

Les étudiants sont autorisés à utiliser une machine à calculer et à consulter les notes de cours, sous forme de transparents distribués par l'enseignant, sans ajout de notes personnelles

I Questions de cours (3 points):

Suite à une enquête sur les conditions de travail sur les salariés du secteur public, vous disposez d'une base de données de 10000 individus pour lesquels les réponses au questionnaire renseignent 10 variables binaires (oui-non) sur des indicateurs d'autonomie du type « vous devez appliquer strictement les consignes » ou « vous pouvez faire varier les délais fixés ».

Expliquer comment vous pouvez construire un indicateur individuel continu de l'autonomie à l'aide d'une analyse factorielle.

Exercice 1 (8 points):

On étudie les performances techniques et environnementales des voitures particulières. L'ensemble des données est constitué de 24 individus, modèles de véhicules, caractérisés par 9 variables quantitatives. Les variables considérées sont : Puissance (en chevaux fiscaux), Cylindrée (en cm³), Longueur (en cm), Largeur (en cm), Poids (en kg), Vitesse maximum (en km/h), Consommation moyenne aux 100 km (en litres d'essence ou gazole), émission (en CO₂) et prix (en euros).

1. On désire faire une analyse en composantes principales en utilisant les 8 premières variables (exceptée le prix). Pourquoi est-il nécessaire ici de remplacer les données d'origine par les données centrées-réduites (1 point) ?
2. Sachant que les variables étudiées sont centrées et réduites, définir et donner la valeur de l'inertie totale du nuage (1,5 point).
3. Compte tenu du tableau des valeurs propres ci-dessous, expliquer pour quelles raisons on trouve ce nombre de valeurs propres. Combien d'axes doit-on garder pour l'analyse ? Quelle sera alors la qualité globale de la représentation du nuage dans le plan constitué des deux premiers axes ? (2 points).

Tableau des valeurs propres

Axes	1	2	3	4	5	6	7	8
Valeurs Propres	5,923	0,992	0,576	0,236	0,154	0,072	0,036	0,011

4. Le tableau ci-dessous donne les coefficients de corrélation entre les huit variables centrées réduites et les deux premières composantes principales. A l'aide de ce tableau, expliquer le type d'effet que nous observons sur le premier axe (1 point).

Corrélations variables / composantes principales		
	1 ^{ère} comp	2 ^{ème} comp
Cylindrée	0.97	-0.17
Puissance	0.95	-0.52
Vitesse	0.94	-0.50
Poids	0.74	0.62
Largeur	0.83	-0.15
Longueur	0.81	0.60
Consommation	0,75	-0,65
Emission	0,78	-0,55

- Expliquer les variables qui déterminent les deux premières composantes principales (précisez clairement les critères utilisés). Interpréter la signification de ces deux composantes (1,5 point)
- Comment peut-on calculer la position d'une variable supplémentaire comme le prix sur le premier axe ? (1 point).
- Les corrélations de la variable prix sont respectivement 0,91 et -0,50. Interprétez le sens économique de ce résultat (1 point).

Exercice 2 (8 points):

Une entreprise de l'industrie agro-alimentaire a réalisé une étude auprès de 245 individus quant à leur appréciation des caractéristiques de production et de conservation de 6 formes de plats cuisinés. Ces six plats sont les conserves « **maison** », « **industrielle** », les plats « **traiteur** », « **surgelé** », « **déshydraté** » et « **frais sous vide** ».

Pour ces six caractéristiques de la variable plat, les répondants ont exprimé s'ils étaient d'accord avec 11 avis : permet de se faire « **plaisir** », bon rapport « **qualité/prix** », permet de manger des « **plats originaux** », offre des plats de « **bonne qualité** », plats de qualité « **savoureux** », plats qu'on peut « **stocker** », « **pratique** » pour préparer un repas, plat « **rapidement prêt** », plat frais, « **naturel** », plats offrant beaucoup de « **recettes** », plat satisfaisant « **tous les goûts** ».

