

Analyse Factorielle des Correspondances Multiples

Introduction

L'analyse factorielle des correspondances est une technique de statistique multidimensionnelle qui sert à représenter graphiquement un tableau croisé. Elle vise à réunir les informations les plus utiles de façon à donner une image claire de l'association de deux variables. Dans l'analyse des correspondances, les lignes représentent les catégories d'une première variable et les colonnes, les catégories d'une deuxième.

En pratique, dans les enquêtes statistiques on prévoit des questionnaires certainement plus de deux questions (à titre d'exemple l'enquête nationale des dépenses de consommation ENDCM :

https://drive.google.com/file/d/12GQqh9sCE6WAMNchdGmqcnsyQW_mPEik/view?usp=sharing)

par conséquent l'A.F.C. est naturellement limitée dans le domaine des traitements des données.

L'intérêt de (A.C.M.) **est de résoudre Le problème** statistique que pose alors ce type de données est l'analyse de la liaison pouvant exister entre un nombre quelconque de variables qualitatives.

L'Analyse des Correspondances Multiple (A.C.M.) est la méthode factorielle de Statistique Descriptive Multidimensionnelle qui permet de traiter ce genre de problématiques.

Objectifs

Mener une ACM provient de la nécessité de répondre aux questions qui vont comme suit :

- ❖ Quels sont les individus qui se ressemblent ?
- ❖ Quels sont ceux qui sont différents existe-t-il des groupes homogènes d'individus ?
- ❖ Existe-t-il des groupes homogènes d'individus ?
- ❖ Est-il possible de mettre en évidence une typologie des individus

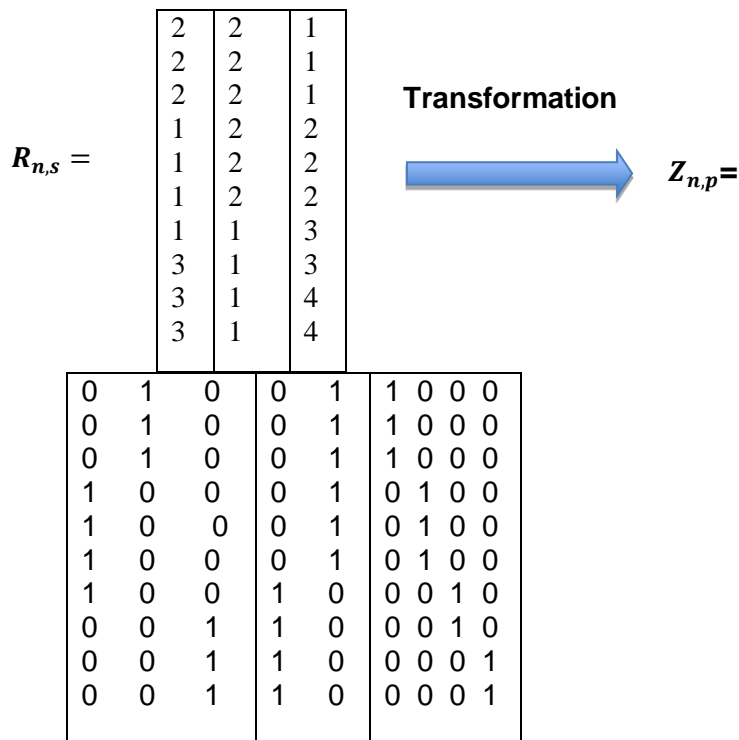
Notations et définitions :

🚦 Tableau de contingence :

Etant donnée qu'on est face à un problème qui présente des variables qualitatives, on procédera dans un premier temps par **codage binaire disjonctif**. Autrement dit, si l'individu possède la caractéristique on affecte 1 et 0 ailleurs.

L'analyse s'applique ensuite sur le tableau disjonctif complet issu de la transformation selon la même thématique de l'AFC c'est-à-dire les profils-lignes et les profils-colonnes. La distance de Khi-2 nous sera un outil puissant afin de déceler les liaisons existantes.

Ici $s=3$ et $p=9$



Construction d'un tableau disjonctif complet

On désigne par **I** l'ensemble des **n** sujets ayant répondu au questionnaire et par **p** le nombre total des modalités des **s** questions.

On a :

$$p = \sum_{q=1}^s p_q$$

On construit, à partir du tableau de données **R**, le tableau **Z** à **n** lignes et **p** colonnes décrivant les **Z** réponses des **n** individus par un codage binaire. Le tableau **Z** est la juxtaposition des sous-tableaux.

✚ **Tableau de BURT-Tableau des faces de l'hypercube de contingence :**

L'ensemble des *modalités* p_q de réponse à une question permet de partitionner l'échantillon en au plus p_q classes. La donnée de deux questions mises sous forme disjonctive complète permet de réaliser deux partitions de l'ensemble des individus enquêtés et l'on obtient un tableau de contingence. L'analyse du tableau croisant les deux partitions peut être généralisée au cas des partitions, **s**'étant un entier supérieur à **2**.

On construit, à partir du tableau disjonctif complet **Z**, le tableau symétrique **B** d'ordre (p,p) qui rassemble les croisements deux à deux de toutes les variables:

$$\mathbf{B} = \mathbf{Z}'\mathbf{Z}$$

B est appelé *tableau de contingence de Burt* associé au tableau disjonctif complet **Z**.

