طريقة التحليل بالمركبات الأساسية ACP:

تستخدم هذه الطريقة في تحليل الجداول الإحصائية بشرط أن تكون كل المتغيرات ذات طبيعة كمية، حيث تهدف هذه الطريقة إلى البحث عن فضاء شعاعي أقل رتبة يتكون عادة من بعدين ويسمح لنا أحسن تمثيل ويحفظ لنا أكبر كمية من المعلومات.

تعريف: تعتبر طريقة التحليل ACP من أقدم طرق تحليل المعطيات، وهي تقنية لتمثيل المعطيات والمعلومات بالاعتماد على بعض الخصائص الجبرية والهندسية، ولم تعرف هذه الطريقة تطورا في الاستعمال إلا بعد التطور الذي شهده الإعلام الآلي.

أهداف طريقة ACP:

- وصف ورسم المعطيات الموجودة في جدول يتكون من أفراد (n) ومتغيرات (p).
- عرض البيانات في فضاء ذو بعد منخفض مع المحافظة على أكبر قدر من المعلومات.
- تحديد العوامل التي تفسر على أفضل نحو تشتت المتغيرات.
- تكوين متغيرات غير مرتبطة خطيا فيما بينها انطلاقا من المتغيرات الأصلية.

شرط استعمال طريقة ACP: أن تكون كل المتغيرات ذات طبيعة كمية.

مبدأ طريقة التحليل بالمركبات الأساسية (ACP):

المبدأ الأساسي لطريقة ACP هو نقل البيانات من المعلم الأصلي ذو مبدأ O. نحو المعلم الجديد ذو المبدأ g. والتحديد المصفوفة X ثم تطبيق طريقة التحليل العاملي العام (AFG) على المصفوفة X على أن يتم استخراج القيم الذاتية من المصفوفة $\left(\frac{X'X}{n}\right)$.

طرق التحليل بالمركبات الأساسية:

1- طريقة التحليل بالمركبات الأساسية المرجحة. (ACP normée)

2- " " " " البسيطة (غير المرجحة).

طريقة التحليل بالمركبات الأساسية المرجحة:

تستعمل هذه الطريقة في حالة عدم تجانس المتغيرات، أي أن تكون متغيرات الجدول الأولي للمعطيات ليس لها نفس وحدة القياس.

الجدول الأولي للمعطيات يتكون من مجموعة من البيانات ذو n فرد و p متغير.

1- تحويل الجدول الأولي إلى الصيغة القياسية:

نقوم بحساب المتوسطات الحسابية لكل متغير حيث كلما كان المتوسط صغيرا فإنه يعبر عن السلسلة.

$$\boxed{\bar{R}_j = \frac{\sum R_{ij}}{n}}$$

-2-

حساب الانحراف المعياري بواسطة القانون التالي:

$$\sigma_R = \sqrt{\frac{1}{n}\sum (R_i-\bar{R}_j)^2}$$

إذا كان عدد الأفراد أقل من 30 نقوم بالقسمة على $(n-1)$ بدلا من (n) أي:

$$\sigma_{R_j} = \sqrt{\frac{1}{n-1}\sum (R_i-\bar{R})^2}$$

عندما يكون الانحراف المعياري أقل قيمة تكون المتغيرة أقل تشتت.

حساب المصفوفة X: يحسب عناصر المصفوفة X حسب القانون التالي:

$$X_{ij} = \frac{R_{ij}-\bar{R}_j}{\sigma_{R_j}}$$

إن الانتقال من الجدول الأولي R إلى المصفوفة X هو بسبب المعلم من المبدأ -O- إلى المبدأ -g-.

