



**Sujet d'examen de 1<sup>ère</sup> session**  
**Semestres 1-3-5**  
**Année universitaire 2013-2014**

**Intitulé de l'épreuve :** Analyse de données  
**Nom de l'enseignant :** Nils Berglund  
Mention / Spécialité / Parcours : M1 ESA  
Année : 2013-2014  
Durée de l'épreuve : 2 heures  
Documents autorisés : Résumé manuscrit de 4 pages A4  
Matériels autorisés : Calculatrice non programmable

**Cadre réservé à l'administration**

Date de l'examen : 18/12/2013  
Lieu(x) : Amphi Cantillon

P1/ 5

**SUJET**

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

**Questions de cours [6 points]**

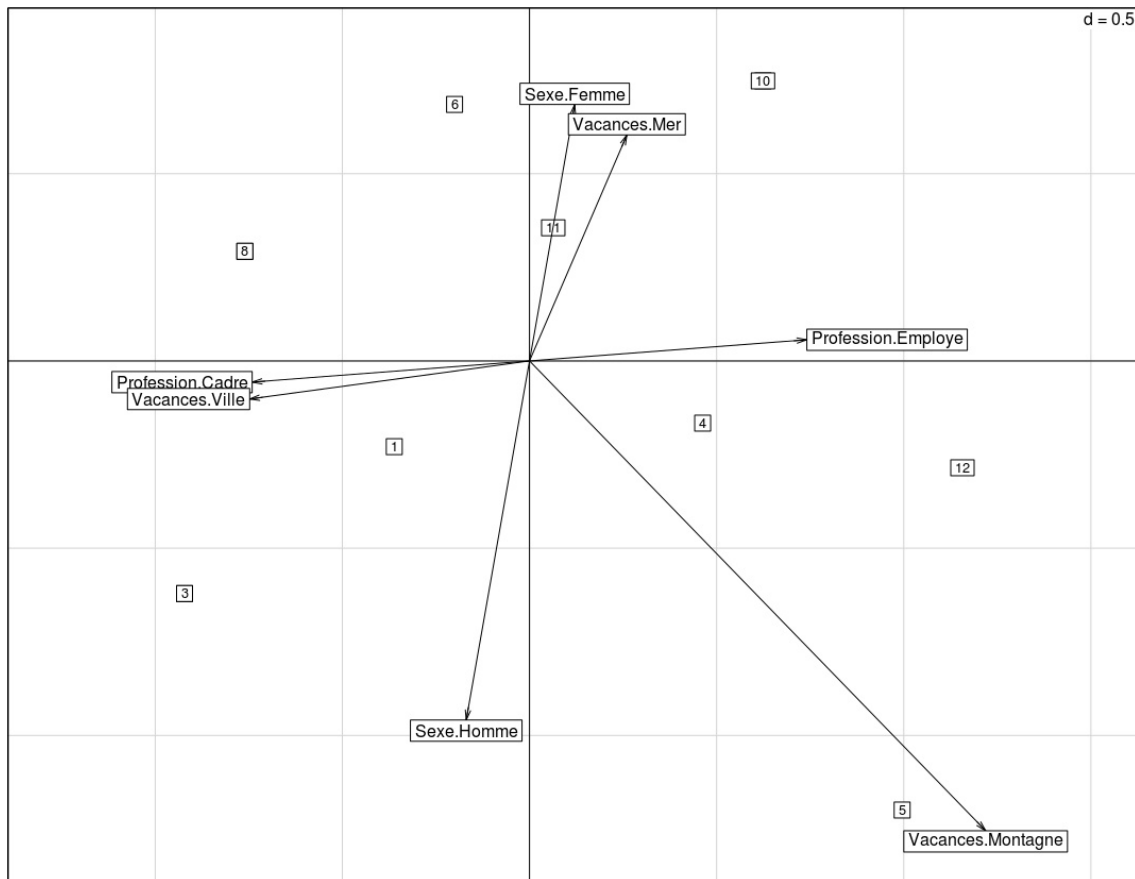
1. Comment est définie la loi du chi-deux?
2. Dans une AFC, pourquoi la matrice  $S = X^T Y^T$  admet-elle toujours la valeur propre 1? Pourquoi ignore-t-on cette valeur propre dans l'analyse?
3. Que signifie la propriété d'équivalence distributionnelle?
4. Quel est l'intérêt des relations quasi-barycentriques?
5. Expliquer la différence entre contributions absolues et relatives dans une AFC.
6. Pourquoi est-il préférable d'éviter les modalités à très faible effectif dans une ACM?

**Exercice 1 [6 points]**

Une agence de voyages a effectué une étude de marché dans une population de 12 personnes. Le tableau suivant représente le sexe de chaque personne interrogée, sa profession et le type de vacances qu'elle préfère (au bord de la mer, visite d'une ville étrangère, à la montagne).

