

Feuille-réponse 4
Quantiles

1 Quartiles et déciles

Former le vecteur y des tailles de `NW.Hnd` du fichier `survey` de la librairie MASS :

```
library(MASS);data(survey)
survey.cc<-survey[complete.cases(survey),]
y<-survey.cc$NW.Hnd
```

1. Indiquer les 3 quartiles Q_1, Q_2, Q_3 de y que vous pouvez obtenir facilement avec la commande `summary(y)`.

	Q_1	Q_2	Q_3
quartile			

2. La commande `sort(y)[ceiling(length(y)/4)]` vous permet de retrouver le premier quartile Q_1 . Quelles commandes permettent de retrouver les deux autres quartiles Q_2 et Q_3 ?

3. Combien y-a-t-il d'individus de l'échantillon dont la valeur est celle de la médiane ? Expliquer.

4. Calculer de même les 9 déciles q_1, \dots, q_9 .

	q_1								q_9
décile									

5. On envisage un nouvel échantillon `yprim` qui prend les même valeurs que y sauf que l'on pose `yprim[37]=10*y[37]` et `yprim[105]=10*y[105]`. Expliquez comment une telle "erreur" aurait pu se produire lors du relevé des mesures y .

6. Comparez les moyennes et écart-types des deux échantillons `yprim` et y .

7. (Re)faire le tableau des déciles q'_k de `yprim`

	q'_1								q'_9
décile									

8. Comparez avec les déciles de y et 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 $z = \text{survey.cc}\$Height$, puis répondre aux mêmes questions :

1. Indiquer les 3 quartiles Q_1, Q_2, Q_3 de z .

	Q_1	Q_2	Q_3
quartile			

2. Combien y-a-t-il d'individus de l'échantillon z dont la valeur est celle de la médiane ? Expliquer.

3. Calculer de même les 9 déciles q_1, \dots, q_9 de z

	q_1								q_9
décile									

4. On envisage un nouvel échantillon z_{prim} qui prend les même valeurs que z sauf que l'on pose $z_{\text{prim}}[20] = 10 * z[20]$ et $z_{\text{prim}}[105] = 10 * z[110]$. Comparez les moyennes et écart-types des deux échantillons z_{prim} et z .

5. (Re)faire le tableau des déciles q'_k pour z_{prim} et comparez avec les déciles de z .

	q'_1								q'_9
décile									

3 Quantiles théoriques (gaussiens) et empiriques

- La commande `Q1=qnorm(0.25,mu,sigma)` permet de calculer le premier quartile (théorique) d'une loi gaussienne $\mathcal{N}(\mu, \sigma)$. Déterminer les 3 quartiles de la loi gaussienne associée aux échantillons y et z et comparez-les avec ceux trouvés pour ces variables. Commenter.
- Dans le cas de la loi théorique, combien y-a-t-il de valeurs de la variable théorique correspondant à la médiane ? Expliquer.
- Quelle commande permet de calculer un vecteur dont les composantes sont les 9 déciles d'une loi gaussienne $\mathcal{N}(\mu, \sigma)$?
- Comparer les déciles théoriques (gaussiens) et les déciles empiriques des échantillons y et z .