Principe

▪ Critère d'ajustement et distance de khi-2 :

Les individus sont tous affectés d'une masse identique $m_i = \frac{1}{n}$
 chacune des modalités j est pondérée par sa fréquence $m_i = \frac{z_j}{ns}$

La distance du *khi-2* appliquée à un tableau disjonctif complet conserve un sens. En effet, dans R^n , la distance entre modalités s'écrit :

$$d^2(j, j') = \sum_{i=1}^n n \left(\frac{z_{ij}}{z_j} - \frac{z_{ij'}}{z_{j'}} \right)^2$$

Ainsi deux modalités choisies par les mêmes individus coïncident. Par ailleurs, les modalités de faible effectif sont éloignées des autres modalités. Dans, la distance entre deux individus i et i' s'exprime par :

$$d^2(j, j') = \frac{1}{s} \sum_{j=1}^p \frac{n}{z_j} \left(\frac{n}{z_{ij}} - z_{i'j} \right)^2$$

Deux individus sont proches s'ils ont choisi les mêmes modalités.

Ils sont éloignés s'ils n'ont pas répondu de la même manière.

■ Axes factoriels et facteurs :

Les transformations opérées sur le tableau des données peuvent s'écrire à partir des trois matrices F , D_n et D_p qui définissent les éléments de base de l'analyse.

- F : d'ordre (n, p) désigne le tableau des fréquences relatives;
- D_n : d'ordre (n, n) est la matrice diagonale dont les éléments diagonaux sont les marges en lignes f_i ;
- D_p : est la matrice diagonale d'ordre (p, p) des marges en colonnes f_j

On a :

$$D_p = \frac{1}{ns} D \text{ De terme général : } f_j = \delta_{ij} \frac{z_j}{ns}$$

$$D_n = \frac{1}{n} I_n \text{ De terme général : } f_i = \frac{\delta_{ij}}{n}$$

Pour trouver les axes factoriels u_α on diagonalise la matrice :

$$S = F' D_n^{-1} F D_p^{-1} = \frac{1}{s} Z' Z D^{-1}$$

De terme général (**attention, s sans indice désigne le nombre de questions dans le questionnaire évoqué au début !**):

$$s_{jj'} = \frac{1}{s_{j'}} \sum_{i=1}^n z_{ij} z_{ij'}$$

Dans R^p , l'équation du $\alpha^{ième}$ axe factoriel u_α est :

$$\frac{1}{s} Z' Z D^{-1} u_\alpha = \lambda_\alpha u_\alpha$$

L'équation du $\alpha^{ième}$ facteur $\varphi_\alpha = D^{-1} u_\alpha$ s'écrit :

$$\frac{1}{s} D^{-1} Z' Z \varphi_\alpha = \lambda_\alpha \varphi_\alpha$$

L'équation du $\alpha^{ième}$ facteur ψ_α dans R^n s'écrit :

$$\frac{1}{s} Z D^{-1} Z' \psi_\alpha = \lambda_\alpha \psi_\alpha$$

Les facteurs φ_α et ψ_α (de norme λ_α) représentent les coordonnées des points-lignes et des points-colonnes sur l'axe factoriel α

Les relations de transition entre les facteurs φ_α et ψ_α sont :

$$\begin{cases} \varphi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} Z D^{-1} Z' \psi_\alpha \\ \psi_\alpha = \frac{1}{s\sqrt{\lambda_\alpha}} Z \varphi_\alpha \end{cases}$$

■ **Facteurs et relations quasi-barycentriques :**

La coordonnée factorielle de l'individu j sur l'axe α est donnée par :

$$\psi_{\alpha j} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^p \frac{z_{ij}}{z_{i.}} \varphi_{\alpha i}$$

Autrement dit :

$$\psi_{\alpha j} = \frac{1}{s\sqrt{\lambda_\alpha}} \sum_{i \in p(i)}^p \varphi_{\alpha i}$$

Où $p(i)$ désigne l'ensemble des modalités choisies par l'individu i .

Au coefficient $\frac{1}{\sqrt{\lambda_\alpha}}$ près, l'individu i se trouve au point moyen du nuage des modalités qu'il a choisies.

De même, la coordonnée de la modalité j sur l'axe α est donnée par :

$$\begin{aligned} \varphi_{\alpha j} &= \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^p \frac{z_{ij}}{z_{.j}} \psi_{\alpha i} \\ \varphi_{\alpha j} &= \frac{1}{z_{.j}\sqrt{\lambda_\alpha}} \sum_{i \in I(j)}^p \psi_{\alpha i} \end{aligned}$$

Avant la dilatation sur l'axe α , la modalité j se trouve au point moyen du nuage des individus qui l'ont choisie comme réponse. Dans le calcul des relations quasi-barycentriques [1.4 - 4], les individus ne sont pas pondérés. Il s'agit de simples calculs de moyennes arithmétiques de coordonnées.

Sous-nuage des modalités d'une même variable : Le nuage des modalités dans R^n peut être décomposé en z sous-espace, le $q^{ème}$ correspondant à l'ensemble des p_q modalités de la variable q . Ces sous nuages ont même centre de gravité G qui est celui du nuage global. En effet, les coordonnées des points du sous-nuage relatif à la variable q sont les colonnes de $Z_q D_q^{-1}$ et les éléments diagonaux de $\frac{1}{n} D_q$ sont les masses p_q relatives des *points* de ce sous-nuage.

Puisque :

$$\sum_{j \in p_q} z_{ij} = 1$$

Alors la $i^{ème}$ composante du centre de gravité du sous-nuage vaut :

$$G_{qi} = \sum_{j \in p_q} \frac{d_{jj}}{n} \frac{z_{ij}}{d_{jj}} = \frac{1}{n} = G_i$$

Où il apparaît que G_{qi} ne dépend pas de q .