	Genre	Profession	Vacances
1	Homme	Cadre	Mer
2	Homme	Cadre	Ville
3	Homme	Cadre	Ville
4	Homme	Employé	Mer
5	Homme	Employé	Montagne
6	Femme	Cadre	Mer
7	Femme	Cadre	Ville
8	Femme	Cadre	Ville
9	Femme	Employé	Mer
10	Femme	Employé	Mer
11	Femme	Employé	Ville
12	Femme	Employé	Montagne

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  $B$ .  
La page suivante montre le résultat d'une ACM effectuée sur ce jeu de données.
4. Discuter les inerties.
5. Interpréter les deux premiers axes factoriels.
6. Que peut-on dire de la modalité "vacances à la montagne"?



```
> acmin$TOT
      inertia      cum      ratio
1 0.5294023 0.5294023 0.3970517
2 0.3547085 0.8841108 0.6630831
3 0.3224688 1.2065796 0.9049347
4 0.1267537 1.3333333 1.0000000
```

```
> acmin$col.abs
      Comp1 Comp2 Comp3
Sexe.Femme      79 2538 1280
Sexe.Homme     110 3553 1792
Profession.Cadre 2546   15   22
Profession.Employe 2546   15   22
Vacances.Mer     262 1408 3980
Vacances.Montagne 2296 2433 1239
Vacances.Ville  2160   40 1666
```

```
> acmin$col.rel
      Comp1 Comp2 Comp3
Sexe.Femme      301 6481 -2972
Sexe.Homme     -301 -6481 2972
Profession.Cadre -8088  -32   42
Profession.Employe 8088   32  -42
Vacances.Mer     712 2568 6600
Vacances.Montagne 4376 -3106 -1439
Vacances.Ville  -5882  -73 -2763
```

**Exercice 2 [8 points]**

Le tableau suivant montre la répartition des tâches quotidiennes au sein du couple Raymonde et Robert Bidochon. Les lignes du tableau correspondent aux différentes tâches. Pour une tâche donnée, la valeur indiquée représente la fréquence d'exécution par Raymonde seule, de manière alternée entre les deux, par Robert seul et par les deux ensemble.

	Femme	En alternance	Mari	Ensemble
Lessive	156	14	2	4
Déjeuner	124	20	5	4
Dîner	77	11	7	13
Petit déjeuner	82	36	15	7
Rangement	53	11	1	57
Vaisselle	32	24	4	53
Courses	33	23	9	55
Conduite	10	51	75	3
Finances	13	13	21	66
Assurances	8	1	53	77
Réparations	0	3	160	2
Vacances	0	1	6	153

1. Interpréter le résultat du test du chi-deux ci-dessous.
2. Discuter les inerties.
3. Pourquoi la variable `ratio` vaut-elle 1.0000000 à la troisième ligne?
4. Interpréter les deux premiers axes factoriels.
5. Quelles sont les tâches exercées majoritairement par Raymonde? Par Robert? Par les deux ensemble?
6. Quelle est la modalité la moins bien représentée? Qu'est-ce que cela signifie?
7. Donner des exemples de tâches dont la répartition au sein du couple est similaire.

```
> summary(as.table(tab))
Number of cases in table: 1648
Number of factors: 2
Test for independence of all factors:
Chisq = 1837, df = 33, p-value = 0

> afcin$TOT
      inertia      cum      ratio
1 0.56476911 0.5647691 0.5066695
2 0.45889432 1.0236634 0.9183558
3 0.09100629 1.1146697 1.0000000
```



> afcin\$row.abs

	Axis1	Axis2
Lessive	1688	645
Dejeuner	1151	542
Diner	503	160
Petit dejeuner	363	408
Rangement	197	269
Vaisselle	49	268
Courses	21	236
Conduite	819	756
Finances	82	545
Assurances	638	396
Reparations	4401	1564
Vacances	88	4212

> afcin\$row.rel

	Axis1	Axis2	con.tra
Lessive	-7116	-2210	1202
Dejeuner	-7117	-2722	819
Diner	-7538	-1944	338
Petit dejeuner	-4416	-4031	417
Rangement	-4676	5174	214
Vaisselle	-1264	5572	198
Courses	-700	6479	150
Conduite	3856	-2890	1077
Finances	1547	8302	270
Assurances	6076	3070	532
Reparations	7451	-2151	2993
Vacances	250	9678	1792

> afcin\$col.abs

	Comp1	Comp2
Femme	4129	1215
En.alternance	37	225
Mari	5742	1698
Ensemble	91	6862

> afcin\$col.rel

	Comp1	Comp2	con.tra
Femme	-7803	-1866	2681
En.alternance	-234	-1152	803
Mari	7993	-1920	3640
Ensemble	161	9825	2875