



1er SEMESTRE 1^{ère} Session 2010-2011

Mention / Parcours / Spécialité : **MASTER MO spécialité SIME : option "Stratégie, Veille d'Entreprise et Technologie de l'Information et de la Communication" (SVETIC) + option "Système d'Information et Pilotage de l'Entreprise" (SIPE)**

Année : **1^{ère}** (Commun SIME)

Intitulé de l'épreuve : Analyse de données

Durée de l'épreuve : 2 heures

Enseignant : N. Berglund

Documents autorisés : Résumé manuscrit de 4 pages A4

Matériels autorisés : Calculatrice non programmable

page 1 / ...⁵

Il sera tenu compte de la qualité de la rédaction.
Les points sont donnés à titre indicatif.

Problème 1 [5 points]

Une urne contient 3 boules noires et 2 boules blanches. On effectue des tirages sans remise.

1. Calculer la probabilité que la première boule tirée soit noire.
2. Calculer la probabilité que la seconde boule tirée soit blanche.
3. Calculer la probabilité que la seconde boule tirée soit blanche sachant que la première boule tirée est noire.
4. Calculer la probabilité que la première boule tirée soit noire sachant que la deuxième boule tirée est blanche.
5. Les tirages sont-ils indépendants?

Problème 2 [5 points]

Une statistique de la pluviométrie annuelle dans 45 agglomérations françaises montre que

- dans 3 agglomérations il est tombé entre 500 et 600 mm;
- dans 6 agglomérations il est tombé entre 600 et 650 mm;
- dans 8 agglomérations il est tombé entre 650 et 700 mm;
- dans 7 agglomérations il est tombé entre 700 et 750 mm;
- dans 8 agglomérations il est tombé entre 750 et 800 mm;
- dans 4 agglomérations il est tombé entre 800 et 900 mm;
- dans 3 agglomérations il est tombé entre 900 et 1000 mm;
- dans 6 agglomérations il est tombé entre 1000 et 1500 mm.

1. Représenter les données sous forme de tableau.
2. Représenter les données sous forme d'histogramme.
3. Tracer sur un graphique séparé la courbe des fréquences cumulées.
4. Déterminer la médiane de la distribution.
5. Calculer des valeurs approchées de la moyenne, de la variance et de l'écart-type de la quantité de pluie.

Problème 3 [5 points]

On lance 2 dés équilibrés. Pour chaque 6 obtenu, on gagne 10 euros. On ne gagne rien pour tout résultat autre que 6.

On considère deux variables aléatoires:

- X égale à la somme gagnée
- Y égale à la différence, en valeur absolue, entre les points du premier et du second dé.

Déterminer

- la loi conjointe de X et Y ,
- les lois (marginales) de X et Y , leur espérance et leur variance,
- la covariance de X et Y ,
- la variance de $X + Y$,
- le coefficient de corrélation de X et Y .

Problème 4 [5 points]

Des experts en gastronomie ont noté 10 cuvées de cidre selon les critères suivants: odeur, contenu en sucre, acidité, amertume, astringence, caractère suffocant, goût piquant, taux d'alcool, parfum, goût fruité. Les résultats sont résumés dans l'échantillon cidre suivant :

```
> tab<-matrix(cidre, ncol=10, byrow=TRUE)
> tab
```

	odeur	sucre	acide	amer	astr.	suff.	piquant	alcool	parfum	fruité
1	2.14	1.86	3.29	2.29	2.00	0.14	2.29	1.86	1.29	1.29
2	2.43	0.79	2.71	2.57	2.00	0.43	2.57	2.86	0.43	0.14
3	2.71	3.14	2.57	2.57	1.43	0.14	2.14	0.86	2.29	1.71
4	3.00	3.71	2.14	2.07	1.57	0.00	1.29	1.00	3.14	3.14
5	3.43	1.29	2.86	3.14	2.17	1.00	1.86	2.86	1.14	0.29
6	3.14	0.86	2.86	3.79	2.57	0.14	1.71	3.29	0.14	0.00
7	3.14	1.14	2.86	2.86	2.00	0.43	1.71	1.86	0.14	0.00
8	2.43	3.71	3.21	1.57	1.71	0.00	1.00	0.57	2.57	2.86
9	5.10	2.86	2.86	3.07	1.79	1.71	0.43	1.43	0.57	2.71
10	3.07	3.14	2.57	3.00	2.00	0.00	0.43	1.29	2.57	3.07

