

DEVOIR ÉCRIT N°2 – JUIN 2008

Sans calculatrice, ni document.

Le sujet comporte 4 pages dont 3 d'annexes.

EXERCICE N°1 : question de cours

On rappelle quelques notations utilisées en AFC et en ACM.

➤ En AFC. On construit la matrice $X^* = \left[\frac{f_{ij}}{\sqrt{f_{i\bullet} f_{\bullet j}}} \right]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ puis on diagonalise la

matrice $T^* = X^* X^*$.

➤ En ACM, on désigne par :

- n : nombre d'individus ou d'unités statistiques
- q : nombre de variables
- m : nombre total de modalités
- $X_j(i)$: valeur de la variable indicatrice X_j pour l'individu i .
- p_j : fréquence d'une colonne par rapport à l'effectif total (égal à nq)

De plus, on pourra noter D le tableau disjonctif complet qui peut donc être considéré comme une matrice de taille $n \times m$.

Montrer que l'inertie totale de l'AFC du tableau disjonctif complet est égal à $\frac{m}{q} - 1$.

EXERCICE N°2 :

Le tableau figurant en annexe donne les résultats concernant des variables de différents types pour certains pays du monde en 2007.

1. Expliquer pourquoi on ne peut pas faire une ACP de ces données. Quel est le nom de la méthode d'analyse des données adéquate dans ce cas ?
2. Construire un tableau de Burt pertinent pour rendre compte de ces données. Quelle est l'inertie totale de ces données ?

EXERCICE N°3 :

On désire effectuer une analyse des résultats du premier tour des élections présidentielles de 2002 en France métropolitaine, par région.

Les données nécessaires sont sur l'annexe.

Certains regroupements de candidats idéologiquement proches ont été effectué..

1. Quel est le nom du type d'analyse de données que nous allons effectuer ? Pourquoi ?
2. Va-t-on utiliser les profils ligne ou les profils colonnes ? Pourquoi ?
3. Quel est le nombre de valeurs propres de la matrice à diagonaliser ?
4. Quelle est la valeur du χ^2 de ce tableau de données ? Faites en une interprétation. On admet que l'effectif total est de 39 100 220. On se contentera d'une extrapolation à partir de la table fournie en annexe.
5. Quels sont les axes à conserver ? Pourquoi ?
6. Quelle est la signification de « cr » ? Quelle est son utilité ?
7. Après avoir sélectionné des régions et des candidats intéressants, faire une tentative d'analyse de ces données.

ANNEXE

EXERCICE N°2

<i>Pays</i>	<i>continent</i>	<i>producteur de pétrole</i>	<i>type de pays</i>
<i>Japon</i>	<i>Asie</i>	<i>Non</i>	<i>Développé</i>
<i>Finlande</i>	<i>Europe</i>	<i>Non</i>	<i>Développé</i>
<i>Russie</i>	<i>Europe</i>	<i>Oui</i>	<i>NPI</i>
<i>Venezuela</i>	<i>Amérique Latine</i>	<i>Oui</i>	<i>PED</i>
<i>Tunisie</i>	<i>Afrique</i>	<i>Non</i>	<i>PED</i>
<i>Mexique</i>	<i>Amérique Latine</i>	<i>Oui</i>	<i>NPI</i>
<i>Chine</i>	<i>Asie</i>	<i>Oui</i>	<i>NPI</i>
<i>Haïti</i>	<i>Amérique Latine</i>	<i>Non</i>	<i>PED</i>
<i>Bangladesh</i>	<i>Asie</i>	<i>Non</i>	<i>PED</i>
<i>Angola</i>	<i>Afrique</i>		<i>PED</i>

EXERCICE N°3

<i>K</i>	<i>ABST</i>	<i>CHIR</i>	<i>EX DRO</i>	<i>DIV GA</i>	<i>DIV DR</i>	<i>ECOL</i>	<i>EX GAU</i>	<i>JOSP</i>	<i>BAYR</i>	<i>ST JO</i>
<i>Al</i>	318667	150816	228226	65708	44667	67226	68612	90502	88528	8403
<i>Aqu</i>	563391	285279	219754	149875	173923	105042	150655	270555	117429	132136
<i>Auv</i>	246066	147064	112678	81949	56680	42340	82301	108558	48439	33559
<i>Bour</i>	298847	153193	169153	86984	65614	49575	83633	130481	52854	32476
<i>Bret</i>	544310	342533	204232	154596	128707	136836	189432	286514	113920	58385
<i>Cent</i>	437553	244240	240509	128822	116371	78454	125028	184173	82415	67135
<i>Champ</i>	253569	125345	151284	55837	50402	37500	67085	88995	41477	26160
<i>Corse</i>	79159	30180	18845	17161	11212	4189	6612	16836	3658	9009
<i>Fra Co</i>	184128	101323	133152	77286	40502	38305	58788	77387	34985	19848
<i>IdF</i>	1919013	838107	662831	551878	245120	336731	322044	679926	298901	44952
<i>Lan</i>	424448	187870	298054	138500	109724	74073	110098	193784	68169	70890
<i>Lim</i>	125841	108259	47900	49471	33354	21388	46397	73486	16556	23705
<i>Lor</i>	470817	206655	265141	102023	61874	75095	128134	162943	70628	23777
<i>M P</i>	435727	237368	237564	152095	135973	104339	140514	270867	86939	93901
<i>N PdC</i>	778723	331826	411786	213845	150907	108927	231442	327347	114197	88987
<i>Nor</i>	588042	329774	295456	154736	145529	105826	194070	243994	103763	87706
<i>PdL</i>	601806	362743	232301	147766	150615	141179	188514	271095	137540	78239
<i>Pic</i>	326365	167134	210718	92509	86767	51103	109901	130520	54292	57835
<i>PC</i>	284327	181423	118908	82631	98919	63266	101552	149108	57349	67925
<i>PACA</i>	772648	404134	579709	228640	181280	141049	166834	277614	130623	92790
<i>RA</i>	962160	450190	593417	296385	209079	221422	241550	363463	204613	81321

