

12/06/07

DE ANALYSE DE DONNÉES

Exercice 1

$$\frac{12,5}{2}$$

Incompl. des échantillons

1)

- 2) la probabilité qu'une personne habite Paris sachant qu'elle est âgée de + de 85 ans est 41,4% 19 ans plus
- la probabilité qu'on ait entre 15 et 19 ans sachant qu'on habite dans les Hauts de Seine est 5,8% 1

Les probabilités provenant des PFL et PFC car ces tableaux nous renseignent sur les probabilités de $X|Y$ et $Y|X$.

- 3) On a choisi la matrice T^* pour faire les calculs car $T^* = X^* \cdot X^*$ ce qui fait qu'on aura une matrice carrée symétrique d'ordre 6. 1
- Si nous avions choisi de travailler avec $W^* = X^* \cdot E \cdot X^*$ on aurait eu une matrice d'ordre 11, ce qui est bien moins évident pour les calculs. 1

T^* possède normalement 6 valeurs propres mais parmi ces valeurs, il y en a qui sont triviales. 1

Les valeurs 0 et 1 ne doivent pas être prises en compte. Il en reste alors 2 : λ_1 et λ_2 .

5. Inertie totale $I = \sum$ des valeurs propres - valeur triviale

$$I = 0,015 + 0,001$$

$$\underline{I = 0,016}$$

$$\chi^2 = I \times N$$

Le calcul au brouillon de l'effectif N me donne

1	105	500
1	227	300
1	382	900
1	338	700
1	134	200
2	125	200
8 313 900		= effectif

8 313 900
0,016
498 824 00
8 313 900.
133 021 400

$$\underline{\chi^2 = 133021,4}$$

Hypothèse : « l'âge et le département d'habitation »

Degré de liberté : $(n-1)(p-1) = (11-1)(6-1) = 10 \times 5 = 50$

Seuil arbitraire : 10%

$$\chi_{0,90}^{\text{théorique}} = 63,2$$

Quoi qu'il arrive, notre hypothèse sera rejetée car $\chi_{\text{th}} < \chi_{\text{obs}}$. Cependant mon χ_{obs} est incroyablement grand donc il se pourrait que j'aie fait une erreur de calcul ou de raisonnement.

Exercice 2

$$M = \begin{bmatrix} 0 & -1 & 3 \\ 2 & 3 & -3 \\ 2 & -1 & 1 \end{bmatrix}$$

Pour déterminer les valeurs propres, je dois calculer $\text{Det } |M - \lambda I| = 0$

$$\begin{vmatrix} 0-\lambda & -1 & 3 \\ 2 & 3-\lambda & -3 \\ 2 & -1 & 1-\lambda \end{vmatrix} = 0 \quad (\Rightarrow) \quad -\lambda \begin{vmatrix} 3-\lambda & -3 \\ -1 & 1-\lambda \end{vmatrix} - 2 \begin{vmatrix} -1 & 3 \\ -1 & 1-\lambda \end{vmatrix} + 2 \begin{vmatrix} -1 & 3 \\ 3-\lambda & -3 \end{vmatrix}$$

$$(\Rightarrow) -\lambda [(3-\lambda)(1-\lambda) - (-1 \times -3)] - 2[-1(1-\lambda) - (3 \times -1)] + 2[(-1 \times -3) - 3(3-\lambda)] = 0$$

$$(\Rightarrow) -\lambda(3 - 3\lambda - \lambda + \lambda^2 - 3) - 2(-1 + \lambda + 3) + 2(3 - 9 + 3\lambda) = 0$$

$$(\Rightarrow) -\lambda(\lambda^2 - 4\lambda) - 2(\lambda + 2) + 2(3\lambda - 6) = 0$$

$$(\Rightarrow) -\lambda^3 + 4\lambda^2 - 2\lambda - 4 + 6\lambda - 12 = 0$$

$$(\Rightarrow) \underline{-\lambda^3 + 4\lambda^2 + 4\lambda - 16 = 0}$$

Je remarque une solution remarquable qui est 2 je peux donc factoriser: $(\lambda - 2)(-\lambda^2 + 6\lambda + 8)$

Je recherche alors des solutions pour λ

$$\Delta = 36 - 4 \times -1 \times 16 = 36 + 64 = 100$$

$$\lambda_1 = \frac{-6 - 10}{-2} = \frac{16}{2} = \underline{8}$$

$$\lambda_2 = \frac{-6 + 10}{-2} = \frac{4}{-2} = \underline{-2}$$

J'en déduis que mes valeurs propres sont $\lambda_1 = 2$ $\lambda_2 = -2$ $\lambda_3 = 0$
Éléments de vérification: Pour une matrice 3×3 je dois avoir 3 valeurs propres.