Les composantes φ_q des modalités d'une variable q (relatives aux facteurs non-triviaux

φ_q) sont centrées puisque ces facteurs correspondent à une analyse du nuage après translation de l'origine en G.

Les facteurs opposent les modalités d'une même variable.

Remarques :

1. Si le tableau disjonctif n'est pas complet (c'est-à-dire si, pour au moins un individu, aucune modalité de réponse à une question n'a été choisie), les modalités d'une même variable ne sont plus centrées sur le centre de gravité du nuage global.
2. Le codage disjonctif complet permet de transformer une variable continue en une variable nominale dont les modalités sont des classes ordonnées. Il est alors utile de tracer la trajectoire qui relie les classes, trajectoire qui peut suggérer des liaisons non linéaires entre cette variable et les axes.

■ **Support du nuage des modalités :**

Les coordonnées des modalités dans R^n sont les colonnes de $Z D^{-1}$. Elles engendrent un sous-espace dont la dimension est le rang de $Z D^{-1}$ de, donc le rang de $Z =$

$[Z_1, \dots, Z_q, \dots, Z_s]$

Tous les sous-espaces engendrés par les Z_q ont en commun la première bissectrice. Le rang maximum de Z est donc :

$$p_1 + (p_2 - 1) + \dots + (p_s - 1) = p - s + 1$$

Le rang maximum de la matrice à diagonaliser 0 sera donc $p - s + 1$. Mais dans l'analyse du nuage par rapport à l'origine 0, la première bissectrice est vecteur propre correspondant à la valeur propre 1 (le nuage est contenu dans le sous-espace D^{-1} orthogonal à la première bissectrice).

Dans l'analyse par rapport au centre de gravité **G**, on trouvera donc $p - s$ valeurs propres non nulles. En choisissant une base dans le support du nuage, on pourra se ramener à la recherche d'éléments propres d'une matrice d'ordre $p - s$.

■ **Meilleure représentation simultanée :**

La présentation de l'analyse des correspondances peut être formulée ici de façon particulière en raison du codage spécifique au tableau disjonctif complet.

Nous cherchons sur un même axe les coordonnées des n individus et des p modalités de façon que :

❖ La coordonnée d'un individu i soit la moyenne arithmétique des coordonnées des modalités qu'il a choisies (à une dilatation P près, que l'on s'efforcera de rendre minimale).

❖ La coordonnée d'une modalité j soit la moyenne arithmétique des coordonnées des individus qui l'ont choisie (à une même dilatation P près). On obtient les relations dites quasi-barycentriques issues de l'analyse du tableau disjonctif complet Z dont le coefficient de dilatation β est de valeur minimale $\beta = \frac{1}{\sqrt{\lambda}}$.

$$\begin{cases} \varphi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} Z D^{-1} Z' \psi_\alpha \\ \psi_\alpha = \frac{1}{s\sqrt{\lambda_\alpha}} Z \varphi_\alpha \end{cases}$$

La représentation simultanée des individus et des modalités est importante pour l'interprétation des résultats. Cependant elle n'est pratiquement pas utilisée, d'une part pour des raisons d'encombrement graphique (on dispose souvent de plusieurs centaines voire de plusieurs milliers d'individus) et d'autre part parce que les individus sont, dans la plupart des applications, anonymes. Ils ne présentent de l'intérêt que par l'intermédiaire de leurs caractéristiques. On peut cependant vouloir projeter les individus sur un plan factoriel afin d'apprécier leur répartition et les zones de densité.

■ Inertie du nuage des modalités et conséquences pratiques :

$$d^2(j, G) = n \sum_{i=1}^n \left(\frac{z_{ij}}{z_j} - \frac{1}{n} \right)^2 = \frac{n}{z_j} - 1$$

La distance d'une modalité au centre de gravité est d'autant plus grande que l'effectif est plus faible.

❖ Inertie d'une modalité :

$$I(j) = m_j d^2(j, G) \text{ avec : } m_j = \frac{z_j}{n}$$

$$I(j) = \frac{1}{s} \left(1 - \frac{z_j}{n} \right)$$

La part d'inertie due à une modalité de réponse est d'autant plus grande que l'effectif dans cette modalité est plus faible.

Le maximum $\frac{1}{s}$ serait atteint par une modalité d'effectif nul. En conséquence, on évite, au moment du codage, les modalités à faibles effectifs susceptibles de perturber les directions des premiers axes factoriels.

❖ Inertie d'une question :

L'inertie de la question q , notée $I(q)$, vaut :

$$I(q) = \sum_{j=1}^{p_q} I(j) = \frac{1}{s} (p_q - 1)$$

Ainsi la part d'inertie due à une question est fonction croissante du nombre de modalités de réponse.

La part minimale $\frac{1}{s}$ correspond aux questions à 2 modalités. D'où l'intérêt d'équilibrer le système des questions, c'est-à-dire le découpage des variables 2 modalités. c'est-à-dire le découpage des variables en modalités, Si on veut faire jouer le même rôle à toutes les questions.

❖ Inertie totale :

On en déduit que l'inertie totale I vaut :

$$I = \sum_q I(q) = \sum_{j=1}^p \frac{z_j}{ns} d^2(j, G)$$

$$I = \frac{p}{s} - 1$$

En particulier, elle vaut 1 dans le cas où toutes les questions ont deux modalités de réponse (cas où $p=2s$).

