



2^e SESSION DU 1^{er} SEMESTRE 2007-2008

Mention / Parcours / Spécialité : **MASTER Sciences du Management**

Année : **1**

Intitulé de l'épreuve : **Analyse de données**

Durée de l'épreuve : **2 heures**

Documents autorisés : **Résumé manuscrit de 4 pages**

Matériels autorisés : **Calculatrice (non programmable)**

page 1 / 4

SUJET

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

Problème 1 [5 points]

Dans une classe de 100 étudiants, une statistique recensant le nombre de SMS mensuels envoyés donne les résultats suivants:

- moins de 10 SMS : 15 étudiants;
- entre 10 et 50 SMS : 10 étudiants;
- entre 50 et 100 SMS : 20 étudiants;
- entre 100 et 150 SMS : 35 étudiants;
- entre 150 et 200 SMS : 15 étudiants;
- plus de 200 SMS : 5 étudiants;

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. En faisant si nécessaire des hypothèses supplémentaires, calculer des valeurs approchées de la moyenne, de la variance et de l'écart-type du nombre de SMS.

Problème 2 [5 points]

Une étude statistique, concernant la présence d'une certaine maladie des voies respiratoires dans une population de 100'000 individus montre que

- 2500 individus sont atteints de la maladie.
- 80% des individus malades sont fumeurs.
- 95% des individus sains sont non fumeurs.

1. Déterminer le nombre de fumeurs de la population.
2. Calculer la probabilité qu'un fumeur soit atteint de la maladie.

Problème 3 [5 points]

On dispose de deux dés truqués.

- Les faces du premier dé sont numérotées 2, 2, 4, 4, 6, 6.
- Les faces du second dé sont numérotées 1, 1, 1, 1, 3, 3.

On jette simultanément les deux dés. Soit X_1 le nombre indiqué par le premier dé, et X_2 le nombre indiqué par le second dé. On considère les variables aléatoires $X = X_1 + X_2$ et $Y = |X_1 - X_2|$.

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]

Un examen d'analyse de données de l'Université de Plaine-sur-Creuse comportait quatre problèmes:

1. Les problèmes 1. et 4. étaient des problèmes de statistique;
2. Les problèmes 2. et 3. étaient des problèmes de probabilités.

