

M1 Économétrie et Statistiques Appliquées

Analyse des Données Qualitatives

Cours en ligne du 1er décembre

Nils Berglund

Institut Denis Poisson, Université d'Orléans



Décembre 2020

ACM : Résumé

- ▷ n : nombre d'individus, s : nombre de variables (questions)
- ▷ p_q : nombres de modalités de la variable q
- ▷ nombre total de modalités : $p = \sum_{i=1}^s p_q$
- ▷ r_{iq} : réponse de l'individu i à la question q
- ▷ Tableau disjonctif : $Z = [Z_1|Z_2|\dots|Z_q|\dots|Z_s]$ matrice $n \times p$
$$z_{ij}^{(q)} = z_{ij} = \begin{cases} 1 & \text{si } r_{iq} = j \\ 0 & \text{sinon} \end{cases}$$
- ▷ Tableau de Burt : $B = Z^T Z$ matrice $p \times p$, avec des blocs $Z_q^T Z_{q'}$
 $Z_q^T Z_q$ est une matrice diagonale d'éléments $z_{\bullet j} = \sum_{i=1}^n z_{ij}$
 $q \neq q'$: $Z_q^T Z_{q'}$ est la table de contingence croisant les variables q et q'

Principe de l'ACM

- ▷ **Normalisation** : chaque ligne a la masse $m_i = \frac{1}{n}$, $\sum_{i=1}^n m_i = 1$

La colonne j a la masse $m_j = \frac{z_{\bullet j}}{ns}$, $\sum_{j=1}^p m_j = 1$

- ▷ On procède comme pour l'AFC avec les matrices

$$F = \frac{1}{ns}Z \quad D_p = \frac{1}{ns}D \quad D_n = \frac{1}{n}I$$

- ▷ Alors $S = \frac{1}{s}BD^{-1}$. Ses valeurs propres sont les inerties λ_α .

Principe de l'ACM

▷ Centre de gravité des colonnes: $G_j = \frac{1}{n}$

$$d^2(j, G) = \frac{n}{z_{\bullet j}} - 1$$

▷ Inertie de la modalité j : $\frac{1}{s} \left(1 - \frac{z_{\bullet j}}{n} \right)$

▷ Inertie de la variable q : $I(q) = \frac{1}{s} (p_q - 1)$

▷ Inertie totale: $I = \sum_{q=1}^s I(q) = \frac{p}{s} - 1$

TD révisions exercice 1

Le tableau suivant représente la couleur des cheveux, celle des yeux, et le sexe dans une population de 12 individus.

	Cheveux	Yeux	Sexe
1	Noir	Bleu	Homme
2	Noir	Bleu	Femme
3	Noir	Brun	Homme
4	Noir	Brun	Homme
5	Brun	Bleu	Femme
6	Brun	Brun	Homme
7	Brun	Brun	Homme
8	Roux	Bleu	Homme
9	Roux	Bleu	Homme
10	Roux	Bleu	Femme
11	Roux	Brun	Femme
12	Roux	Brun	Femme

1. Déterminer les tableaux de contingence des variables croisées deux à deux.
2. Représenter les données sous forme de tableau disjonctif Z et calculer ses marges.
3. Calculer le tableau de Burt $B = Z^T Z$.

TD révisions exercice 1

TD révisions exercice 2

Un fabricant de téléphones portables a effectué une étude de marché dans une population de 12 étudiants. Le tableau suivant représente la couleur du téléphone portable, le système d'exploitation de l'ordinateur de chaque étudiant, ainsi que la faculté où il est inscrit.

