

Aucun support autorisé.
Réponses sur les feuilles sujet uniquement

Q1. QCM

Mettez une croix en face de chaque affirmation correcte.

La réponse à une question sera considérée correcte si toutes et uniquement les affirmations correctes sont identifiées. Pour chaque question, il peut y avoir plusieurs affirmations correctes.

A. Pour réaliser une ACP, les valeurs propres λ_i et vecteurs propres V_i :	
<input type="checkbox"/>	Sont calculées à partir de la matrice de variance / covariance
<input type="checkbox"/>	Sont calculées à partir de la matrice de corrélation
<input type="checkbox"/>	Répondent à une formule de la forme $A V_i = \lambda_i V_i$
<input type="checkbox"/>	Répondent à une formule de la forme $\lambda_i A = A V_i$
B. Pour le calcul des valeurs propres et des vecteurs propres :	
<input type="checkbox"/>	Il est nécessaire d'utiliser les données centrées et réduites
<input type="checkbox"/>	Il est nécessaire d'utiliser les données initiales
<input type="checkbox"/>	Il est possible d'utiliser indifféremment les données initiales ou les données centrées et réduites
C. L'utilisation des données « réduites » :	
<input type="checkbox"/>	A pour seul objectif de faciliter la présentation des résultats
<input type="checkbox"/>	Est nécessaire quelles que soient les données initiales
<input type="checkbox"/>	Est nécessaire uniquement si les données initiales ont des échelles de valeurs différentes
<input type="checkbox"/>	Aucune affirmation ci-dessus n'est correcte
D. Le facteur de corrélation entre deux variables x et y	
<input type="checkbox"/>	A une valeur comprise entre -1 et +1
<input type="checkbox"/>	A une valeur comprise entre 0 et 1
<input type="checkbox"/>	Est égal à 0 si les variables x et y sont linéairement liées
<input type="checkbox"/>	Rien n'est correct dans tout cela !
E. Soit deux variables qualitatives X (n valeurs possibles) et Y (m valeurs possibles).	
<input type="checkbox"/>	Pour effectuer une analyse des liens éventuels entre ces deux variables, il est nécessaire d'avoir $n = m$
<input type="checkbox"/>	Pour effectuer une analyse des liens éventuels entre ces deux variables, il est souhaitable d'avoir $n = m$
<input type="checkbox"/>	Si $n \neq m$, alors on va choisir la variable ayant le plus de valeurs possibles (ex. X si $n > m$), et on va réduire ce nombre en remplaçant plusieurs variables par leur somme.
<input type="checkbox"/>	Que n et m soient égaux ou non, on peut toujours effectuer une analyse
<input type="checkbox"/>	Rien de tout cela n'est vrai
F. En ACP, une composante principale	
<input type="checkbox"/>	Est calculée à partir des données initiales ou centrées réduites, en les multipliant par une valeur propre
<input type="checkbox"/>	Est calculée à partir des données initiales ou centrées réduites, en les multipliant par un vecteur propre
<input type="checkbox"/>	Peut être calculée à partir d'une des deux méthodes ci-dessus, selon le résultat attendu
<input type="checkbox"/>	Rien de tout cela
G. La moyenne d'une variable quantitative centrée et réduite	
<input type="checkbox"/>	Est égale à -1 ou +1
<input type="checkbox"/>	Dépend de l'amplitude des valeurs disponibles dans les données initiales
<input type="checkbox"/>	Est égale à 0
<input type="checkbox"/>	On ne peut rien dire avant d'avoir fait le calcul

H. Soit deux variables qualitatives X et Y. Soit le tableau des effectifs et des effectifs marginaux correspondant :

		<u>Variable Y</u>			
		1	j	m	
<u>Variable X</u>	1				
	i		z_{ij}		$z_{i.}$
	n				
			$z_{.j}$		z
		Nombre d'individus ayant répondu 'j' à la question Y			Nombre total d'individus ayant répondu

La valeur de z est la somme des valeurs $z_{i.}$

La valeur de z est la somme des valeurs $z_{i.}$ plus la somme des valeurs $z_{.j}$

La valeur de z est la somme des valeurs z_{ij}

La formule est plus compliquée que tout cela...

I. Le tableau du Khi 2 a pour utilité

De déterminer sans erreur possible si deux variables sont indépendante

De déterminer avec quel risque d'erreur on peut affirmer que deux variables sont indépendantes

S'utilise dans le cas de variables quantitatives

S'utilise dans le cas de variables qualitatives

J. Pour la sélection des composantes principales à utiliser lors d'une ACP, on retient celles correspondant

Aux valeurs propres pris dans l'ordre décroissant de leurs valeurs

Aux valeurs propres pris dans l'ordre croissant de leurs valeurs

Aux valeurs extrêmes (plus grande, plus petite) des valeurs propres

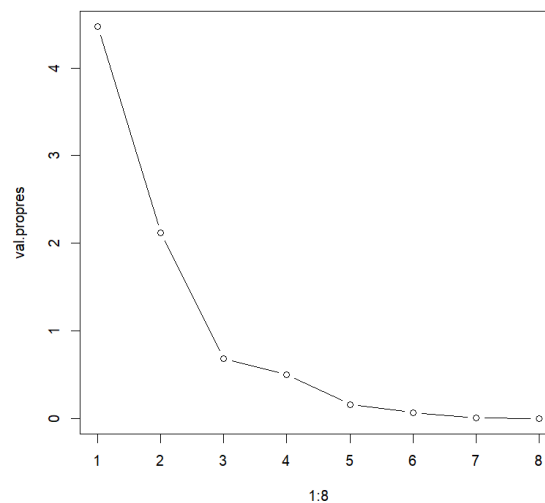
Aux sommes les plus grandes des coefficients contenus dans les vecteurs propres

Q2. Valeurs / vecteurs propres

Soit le tableau de valeurs propres suivant :

Comp . 1	Comp . 2	Comp . 3	Comp . 4	Comp . 5	Comp . 6	Comp . 7	Comp . 8
4 . 47	2 . 11	0 . 68	0 . 50	0 . 15	0 . 06	0 . 009	0 . 000006

et la courbe « en coude » correspondante :



Expliquez comment ces informations sont utilisées pour déterminer combien et quelles composantes principales seront retenues pour l'étude des données.

Réponse Q2

Q3. Fréquences théoriques

En reprenant la notation du tableau en Q1-H, donnez la formule de la fréquence théorique pour « $X=i$ / $Y=j$ ».

Q4. Interprétation d'une composante principale

Soit le vecteur propre ci-contre, pour lequel on donne les coefficients associés à chaque variable (notes dans différentes matières à un examen).

Comment peut-on interpréter les valeurs de chaque individu pour cette composante principale (valeur positive élevée, valeur négative basse).

Expliquez comment vous procédez pour arriver à votre interprétation.

	Vecteur
Mathématiques	+0,98
Littérature	-0,76
Gymnastique	-0,02
Langue vivante	-0,56
Physique	+0,34
Philosophie	+0,05
Informatique	+0,86