L'inertie totale dépend uniquement du nombre de variables et de modalités et non des liaisons entre les variables. C'est une quantité qui, dans le cadre de l'analyse des correspondances multiples (comme dans celui de l'analyse en composantes principales normée), n'a pas de signification statistique.

■ Règles d'interprétation :

Dire qu'il existe des affinités entre réponses, c'est dire aussi qu'il existe des individus qui ont choisi simultanément toutes ou presque toutes ces réponses.

L'analyse des correspondances multiples met alors en évidence des types d'individus ayant des profils semblables quant aux attributs choisis pour les décrire. Compte tenu des distances entre les éléments du tableau disjonctif complet et des relations barycentriques particulières, on exprime :

- ❖ **La proximité entre individus en termes de ressemblances** : deux individus se ressemblent s'ils ont choisi globalement les mêmes modalités.
- ❖ **La proximité entre modalités de variables différentes en termes d'association** : ces modalités correspondent aux points moyens des individus qui les ont choisies et sont proches parce qu'elles concernent globalement les mêmes individus ou des individus semblables.
- ❖ **La proximité entre deux modalités d'une même variable en termes de ressemblance** : par construction, les modalités d'une même variable s'excluent. Si elles sont proches, cette proximité s'interprète en termes de ressemblance entre les groupes d'individus qui les ont choisies (vis-à-vis d'autres variables actives de l'analyse).

Les règles d'interprétation des résultats (coordonnées, contributions, cosinus carrés) concernant les éléments actifs d'une analyse des correspondances multiples **sont sensiblement les mêmes** que celles d'une analyse des correspondances simple. On calcule la contribution et la qualité de représentation de chaque modalité et de chaque individu, si ceux-ci ne sont pas anonymes pour l'analyse. Cependant, la notion de variable doit être prise en compte au moment de l'interprétation, ceci au travers de ses modalités. Compte tenu de la décomposition de l'inertie du nuage des modalités, on calcule la contribution d'une variable au facteur a en sommant les contributions de ses modalités sur ce facteur. On repère ainsi, en plus des modalités responsables des axes factoriels, les variables qui ont participé à la définition du facteur. On obtient un indicateur de liaison entre la variable et le facteur.

Application sur le logiciel R :

1-Présentation de la base des données :

Autant qu'un package R dédié à la statistique exploratoire multidimensionnelle. FactoMiner, comporte une base de données "data(poison)" constituée de 55 lignes et 15 colonnes, faisant référence à une enquête réalisée sur un échantillon d'enfants du primaire ayant souffert d'une intoxication alimentaire ces derniers ont été interrogés sur leurs symptômes et sur ce qu'ils ont mangé.

2-visualisations des valeurs propres :

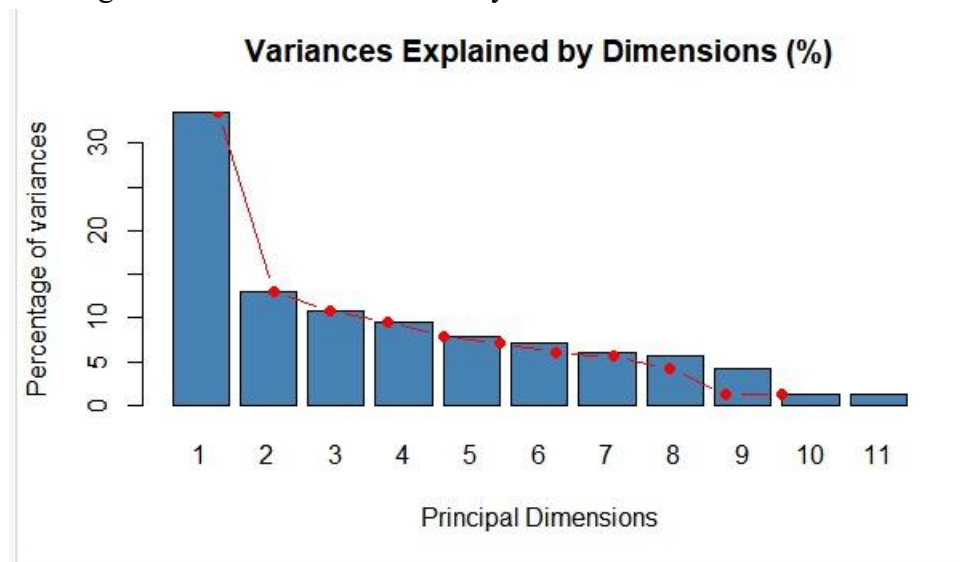
Le graphique ci-dessous montre le pourcentage de la variance expliqué par chaque axe.

Les deux premiers axes de l'ACM expliquent 46.44% de l'inertie totale du jeu de données en question ; Cela signifie que 46.44% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan.

C'est un pourcentage relativement moyen, et le premier plan représente donc seulement une part de la variabilité contenue dans l'ensemble du jeu de données actif.

Cette valeur est supérieure à la valeur référence de 33.27%, la variabilité expliquée par ce plan est donc significative (cette inertie de référence est le quantile 0.95 de la distribution des pourcentages d'inertie obtenus en simulant 864 jeux de données aléatoires de dimensions comparables sur la base d'une distribution uniforme).

Du fait de ces observations, on préconise de considérer également les dimensions supérieures ou égales à la troisième dans l'analyse.



