـــحيحة التـــي يفتـــرض انتمـــائهم إليهـــا وفقـــا للمعـــايير او القياســـات التي تم الحصول عليها من المشاهدات المعلومة سابقا.

2. قوة دالة التمييز:

تعتمـــد دالـــة التمييـــز علـــي ســـلامة توزيعهـــا للمشـــاهدات (المفـــردات) علـــي المجموعـــات الصـــحيحة (أي التـــي تنتمـــي لهـــا المشـــاهدات فعليـــاً) وعليـــه فـــإن قـــوه داله التمييز ترتبط طرديا بالتوزيع السليم للمشاهدات.

3. أنواع دوال التمييز:

- ✓ دالة ال تمييز الخطية
- ✓ دالة التمييز غير الخطية

- ✓ دالة التمييز اللوجستية

4. تقدير معلمات الدالة التميزية الخطية:

لتقدير معلمات الدالة التمييزية في حالة وجود مجموعتين نتبع الخطوات الآتية:

1- يجاد متوسط كل متغير في كل مجموعة و كالآتي:

المجموعة الأولى:

$$\bar{X}_1^{(1)} = \sum_{i=1}^{n1} \frac{X_{1i}}{n_1}$$

$$\bar{X}_2^{(1)} = \sum_{i=1}^{n1} \frac{X_{2i}}{n_1}$$

$$\bar{X}_k^{(1)} = \sum_{i=1}^{n1} \frac{X_{ki}}{n_1}$$

حيـــــث $\bar{X}_i^{(1)}$ يمثل متوسط المتغير في المجموعة الأولى.

بنفس الطريقة للمجموعة الثانية:

$$\bar{X}_1^{(2)} = \sum_{i=1}^{n1} \frac{X_{1i}}{n_2}$$

$$\bar{X}_2^{(2)} = \sum_{i=1}^{n2} \frac{X_{2i}}{n_2}$$

$$\vdots$$

$$\bar{X}_k^{(2)} = \sum_{i=1}^{n2} \frac{X_{ki}}{n_2}$$

حيــث $\bar{X}_i^{(2)}$ يمثل متوسط المتغير في المجموعة الثانية:

2/ إيجـــاد المســـافة بـــين المتغيـــرينdi مـــن خـــلال إيجـــاد الفـــرق بـــين متوســـطي كـــل متغيرين من المجموعتين.

$$d_i = \bar{X}_{i(1)} - \bar{X}_{i(2)}\ , i = 1,2,3 \dots, k$$

$$d_i = \bar{X}_{1(1)} - \bar{X}_{1(2)}$$
$$d_i = \bar{X}_{2(1)} - \bar{X}_{2(2)}$$
$$\vdots$$
$$d_k = \bar{X}_{k(1)} - \bar{X}_{k(2)}$$

يتم وضع هذه الفروق في شكل شعاع عمودي ويرمز له بالرمز d

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix}$$

3/ حساب مجموع المربعات لكل متغير في كل مجموعة و كذلك مجموعة حاصل ضرب كل متغيرين داخل كل مجموعة كالآتي:

$$SS_{x_1}^{(1)} = SS_{11}^{(1)} = \sum_{i=1}^{n1} X_{1i}^2 - \frac{\left(\sum_{i=1}^{n1} X_{1i}\right)^2}{n1}$$

$$\vdots$$

$$SS_{kk}^{(1)} = \sum_{i=1}^{n1} X_{ki}^2 - \frac{\left(\sum_{i=1}^{n1} X_{ki}\right)^2}{n_1}$$

نتابع الحساب بحيث

$$SS_{ii}^{(1)} = \sum_{i=1}^{n1} X_{ii}^2 - \frac{\left(\sum_{i=1}^{n1} X_{ii}\right)^2}{n_1} \quad i = 1,2, \dots \dots, n1$$

يمثل مجموع مربعات المتغير (i) في المجموعة الأولي

بنفس الطريقة للمجموعة الثانية بحيث

$$SS_{ii}^{(2)} = \sum_{i=1}^{n2} X_{ii}^2 - \frac{\left(\sum_{i=1}^{n2} X_{ii}\right)^2}{n_2} \quad i = 1,2, \dots \dots, n2)$$

يمثل مجموع مربعات المتغير (i) في المجموعة الثانية

ويحسب حاصل ضرب المتغيرين بالعلاقة التالية:

$$S_{ii} = \sum X_{ij} - \frac{(\sum X_i)(\sum X_j)}{n}$$

4/ حساب مصفوفة التباين و التباين المشترك للمجموعتين

- حساب التباين

$$V_{ii} = \frac{S_{i(1)} + S_{i(2)}}{n_1 + n_2 - 2}$$

- حساب التغاير

$$V_{ii} = \frac{S_{ij(1)} + S_{ij(2)}}{n_1 + n_2 - 2}$$

- حساب مصفوفة التباين و التباين (S_{p}^2) المشترك داخل المجموعتين

$$S_p^2 = \frac{[(n_1 - 1)S_{(1)} + (n_2 - 1)S_{(2)}]}{n_1 + n_2 - 2}$$

وعليه من المعادلات الأخيرة فإن مصفوفة التباين و التباين المشترك تكون بالشكل التالي:

$$V = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1k} \\ V_{21} & V_{22} & \cdots & V_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ V_{k1} & V_{k2} & \cdots & V_{kk} \end{bmatrix}$$

ومـــن المعلـــوم أن مصـــفوفة التبـــاين والتبـــاين المشـــترك مربعـــة و متماثلـــة قطرهـــا الرئيسي يمثل التباينات و عناصرها الأخرى تمثل التغايرات مشتركة.

5/ حساب قيم معاملات الدالة التميزية الخطية

تحسب معاملات الدالة التميزية الخطية من المعادلة التالية

$$\begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1k} \\ V_{21} & V_{22} & \cdots & V_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ V_{k1} & V_{k2} & \cdots & V_{kk} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix}$$

وعليه فإن معاملات الدالة التميزية الخطة تكتب كالتالي:

$$\alpha = V^{-1} d$$

5. الاهمية النسبية للمتغيرات

اهميـــة التحليـــل التمييـــزي هـــو أنـــه يســـمح بالمقارنـــة بـــين المتغيـــرات المســـتقلة مـــن حيـــث أهميتهـــا فـــي عمليـــة التمييـــز ، و تحســـب الاهميـــة النســـبية للمتغيـــرات مـــن الصيغة التالية:

$$\alpha_j^* = \alpha_j \sqrt{V_{jj}}$$

بحيث: $j = 1,2, \dots \dots \dots \dots \dots . . , k$

نقـــوم بمقارنـــة القـــيم الناتجـــة مـــن حســـاب α_j^* وتـــرتيهم تنازليـــا ، وعليـــه فـــإن أكبـــر قيمـــة مـــن جملـــة القيمـــة هـــو أهـــم متغيـــر لـــه القـــدرة علـــى عمليـــة التمييـــز بـــين المجموعتين ويليها ثاني أكبر قيمة له القدرة على التمييز إلى اخر قيمة.

6. اختبار قدرة دالة التمييز في التمييز بين مجموعتين:

بعـــد حســـاب الدالـــة التمييزيـــة فـــأن اختبـــار قـــدرتها علـــى التمييـــز بـــين مجمـــوعتين تعتبـــر مرحلـــة ذات أهميـــة بالغـــة فـــي التحليـــل التمييـــزي ،ويـــتم اختبـــار قـــدرة الدالـــة على التمييز بين مجموعتين بإختبار ستيودنت t أو اختبار فيشر F:

1.6. إختبار ستيودنت t لقدرة دالة التمييزية

1/ نحســـب القـــيم التمييزيـــة لكـــل مشـــاهدة فـــي كـــل مجموعـــة و ذلـــك بتعـــويض قـــيم المتغيرات Xi المستقلة.