Valeurs propres

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9
0,013	0,009	0,003	0,002	0,001	0,001	0,000	0,000	0,000

Composantes principales

	<i>Axe 1</i>			<i>Axe 2</i>			<i>1&2</i>
	<i>coord.(y)</i>	<i>qlt</i>	<i>cr</i>	<i>coord.(y)</i>	<i>qlt</i>	<i>cr</i>	<i>cum.</i>
<i>Alsace</i>	-0,220	0,556	0,105	0,115	0,151	0,042	0,707
<i>Aquitaine</i>	0,203	0,801	0,172	-0,027	0,014	0,004	0,814
<i>Auvergne</i>	0,060	0,332	0,007	-0,043	0,172	0,005	0,504
<i>Bourgogne</i>	-0,007	0,018	0,000	0,030	0,368	0,003	0,386
<i>Bretagne</i>	0,066	0,148	0,018	-0,130	0,568	0,103	0,716
<i>Centre</i>	0,060	0,621	0,012	0,033	0,189	0,005	0,810
<i>Champagne</i>	-0,038	0,107	0,003	0,081	0,477	0,016	0,585
<i>Corse</i>	-0,025	0,004	0,000	-0,105	0,073	0,006	0,077
<i>Franche comté</i>	-0,050	0,111	0,004	0,088	0,338	0,017	0,450
<i>Ile de France</i>	-0,155	0,517	0,273	-0,143	0,439	0,339	0,956
<i>Languedoc</i>	0,030	0,036	0,003	0,135	0,717	0,086	0,754
<i>Limousin</i>	0,167	0,368	0,029	-0,123	0,198	0,023	0,566
<i>Lorraine</i>	-0,128	0,577	0,050	0,042	0,062	0,008	0,639
<i>Midi Pyrénées</i>	0,153	0,731	0,086	0,005	0,001	0,000	0,732
<i>Nord Pas de Calais</i>	0,001	0,000	0,000	0,036	0,132	0,010	0,132
<i>Normandie</i>	0,074	0,630	0,024	0,008	0,007	0,000	0,637
<i>Pays de Loire</i>	0,078	0,270	0,027	-0,093	0,385	0,057	0,655
<i>Picardie</i>	0,074	0,251	0,013	0,115	0,611	0,048	0,861
<i>Poitou Charente</i>	0,223	0,948	0,116	-0,039	0,029	0,005	0,977
<i>PACA</i>	-0,058	0,110	0,019	0,152	0,752	0,194	0,863
<i>Rhône Alpes</i>	-0,074	0,424	0,038	0,053	0,216	0,029	0,640

	<i>Axe 1</i>				<i>Axe 2</i>		<i>1&2</i>
	<i>coord.(f)</i>	<i>qlt</i>	<i>cr</i>		<i>qlt</i>	<i>cr</i>	<i>cum.</i>
<i>ABST</i>	-0,078	0,542	0,125	-0,046	0,185	0,062	0,728
<i>CHIR</i>	0,020	0,037	0,004	-0,062	0,370	0,059	0,408
<i>EXT DRO</i>	-0,101	0,181	0,107	0,214	0,814	0,700	0,995
<i>DIV GAU</i>	-0,051	0,147	0,015	-0,040	0,094	0,014	0,241
<i>DIV DRO</i>	0,184	0,860	0,150	0,046	0,053	0,014	0,913
<i>ECOLO</i>	-0,037	0,059	0,005	-0,070	0,208	0,028	0,267
<i>EXT GAU</i>	0,102	0,376	0,057	0,003	0,000	0,000	0,376
<i>JOSP</i>	0,071	0,346	0,043	-0,069	0,327	0,059	0,674
<i>BAYR</i>	-0,039	0,059	0,006	-0,026	0,026	0,004	0,086
<i>ST JO</i>	0,461	0,885	0,490	0,134	0,075	0,061	0,960