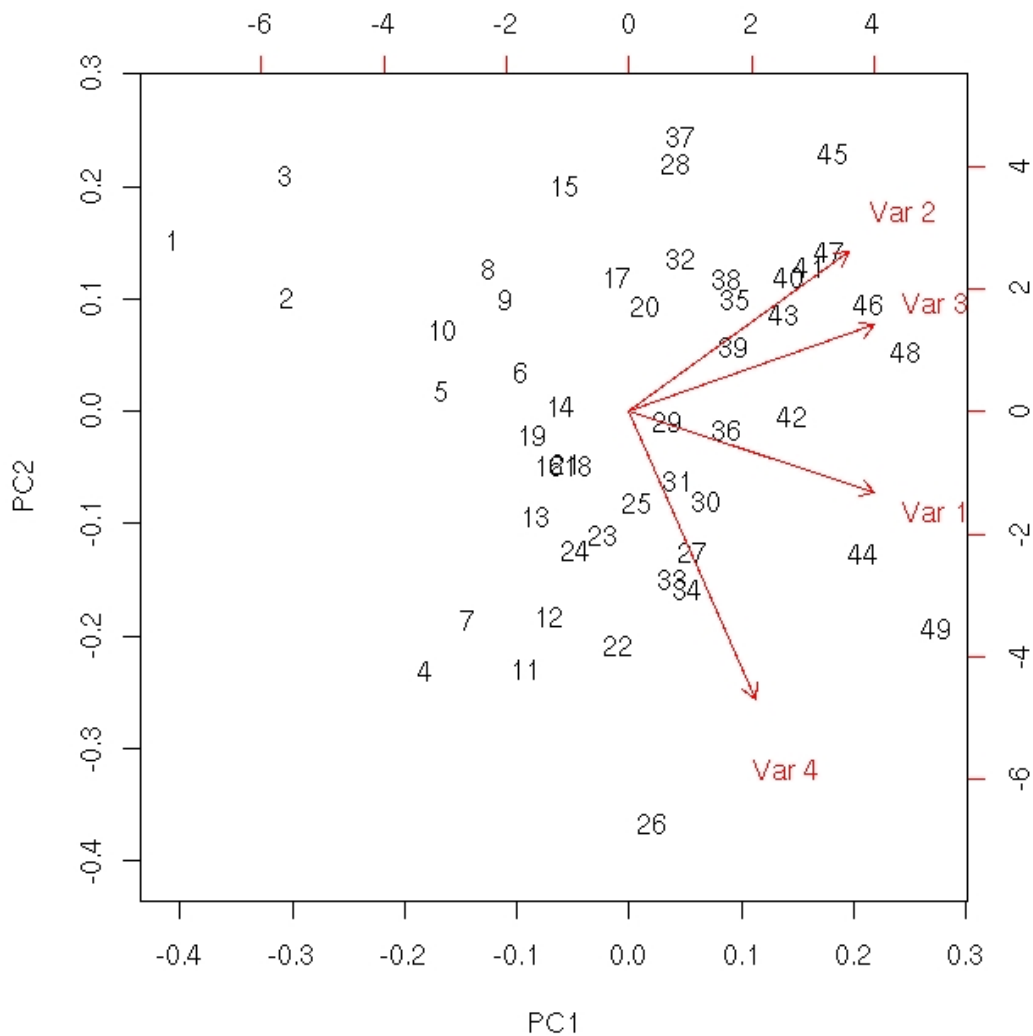
Chaque problème comptait 5 points. Le nombre de points obtenus par 50 étudiants se répartit comme suit:

Pb 1	Pb 2	Pb 3	Pb 4
1	0	0	0
2.5	0	0	0
1.5	2	0	0
3	0	0	3
3	0	2	1
5	0	1	0
3.5	0	0.5	2.5
3.5	0	3	0
4	0	2.5	0
2.5	1	2	1
4.5	0	0	2.5
4	0	1.5	2.5
3.5	0	2.5	2
4	0	3	1
3.5	2	3	0
3.5	2	1	2
5	2	1.5	0
4	0	3	1.5
3	2	1.5	2
5	0	4	0
3	3	1	2.5
4.5	2	0	3
4	2.5	0.5	2.5
2.5	0	4.5	3
4	0	4	2
5	1	0	4
5	0	3.5	2
4.5	4	2	0
4	3	2	2
5	2	2	2
4	0	5	2
3.5	2	5	1
3.5	0	5	3
4	1	3.5	3
4.5	2.5	4	1
4.5	3	2.5	2
2.5	4	5	1
4	5	2	1.5
3.5	3	4.5	2
5	4	3	1
5	4	3.5	1
4.5	2	5	2
4	5	3	2
5	2	5	3
4	5	5	1
5	4	4.5	1.5
3.5	5	5	2
5	4	5	2
5	3	5	4

Une analyse avec le logiciel R livre les résultats suivants (les données du tableau précédent ont d'abord été stockées dans une variable exos) :

```
> tnotes<-matrix(exos, ncol=4, byrow=TRUE)
> cor(tnotes)
      [,1]      [,2]      [,3]      [,4]
[1,] 1.0000000 0.217926782 0.2717258 0.211609734
[2,] 0.2179268 1.000000000 0.2934095 0.007787658
[3,] 0.2717258 0.293409516 1.0000000 0.075471093
[4,] 0.2116097 0.007787658 0.0754711 1.000000000

> prcomp(tnotes,scale=TRUE)->acp
> acp
Standard deviations:
[1] 1.2547994 1.0164243 0.8389642 0.8297586
> biplot(acp)
```



Interpréter ces résultats.