DEVOIR ÉCRIT N°2

Les documents et la calculatrice sont interdits.
 Vous indiquerez votre groupe de TD sur votre copie.

EXERCICE N°1 :

Voici les effectifs de population par classe d'âge relevés par le recensement de 1999 dans six départements de la région parisienne. Les classes d'âge sont exprimées en années.

K	Paris-75	Esson-91	HdeS-92	S-St-D-93	VdeM-94	Vd'O-95
0-4	110300	81900	99600	104000	85300	82000
5-9	91900	85300	87700	105000	81500	91000
10-14	88500	79500	79700	96800	73300	87900
15-19	109900	76000	82800	93900	73500	81900
20-24	191900	77400	106400	102900	87500	76700
25-54	1016000	524900	654800	631400	565600	504500
55-59	101700	52400	65000	57200	55100	43600
60-64	86900	45500	58800	51900	51100	38300
65-74	153700	65400	103400	80100	85200	59300
75-84	103500	31500	59500	39700	46500	28300
>85	70900	14400	31200	20000	22600	12000

Matrice F des fréquences

F	Paris-75	Esson-91	HdeS-92	S-St-D-93	VdeM-94	Vd'O-95
0-4	0.013	0.010	0.012	0.012	0.010	0.010
5-9	0.011	0.010	0.010	0.012	0.010	0.011
10-14	0.011	0.009	0.009	0.012	0.009	0.010
15-19	0.013	0.009	0.010	0.011	0.009	0.010
20-24	0.023	0.009	0.013	0.012	0.010	0.009
25-54	0.121	0.062	0.078	0.075	0.067	0.060
55-59	0.012	0.006	0.008	0.007	0.007	0.005
60-64	0.010	0.005	0.007	0.006	0.006	0.005
65-74	0.018	0.008	0.012	0.010	0.010	0.007
75-84	0.012	0.004	0.007	0.005	0.006	0.003
>85	0.008	0.002	0.004	0.002	0.003	0.001

Profils-lignes et profils-colonnes

PFL	Paris-75	Esson-91	HdeS-92	S-St-D-93	VdeM-94	Vd'O-95
0-4	0.196	0.145	0.177	0.185	0.151	0.146
5-9	0.169	0.157	0.162	0.194	0.150	0.168
10-14	0.175	0.157	0.158	0.191	0.145	0.174
15-19	0.212	0.147	0.160	0.181	0.142	0.158
20-24	0.299	0.120	0.166	0.160	0.136	0.119
25-54	0.261	0.135	0.168	0.162	0.145	0.129
55-59	0.271	0.140	0.173	0.153	0.147	0.116
60-64	0.261	0.137	0.177	0.156	0.154	0.115
65-74	0.281	0.120	0.189	0.146	0.156	0.108
75-84	0.335	0.102	0.193	0.128	0.150	0.092
>85	0.414	0.084	0.182	0.117	0.132	0.070

PFC	Paris-75	Esson-91	HdeS-92	S-St-D-93	VdeM-94	Vd'O-95
0-4	0.052	0.072	0.070	0.075	0.070	0.074
5-9	0.043	0.075	0.061	0.076	0.066	0.082
10-14	0.042	0.070	0.056	0.070	0.060	0.080
15-19	0.052	0.067	0.058	0.068	0.060	0.074
20-24	0.090	0.068	0.074	0.074	0.071	0.069
25-54	0.478	0.463	0.458	0.457	0.461	0.456
55-59	0.048	0.046	0.045	0.041	0.045	0.039
60-64	0.041	0.040	0.041	0.038	0.042	0.035
65-74	0.072	0.058	0.072	0.058	0.069	0.054
75-84	0.049	0.028	0.042	0.029	0.038	0.026
>85	0.033	0.013	0.022	0.014	0.018	0.011