نحسب القيمة التمييزية للمشاهدة الأولي في المجموعة الأولي :

$$L_{1(1)} = \alpha_1 X_{11} + \alpha_2 X_{21} + \cdots + \alpha_k X_{k1}$$

نحسب القيمة التمييزية للمشاهدة الثانيةفي المجموعة الأولي :

$$L_{2(1)} = \alpha_1 X_{12} + \alpha_2 X_{22} + \cdots + \alpha_k X_{k2}$$

بنفس الطريقة لباقي مشاهدات المجموعة الاولى

2/ نقوم بحساب الوسط الحسابي للقيم التمييزية لكل مجموعة :

- الوسط الحسابي للقيم التمييزية للمجموعة الاولى

$$\bar{L}_{(1)} = \frac{\sum_{i=1}^{n1} L_i^{(1)}}{n_1}$$

حيث $i = 1,2,3 \ldots\ldots.n_1$

- الوسط الحسابي للقيم التمييزية للمجموعة الثانية

$$\bar{L}_{(2)} = \frac{\sum_{i=1}^{n2} L_i^{(2)}}{n_2}$$

حيث $i = 1,2,3 \ldots\ldots.n_2$

نقوم باختبار الفرق بين متوسط المجموعتين ، ونستخدم لذلك الأختبار الإحصائي (t) للمقارنة بين المتوسطات الحسابية للمجموعات و ذلك لبيان أهمية دالة التصنيف ، ونقوم بإجراء الاختبار الإحصائي كالتالي:

بحيث فرضيات الاختبار تشكل كالتالي:

$$\begin{cases} H_0: \text{الدالة ليس لها القدرة على عملية التميز} \\ H_1: \text{الدالة لها القدرة على عملية التمييز} \end{cases}$$

ويتم اختبار هذه الفرضية من خلال اختبار الفرق بين متوسطي القيم التميزية للمجموعتين ، وتكتب الفرضية كالتالي:

$$\begin{cases} H_0: ML_1 = ML_2 \\ H_1: ML_1 \neq ML_2 \end{cases}$$

$$t = \frac{\bar{L}_{(1)} - \bar{L}_{(2)}}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

نقارن قيمة t المحسوبة مع الجدولية $t_{(n_1+n_2-2,\frac{\alpha}{2})}$

من الفرضية الصفرية متوسط القيمة التمييزية للمجموعة الأولى لا يختلف إحصائياً عن متوسط القيم التمييزية للمجموعة الثانية مع العلم بأن القيم التمييزية للمجموعتين حسبت من دالة واحدة ، فإذا قبلت هذه الفرضية فذلك يعني أن نمط القيم التمييزية في المجموعتين متشابهة و هذا يعن عدم قدرة الدالة التمييزية على التمييز ، أما إذا رفضة الفرضية الصفرية فهذا يعني قدرة الدالة التمييزية على التميز.

2.6. اختبار فيشر F لقدرة دالة التمييزية:

يتم صياغة الفرضيات كالتالي:

$$\begin{cases} H_0: \text{الدالة ليس لها القدرة على عملية التمييز} \\ H_1: \text{الدالة لها القدرة على عملية التمييز} \end{cases}$$

ويتم اختبار هذه الفرضية من خلال اختبار الفرق بين متوسطي القيم التميزية للمجموعتين ، وتكتب الفرضية كالتالي:

$$\begin{cases} H_0: ML_1 = ML_2 \\ H_1: ML_1 \neq ML_2 \end{cases}$$

نقوم بتكوين جدول تحليل التباين كالتالي:

S.O.V	d.f	S.S	M.S	F
Between X	k-1	SSB	MSB	$\frac{MSB}{MSE}$
Within X	n_1+n_2-k-1	SSE	MSE	
Total	n_1+n_2-2			

ونقوم بالحسابات التالية

- مجموع المربعات داخل المتغيرات

$$SSE = D^2$$

$$D^2 = \alpha_1 d_1 + \alpha_2 d_2 + \cdots + \alpha_k d_k$$

- مجموع المربعات بين المتغيرات المستقلة

$$SSB = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} (D^2)^2)$$

- مجموع المربعات الكلي

$$SST = SSB + SSE$$

نجد قيمة F المحسوبة كالآتي

$$F = \frac{n_1 + n_2 - m - 1}{(n_1 + n_2 - 2)m} \quad T^2$$

بحيث m تمثل درجة حرية المتغيرات

أو من

$$F = \frac{MSB}{MSE} = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D^2$$

بقارنـــة قيمـــة F المحســـوبة مـــع الجدوليـــة $F_{(n_1+n_2-k-1,\alpha)}$ فـــإذا كانـــت القيمـــة المحســـوبة أكبـــر مـــن الجدوليـــة نـــرفض الفرضـــية الصـــفرية ونقبـــل الفرضـــية البديلة أي أن الدالة التمييزية لها القدرة على التمييز.

7. النقطة الفاصلة

النقطـــة الفاصـــلة تمثـــل الحـــد الفاصـــل الـــذي يفصـــل بـــين المجمـــوعتين، وتســـتخدم لغرض تصنيف المشاهدات إلى المجموعة الأقرب لها ،

فـــإذا كانـــت قيمـــة الدالـــة بعـــد تعـــويض قيمـــة المشـــاهدة فيهـــا أكبـــر مـــن الصـــفر فالمشـــاهدة تصـــنف فـــي المجموعـــة الأولـــى ، أمـــا إذا كانـــت قيمـــة الدالـــة أصـــغر مـــن الصـــفر فيـــتم تصـــنيف المشـــاهدة فـــي المجموعـــة الثانيـــة ، ولكـــن عنـــدما تكـــون الدالـــة التمييزيـــة للمجموعـــة الأولـي أكبـــر مـــن قيمـــة الدالـــة التمييزيـــة فـي المجموعـــة الثانية فإن المشاهدة تصنف في المجموعة الأولى كما يلي:

$$L > \frac{1}{2}\left(\bar{L}_1^{(1)} + \bar{L}_2^{(2)}\right)$$

تصنف المشاهدة في المجموعة الثانية إذا:

$$L < \frac{1}{2}\left(\bar{L}_1^{(1)} + \bar{L}_2^{(2)}\right)$$

بفرض أن:

$$\alpha_0 = \frac{1}{2}\left(\bar{L}_1^{(1)} + \bar{L}_2^{(2)}\right)$$

للحصـــول علـى دالـــة التصـــنيف L* نقـــوم بـــدمج النقطـــة الفاصـــلة 0α مـــع الدالـــة التميزية فنحصل على:

$$L^* = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \cdots + \alpha_k x_k$$

يتم تصنيف المشاهدات إلى المجموعات حسب العلاقات التالية:

$$\begin{cases} L^* > \text{تصنف المشاهد في المجموعة الاولى} \\ L^* < \text{تصنف المشاهد في المجموعة الثانية} \\ L^* = 0 \ \text{لا يمكن إجراء التصنيف} \end{cases}$$

8. نسبة الخطأ:

عنـــد تحديـــد النقطـــة الفاصـــلة بـــين المجمـــوعتين فإنـــه قـــد يكـــون هنـــاك تصـــنيف غيـــر صـــحيح عنـــد اســـتعمال دالـــة L* التمييـــز وقـــد تصـــنف مشـــاهدة معينـــة فـــي المجموعة الاولى بينما تعود في الحقيقة إلى المجموعة الثانية والعكس.وهناك نوعان من أخطاء التصنيف