	Couleur	Système	Faculté
1	Noir	Windows	Sciences
2	Noir	Linux	DEG
3	Argenté	Windows	Sciences
4	Argenté	Mac	DEG
5	Argenté	Linux	Sciences
6	Argenté	Linux	Sciences
7	Argenté	Linux	Sciences
8	Bleu ciel	Mac	DEG
9	Bleu ciel	Linux	Sciences
10	Vert menthe	Mac	Sciences
11	Vert menthe	Mac	DEG
12	Vert menthe	Mac	DEG

TD révisions exercice 2

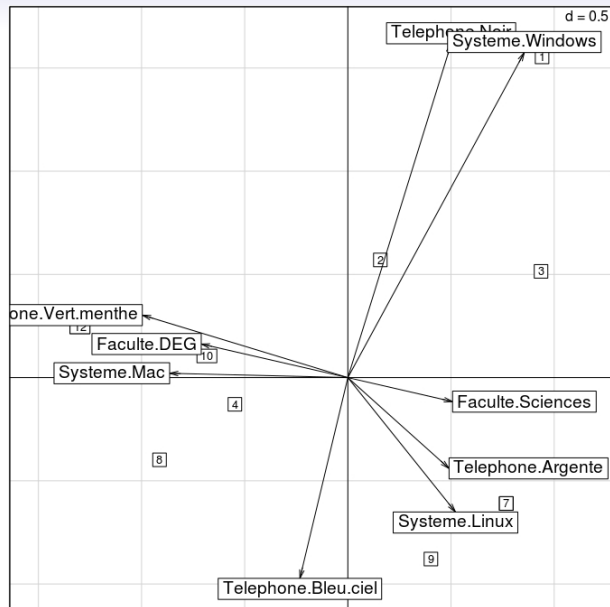
1. Déterminer les tableaux de contingence des variables croisées deux à deux.
2. Représenter les données sous forme de tableau disjonctif Z et calculer ses marges.
3. Calculer le tableau de Burt $B = Z^T Z$.

TD révisions exercice 2

4. Au vu du résultat de l'ACM ci-dessous, quelle stratégie conseilleriez-vous au vendeur participant au forum en faculté DEG?

```
> acmin$TOT
      inertia      cum      ratio
1 0.72427724 0.7242772 0.3621386
2 0.43333543 1.1576127 0.5788063
3 0.35498491 1.5125976 0.7562988
4 0.26414381 1.7767414 0.8883707
5 0.19115518 1.9678966 0.9839483
6 0.03210343 2.0000000 1.0000000
```

TD révisions Exercice 3



TD révisions Exercice 3

```
> acmin$row.abs
```

	Axis1	Axis2	Axis3
1	1020	4643	267
2	28	626	2283
3	1009	510	969
4	345	32	174
5	678	715	258
6	678	715	258
7	678	715	258
8	959	305	2169
9	186	1485	1908
10	537	21	995
11	1941	116	230
12	1941	116	230

```
> acmin$row.rel
```

	Axis1	Axis2	Axis3	con.tra
1	2481	6761	-319	1488
2	95	1251	-3741	1083
3	3697	1117	1741	988
4	-2142	-119	528	583
5	5032	-3174	938	488
6	5032	-3174	938	488
7	5032	-3174	938	488
8	-3205	-610	-3554	1083
9	683	-3257	-3428	988
10	-2738	65	2486	710
11	-8724	313	507	806
12	-8724	313	507	806

TD révisions Exercice 3

```
> acmin$col.abs
```

	Comp1	Comp2	Comp3
Comp3 con.tra			
Telephone.Argente	768	617	1607
Telephone.Bleu.ciel	69	1207	3825
Telephone.Noir	320	3338	1931
Telephone.Vert.menthe	1900	176	1199
Systeme.Linux	868	1352	352
Systeme.Mac	2380	1	152
Systeme.Windows	933	3165	103
Faculte.DEG	1611	85	485
Faculte.Sciences	1151	60	347

```
> acmin$col.rel
```

	Comp1	Comp2	Comp3	con.tra
Telephone.Argente	2859	-1374	2934	972
Telephone.Bleu.ciel	-180	-1882	-4888	1389
Telephone.Noir	833	5207	-2467	1389
Telephone.Vert.menthe	-5503	304	1702	1250
Systeme.Linux	3234	-3013	-642	972
Systeme.Mac	-8867	3	278	972
Systeme.Windows	2434	4937	131	1389
Faculte.DEG	-6001	189	-886	972
Faculte.Sciences	6001	-189	886	694

Interprétation des axes

Notes