3-Coordonnées et contributions des variables :


```
> res.var$coord
```

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
Nausea_n	0.26739087	0.121390290	-0.265583253	0.03376130	0.073704999
Nausea_y	-0.95815062	-0.434981874	0.951673323	-0.12097801	-0.264109581
Vomit_n	0.47902794	-0.409194649	0.084492799	0.27361142	0.052452504
Vomit_y	-0.71854191	0.613791974	-0.126739198	-0.41041713	-0.078678757
Abdo_n	1.31802207	-0.035745005	-0.005094243	-0.15360951	-0.069869870
Abdo_y	-0.64119993	0.017389462	0.002478280	0.07472895	0.033990747
Fever_n	1.17183098	-0.174895110	0.097275290	-0.18967098	-0.018478570
Fever_y	-0.66961770	0.099940063	-0.055585880	0.10838342	0.010559183
Diarrhea_n	1.18282245	-0.002756840	-0.082979839	-0.24123007	-0.104910816
Diarrhea_y	-0.67589854	0.001575337	0.047417051	0.13784575	0.059949038
Potato_n	-0.70730851	-2.619110755	2.138637690	0.37461658	1.355830579
Potato_y	0.04080626	0.151102544	-0.123382944	-0.02161249	-0.078220995
Fish_n	-0.62079031	-1.213389338	-3.691000244	5.56976822	0.154864481
Fish_y	0.01149612	0.022470173	0.068351856	-0.10314386	-0.002867861
Mayo_n	1.31277835	0.394914774	0.417481883	0.37330832	-0.316756981
Mayo_y	-0.29172852	-0.087758839	-0.092773752	-0.08295741	0.070390440
Courg_n	-0.39073240	2.112852279	0.728297960	0.32368182	1.125993240
Courg_y	0.03907324	-0.211285228	-0.072829796	-0.03236818	-0.112599324
Cheese_n	1.15462219	0.605560528	0.283133231	0.36031062	1.681357074
Cheese_y	-0.16838240	-0.088310910	-0.041290263	-0.05254530	-0.245197907
Icecream_n	0.61038541	1.340331612	1.917576423	1.71261979	-1.458589548
Icecream_y	-0.04787337	-0.105124048	-0.150398151	-0.13432312	0.114399180

```
> res.var$contrib
```

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
Nausea_n	1.515868554	8.110001e-01	4.670018e+00	0.08449397	4.897791e-01
Nausea_y	5.431862319	2.906084e+00	1.673423e+01	0.30277007	1.755042e+00
Vomit_n	3.733666829	7.072263e+00	3.627455e-01	4.25893721	1.903638e-01
Vomit_y	5.600500244	1.060839e+01	5.441183e-01	6.38840581	2.855456e-01
Abdo_n	15.417636578	2.943661e-02	7.192511e-04	0.73219636	1.842427e-01
Abdo_y	7.500471849	1.432051e-02	3.499060e-04	0.35620363	8.963157e-02
Fever_n	13.541285078	7.830146e-01	2.913961e-01	1.24036823	1.431873e-02
Fever_y	7.737877188	4.474369e-01	1.665121e-01	0.70878185	8.182133e-03
Diarrhea_n	13.796503952	1.945529e-04	2.120430e-01	2.00637332	4.615390e-01
Diarrhea_y	7.883716544	1.111731e-04	1.211674e-01	1.14649904	2.637366e-01
Potato_n	0.740012332	2.633986e+01	2.112732e+01	0.72579521	1.156299e+01
Potato_y	0.042693019	1.519608e+00	1.218884e+00	0.04187280	6.670953e-01
Fish_n	0.190015838	1.884451e+00	2.097668e+01	53.48021026	5.028535e-02
Fish_y	0.003518812	3.489724e-02	3.884571e-01	0.99037426	9.312103e-04
Mayo_n	8.497335124	1.996141e+00	2.683638e+00	2.40244916	2.103731e+00
Mayo_y	1.888296694	4.435870e-01	5.963641e-01	0.53387759	4.674958e-01
Courg_n	0.376381955	2.856885e+01	4.083541e+00	0.90307845	1.329167e+01
Courg_y	0.037638196	2.856885e+00	4.083541e-01	0.09030784	1.329167e+00
Cheese_n	4.601270591	3.285471e+00	8.640292e-01	1.56664651	4.149118e+01
Cheese_y	0.671018628	4.791313e-01	1.260043e-01	0.22846928	6.050798e+00
Icecream_n	0.734798428	9.197484e+00	2.264718e+01	20.22556991	1.784284e+01
Icecream_y	0.057631249	7.213713e-01	1.776249e+00	1.58631921	1.399439e+00