- خطأ التصنيف الظاهري:

و يحسب من جدول التصنيف التالي:

المجموعة	تابع المجموعة الاولى	تابع المجموعة الثانية	مجموع
الاولى	n_{11}	n_{12}	n_1
الثانية	n_{21}	n_{22}	n_2

n_{11}: عدد المشاهدات من المجموعة الأولى والتي تم تصنيفها في نفس المجموعة وبالتالي صنفت بطريقة صحيحة

n_{12}: عدد المشاهدات من المجموعة الأولى والتي تم تصنيفها خطأ في المجموعة الثانية

n_{21}: عدد المشاهدات التي تنتمي إلى المجموعة الثانية وتم تصنيفها خطأ في المجموعة الأولى

n_{22}: عدد المشاهدات في المجموعة الثانية التي تم تصنيفها في نفس المجموعة و صنفت بطريقة صحيحة

ونحسب الخطأ الظاهري كما يلي:

P_{12}: نسبة المشاهدات التي تنتمي للمجموعة الأولى وصنفت خطأ في المجموعة الثانية

$$P_{12} = \frac{n_{12}}{n_1}$$

P_{21}: نسبة المشاهدات التي تنتمي للمجموعة الثانية وصنفت خطأ في الأولى

$$P_{21} = \frac{n_{21}}{n_2}$$

ونحسب معدل الخطأ الظاهري كالتالي:

$$\frac{n_{12} + n_{21}}{n_1 + n_2}$$

- خطأ التصنيف الحقيقي

ويعتمد في تحديد احتمال خطأ التصنيف على D^2 إحصائية مهلونوبيس (Mahalnobis)

$$P_1 = P_2 = F\left(-\frac{\sqrt{D^2}}{2}\right)$$

وكلما كانت صغيرة D^2 فهو دليل على صغر خطأ التصنيف والكفاءة دالة التمييز والعكس.

المحور الرابع: أمثلة تطبيقية

1. مثال تطبيقي للتحليل بطريقة مركبات الأساسية باستخدام SPSS

كلف مختبر صيدلاني بإجراء مسح لتحليل معايير اختيار معجون الأسنان.

طُلب من 40 شخص تقييم أهمية الخصائص التالية المنسوبة إلى معجون الأسنان:

X1= الجنس

X2= العمر

X3=يعطي نفسا لطيفا

X4=يبيض الأسنان

X5=يقوي اللثة

X6=مذاق جيد

X7=يمنع التسوس

الجدول التالي يبين معطيات المبحوثين:

	الجنس	العمر	يعطي نفسا لطيفا	يبيض الأسنان	يقوي اللثة	مذاق جيد	يمنع التسوس
1	1	3	1	1	3	1	2
2	2	1	2	2	1	2	1
3	4	5	4	4	5	4	2
4	1	2	1	4	1	1	1
5	5	1	5	5	1	5	2
6	2	2	2	2	2	2	1

7	3	2	3	3	2	3	1
8	4	1	4	4	1	4	1
9	1	2	3	3	2	3	1
10	5	2	1	5	2	5	1
11	4	5	4	1	5	3	2
12	2	5	3	2	2	3	1
13	1	3	1	1	3	1	2
14	3	4	3	3	4	3	2
15	4	5	4	4	5	4	2
16	2	4	2	2	4	2	2
17	1	1	1	1	1	1	1
18	2	3	2	2	5	2	2
19	3	2	3	3	2	3	1
20	4	2	2	2	2	2	1
21	2	1	2	2	1	2	1
22	2	2	2	2	2	2	1
23	3	2	1	1	1	1	1
24	2	3	2	2	3	2	2
25	5	2	5	5	2	5	1
26	2	1	2	2	1	2	1
27	2	2	2	2	2	2	1
28	5	2	3	5	2	3	1
29	2	1	2	2	1	2	1
30	3	2	3	3	2	3	1
31	4	2	4	4	2	4	1
32	5	1	5	5	1	5	1
33	4	3	4	4	2	4	1
34	1	2	1	1	2	1	2
35	3	3	2	3	1	3	2
36	2	2	2	2	2	2	2
37	1	2	1	1	3	1	1
38	5	2	3	5	2	3	2
39	1	3	1	1	3	1	2
40	1	1	1	1	2	1	1

بعد عملية إدخال بيانات الجدول السابق في برنامج SPSS نحصل على المخرجات التالية:

جدول الإحصاء الوصفي Statistiques descriptives

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
نفس_منعش	2.25	1.214	40
بيض_الاسنان	2.68	1.385	40

يقوي_اللثة	2.48	1.240	40
مذاق_جيد	2.35	1.189	40
يمنع_التسوس	2.73	1.396	40

يقدم وصف للمتغيرات المستعملة من خلال مؤشرات الاحصاء الوصفي والمتمثلة في المتوسط و الانحراف المعياري. ونلاحظ أن متوسط الافراد الذين يعتقدون أن معجون الاسنان هذا يعطي نفس منعش يساوي 2.25 بإنحراف معياري يساوي 1.214، وايضا الذين يرون أنه يمنع التسوس2.73 بتشتت 1.396.

جدول مصفوفة الإرتباط Matrice de corrélation

Correlation Matrix[a]

		نفس_منعش	يبض_الاسنان	يقوي_اللثة	مذاق_جيد	يمنع_التسوس
Correlation	نفس_منعش	1.000	-.133-	.106	.808	.011
	يبض_الاسنان	-.133-	1.000	.690	-.069-	.802
	يقوي_اللثة	.106	.690	1.000	.163	.744
	مذاق_جيد	.808	-.069-	.163	1.000	.060
	يمنع_التسوس	.011	.802	.744	.060	1.000
Sig. (1-tailed)	نفس_منعش		.206	.257	.000	.472
	يبض_الاسنان	.206		.000	.335	.000
	يقوي_اللثة	.257	.000		.158	.000
	مذاق_جيد	.000	.335	.158		.358
	يمنع_التسوس	.472	.000	.000	.358	

a. Determinant = .046

نذكر هنا أن مصفوفة الارتباط هي مصفوفة متناظرة وبالتالي يمكن قراءة قيم الارتباط بين المتغيرات إما من المثلث العلوي أو السفلي.

كما أن الجدول يعطي المعنوية لكل قيمة ارتباط من خلال الجزء Sig. (1-tailed).

مـــن مصـــفوفة الارتبـــاط نلاحـــظ أن عـــدد قـــيم الارتبـــاط الاكبـــر مـــن 0.5 يســـاوي 5 وهـــو متســـاوي مـــع عـــدد قـــيم الارتبـــاط الاقـــل مـــن 0.5 وبالتـــالي يمكـــن متابعـــة التحليل.

ونلاحـــظ أن قيمـــة الارتبـــاط بـــين (نفـــس مـــنعش) والمتغيـــر (مـــذاق جيـــد) يســـاوي 0.808 وهـــو ارتبـــاط قـــوي حيـــث يجتمـــع أفـــراد العينـــة علـــى أن الـــنفس المـــنعش للمعجـــون لابـــد أن يرافقـــه المـــذاق الجيـــد، كمـــا نلاحـــظ ان قيمـــة (Sig.(1-tailed تســـاوي 0.00 وهـــي أكبـــر مـــن 0.05 مســـتوى المعنويـــة المفتـــرض، وبالتـــالي فإننـــا نقبـــل الفرضـــية الصـــفرية ونـــرفض الفرضـــية البديلـــة،اي أن الارتبـــاط بـــين المتغيـــر (نفس منعش) والمتغير (مذاق جيد) معنوي.

ونلاحـــظ أن قيمـــة الارتبـــاط بـــين (يـــبض الاســـنان) والمتغيـــر (يقـــوي اللثـــة) يســـاوي 0.690 وهـــو ارتبـــاط جيـــد حيـــث يجتمـــع أفـــراد العينـــة علـــى أن عمـــل المعجـــون فـــي تبـــيض الاســـنان مـــن شـــأنه تقويـــة اللثـــة، كمـــا نلاحـــظ ان قيمـــة (Sig.(1-tailed تســـاوي 0.00 وهـــي أكبـــر مـــن 0.05 مســـتوى المعنويـــة المفتـــرض، وبالتـــالي فإننـــا نقبـــل الفرضـــية الصـــفرية ونـــرفض الفرضـــية البديلـــة،اي أن الارتبـــاط بـــين المتغيـــر (يبض الاسنان) والمتغير (يقوي اللثة) معنوي.

ونلاحـــظ أن قيمـــة الارتبـــاط بـــين المتغير(يمنـــع التســـوس) والمتغيـــر (يـــبض الاســـنان) والمتغيـــر (يقـــوي اللثـــة) يســـاوي 0.802 و0.744 علـــى التـــوالي وهـــي ارتباطـــات جيـــدة حيـــث يجتمـــع أفـــراد العينـــة علـــى أن عمـــل المعجـــون فـــي القضـــاء فـــي منـــع التســـوس يســـاهم فـــي تبـــيض الاســـنان و تقويـــة اللثـــة، كمـــا نلاحـــظ ان قيمـــة (Sig.(1-tailed تســـاوي 0.00 وهـــي أكبـــر مـــن 0.05 مســـتوى المعنويـــة المفتـــرض، وبالتـــالي فإننـــا نقبـــل

الفرضـــية الصـــفرية ونـــرفض الفرضـــية البديلـــة،اي أن الارتبـــاط بـــين المتغيـــر (يمنـــع التسوس) والمتغير (يبيض الاسنان) والمتغير (يقوي اللثة) معنوي.

كمـــا يشـــير محـــدد المصـــفوفة الـــذي يســـاوي 0.046 وهـــوأكبر مـــن 0.00001 إلـــى عـــدم وجـــود مشـــكل الارتبـــاط الـــذاتي وجـــود الحـــد الادنـــى مـــن البيانـــات أي أن المصفوفة ليست مصفوفة فارغة.

Inverse of Correlation Matrix

	منعش_نفس	الاسنان_يبض	اللثة_يقوي	جيد_مذاق	التسوس_يمنع
منعش_نفس	2.960	.450	-.125-	-2.331-	-.162-
الاسنان_يبض	.450	3.222	-.799-	.109	-2.000-
اللثة_يقوي	-.125-	-.799-	2.511	-.291-	-1.209-
جيد_مذاق	-2.331-	.109	-.291-	2.940	-.019-
التسوس_يمنع	-.162-	-2.000-	-1.209-	-.019-	3.506

جدول مؤشر KMO واختبار بارتلت Bartlett's Test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.654
Bartlett's Test of Sphericity	Approx. Chi-Square	112.262
	df	10
	Sig.	.000

مـــن خـــلال قيـــاس مؤشـــر KMO الـــذي علـــى أساســـه نســـتدل علـــى مـــدى كفايـــة عـــدد أفـــراد العينـــة ويجـــب أن تكـــون قيمتـــه أكبـــر مـــن 0.50 حتـــى تكـــون العينـــة كافيـــة وهذا شرط أساسي يجب تحقيقه،

ومـــن الملاحـــظ أن مؤشـــر KMO يســـاوي0.654 وبتـــالى فـــإن حجـــم العينـــة كـــافي لإجراء الدراسة.

اختبـــــار Bartlett لدائريــــة Sphericity فهـــو مؤشـــر للعلاقـــة بـــين المتغيـــرات ويجـــب أن يكـــون مســـتوى الدلالـــة لهـــذه العلاقـــة أقـــل مـــن 0.05 وذلـــك حتـــى نســـتطيع التأكيـــد علـــى أن هـــذه العلاقـــة دالـــة إحصـــائيا. ويظهـــر مـــن الجـــدول أن قيمـــة إختبـــار Bartlett تســـاوي 112.262 وهـــي دالـــة احصـــائيا حيـــث أن .Sig أقـــل مـــن 0.05 ومنـــه نـــرفض الفرضـــية الصـــفرية و نقبـــل الفرضـــية البديلـــة أي أنـــه يوجـــد معاملات ارتباط غير معدومة.

Anti-image Matrices

		منعش_نفس	الاسنان_يبض	اللثة_يقوي	جيد_مذاق	التسوس_يمنع
Anti-image Covariance	منعش_نفس	.338	.047	-.017-	-.268-	-.016-
	الاسنان_يبض	.047	.310	-.099-	.011	-.177-
	اللثة_يقوي	-.017-	-.099-	.398	-.039-	-.137-
	جيد_مذاق	-.268-	.011	-.039-	.340	-.002-
	التسوس_يمنع	-.016-	-.177-	-.137-	-.002-	.285
Anti-image Correlation	منعش_نفس	.512[a]	.146	-.046-	-.790-	-.050-
	الاسنان_يبض	.146	.715[a]	-.281-	.035	-.595-
	اللثة_يقوي	-.046-	-.281-	.805[a]	-.107-	-.408-
	جيد_مذاق	-.790-	.035	-.107-	.519[a]	-.006-
	التسوس_يمنع	-.050-	-.595-	-.408-	-.006-	.697[a]

a. Measures of Sampling Adequacy(MSA)

جدول جودة التمثيل Communalities

Communalities

	Initial	Extraction
نفس_منعش	1.000	.903
يبض_الاسنان	1.000	.856
يقوي_اللثة	1.000	.807
مذاق_جيد	1.000	.901
يمنع_التسوس	1.000	.871

Extraction Method: Principal Component Analysis.

من جدول نوعية التمثيل فنلاحظ أن تمثيل كل المتغيرات أكبر من 0.4 وبالتالي كل المتغيرات تدخل في الدراسة و لا يتم استبعاد أي متغيرة.

كما نلاحظ أن كل المتغيرات ذات تمثيل عالي حيث كانت نسبة الاستخراج الاصغر 0.807 أي 80.70% للمتغير (يقوي اللثة).

جدول التباين الكلي المشروح Total Variance Explained

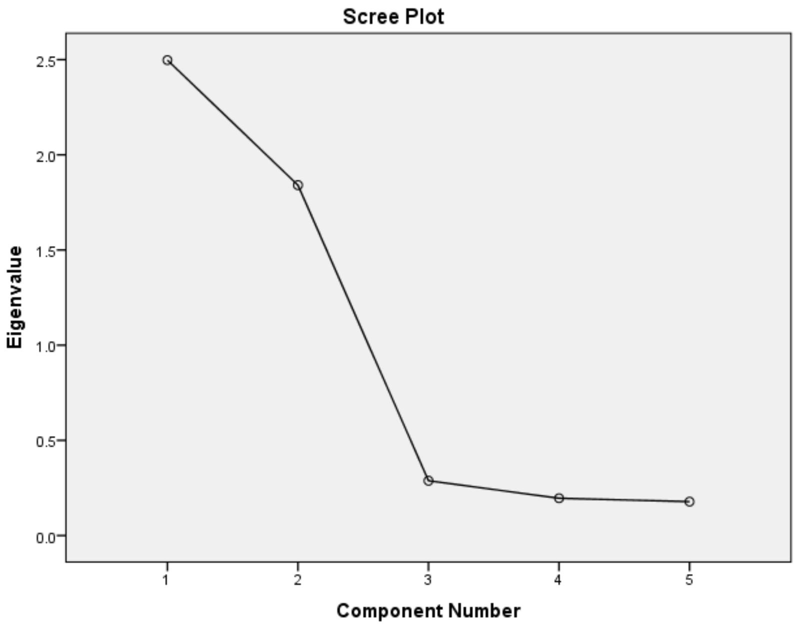
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.498	49.950	49.950	2.498	49.950	49.950	2.493	49.863	49.863
2	1.841	36.813	86.763	1.841	36.813	86.763	1.845	36.900	86.763
3	.288	5.759	92.522						
4	.196	3.919	96.441						
5	.178	3.559	100.000						

Extraction Method: Principal Component Analysis.

نلاحظ من الجدول أنه توجد قيمتين ذاتيتين أكبر من الواحد، وعليه سيتم اعتماد هاتين القيمتين وذلك حسب محك كايز وإقصاء باقي القيم الذاتية، وبالتالي يتم اختزال البيانات في محورين حيث يختزل الاول نسبة 49.86% من البيانات ويختزل المحور الثاني 36.90% من البيانات ومعا يتحدان في اختزال ما نسبته 86.76% من مجمل البيانات وهي نسبة جد مقبولة حيث 13.34% فقط من البيانات لم نتمكن من تمثيلها.

ويمكن من الشكل الموالي إلقاء نظرة على انحدار القيم الذاتية



حيــث نلاحــظ أن بعــد القيمتــين الاولــى والثانيــة ينحــدر خــط القــيم الذاتيــة بشــكل جــد معنــوي ثــم يأخــذ الشــكل الافقــي مــع القــيم الثلاثــة المتبقيــة، وهــو يــدعم اعتمادنا في عملية التحليل على القيمتين الاول والثانية.

احداثيات المتغيرات قبل وبعد التدوير:

مصفوفة الاحداثيات قبل التدوير:

Component Matrix[a]

	Component	
	1	2
نفس_منعش		.949
يبض_الاسنان	.900	
يقوي_اللثة	.895	
مذاق_جيد		.941
يمنع_التسوس	.932	

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

المحـــــــــــــور الاول: نلاحـــظ أن أن المحـــور الاول يحتـــوي علـــى المتغيـــرات يبـــيض الاســـنان والمتغيـــر يقـــوي اللثـــة والمتغيـــر يمنـــع التســـوس، وهـــي متغيـــرات تعكـــس الجانـــب الصـــحي لمعجـــون الاســـنان وبالتـــالي يمكـــن أن نســـمي هـــذا المحـــور بـــ عوامـــل صحية.

المحـــــــــــور الثـــــــــاني: نلاحـــظ أن أن المحـــور الثـــاني يحتـــوي علـــى المتغيـــرات نفـــس مـــنعش، ومـــذاق جيـــد وهـــذه المتغيـــرات تعكـــس الجوانـــب الحســـية وبالتـــالي يمكـــن أن نسمي الحور بعوامل حسية

Reproduced Correlations

		منعش_نفس	الاسنان_يبض	اللثة_يقوي	جيد_مذاق	التسوس_يمنع
Reproduced Correlation	منعش_نفس	.903[a]	-.151-	.127	.900	.003
	الاسنان_يبض	-.151-	.856[a]	.788	-.090-	.850
	اللثة_يقوي	.127	.788	.807[a]	.186	.830
	جيد_مذاق	.900	-.090-	.186	.901[a]	.065
	التسوس_يمنع	.003	.850	.830	.065	.871[a]
Residual[b]	منعش_نفس		.018	-.020-	-.092-	.009
	الاسنان_يبض	.018		-.099-	.021	-.048-
	اللثة_يقوي	-.020-	-.099-		-.023-	-.085-
	جيد_مذاق	-.092-	.021	-.023-		-.005-
	التسوس_يمنع	.009	-.048-	-.085-	-.005-	

Extraction Method: Principal Component Analysis.

a. Reproduced communalities

b. Residuals are computed between observed and reproduced correlations. There are 3 (30.0%) nonredundant residuals with absolute values greater than 0.05.

مصفوفة الاحداثيات بعد التدوير:

Rotated Component Matrix[a]

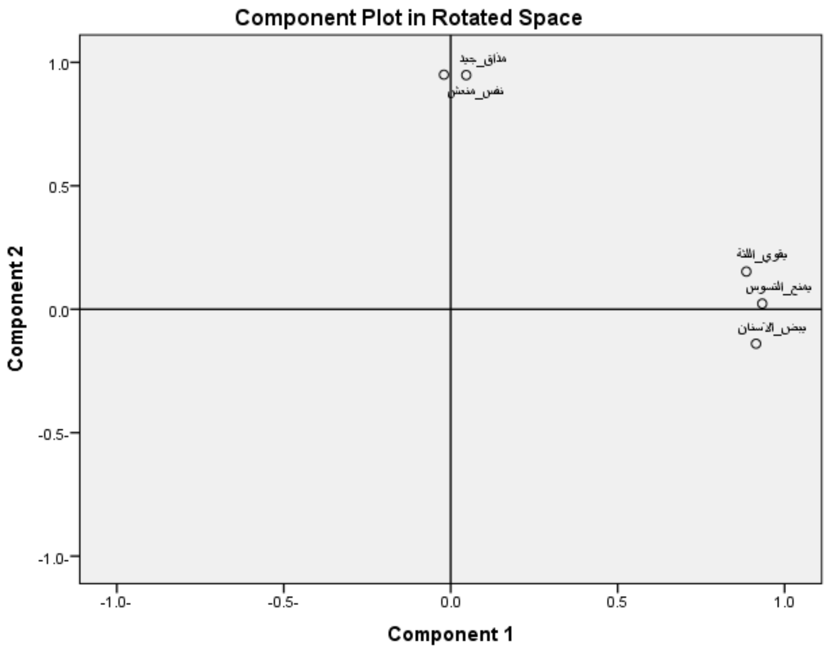
	Component	
	1	2
منعش_نفس		.950
الاسنان_يبض	.914	
اللثة_يقوي	.885	

جيد_مذاق		.948
التسوس_يمنع	.933	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

تسمح عملية التدوير بإختيار الموقع الانسب لتمثيل المتغيرات، ونلاحظ أنه بعد عملية التدوير تعزز تمثيل المحاور ولو بشكل طفيف

شكل المكونات بعد التدوير

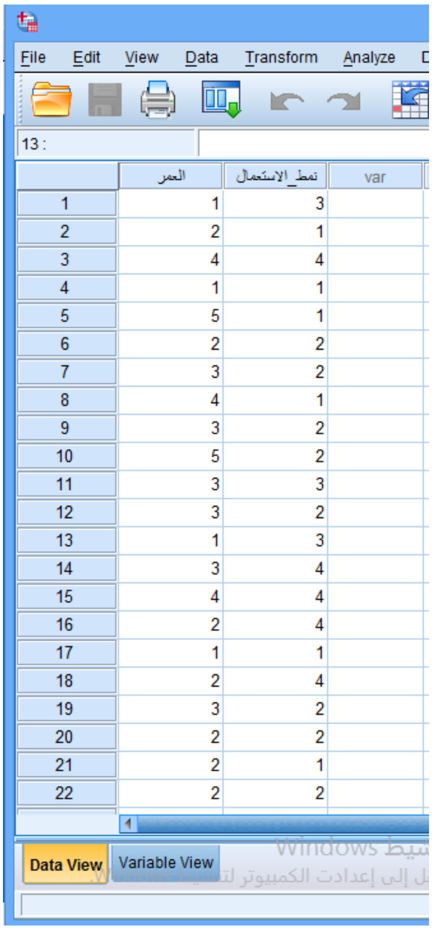


نلاحظ من الشكل السابق أن المتغيرات تنقسم إلى مجموعتين متمايزتين الاولى على المحور الاول والتي تظم ثلاثة متغيرات والمجموعة الثانية ممثلة على المحور الثاني.

