



FACULTÉ DE DROIT  
ÉCONOMIE & GESTION

**Sujet d'examen des semestres impairs 1<sup>ère</sup> session  
Année universitaire 2022-2023**

<b>Intitulé de l'épreuve :</b>	Analyse de données qualitatives
<b>Nom de l'enseignant :</b>	Nils Berglund
<b>Année :</b>	M1 ESA
<b>Mention / Spécialité / Parcours :</b>	2022-2023
<b>Durée de l'épreuve :</b>	2 heures
<b>Documents autorisés :</b>	Résumé manuscrit de 4 pages A4
<b>Matériels autorisés :</b>	Calculatrice non programmable
	15/12/2022

P1/5

---

**SUJET**

Les téléphones portables, ordinateurs et tablettes doivent être éteints durant l'examen.  
Il sera tenu compte de la qualité de la rédaction.  
Les points sont donnés à titre indicatif.

**Problème 1 [7 points]**

Le tableau de contingence suivant montre les effectifs pour un jeu de données croisant deux variables qualitatives. Les variables ont respectivement 4 et 3 modalités. Celles de la première sont notées A, B, C, D, alors que la celles de la seconde sont notées a, b, c.

	a	b	c
A	100	0	50
B	60	60	30
C	80	480	40
D	160	60	80

1. Donner la définition de la propriété d'équivalence distributionnelle. Utiliser cette propriété afin de simplifier le tableau de contingence ci-dessus.
2. Illustrer la propriété d'équivalence distributionnelle dans un cas particulier (pour un choix particulier de lignes ou de colonnes).
3. Effectuer un test d'indépendance du chi-deux *sur le tableau simplifié*, en expliquant les étapes principales du calcul. Peut-on rejeter l'hypothèse d'indépendance à un ou plusieurs seuils indiqués ci-dessous ?

Tableau de valeurs de la loi du  $\chi^2$  :

Degrés de liberté	90%	95%	97.5%	99%	99.9%
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.833	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322
8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588
11	17.275	19.675	21.920	24.725	31.264
12	18.549	21.026	23.337	26.217	32.910

4. Calculer, *pour le tableau simplifié*, les matrices  $X$  et  $Y$  des profils-ligne et des profils-colonne. Il est conseillé d'écrire ces matrices sous forme de fractions.  
Déterminer, avec un minimum de calculs, les valeurs propres des matrices  $S = X^T Y^T$  et  $T = Y^T X^T$ .
5. Que peut-on dire sur les nuages de points (quelle est la dimension des espaces dans lesquels ils se trouvent) ? Expliquer, sans faire de calcul, comment on détermine leur forme.

**Problème 2 [6 points]**

On considère la matrice

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}.$$

1. Déterminer les valeurs propres de  $A$ .
2. Déterminer, pour chaque valeur propre de  $A$ , un vecteur propre de norme 1.
3. Soit  $\mathcal{E}_c$  l'ellipse d'équation

$$\begin{bmatrix} x & y \end{bmatrix} A^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = c^2 \quad (1)$$

Que vaut son demi-grand axe ? Son demi-petit axe ?

4. Calculer les points d'intersection de l'ellipse avec les axes de coordonnées, c'est-à-dire les points de l'ellipse de la forme  $(x, 0)$  et  $(0, y)$  avec  $x, y \in \mathbb{R}$ .
5. Esquisser l'ellipse  $\mathcal{E}_c$  dans le plan  $(x, y)$  pour  $c = 1$  et  $c = \frac{1}{2}$ .
6. Si  $A$  est la matrice de covariance d'un jeu de données, que représentent les ellipses  $\mathcal{E}_c$  ?

### Problème 3 [7 points]

Le tableau suivant montre le nombre de brevets en télécommunication déposés dans 9 pays, de 1980 à 1986.

	1986	1985	1984	1983	1982	1981	1980
USA	986	774	711	591	467	404	258
Japon	653	552	361	307	195	129	43
Allemagne	405	357	347	254	294	313	208
France	189	158	200	171	153	184	147
GB	204	182	158	137	92	86	67
Italie	31	28	28	21	22	29	15
Pays-Bas	64	59	61	64	33	30	11
Suède	25	19	31	25	15	12	13
Suisse	23	34	19	30	17	30	15

1. Expliquer les lignes ci-dessous fournies par le logiciel R.

```
> summary(as.table(tab))
Number of cases in table: 11541
Number of factors: 2
Test for independence of all factors:
Chisq = 516.4, df = 48, p-value = 9.223e-80
```

Les pages suivantes montrent le résultat d'une AFC effectuée sur ces données.

2. Discuter les inerties.
3. Y a-t-il des points mal représentés sur le plan des deux premiers axes factoriels ? Et dans l'espace des trois premiers axes factoriels ?
4. Interpréter, dans la mesure du possible, les trois premiers axes factoriels.
5. Citer au moins deux exemples de groupes de pays dont l'évolution du nombre de brevets est similaire.
6. Quels sont les pays qui ont connu la plus forte croissance de dépôts de brevets ? La plus faible croissance ? Une décroissance ?
7. Quelle est la particularité de la Suisse ?