4-Coordonnées et contributions des individus :

```
> res.ind$coord
```

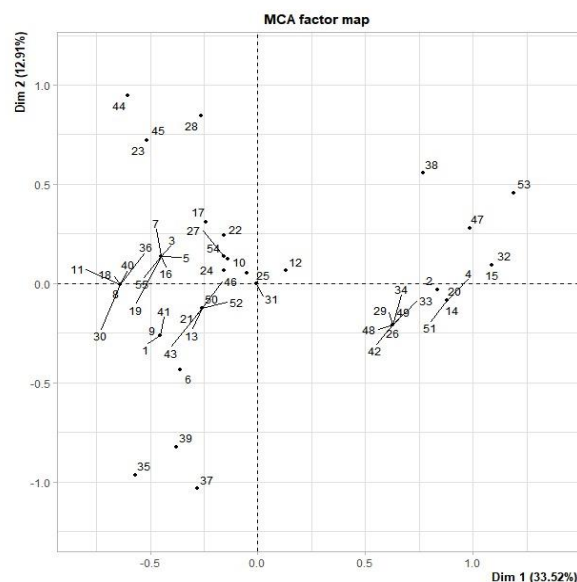
	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
Nausea_n	0.26739087	0.121390290	-0.265583253	0.03376130	0.073704999
Nausea_y	-0.95815062	-0.434981874	0.951673323	-0.12097801	-0.264109581
Vomit_n	0.47902794	-0.409194649	0.084492799	0.27361142	0.052452504
Vomit_y	-0.71854191	0.613791974	-0.126739198	-0.41041713	-0.078678757
Abdo_n	1.31802207	-0.035745005	-0.005094243	-0.15360951	-0.069869870
Abdo_y	-0.64119993	0.017389462	0.002478280	0.07472895	0.033990747
Fever_n	1.17183098	-0.174895110	0.097275290	-0.18967098	-0.018478570
Fever_y	-0.66961770	0.099940063	-0.055585880	0.10838342	0.010559183
Diarrhea_n	1.18282245	-0.002756840	-0.082979839	-0.24123007	-0.104910816
Diarrhea_y	-0.67589854	0.001575337	0.047417051	0.13784575	0.059949038
Potato_n	-0.70730851	-2.619110755	2.138637690	0.37461658	1.355830579
Potato_y	0.04080626	0.151102544	-0.123382944	-0.02161249	-0.078220995
Fish_n	-0.62079031	-1.213389338	-3.691000244	5.56976822	0.154864481
Fish_y	0.01149612	0.022470173	0.068351856	-0.10314386	-0.002867861
Mayo_n	1.31277835	0.394914774	0.417481883	0.37330832	-0.316756981
Mayo_y	-0.29172852	-0.087758839	-0.092773752	-0.08295741	0.070390440
Courg_n	-0.39073240	2.112852279	0.728297960	0.32368182	1.125993240
Courg_y	0.03907324	-0.211285228	-0.072829796	-0.03236818	-0.112599324
Cheese_n	1.15462219	0.605560528	0.283133231	0.36031062	1.681357074
Cheese_y	-0.16838240	-0.088310910	-0.041290263	-0.05254530	-0.245197907
Icecream_n	0.61038541	1.340331612	1.917576423	1.71261979	-1.458589548
Icecream_y	-0.04787337	-0.105124048	-0.150398151	-0.13432312	0.114399180

```
> res.ind$contrib
```

	Dim 1	Dim 2	Dim 3	2.856885e+01	4.083541e+00	0.90307845
Courg_y	0.037638196	2.856885e+00	4.083541e-01	0.09030784	1.329167e+00	
Cheese_n	4.601270591	3.285471e+00	8.640292e-01	1.56664651	4.149118e+01	
Cheese_y	0.671018628	4.791313e-01	1.260043e-01	0.22846928	6.050798e+00	
Icecream_n	0.734798428	9.197484e+00	2.264718e+01	20.22556991	1.784284e+01	
Icecream_y	0.057631249	7.213713e-01	1.776249e+00	1.58631921	1.399439e+00	

5-Carte des individus :

La dimension 1 oppose des individus tels que 4, 14, 20, 51, 32, 15, 26, 29, 53 et 2 (à droite du graphe, caractérisés par une coordonnée fortement positive sur l'axe) à des individus comme 23, 45, 28, coordonnée sur l'axe.



✚ **Le groupe** auquel les individus **4, 14, 20, 51, 26, 29 et 2** appartiennent (caractérisés par une coordonnée positive sur l'axe) partage :

- Une forte fréquence des modalités Diarrhae=Diarrhea_n, Abdominals=Abdo_n, Fever=Fever_n, Vomiting=Vomit_n et Nausea=Nausea_n (du plus commun au plus rare).
- Une faible fréquence des modalités Abdominals=Abdo_y, Diarrhae=Diarrhea_y, Fever=Fever_y, Vomiting=Vomit_y et Nausea=Nausea_y (du plus rare au plus commun).

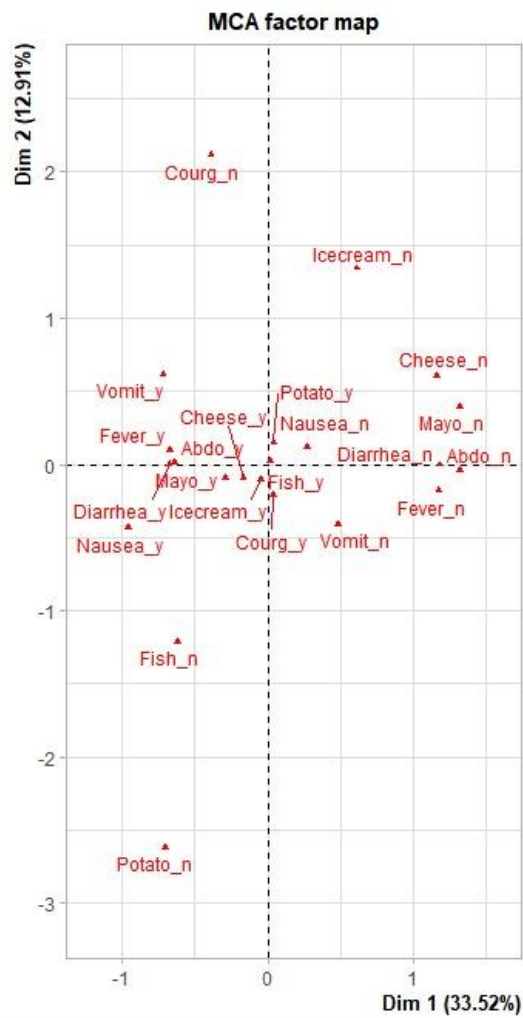
✚ **Le groupe** auquel les individus **32, 15, 53 et 47** appartiennent (caractérisés par une coordonnée positive sur l'axe) partage :