2. **مثال تطبيق التحليل** العاملي المتناظر باستخدام SPSS:

تـم طـرح سـؤالين علـى مجموعـة مـن المنتسـبين إلـى بريـد الجزائـر حـول طريقـة استـخدامهم للبطاقـة الـدفع الالكترونـي "الذهبيـة " وتمحـورت الاسـئلة حـول الفئة العمرية ونمط الاستخدام.

بعد ادخال البيانات في برنامج SPSS كالتالي:



	العمر	نمط_الاستعمال	var
1	1	3	
2	2	1	
3	4	4	
4	1	1	
5	5	1	
6	2	2	
7	3	2	
8	4	1	
9	3	2	
10	5	2	
11	3	3	
12	3	2	
13	1	3	
14	3	4	
15	4	4	
16	2	4	
17	1	1	
18	2	4	
19	3	2	
20	2	2	
21	2	1	
22	2	2	

نحصل على المخرجات التالية:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
الاستعمال_نمط * العمر	85	100.0%	0	0.0%	85	100.0%

نلاحظ أنه تم معالجة 85 حالة ولا توجد حالات مفقودة أو غير مدرجة.

جدول اختبار التوافق:

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Nominal by Nominal	Phi	.733			.000
	Cramer's V	.423			.000
	Contingency Coefficient	.591			.000
Ordinal by Ordinal	Kendall's tau-b	-.011-	.093	-.118-	.906
	Kendall's tau-c	-.011-	.091	-.118-	.906
	Gamma	-.014-	.120	-.118-	.906
	Spearman Correlation	-.016-	.113	-.150-	.881[c]
Interval by Interval	Pearson's R	-.010-	.108	-.090-	.929[c]
Measure of Agreement	Kappa	.050	.063	.865	.387
N of Valid Cases		85			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

c. Based on normal approximation.

نشير هنا إلى أنه لا يمكن استعمال معامل الارتباط لبيرسن لأن المتغيرات نوعية وبالتالي نلجأ الي استعمال اختبار التوافق Contingency Coefficient والذي يسمح بإختبار الفرضية التالية:

H_0: لا يوجد ارتباط بين الفئة العمرية و استعمال البطاقة الذهبية

H_1: يوجد ارتباط بين الفئة العمرية و استعمال البطاقة الذهبية

ومن الجدول نجد أن قيمة معامل التوافق تساوي 0.591 وهي قيمة مقبولة، وذات معنوية ذلك لأن مستوي الدلالة يساوي0.00 وهو أقل من 0.05 وبالتالي نرفض الفرضية الصفرية التي مفادها أنه لا يوجد ارتباط بين الفئة العمرية واستعمال البطاقة الذهبية، ونقبل الرضية البديلة التي تؤكد على وجود رتباط بين الفئة العمرية واستعمال البطاقة الذهبية.

الجدول المتقاطع

يوضح الجدول المتقاطع النسب بالنسبة للأسطر والأعمدة والعينة

العمر * الاستعمال_نمط Crosstabulation

			الاستعمال_نمط				
			بطريقة منتظمة	غالبا	في المناسبات	لا استعملها	Total
العمر	20-30	Count	3	4	12	0	19
		العمر% within	15.8%	21.1%	63.2%	0.0%	100.0%
		نمط_الاستعمال% within	27.3%	11.1%	66.7%	0.0%	22.4%
		% of Total	3.5%	4.7%	14.1%	0.0%	22.4%
	30-40	Count	4	10	3	8	25
		العمر% within	16.0%	40.0%	12.0%	32.0%	100.0%
		نمط_الاستعمال% within	36.4%	27.8%	16.7%	40.0%	29.4%
		% of Total	4.7%	11.8%	3.5%	9.4%	29.4%
	40-50	Count	1	14	3	4	22
		العمر% within	4.5%	63.6%	13.6%	18.2%	100.0%
		نمط_الاستعمال% within	9.1%	38.9%	16.7%	20.0%	25.9%
		% of Total	1.2%	16.5%	3.5%	4.7%	25.9%
	50-60	Count	1	4	0	8	13
		العمر% within	7.7%	30.8%	0.0%	61.5%	100.0%
		نمط_الاستعمال% within	9.1%	11.1%	0.0%	40.0%	15.3%
		% of Total	1.2%	4.7%	0.0%	9.4%	15.3%
	60 من أكبر	Count	2	4	0	0	6
		العمر% within	33.3%	66.7%	0.0%	0.0%	100.0%
		نمط_الاستعمال% within	18.2%	11.1%	0.0%	0.0%	7.1%
		% of Total	2.4%	4.7%	0.0%	0.0%	7.1%

Total	Count	11	36	18	20	85
	% within العمر	12.9%	42.4%	21.2%	23.5%	100.0%
	% within نمط_الاستعمال	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	12.9%	42.4%	21.2%	23.5%	100.0%

نأخذ كمثال الفئة العمرية التالية

40-50	Count	1	14	3	4	22
	% within العمر	4.5%	63.6%	13.6%	18.2%	100.0%
	% within نمط_الاستعمال	9.1%	38.9%	16.7%	20.0%	25.9%
	% of Total	1.2%	16.5%	3.5%	4.7%	25.9%
50-60	Count	1	4	0	8	13

الفئة العمرية تظم الاشخاص الذين تتراوح اعمارهم من 40 إلى 50 سنة حيث نجد:

- ✓ استخدام البطاقة الذهبية بصفة منتظمة :

- شخص واحد يستخدم البطاقة الذهبية بصفة منتظمة
- 4.5% من الفئة العمرية 40-50 يستخدمون البطاقة الذهبية بصفة منتظمة
- 9.1% من الاشخاص الذين يستعملون البطاقة الذهبية بصفة منتظمة هم من الفئة العمرية 40-50 سنة.
- 1.2% هم اشخاص من الفئة العمرية 40-50 و يستعملون البطاقة الذهبية بصفة منتظمة.

- ✓ استخدام البطاقة الذهبية في غالب الاحيان:

- 14 شخص يستخدم البطاقة الذهبية في غالب الاحيان
- 63.6% من الفئة العمرية 40-50 يستخدمون البطاقة الذهبية في غالب الاحيان

- 38.9% من الاشخاص الذين يستعملون البطاقة الذهبية في غالب الاحيان هم من الفئة العمرية 40-50 سنة.

- 16.5% هم اشخاص من الفئة العمرية 40-50 و يستعملون البطاقة الذهبية في غالب الاحيان.

