Université de Nice — L1MASS, année 2014-2015
Département de Mathématiques — Statistique
NOM : Date : .
PRENOM : Groupe : .
**Salle et machine** : .

**Feuille-réponse 4**
**Déciles, quantiles**

# 1 Analogue du cours 04

Former le vecteur `y` des tailles de `NW.Hnd` du fichier `survey`

1. Calculer les quartiles $Q_1$, $Q_2$, $Q_3$ et donner ici votre code

2. Donner ci-dessous les quartiles $Q_1$, $Q_2$, $Q_3$

| | $Q_1$ | $Q_2$ | $Q_3$ |
|---|---|---|---|
| quartile | | | |

3. Calculer les déciles $q_1$, ..., $q_9$ et donner ici votre code

4. Donner ci-dessous les déciles $q_1$, ..., $q_9$

| | $q_1$ | | | | | | | | $q_9$ |
|---|---|---|---|---|---|---|---|---|---|
| décile | | | | | | | | | |

5. Pour chaque classe $C_k$ des $x_i$ tels que $q_{k-1} < x_i \leq q_k$ déterminer la proportion $p_k = e_k/n$ de la population totale $n =$`length(y)` contenue dans cette classe, de façon à déterminer la plus grande et la plus petite proportion. Expliquez comment vous procédez :

6. Donner ici vos résultats

| $p_1$ | | | | | | | | | $p_{10}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |

7. Expliquez quelle valeur on attend et commentez les écarts les plus importants

8. On envisage l'échantillon `yprim=y` mais dans le quel on a posé `yprim[37]=10*y[37]` et `yprim[105]=10*y[105]`. Expliquez comment une telle "erreur" aurait pu se produire

9. Comparez les moyennes et écart-types des deux échantillons.

10. (Re)faire le tableau des déciles $q'_k$ pour `yprim`

| | $q'_1$ | | | | | | | | $q'_9$ |
|---|---|---|---|---|---|---|---|---|---|
| décile | | | | | | | | | |

11. Commentez votre résultat dans l'esprit de la remarque *la médiane est moins sensible que la moyenne aux données aberrantes.*

## 2 Mêmes questions pour les statures

Reprendre l'étude ci-dessus pour les statures `Height`. Pour cela définir la variable `z=survey.cc$Height`, puis répondre aux mêmes questions :

1. Calculer les quartiles $Q_1$, $Q_2$, $Q_3$ et donner ici votre code

2. Donner ci-dessous les quartiles $Q_1$, $Q_2$, $Q_3$

| | $Q_1$ | $Q_2$ | $Q_3$ |
|---|---|---|---|
| quartile | | | |

3. Calculer les déciles $q_1$, ..., $q_9$ et donner ici votre code

4. Donner ci-dessous les déciles $q_1$, ..., $q_9$

| | $q_1$ | | | | | | | | $q_9$ |
|---|---|---|---|---|---|---|---|---|---|
| décile | | | | | | | | | |

# 3 Un modèle et un échantillon gaussien pour les statures

On pose $\mu =$ `muz` = moyenne de `z` et $\sigma$=`sigmaz`= écart-type de `z`. Simuler un échantillon `z.norm` de même taille que `z` et de loi normale $\mathcal{N}(\mu, \sigma)$. **Important :** De façon à pouvoir reproduire vos données et avoir un échantillon différent de celui des machines voisines, utilisez la commande `set.seed` *juste avant* la commande `r.norm` de la manière suivante :

`set.seed(sssmm)` avec `sss`= numéro de la salle et `mm`= numéro de votre machine.

```
z.norm=rnorm(length(z),mean(z),sd(z))
```

Comparez les valeurs pour deux exécutions successives de `z.norm` selon que vous ayez fait précéder la deuxième de `set.seed` ou non. Veillez à bien faire cette succession pour l'échantillon que vous utilisez. Si vous devez réinitialiser votre `R`, il est important de recommencer ainsi.

1. Valeurs obtenue de `z.norm[1]` selon que vous avez fait précédé la deuxième aussi de `set.seed(sssmm)` ou non :

2. Que valent $\mu$ et $\sigma$ ? Quels sont les trois quartiles (théoriques) pour la loi $\mathcal{N}(\mu, \sigma)$ ?

3. Quelles sont les valeurs [1] $Q_1$ et $Q_3$ des 1er et 3ème quartile de l'échantillon `z.norm`.

4. Quelle est la médiane $Q_2$ et la moyenne de `z.norm`

5. Quelle est la longueur maximale des moustaches ?

6. `z.norm` a-t-il des valeurs exceptionnelles ; si oui, lesquelles ?

7. Quelle est l'intervalle des valeurs adjacentes ?

1. Indication : utilisez `b<-boxplot(z.norm)`