- Une forte fréquence des modalités Cheese=Cheese_n, Mayo=Mayo_n, Icecream=Icecream_n, Diarrhae=Diarrhea_n et Abdominals=Abdo_n (du plus commun au plus rare).
- Une faible fréquence des modalités Cheese=Cheese_y, Mayo=Mayo_y, Icecream=Icecream_y, Abdominals=Abdo_y et Diarrhae=Diarrhea_y (du plus rare au plus commun).

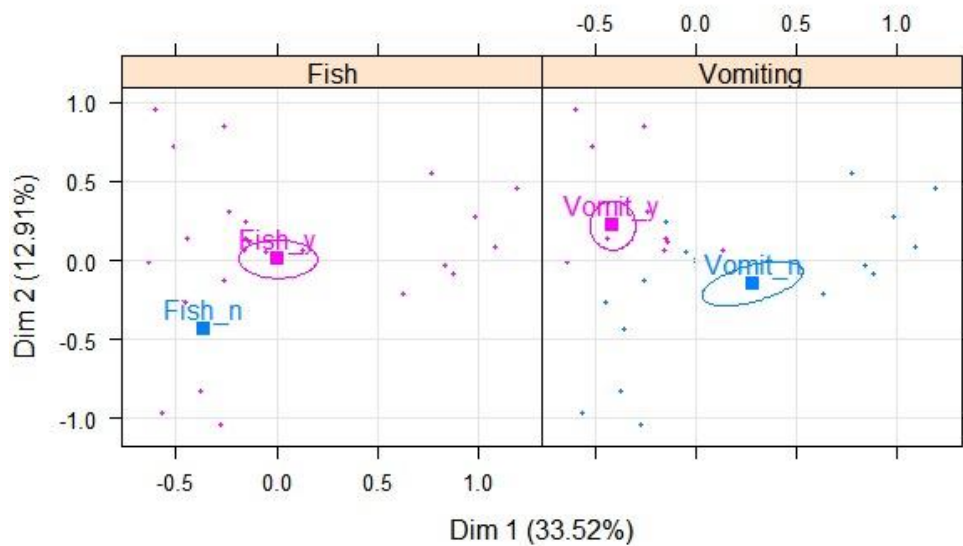
✚ **Le groupe 3** (caractérisés par une coordonnée négative sur l'axe) partage :

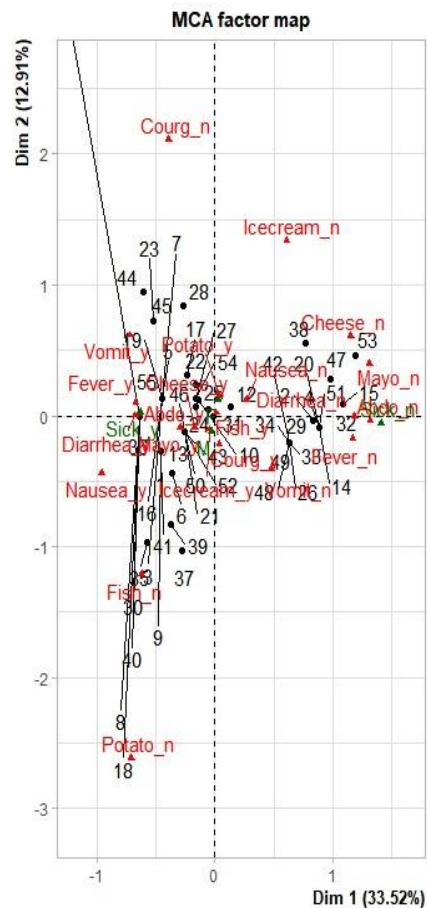
- une forte fréquence des modalités Vomiting=Vomit_y, Fever=Fever_y, Diarrhae=Diarrhea_y, Abdominals=Abdo_y, Mayo=Mayo_y et Nausea=Nausea_y (du plus commun au plus rare).
- une faible fréquence des modalités Vomiting=Vomit_n, Fever=Fever_n, Abdominals=Abdo_n, Diarrhae=Diarrhea_n, Mayo=Mayo_n et Nausea=Nausea_n (du plus rare au plus commun).

