الحل النموذجي للمتحان
مادة تحليل المعطيات

تمرين (1) (05 ن)

1 – **أركان الاستبيان الاستبيان:** الباحث، الموضوع، العينة، الأسئلة أو العبارات، (1 ن)

2 – **استخدامات التحليل العاملي (استكشافي وتوكيدي):** فهم أفضل للبيانات، صياغة الفرضيات، تحليل الارتباط بين المتغيرات، الكشف عن المتغيرات الكامنة، تعديل النموذج وضبطه، بناء مؤشرات توليفية، .. (1 ن)

3– أ- **حساب الجذور المميزة λ_i للمصفوفة:** (λ_2 =20، λ_2=0، λ_1= 10).مع التعليل استنادا لخاصيتي الرتبة والأثر والمنقول. (1 ن)

ب– المتجهات المميزة للمصفوفة **ومقلوبها متساوية.** (1 ن)

4– المراحل العملية لطريقة ACP: أ\ جدول البيانات الخام، ب/ بيانات مركز الثقل G، ج/ جدول البيانات الممركزة Xc، د/ مصفوفة التباينات والتباينات المشتركة Σ، هـ/ حساب وترتيب λ_i لـ Σ، و/ حساب متجهاتها، ز/ تحديد المحاور والمركبات الأساسية،ح/ تحليل الارتباطات بين المركبات الأساسية والمتغيرات الأولية، ط/ تعريف المركبات الأساسية .... (1 ن)

تمرين (2) (08 ن)

1 – إكمال المصفوفة: عناصر قطر المصفوفة R 1 والبقية **بالتناظر** (**1 ن**)، (**1 ن**)

R مصفوفة **معاملات الارتباط** بين **المتغيرات الأولية V_j**

Σ مصفوفة **التباين والتباين المشترك** بين **المتغيرات الأولية V_j**

$$R_{(V_1,V_2)} = cov(V_1,V_2) / \sigma v1\ \sigma v2$$

$$\Sigma = \begin{pmatrix} 15 & 16 & -12 \\ 16 & 20 & 4 \\ -12 & 4 & 25 \end{pmatrix} \quad R = \begin{pmatrix} 1 & 0.924 & -0.62 \\ 0.924 & 1 & 0.18 \\ -0.62 & 0.18 & 1 \end{pmatrix}$$

2– استنتاج قيمة عزم العطالة الكلي: $I_G = Tr(\Sigma) \Rightarrow \mathbf{I_G = 60}$ (تحليل المركبات الأساسية البسيط.) (**2 ن**)

3– حساب عزم العطالة المحمول بكل محور: (**2 ن**)

$\tau i = I_{\Delta i} / I_G \quad \Rightarrow \quad I_{\Delta i} = \tau i\ I_G$

$r_1 = 50\%, r_2 = 35\%, r_3 = 15\%$

(0.5 ن:) $\mathbf{I_{\Delta 1}} = 30$, ***(1 ن:)*** $\mathbf{I_{\Delta 2}} = 21$, ***(0.5 ن:)*** $\mathbf{I_{\Delta 3}} = 09$

4– عدد المحاور التي يمكن أخذها للدراسة: 3 أي ($\Delta 1$ و $\Delta 2$ و$\Delta 3$) لأن $\sum cr_j \geq 90\%$ **طريقة ACP البسيط.** (**1 ن**)

5– المؤشرات الأخرى المستخدمة للتحليل: معيار الفرق الأول، ومعيار اختبار الحصاة. (**1 ن**)

6– أكثر المتغيرات القديمة والجديدة تفسيرا لعزم العطالة الكلي I_G.

المتغير الجديد Z_1 ذو أكبر تباين من بين المتغيرات الجديدة،

والمتغير V_3 هو الأكثر مساهمة من بين المتغيرات الأولية لأن له أكبر تباين.