Le tableau de contingence est le suivant :

	maison	industrielle	traiteur	surgelé	déshydraté	frais sous vide
Plaisir	223	91	211	159	47	83
Qualité/prix	201	145	145	137	59	71
Originaux	201	93	93	176	39	7
Bonne qualité	149	39	39	181	22	76
Savoureux	225	74	74	145	27	83
Stocker	37	218	218	189	162	44
Pratique	20	221	221	216	162	142
Rapidement prêt	27	228	228	186	189	169
Naturel	221	59	59	167	22	86
Recettes	203	108	108	169	39	66
Tous les goûts	221	127	127	172	49	83

1. Rappeler comment on construit le tableau des profils-lignes et des profils colonnes. On a extrait ci-dessous deux parties du tableau des profils-lignes et des profils-colonnes. Définir ce que signifie la valeur observée à l'intersection des modalités savoureux et traiteur dans ces deux tableaux. (2 points).

Tableau des profils-lignes

	maison	industrielle	traiteur	surgelé	déshydraté	frais sous vide
Savoureux	0,358	0,118	0,118	0,231	0,043	0,132
Stocker	0,043	0,251	0,251	0,218	0,187	0,051
Pratique	0,020	0,225	0,225	0,220	0,165	0,145

Tableau des profils-colonnes

	maison	industrielle	traiteur	surgelé	déshydraté	frais sous vide
Savoureux	0,130	0,053	0,049	0,076	0,033	0,091
Stocker	0,021	0,155	0,143	0,100	0,198	0,048
Pratique	0,012	0,156	0,145	0,114	0,198	0,156

2. On réalise une analyse factorielle des correspondances. Pourquoi le tableau des valeurs propres et de l'inertie ci-dessous ne présente-t-il que cinq valeurs propres ? D'après ce tableau, combien d'axes devrions-nous retenir pour l'interprétation ? (2 points)

Tableau des valeurs propres

Axes	Valeur propre	Pourcentage expliqué	Pourcentage cumulé
1	0.1495	78.31	78.31
2	0.0280	14.67	92.98
3	0.0088	4.65	97.63
4	0.0038	2.02	99.65
5	0.0007	0.35	100
Total	0.1908	100	

3. Quelles sont les avis sur les plats qui caractérisent les deux premiers axes principaux (vous justifierez précisément votre réponse à l'aide du tableau d'aide à l'interprétation pour les modalités ligne actives) (1.5 points).

Avis	Poids	Axe 1			Axe 2		
		Contribution	Coordonnées	Cos2	Contribution	Coordonnées	Cos2
Plaisir	9,84	5,19	0,28	0,75	6,15	0,13	0,17
Qualité/prix	8,06	0,14	0,05	0,02	23,32	-0,28	0,74
Originaux	9,34	4,01	0,25	0,81	2,50	0,09	0,09
Bonne qualité	6,49	3,23	0,27	0,41	3,50	-0,12	0,08
Savoureux	8,89	8,05	0,37	0,91	3,02	0,10	0,06
Stocker	8,00	28,75	-0,73	0,82	30,24	-0,33	0,16
Pratique	11,06	18,15	-0,49	0,90	8,84	0,15	0,08
Rapidement prêt	11,83	19,93	-0,50	0,83	19,43	0,22	0,15
Naturel	8,15	7,68	0,38	0,92	0,77	-0,05	0,02
Recettes	8,69	2,74	0,22	0,86	1,37	-0,07	0,08
Tous les goûts	9,64	2,13	0,18	0,84	0,84	-0,05	0,06

4. Quels sont les types de plats qui caractérisent les deux premiers axes principaux (vous justifierez précisément votre réponse à l'aide du tableau d'aide à l'interprétation pour les modalités colonne actives) (1.5 points).

Plat	Poids	Axe 1			Axe 2		
		Contribution	Coordonnées	Cos2	Contribution	Coordonnées	Cos2
Maison	20,9	41,51	0,54	0,91	18,67	-0,16	0,08
Conserve	16,97	18,76	-0,41	0,87	6,48	-0,10	0,06
Traiteur	17,45	5,39	0,21	0,32	57,26	0,30	0,63
Surgelé	22,95	0,02	-0,01	0,01	6,52	-0,09	0,28
Déshydraté	9,88	33,88	-0,72	0,99	0,03	-0,01	0,00
Frais sous vide	11,84	0,44	-0,07	0,40	11,04	0,16	0,40

5. Interprétez les deux premiers axes (1 point).