مثال 1: ليكن الجدول التالي:

المتغيرات / الأفراد	X_1	X_2	X_3
1	12	11	11
2	10	11	12
3	13	12	13
4	10	9	13
5	12	9	12
6	9	8	11
الوسط الحسابي $\bar{R}_j$	11	10	12
الانحراف المعياري σ_j	1,549	1,549	0,894

حيث: $\bar{R}_j = \frac{\sum R_i}{n}$

$$\sigma_{R_j} = \sqrt{\frac{1}{n-1}\sum(R_i-\bar{R}_j)^2}$$

- نلاحظ أن المتغير X_3 هو الذي يملك أصغر انحراف معياري إذن فهو أقل تشتت ( ) وأكثر استقرارا من المتغيرين X_1 و X_2.
- المتغيران X_1 و X_2 يملكان أكبر قيمة في انحراف المعياري أي أنهما أكثر تشتت وأقل استقرارا من المتغير X_1.

نقوم بحساب المصفوفة X وذلك بإيجاد عناصرها بواسطة القانون التالي:

$$X_{ij} = \frac{R_{ij} - \bar{R}_j}{\sigma_{R_j}}$$

بتطبيق القانون نتحصل على المصفوفة X :

$$X_{6,3} = \begin{pmatrix} 0.645 & 0.645 & -1.118 \\ -0.645 & 0.645 & 0.00 \\ 1.290 & 1.290 & 1.118 \\ -0.645 & -0.645 & 1.118 \\ 0.645 & -0.645 & 0.00 \\ -1.290 & -1.290 & -1.118 \end{pmatrix}$$

2- حساب مصفوفة الارتباطات C_{pp} :

هي مصفوفة مربعة ذات p سطر و p عمود متناظرة مجموع عناصر قطرها يساوي p ، وحسب العلاقة التالية :

$$\boxed{C_{pp} = \frac{X'X}{n-1}} = \begin{pmatrix} 1 & r_{1,2} & \cdots & r_{(1,p)} \\ \vdots & 1 & & \\ r_{(j,1)} & r_{(j,j)} & \cdots & r_{(j,p)} \\ \vdots & & & \vdots \\ r_{(p,1)} & \cdots & \cdots & 1 \end{pmatrix}$$

$$\boxed{\begin{array}{l} C_{pp} = \dfrac{X'X}{n} \\ C_{pp} = \dfrac{X'X}{n-1} \end{array}} \left\{ \begin{array}{l} \text{...... إذا كانت العينة أكبر من 30} \\ \text{" " " أصغر من 30} \end{array} \right.$$

مع :

$$-1 \leq r_{ij} \leq 1$$

-5-

بالرجوع الى المثال السابق :

$$C_{3\times3} = \frac{X'X}{n-1} = \frac{1}{5}\begin{pmatrix}5 & 3,388 & 1,422\\ 3,388 & 5 & 1,422 \\ 1,422 & 1,422 & 5\end{pmatrix}$$

$$C_{3.3} = \begin{pmatrix}1 & 0.667 & 0.289\\ 0.667 & 1 & 0.289\\ 0.289 & 0.289 & 1\end{pmatrix}$$

معاملات الارتباطات بين المتغيرات .

$$\left\{\begin{array}{l} \%\,66.7 = 0.667 = r_{2,1} = r_{1,2} \\ \%\,28.9 = 0.289 = r_{3,1} = r_{1,3} \\ \%\,28.9 = 0.289 = r_{3,2} = r_{2,3} \end{array}\right.$$

3- حساب القيم الذاتية للمصفوفة C : نقوم بحساب القيم الذاتية للمصفوفة C وترتيبها ترتيبا تنازليا حسب القانون التالي :

$$\boxed{|C - \lambda I_p| = 0}$$

خصائص القيم الذاتية :

- مجموع القيم الذاتية يساوي مجموع عناصر قطر المصفوفة C
- يسمى عدد القيم الذاتية الغير معدوم رتبة المصفوفة C
- يسمى مجموع القيم الذاتية بالكثافة الكلية للبيانات I_p .

$$\sum_{\alpha=1}^{p} \lambda_\alpha = Trace(C) = p$$

$$I_p = \sum_{\alpha=1}^{p} \lambda_\alpha$$

-5-

بالرجوع إلى المثال السابق :