تطبيق AFC بالخطوات التالية:

الجدول المعطيات الاساسية:

Correspondence Table

العمر	الاستعمال_نمط بطريقة منتظمة	غالبا	في المناسبات	لا استعملها	Active Margin
20-30	3	4	12	0	19
30-40	4	10	3	8	25
40-50	1	14	3	4	22
50-60	1	4	0	8	13
60 من أكبر	2	4	0	0	6
Active Margin	11	36	18	20	85

حيث يوضح مكونات عناصر الاسطر ومكونات عناصر الاعمدة والتكرارات .

جدول التكرارات النسبية للأسطر

Row Profiles

العمر	الاستعمال_نمط منتظمة بطريقة	غالبا	المناسبات في	استعملها لا	Active Margin
20-30	.158	.211	.632	.000	1.000
30-40	.160	.400	.120	.320	1.000
40-50	.045	.636	.136	.182	1.000
50-60	.077	.308	.000	.615	1.000
60 من أكبر	.333	.667	.000	.000	1.000
Mass	.129	.424	.212	.235	

جدول التكرارات النسبية للأعمدة

Column Profiles

العمر	الاستعمال_نمط منتظمة بطريقة	غالبا	المناسبات في	استعملها لا	Mass
20-30	.273	.111	.667	.000	.224
30-40	.364	.278	.167	.400	.294
40-50	.091	.389	.167	.200	.259
50-60	.091	.111	.000	.400	.153
60 من أكبر	.182	.111	.000	.000	.071
Active Margin	1.000	1.000	1.000	1.000	

جدول العوامل المستخرجة:

Summary

Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia: Accounted for	Proportion of Inertia: Cumulative	Confidence Singular Value: Standard Deviation	Confidence Singular Value: Correlation 2
1	.613	.376			.700	.700	.077	.446
2	.343	.118			.219	.918	.090	
3	.209	.044			.082	1.000		
Total		.538	45.689	.000[a]	1.000	1.000		

a. 12 degrees of freedom

حيث نلاحظ أنه تم تلخيص البيانات الاساسية في ثلاثة محاور المحور الاول يختزل 70% من نسبة التشتت أو من التباين الكلي، المحور الثاني يختزل 21.9% والمحور الثالث يختزل 8.2%.

ويوضح الجدول قيمت كاي مربع حيث يقيس وجود علاقة بين العمر واستخدام البطاقة من عدمه ، وقد بلغت قيمة كاي مربع 45.86 وهي قيمة مقبولة، وذات معنوية ذلك لأن مستوي الدلالة يساوي0.00 وهو أقل من

0.05 وبالتـــالي نـــرفض الفرضـــية الصـــفرية التـــي مفادهـــا أنـــه لا يوجـــد ارتبـــاط بـــين الفئـــة العمريـــة واســـتعمال البطاقـــة الذهبيـــة، ونقبـــل الرضـــية البديلـــة التـــي تؤكـــد على وجود رتباط بين الفئة العمرية واستعمال البطاقة الذهبية.

جدول خصائص الاسطر

Overview Row Points[a]

		Score in Dimension			Contribution				
					Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
العمر	Mass	1	2	Inertia	1	2	1	2	Total
20-30	.224	-1.365-	-.334-	.264	.679	.073	.967	.032	1.000
30-40	.294	.311	-.090-	.023	.046	.007	.753	.035	.788
40-50	.259	.190	.508	.052	.015	.195	.110	.441	.551
50-60	.153	1.016	-.846-	.134	.257	.319	.720	.279	1.000
60 من أكبر	.071	.129	1.405	.064	.002	.406	.011	.746	.757
Active Total	1.000			.538	1.000	1.000			

a. Symmetrical normalization

✓ المساهمة في عطالة الابعاد:

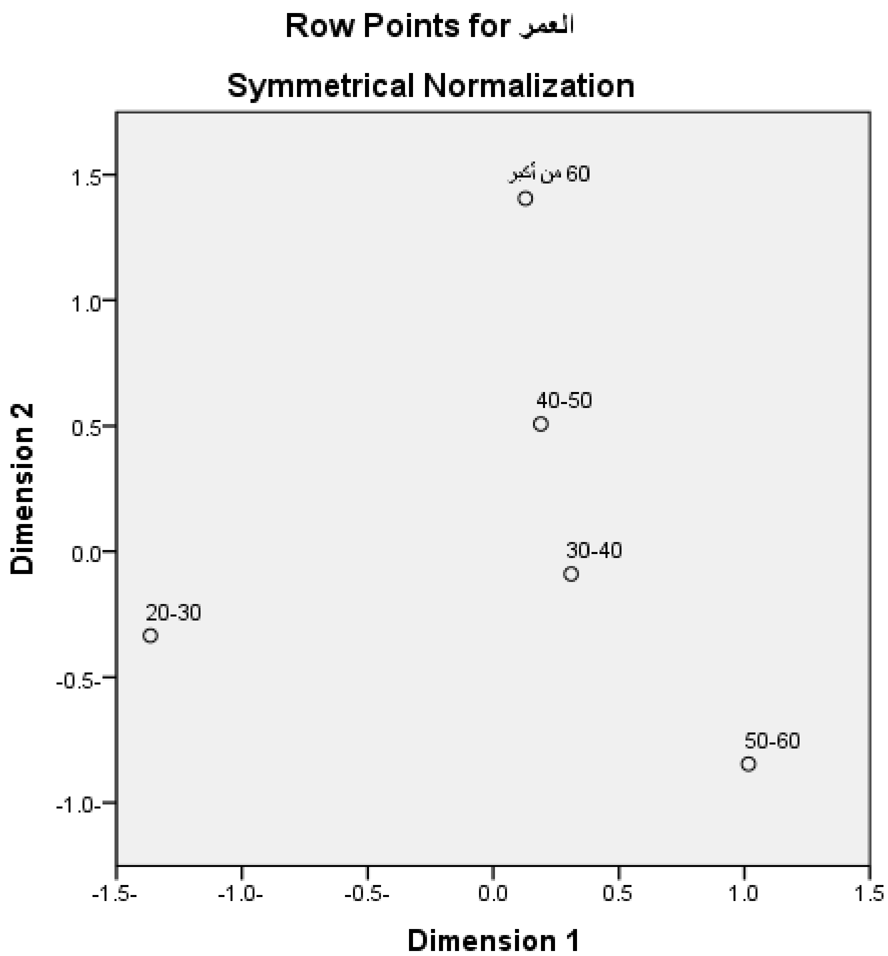
- نلاحـــظ أن أكبـــر مســـاهمة للعطالـــة فـــي المحـــور الاول هـــي للفئـــة العمريـــة 20-30سنة

- نلاحـــظ أن أكبـــر مســـاهمة للعطالـــة فـــي المحـــور الثـــاني هـــي للفئـــة العمريـــة أكبر من 60سنة

✓ المساهمة في عطالة النقاط:

- يساهم المحور الاول في عطالة 96.7% بالنسبة للفئة العمرية 20-30 و75.3% بالنسبة للفئة العمرية30-40 و72% بالنسبة للفئة العمرية 50-60 سنة.
- يساهم المحور الثاني في عطالة 74.6% بالنسبة للفئة العمرية أكبر من 60 سنة.