La matrice T^*

Y'Y	Paris-75	Esson-91	HdeS-92	S-St-D-93	VdeM-94	Vd'O-95
Paris-75	0.261	0.181	0.208	0.200	0.192	0.177
Esson-91	0.181	0.136	0.151	0.150	0.141	0.135
HdeS-92	0.208	0.151	0.170	0.167	0.158	0.149
S-St-D-93	0.200	0.150	0.167	0.166	0.155	0.149
VdeM-94	0.192	0.141	0.158	0.155	0.146	0.139
Vd'O-95	0.177	0.135	0.149	0.149	0.139	0.135

Valeurs propres

λ_1	λ_2	λ_3	λ_4	λ_5
0.015	0.001	0.000	0.000	0.000

Coordonnées des variables dans les axes 1-2

	Axe 1				Axe 2			
	coord.(ψ)	qlt	cr	cum.	coord.(ψ)	qlt	cr	cum.
0-4	0.127	0.889	0.072	0.889	0.039	0.083	0.161	0.972
5-9	0.216	0.998	0.202	0.998	0.007	0.001	0.005	0.999
10-14	0.215	0.991	0.185	0.991	-0.016	0.006	0.025	0.996
15-19	0.121	0.954	0.060	0.954	-0.025	0.040	0.061	0.994
20-24	-0.099	0.831	0.050	0.831	-0.041	0.144	0.209	0.975
25-54	-0.016	0.762	0.008	0.762	-0.007	0.143	0.036	0.905
55-59	-0.054	0.684	0.009	0.684	0.014	0.046	0.014	0.730
60-64	-0.040	0.474	0.004	0.474	0.036	0.367	0.080	0.841
65-74	-0.104	0.788	0.047	0.788	0.053	0.201	0.287	0.989
75-84	-0.237	0.970	0.138	0.970	0.034	0.020	0.068	0.990
>85	-0.406	0.988	0.224	0.988	-0.041	0.010	0.055	0.998

	Axe 1				Axe 2			
	coord.(ϕ)	qlt	cr	cum.	coord.(ϕ)	qlt	cr	cum.
75	-0.183	0.987	0.566	0.987	-0.021	0.013	0.175	1.000
91	0.102	0.934	0.094	0.934	0.000	0.000	0.000	0.934
92	-0.033	0.429	0.013	0.429	0.037	0.528	0.368	0.957
93	0.098	0.959	0.106	0.959	-0.012	0.015	0.039	0.974
94	0.008	0.051	0.001	0.051	0.033	0.900	0.249	0.951
95	0.159	0.960	0.221	0.960	-0.028	0.031	0.169	0.990

Questions :

1. Compléter les tableaux des profils-lignes et profils-colonnes avec les profils moyens. Que peut-on en dire?
2. Quelle est la probabilité qu'une personne habite Paris si elle est âgée de plus de 85 ans? Et sachant qu'on habite dans les Hauts de Seine, quelle la probabilité que la personne ait entre 15 et 19 ans?
3. Pourquoi a-t-on choisit la matrice T^* pour les calculs? Combien possède-t-elle de valeurs propres?
4. Montrer que les valeurs propres non nulles de T^* sont aussi valeurs propres de W^* .
5. Déterminer l'inertie totale. En déduire la valeur du χ^2 de cet échantillon. On donne en annexe la table du Khi-Deux. Que peut on dire de l'hypothèse: « l'âge et le département d'habitation sont indépendants »?
6. Quels sont les variables à ne pas prendre en compte dans l'interprétation? Pourquoi?
7. Représenter les variables dans le plan 1-2 et donner une interprétation des axes.