وبحساب القيم الذاتية انطلاقا من العلاقة نجد :

$\lambda_1 = 1.86$ ، $\lambda_2 = 0.81$ ، $\lambda_3 = 0.33$

مع $\sum_{\alpha=1}^{3} \lambda_\alpha = 3$

4 ـ حساب الأشعة الذاتية للمصفوفة C :

نحسب الأشعة الذاتية للمصفوفة C بطريقتين هما :

أولا : $\boxed{(C - \lambda_\alpha I_p) u_\alpha = 0_p}$

ثانيا : $\boxed{\|u_\alpha\| = 1 \Rightarrow \sqrt{a_1^2 + a_2^2 + \dots + a_p^2} = 1}$

أي طويلة الشعاع u_α تساوي الواحد .

بالرجوع إلى المثال السابق :

لدينا : الشعاع $u_1 = \begin{pmatrix} a \\ a \\ 0.429 \end{pmatrix}$ وبحساب طويلة الشعاع u_1

نجد : $u_1 = \begin{pmatrix} 0.639 \\ 0.639 \\ 0.429 \end{pmatrix}$

بالنسبة للشعاع u_2 فهو يكتب بالشكل : $u_2 = \begin{pmatrix} -0.303 \\ 0.303 \\ b \end{pmatrix}$ وبحساب

طويلة الشعاع u_2 نجد $u_2 = \begin{pmatrix} -0.303 \\ 0.303 \\ 0.903 \end{pmatrix}$

ـ الشعاع u_3 فهو يكتب : $u_3 = \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix}$ وبحساب الطويلة نجد

$u_3 = \begin{pmatrix} 0.707 \\ -0.707 \\ 0 \end{pmatrix}$

-7-

5 - التحليل في ACP:

نقوم أولا بحساب نسبة التمثيل بواسطة العلاقة التالية:

$$t_\alpha = \frac{\lambda_\alpha}{J_\alpha} \times 100\%$$

ثم نرسم جدول نسب التمثيل:

المثال السابق:

جدول نسب التمثيل:

المحاور	القيم الذاتية	نسبة التمثيل	النسب التجميعية
1	1.86	$t_1 = \frac{1.86}{3} \times 100 = 62\%$	62%
2	0.81	$t_2 = 27\%$	89%
3	0.33	$t_3 = 11\%$	

نلاحظ من الجدول أن 62% من البيانات الجدول ممثلة في المحور الأول و 27% من بيانات الجدول ممثلة في المحور الثاني و 11% من البيانات ممثلة في المحور 3.

89% من بيانات الجدول الأولي للمعطيات ممثلة في المستوى الأول وهي كافية ومعبرة بيانيا ويمكن الاعتماد عليها في التحليل والدراسة.

- ايجاد احداثيات الافراد

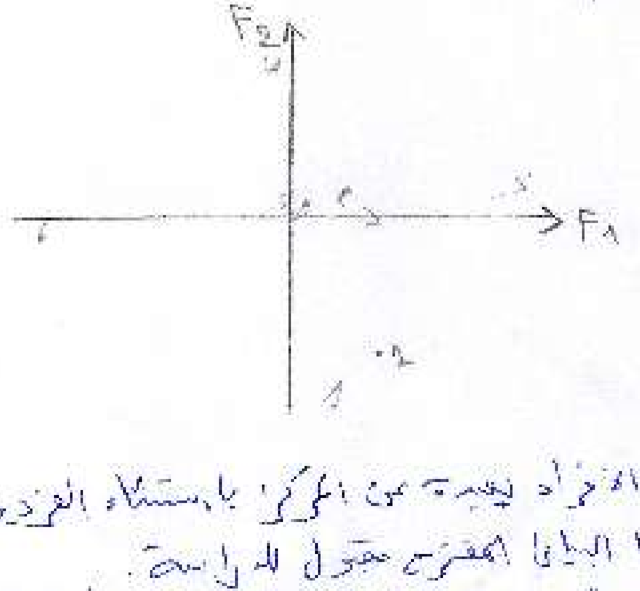
لدينا $F_\alpha(i)$ فهو يمثل احداثيات الافراد على المحور α ويحسب حسب العلاقة التالية:

$$\boxed{F_\alpha(i) = X M_\alpha}$$

بالرجوع الى مثالنا السابق :