6-Carte des variables :



7- Représentation simultanée :





8-Qualité de la représentation des variables :

```
> res.var$cos2
```

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
Nausea_n	0.256200730	5.280258e-02	2.527485e-01	0.004084375	0.0194661965
Nausea_y	0.256200730	5.280258e-02	2.527485e-01	0.004084375	0.0194661965
Vomit_n	0.344201648	2.511604e-01	1.070855e-02	0.112294813	0.0041268978
Vomit_y	0.344201648	2.511604e-01	1.070855e-02	0.112294813	0.0041268978
Abdo_n	0.845115652	6.215864e-04	1.262496e-05	0.011479077	0.0023749291
Abdo_y	0.845115652	6.215864e-04	1.262496e-05	0.011479077	0.0023749291
Fever_n	0.784678768	1.747903e-02	5.407133e-03	0.020557189	0.0001951186
Fever_y	0.784678768	1.747903e-02	5.407133e-03	0.020557189	0.0001951186
Diarrhea_n	0.799467973	4.342953e-06	3.934659e-03	0.033252541	0.0062893025
Diarrhea_y	0.799467973	4.342953e-06	3.934659e-03	0.033252541	0.0062893025
Potato_n	0.028862615	3.957543e-01	2.638714e-01	0.008096399	0.1060544169
Potato_y	0.028862615	3.957543e-01	2.638714e-01	0.008096399	0.1060544169
Fish_n	0.007136678	2.726507e-02	2.522867e-01	0.574487370	0.0004441298
Fish_y	0.007136678	2.726507e-02	2.522867e-01	0.574487370	0.0004441298
Mayo_n	0.382974888	3.465726e-02	3.873136e-02	0.030968690	0.0222966634
Mayo_y	0.382974888	3.465726e-02	3.873136e-02	0.030968690	0.0222966634
Courg_n	0.015267181	4.464145e-01	5.304179e-02	0.010476992	0.1267860776
Courg_y	0.015267181	4.464145e-01	5.304179e-02	0.010476992	0.1267860776
Cheese_n	0.194418058	5.347760e-02	1.169065e-02	0.018932629	0.4122652350
Cheese_y	0.194418058	5.347760e-02	1.169065e-02	0.018932629	0.4122652350
Icecream_n	0.029221204	1.409011e-01	2.883999e-01	0.230044435	0.1668614487
Icecream_y	0.029221204	1.409011e-01	2.883999e-01	0.230044435	0.1668614487

9-Qualité de la représentation des individus :

```
> res.ind$cos2
```

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
Nausea_n	0.256200730	5.280258e-02	2.527485e-01	0.004084375	0.0194661965
Nausea_y	0.256200730	5.280258e-02	2.527485e-01	0.004084375	0.0194661965
Vomit_n	0.344201648	2.511604e-01	1.070855e-02	0.112294813	0.0041268978
Vomit_y	0.344201648	2.511604e-01	1.070855e-02	0.112294813	0.0041268978
Abdo_n	0.845115652	6.215864e-04	1.262496e-05	0.011479077	0.0023749291
Abdo_y	0.845115652	6.215864e-04	1.262496e-05	0.011479077	0.0023749291
Fever_n	0.784678768	1.747903e-02	5.407133e-03	0.020557189	0.0001951186
Fever_y	0.784678768	1.747903e-02	5.407133e-03	0.020557189	0.0001951186
Diarrhea_n	0.799467973	4.342953e-06	3.934659e-03	0.033252541	0.0062893025
Diarrhea_y	0.799467973	4.342953e-06	3.934659e-03	0.033252541	0.0062893025
Potato_n	0.028862615	3.957543e-01	2.638714e-01	0.008096399	0.1060544169
Potato_y	0.028862615	3.957543e-01	2.638714e-01	0.008096399	0.1060544169
Fish_n	0.007136678	2.726507e-02	2.522867e-01	0.574487370	0.0004441298
Fish_y	0.007136678	2.726507e-02	2.522867e-01	0.574487370	0.0004441298
Mayo_n	0.382974888	3.465726e-02	3.873136e-02	0.030968690	0.0222966634
Mayo_y	0.382974888	3.465726e-02	3.873136e-02	0.030968690	0.0222966634
Courg_n	0.015267181	4.464145e-01	5.304179e-02	0.010476992	0.1267860776
Courg_y	0.015267181	4.464145e-01	5.304179e-02	0.010476992	0.1267860776
Cheese_n	0.194418058	5.347760e-02	1.169065e-02	0.018932629	0.4122652350
Cheese_y	0.194418058	5.347760e-02	1.169065e-02	0.018932629	0.4122652350
Icecream_n	0.029221204	1.409011e-01	2.883999e-01	0.230044435	0.1668614487
Icecream_y	0.029221204	1.409011e-01	2.883999e-01	0.230044435	0.1668614487

10-Résultats des variables quantitatives supplémentaires :

```
> res.mca$quali.sup
```

\$coord

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
sick_n	1.40983382	-0.05780887	0.03535043	-0.12299072	-0.05644922
sick_y	-0.63071513	0.02586186	-0.01581467	0.05502217	0.02525360
F	0.02736360	0.11839736	0.08920658	-0.03662471	-0.03965681
M	-0.02837707	-0.12278244	-0.09251053	0.03798119	0.04112558

\$cos2

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
sick_n	0.8892035226	0.001495045	0.0005590553	0.006767216	0.001425546
sick_y	0.8892035226	0.001495045	0.0005590553	0.006767216	0.001425546
F	0.0007764989	0.014537117	0.0082525484	0.001391050	0.001630909
M	0.0007764989	0.014537117	0.0082525484	0.001391050	0.001630909

\$v.test

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
sick_n	6.9294293	-0.2841346	0.1737498	-0.6045078	-0.2774518
sick_y	-6.9294293	0.2841346	-0.1737498	0.6045078	0.2774518
F	0.2047705	0.8860047	0.6675609	-0.2740743	-0.2967644
M	-0.2047705	-0.8860047	-0.6675609	0.2740743	0.2967644

\$eta2

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
sick	0.8892035226	0.001495045	0.0005590553	0.006767216	0.001425546
Sex	0.0007764989	0.014537117	0.0082525484	0.001391050	0.001630909