$\chi^2_{0,995}$	$\chi^2_{0,99}$	$\chi^2_{0,975}$	$\chi^2_{0,95}$	$\chi^2_{0,90}$	$\chi^2_{0,75}$	$\chi^2_{0,50}$	$\chi^2_{0,25}$	$\chi^2_{0,10}$	$\chi^2_{0,05}$	$\chi^2_{0,025}$	$\chi^2_{0,01}$	$\chi^2_{0,005}$
1	7,88	6,63	5,02	3,84	2,71	1,32	0,455	0,102	0,0158	0,0010	0,0002	0,0000
2	10,6	9,21	7,38	5,99	4,61	2,77	1,39	0,575	0,211	0,103	0,0506	0,0201
3	12,8	11,3	9,35	7,81	6,25	4,11	2,37	1,21	0,584	0,352	0,216	0,115
4	14,9	13,3	11,1	9,49	7,78	5,39	3,36	1,92	1,06	0,711	0,484	0,297
5	16,7	15,1	12,8	11,1	9,24	6,63	4,35	2,67	1,61	1,15	0,831	0,554
6	18,5	16,8	14,4	12,6	10,6	7,84	5,35	3,45	2,20	1,64	1,24	0,872
7	20,3	18,5	16,0	14,1	12,0	9,04	6,35	4,25	2,83	2,17	1,69	1,24
8	22,0	20,1	17,5	15,5	13,4	10,2	7,34	5,07	3,49	2,73	2,18	1,65
9	23,6	21,7	19,0	16,9	14,7	11,4	8,34	5,90	4,17	3,33	2,70	2,09
10	25,2	23,2	20,5	18,3	16,0	12,5	9,34	6,74	4,87	3,94	3,25	2,56
11	26,8	24,7	21,9	19,7	17,3	13,7	10,3	7,58	5,58	4,57	3,82	3,05
12	28,3	26,2	23,3	21,0	18,5	14,8	11,3	8,44	6,30	5,23	4,40	3,57
13	29,8	27,7	24,7	22,4	19,8	16,0	12,3	9,30	7,04	5,89	5,01	4,11
14	31,3	29,1	26,1	23,7	21,1	17,1	13,3	10,2	7,79	6,57	5,63	4,66
15	32,8	30,6	27,5	25,0	22,3	18,2	14,3	11,0	8,55	7,26	6,26	5,23
16	34,3	32,0	28,8	26,3	23,5	19,4	15,3	11,9	9,31	7,96	6,91	5,81
17	35,7	33,4	30,2	27,6	24,8	20,5	16,3	12,8	10,1	8,67	7,56	6,41
18	37,2	34,8	31,5	28,9	26,0	21,6	17,3	13,7	10,9	9,39	8,23	7,01
19	38,6	36,2	32,9	30,1	27,2	22,7	18,3	14,6	11,7	10,1	8,91	7,63
20	40,0	37,6	34,2	31,4	28,4	23,8	19,3	15,5	12,4	10,9	9,59	8,26
21	41,4	38,9	35,5	32,7	29,6	24,9	20,3	16,3	13,2	11,6	10,3	8,90
22	42,8	40,3	36,8	33,9	30,8	26,0	21,3	17,2	14,0	12,3	11,0	9,54
23	44,2	41,6	38,1	35,2	32,0	27,1	22,3	18,1	14,8	13,1	11,7	10,2
24	45,6	43,0	39,4	36,4	33,2	28,2	23,3	19,0	15,7	13,8	12,4	10,9
25	46,9	44,3	40,6	37,7	34,4	29,3	24,3	19,9	16,5	14,6	13,1	11,5
26	48,3	45,6	41,9	38,9	35,6	30,4	25,3	20,8	17,3	15,4	13,8	12,2
27	49,6	47,0	43,2	40,1	36,7	31,5	26,3	21,7	18,1	16,2	14,6	12,9
28	51,0	48,3	44,5	41,3	37,9	32,6	27,3	22,7	18,9	16,9	15,3	13,6
29	52,3	49,6	45,7	42,6	39,1	33,7	28,3	23,6	19,8	17,7	16,0	14,3
30	53,7	50,9	47,0	43,8	40,3	34,8	29,3	24,5	20,6	18,5	16,8	15,0
40	66,8	63,7	59,3	55,8	51,8	45,6	39,3	33,7	29,1	26,5	24,4	22,2
50	79,5	76,2	71,4	67,5	63,2	56,3	49,3	42,9	37,7	34,8	32,4	29,7
60	92,0	88,4	83,3	79,1	74,4	67,0	59,3	52,3	46,5	43,2	40,5	37,5
70	104,2	100,4	95,0	90,5	85,5	77,6	69,3	61,7	55,3	51,7	48,8	45,4
80	116,3	112,3	106,6	101,9	96,6	88,1	79,3	71,1	64,3	60,4	57,2	53,5
90	128,3	124,1	118,1	113,1	107,6	98,6	89,3	80,6	73,3	69,1	65,6	61,8
100	140,2	135,8	129,6	124,3	118,5	109,1	99,3	90,1	82,4	77,9	74,2	70,1

TABLE DU KHI-DEUX

EXERCICE N°2

1. Déterminer les valeurs propres de la matrice $M = \begin{bmatrix} 0 & -1 & 3 \\ 2 & 3 & -3 \\ 2 & -1 & 1 \end{bmatrix}$
2. Citer trois éléments permettant de contrôler la justesse de votre résultat.