نعطي احداثيات الافراد على المحاور في الجدول التالي :

الافراد \ المحاور	F_1	F_2	F_3
1	0,388	-1,535	0,000
2	0,000	0,000	-1,000
3	2,332	0,249	0,000
4	-0,378	1,535	0,000
5	0,000	0,000	1,000
6	-2,339	-0,219	0,000



نلاحظ ان الافراد بعيدة عن المركز باستثناء الفردين ② ⑤ وبالتالي فلا يمكن تمثيلها البيانات المعروضة حول الدراسة.
نلاحظ ان الفرد ③ بالقيم الموجبة والفرد ⑥ بالقيم السالبة على المحور ①
الفرد ④ بالقيم الموجبة والفرد ① بالقيم السالبة بالنسبة للمحور ②

- نسب تمثيل الأفراد في المركبات الأساسية (جودة تمثيل الفرد i).

تحسب بواسطة القانون التالي:

$$\boxed{\cos^2\theta_{i\alpha} = \frac{F_\alpha^2(i)}{\sum_{\alpha=1} F_\alpha^2(i)}}$$

حيث: $0 \le \cos^2\theta_{i\alpha} \le 1$

$$\sum \cos^2\theta_{i\alpha} = 1$$

- نسب مساهمة الأفراد في تمثيل المحاور:

تحسب نسبة مساهمة الفرد في تمثيل المحور بالعلاقة التالية:

$$\boxed{C_i^\alpha = \frac{F_\alpha^2(i)}{(n-1)\lambda_\alpha}} \quad \text{مع} \quad \sum C_i^\alpha = 1$$

بالرجوع إلى مثالنا السابق، الجدول التالي يعطينا نسب تمثيل الأفراد في المركبات الأساسية ونسب مساهمة الأفراد في تمثيل المحاور:

الأفراد	F_1	F_2	F_3	$\cos^2\theta_{i1}$	$\cos^2\theta_{i2}$	$\cos^2\theta_{i3}$	C_i^1	C_i^2	C_i^3
1	0.378	-1.535	0	0.057	0.943	0.00	0.012	0.48	0
2	0.00	0.00	-1	0.00	0.00	1	0	0	0.50
3	2.332	0.249	0	0.989	0.011	0	0.487	0.035	0
4	-0.378	1.535	0	0.057	0.943	0	0.014	0.48	0
5	0.00	0.00	1	0.00	0.00	1	0	0	0.50
6	-2.332	-0.249	0	0.989	0.011	0	0.487	0.035	0
$\sum$	0	0	0				1	1	1

-10-

6- التحليل في IR :

- احداثيات المتغيرات على المحاور :

نحسب احداثيات المتغيرات على المحاور بواسطة القانون التالي

$$\boxed{G_\alpha = \sqrt{\lambda_\alpha}\, M_\alpha}$$

في مثالنا السابق

$$G_1 = \sqrt{\lambda_1}\, M_1 = \begin{pmatrix} 0.871 \\ 0.871 \\ 0.871 \end{pmatrix} \; ; \; G_2 = \sqrt{\lambda_2}\, M_2 = \begin{pmatrix} -0.272 \\ -0.272 \\ 0.812 \end{pmatrix}$$

$$G_3 = \sqrt{\lambda_3}\, M_3 = \begin{pmatrix} 0.408 \\ -0.408 \\ 0.00 \end{pmatrix}$$