(**0.5**)

7– كتابة المصفوفة R_Z (بين المتغيرات الجديدة) (**0.5**)

$$R_Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

1 - **تحليل العلاقة بين كل من المتغير V3 وبين المركبين الأساسيين Z1 و Z2** **(1 ن)**

من خلال إسقاط النقط الممثلة للمتغيرات V على المحاور الممثلة للمتغيرات الجديدة Z نجد قيمة معامل الارتباط بين المتغيرين:

V3 له **ارتباط خطي ضعيف سالب** مع **Z1** (**r→−0**)، وله **ارتباط خطي قوي موجب** مع **Z2** (**r→+1**)،

V1 له **ارتباط خطي قوي موجب** مع **Z1** (**r→+1**)، وله **ارتباط خطي ضعيف موجب** مع **Z2** (**r→+0**)،

V4 له **ارتباط خطي قوي موجب** مع **Z1** (**r→+1**)، وله **ارتباط خطي ضعيف سالب** مع **Z2** (**r→−0**)،

V2 له **ارتباط خطي قوي سالب** مع **Z1** (**r→−1**)، وله **ارتباط خطي ضعيف موجب** مع **Z2** (**r→+0**)،

- **لا يمكن تحليل ارتباط المتغير V5 لأنه غير ممثل جيدا على الفضاء الجزئي (2،1).**

2 -**تحليل نوعية تمثيل المتغيرات الأولية على الفضاء الجزئي (1، 2):** **(1 ن)**

- **نقول أن متغير V_j ممثل جيدا على الفضاء الجزئي (1، 2)** الذي محاوره من المتغيرات الجديدة فقط إذا كان **معامل ارتباطه** مع هذه المتغيرات **الجديدة قريبا من 1**، وهذا يحصل إذا كان **Cos^2 الزاوية بين الفضاء الجزئي والشعاع $\overrightarrow{GV_j}$ قريبا من 1.** ***(0.5 ن):***
- V_1 و V_2 و V_3 و V_4 **ممثلة جيدا** (لأنها قريبة من حافة دائرة الارتباط)، أما V_5 فهو **غير ممثل جيدا** (بعيدة عن الحافة). ***(0.5 ن):***

3 -**العلاقة بين المتغيرات القديمة فيما بينها:** **(2 ن)**

- يشترط أن يكون المتغير ممثلا جيدا لكي يمكن تحديد علاقته بغيره من المتغيرات القديمة.
- **معامل الارتباط** بين متغيرين قديمين **يساوي** إلى **جيب تمام** الزاوية بين النقطتين الممثلتين لهما على الفضاء الجزئي (1، 2). ***(1 ن):***
- $\cos\alpha = 1 \Rightarrow \alpha = 0°$. $\cos\alpha = -1 \Rightarrow \alpha = 180°$. $\cos\alpha = 0 \Rightarrow \alpha = 90°$
- **V_1 و V_4** مرتبطان ارتباطا **إيجابيا قويا**. (**قريبين**) ويرتبط كل منهما ارتباطا **سلبيا قويا** مع **V_2**. (**زاوية مستقيمة**)
- كل من المتغيرات **V_1 و V_2 و V_4 لا يرتبط خطيا** مع المتغير **V_3**. (**زاوية قائمة**) ***(1 ن):***

4 -**تفسير التعاكس في الإشارة بين V_1 وV_3 :** **(1 ن)**

اتجاه المحاور اختياري، لذا أهمية الإشارة ضعيف نسبيا والأهم هو تموقع النقط الممثلة للمتغيرات بالنسبة لبعضها البعض، حيث نجد أن هذين المتغيرين غير مرتبطين خطيا كما يتضح من جواب السؤال السابق (آخر سطر).

5- **تحليل حدود الاستقراء أو التنبؤ** **(2 ن)**

- يلاحظ أن طويلة الأشعة $\|\overrightarrow{GU_i}\|^2$ لأول وآخر الأفراد أي بداية ونهاية فترة الدراسة تكون أكبر منها في بقية فترات الدراسة، مما يعني أن هذه الفترة ممثلة بشكل سيء في التحليل، وهذا ما يجعل القيام بالتنبؤ والاستقراء استنادا إلى نتائج هذه الفترة غير مجد.