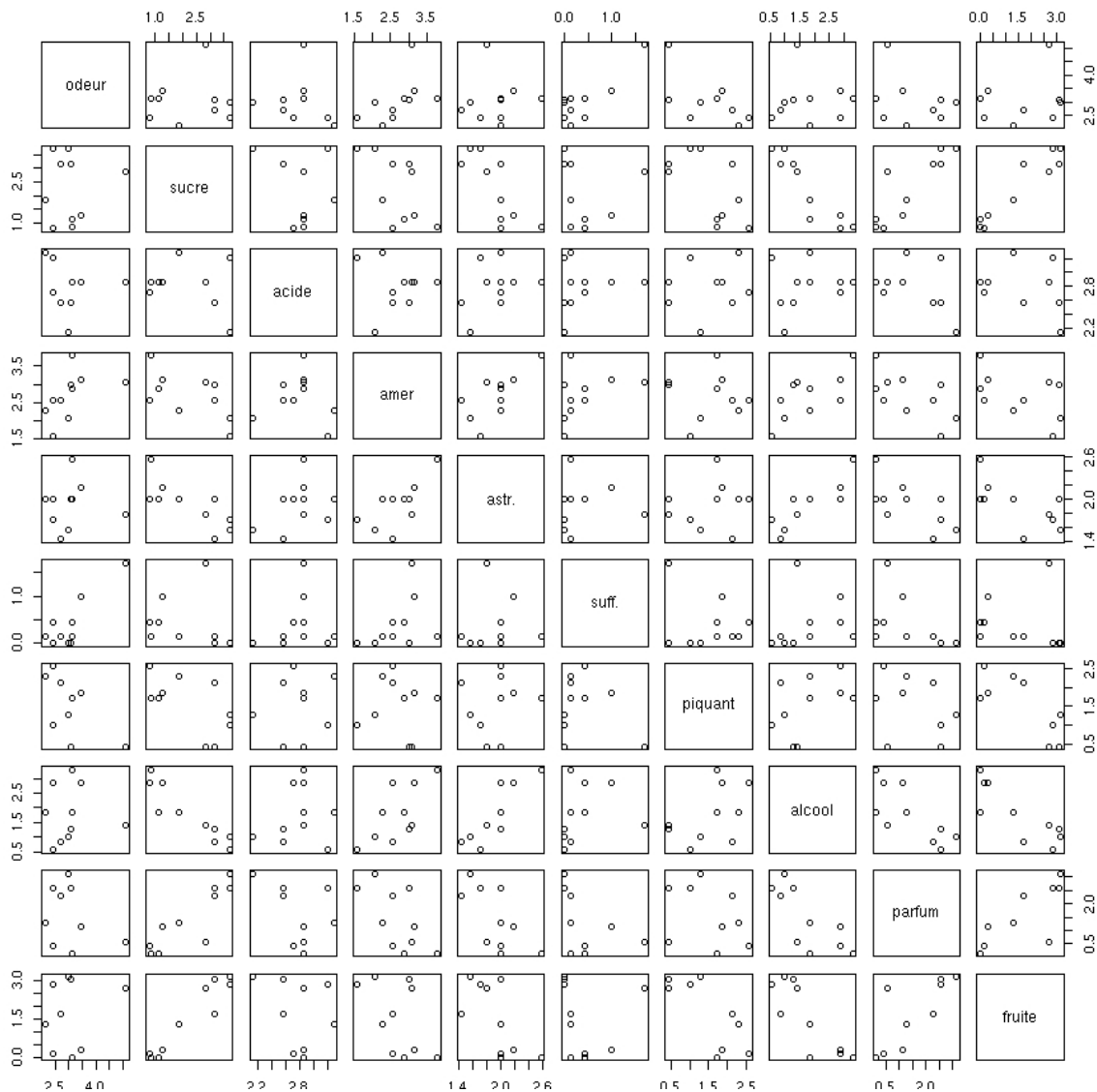
Interpréter les résultats suivants, fournis par une analyse avec le logiciel R :

```
> cor(tab)
```

	odeur	sucre	acide	amer	astr.	suff.
odeur	1.00000000	0.07520193	-0.15656878	0.49256395	0.03665736	0.84451794
sucre	0.07520193	1.00000000	-0.28738736	-0.59547443	-0.77049690	-0.19449458
acide	-0.15656878	-0.28738736	1.00000000	-0.08268896	0.33851199	0.13581213
amer	0.49256395	-0.59547443	-0.08268896	1.00000000	0.71017744	0.38248505
astr.	0.03665736	-0.77049690	0.33851199	0.71017744	1.00000000	0.07107982
suff.	0.84451794	-0.19449458	0.13581213	0.38248505	0.07107982	1.00000000
piquant	-0.60522310	-0.60896335	0.14069087	-0.03400280	0.14343210	-0.23270753
alcool	0.02669909	-0.92277479	0.15325118	0.69539761	0.85907729	0.22306947
parfum	-0.29026182	0.86969292	-0.39714653	-0.63244799	-0.66144035	-0.50140723
fruité	0.18227871	0.95316429	-0.27415640	-0.50427150	-0.64406541	-0.09885815

	piquant	alcool	parfum	fruité
odeur	-0.6052231	0.02669909	-0.2902618	0.18227871
sucre	-0.6089634	-0.92277479	0.8696929	0.95316429
acide	0.1406909	0.15325118	-0.3971465	-0.27415640
amer	-0.0340028	0.69539761	-0.6324480	-0.50427150
astr.	0.1434321	0.85907729	-0.6614403	-0.64406541
suff.	-0.2327075	0.22306947	-0.5014072	-0.09885815
piquant	1.0000000	0.48166080	-0.3324681	-0.72801187
alcool	0.4816608	1.00000000	-0.7590623	-0.83425684
parfum	-0.3324681	-0.75906235	1.0000000	0.79646537
fruité	-0.7280119	-0.83425684	0.7964654	1.00000000

```
> plot(data.frame(tab))
```



```
> acp<-prcomp(tab, scale=TRUE)
```

```
> summary(acp)
```

Importance of components:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Standard deviation	2.270	1.582	1.047	0.9135	0.4401	0.3739	0.22162	0.15595
Proportion of Variance	0.515	0.250	0.110	0.0834	0.0194	0.0140	0.00491	0.00243
Cumulative Proportion	0.515	0.766	0.875	0.9587	0.9781	0.9921	0.99701	0.99944
	PC9	PC10						
Standard deviation	0.07477	1.57e-16						
Proportion of Variance	0.00056	0.00e+00						
Cumulative Proportion	1.00000	1.00e+00						

```
> biplot(acp)
```