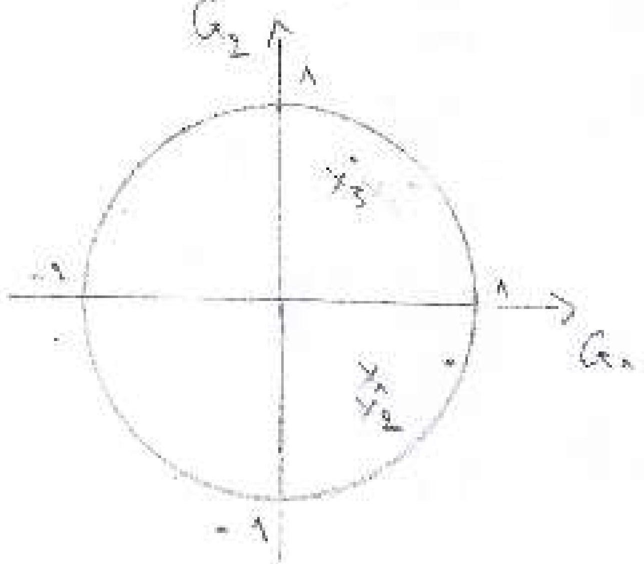
ونلخص الاحداثيات في الجدول التالي :

المتغيرات	G_1	G_2	G_3
X_1	0.871	-0.272	0.408
X_2	0.871	-0.272	-0.408
X_3	0.871	0.812	0.00

تمثل المتغيرات في دائرة نصف قطرها عبارة عن واحد (1) تسمى دائرة الارتباطات او دائرة الترابط .

جودة التمثيل : كل المتغيرات التي تكون بعيدة عن مركز الدائرة وقريبة من المحيط هي ممثلة في الدراسة .

-11-



- علاقة المتغيرات مع المحاور:

نلاحظ أن المتغيرين X_1 و X_2 لهما ارتباط قوي وموجب مع المحور ①

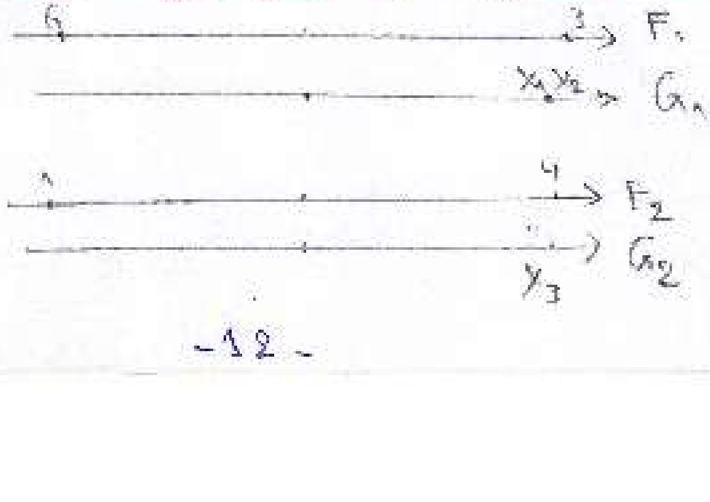
أما المتغير X_3 له ارتباط قوي وموجب مع المحور ②

- علاقة المتغيرات فيما بينها:

نلاحظ أن المسافة بين المتغيرين X_1 و X_2 ضعيفة (معدومة) مما يعني أنه هناك ارتباط قوي وموجب فيما بينهما.

نلاحظ أن المسافة بين (X_1 X_2) من جهة و X_3 من جهة أخرى بعيدة مما يعني أنهما مستقلان عن بعضهما

8- علاقة التمثيل البياني للمتغيرات مع الأفراد:



- بالنسبة للمحور الأول:

نلاحظ أن الفرد (3) له أكبر قيمة في المتغيرين x_1 و x_2
والفرد (6) له أقل قيمة في المتغيرين x_1 و x_2

- بالنسبة للمحور الثاني:

نلاحظ أن الفرد (5) له أكبر قيمة في المتغير x_3 ،
والفرد (9) له أقل قيمة في المتغير x_3 .