كما هو موضح في الشكل التالي:



-
-
-

جدول خصائص الاعمدة

Overview Column Points[a]

نمط_الاستعمال	Mass	Score in Dimension		Inertia	Contribution				
					Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
		1	2		1	2	1	2	Total
بطريقة منتظمة	.129	-.206-	.294	.044	.009	.033	.077	.088	.165
غالبا	.424	.221	.576	.067	.034	.409	.191	.722	.912
في المناسبات	.212	-1.348-	-.447-	.252	.627	.123	.936	.058	.993
لا استعملها	.235	.927	-.796-	.175	.330	.435	.708	.292	1.000
Active Total	1.000			.538	1.000	1.000			

a. Symmetrical normalization

✓ المساهمة في عطالة الابعاد:

- نلاحظ أن أكبر مساهمة للعطالة في المحور الاول هي استخدام البطاقة في المناسبات

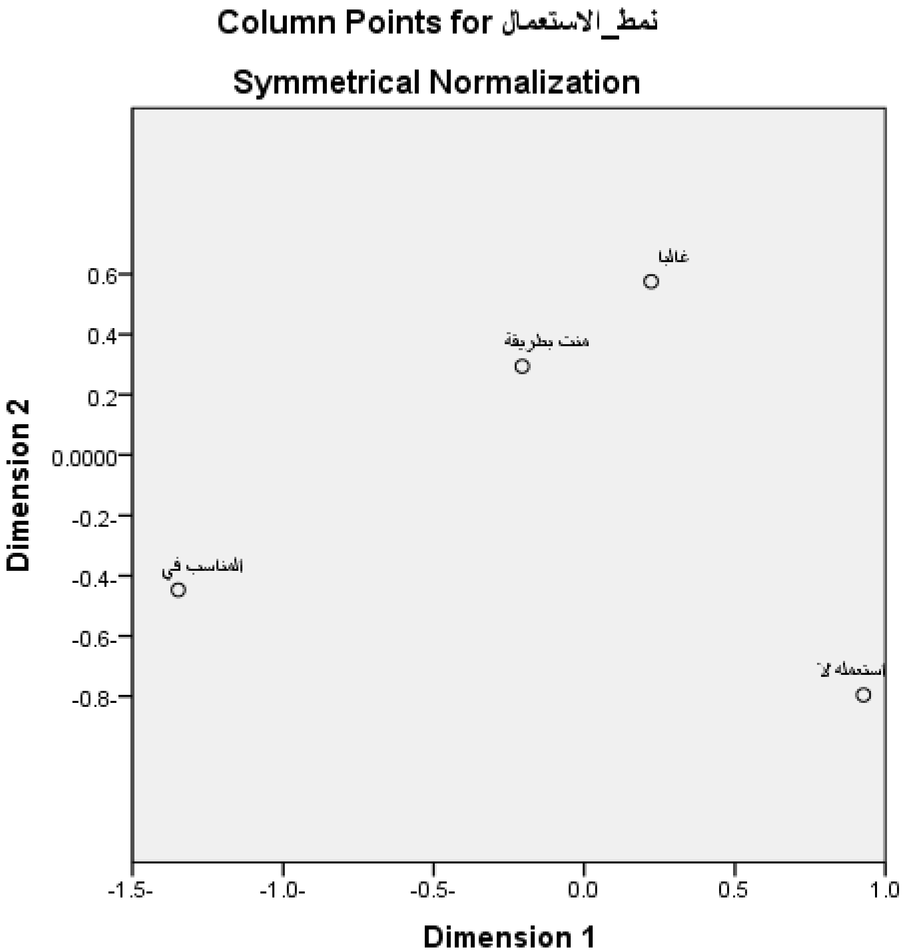
- نلاحظ أن أكبر مساهمة للعطالة في المحور الثاني عدم استخدام البطاقة

✓ المساهمة في عطالة النقاط:

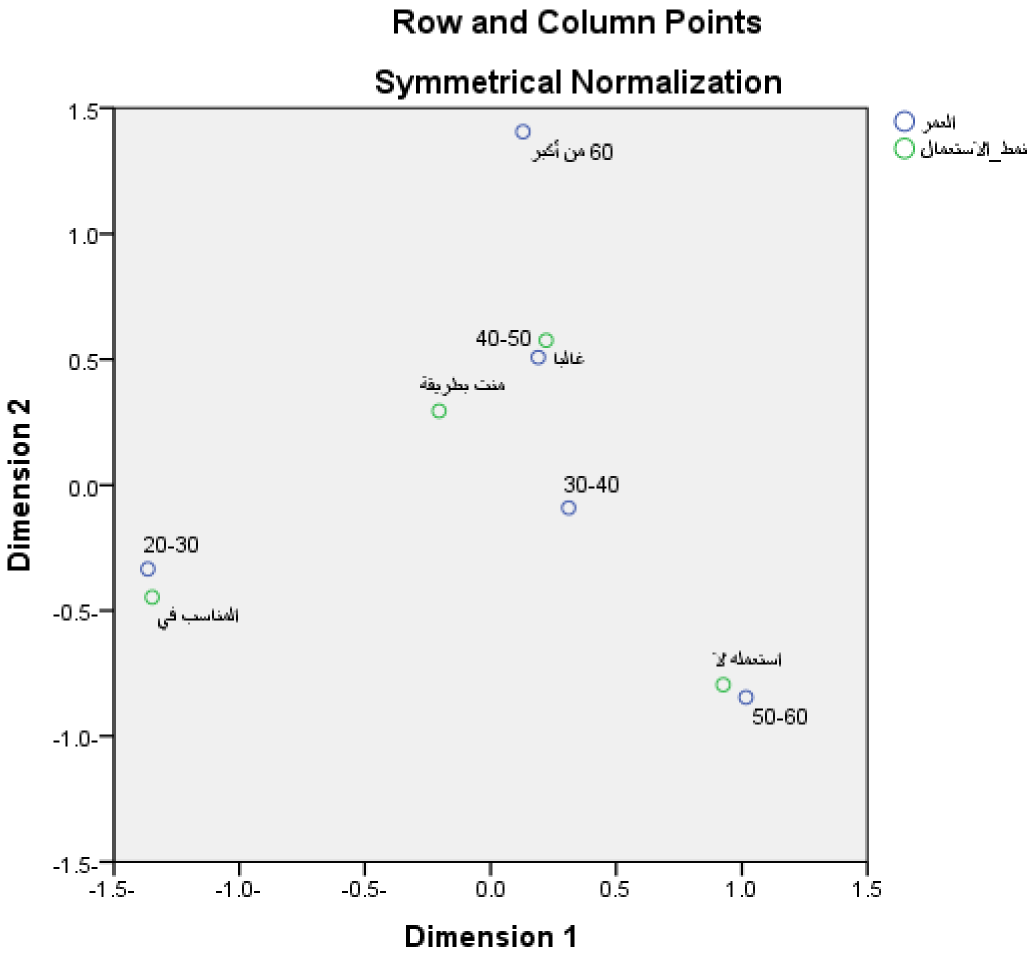
- يساهم المحور الاول في عطالة 93.6% بالنسبة لاستخدام البطاقة في المناسبات و70.8% بالنسبة لعدم استخدام البطاقة.

- يساهم المحور الثاني في عطالة 72.2% بالنسبة لاستخدام البطاقة غالبا.

كما هو موضح في الشكل التالي:



تموقع نقاط الاسطر والأعمدة:



مــن الشــكل الســابق الــذي يوضــح تموقــع نقــاط الاســطر والأعمــدة نســتنتج ثلاثــة مجموعات

- المجموعــة الاولــى: حيــث تظــم الاشــخاص الــذين لايســتعملون البطاقــة والفئة العمرية 50-60.
- المجمــوع الثانيــة: تظــم الاشــخاص الــذين يســتعملون البطاقــة فــي المناسبات والفئة العمرية 20-30 سنة.

- المجموعة الثالثة: وتضم الاشخاص الذين يستعملون البطاقة في غالب الاحيان أو بطريقة منتظمة والفئتين العمريتين 30-40 و40-50 سنة