Inerties:

```
> afc <- dudi.coa(tab, scan = FALSE, nf=4)
> afcin <- inertia.dudi(afc, col.inertia=T, row.inertia=T)
> afcin$TOT
      inertia      cum      ratio
1 3.981921e-02 0.03981921 0.8898905
2 2.643138e-03 0.04246234 0.9489601
3 1.500220e-03 0.04396256 0.9824874
4 6.113650e-04 0.04457393 0.9961504
5 1.279688e-04 0.04470190 0.9990103
6 4.428695e-05 0.04474619 1.0000000
```

Contributions absolues et relatives des lignes:

```
> afcin$row.abs
      Axis1 Axis2 Axis3 Axis4
USA      121   186   935   545
Japon   4867   613   190   299
Allemagne 1865  2773    6   673
France  2697   407    20  1320
GB        10   436    0  2961
Italie   177   216   289   613
Pays-Bas  47  3268  1077  3182
Suede    34  2092   648    48
Suisse   181    9  6834   359

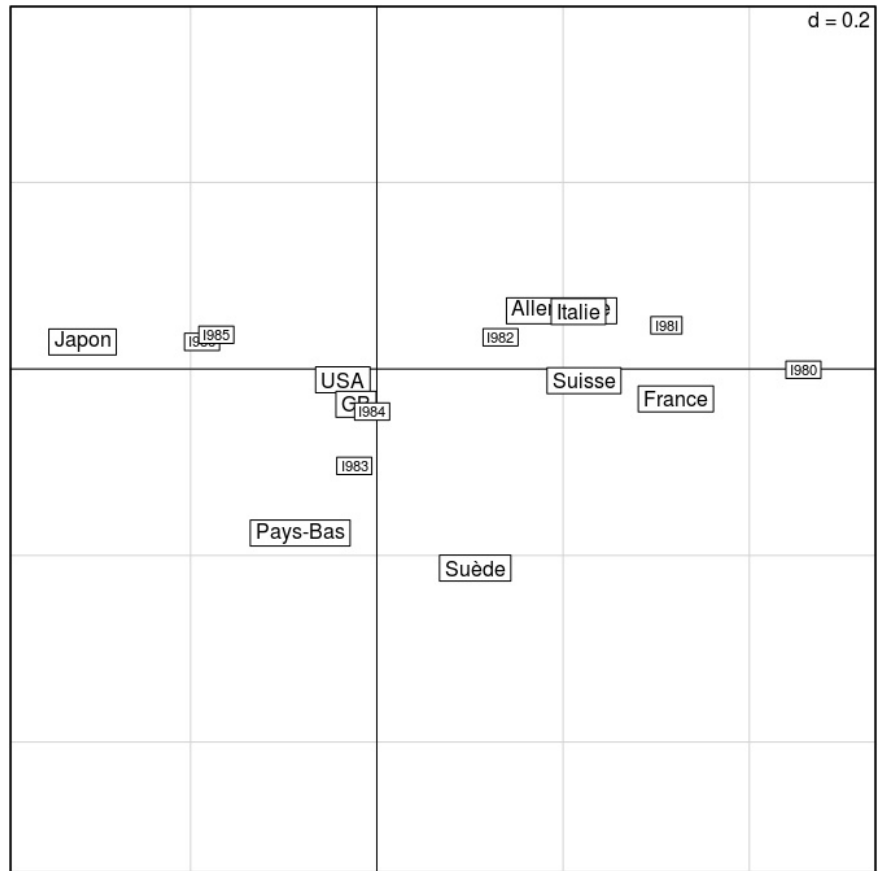
> afcin$row.rel
      Axis1 Axis2 Axis3 Axis4
USA     -6510  -663  1889  449
Japon   -9892    83   -15   -9
Allemagne 9024   891    1   50
France  9812   -98    3  -74
GB     -1112 -3215   -1 -5053
Italie  8047   651  -494  427
Pays-Bas -1335 -6103 -1142 1374
Suede   1668 -6818  1199  -36
Suisse  4057   -13 -5774 -124
```

Contributions absolues et relatives des colonnes:

```
> afcin$col.abs
      Comp1 Comp2 Comp3 Comp4
1986  1978   710  1209   57
1985  1398   950  1903  1081
1984    1  1312  1720   482
1983   20  5667  1428    0
1982   495   487   558  2828
1981  2560   873  2352   763
1980  3548    0   830  4789

> afcin$col.rel
      Comp1 Comp2 Comp3 Comp4
1986  -9495   226   219   -4
1985  -8975   405  -460  -107
1984   -62 -5085  3782  432
1983  -450 -8287 -1185    0
1982  8292   542   352   727
1981  9406   213  -326   43
1980  9713    0    86  -201
```

Composantes 1 et 2:



Composantes 1 et 3:

