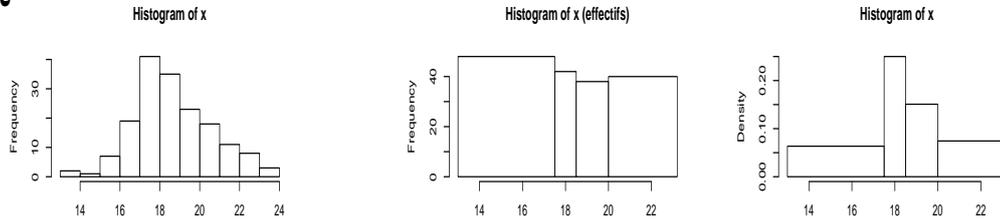


Cours 03 Quartiles et boîtes à moustaches

1 Quartiles



Le premier quartile d'un échantillon est le caractère (la valeur) de l'individu tel qu'un quart des individus de cet échantillon ont un caractère inférieur ou égal. Si x est cet échantillon, il s'agit donc de `sort(x)[ceiling(length(x)/4)]`. Pour le deuxième quartile (ou médiane) c'est la moitié des individus qui ont un caractère inférieur ou égal, et ce sont les trois-quarts des individus pour le troisième quartile (remplacer `length(x)/4` par `length(x)/2` pour la médiane, et par `length(x)*3/4` pour le troisième quartile).

Par exemple, si x est l'échantillon 11 1 5 12 5 3 7 1 18 3 14 7 alors en triant cet échantillon avec la commande `sort(x)` on voit les trois quartiles sont les valeurs en gras : 1 1 **3** 3 5 **5** 7 7 **11** 12 14 18.

2 Boîtes à moustaches (boxplot)

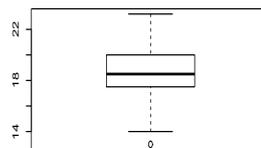


FIGURE 1 – La boîte qui s'étend de Q_1 à Q_3 et ses moustaches de longueur maximale $1.5 \cdot I$ avec $I = Q_3 - Q_1$.

Désignons par Q_1 , Q_2 , et Q_3 les trois quartiles d'un échantillon x . L'*intervalle interquartile* I est la longueur $Q_3 - Q_1$. Les *valeurs exceptionnelles* (ou *outliers*) de l'échantillon x sont les valeurs de cet échantillon qui ne sont pas dans l'intervalle $[Q_1 - I * 1.5, Q_3 + I * 1.5]$. Les moustaches ajoutées à la boîte allant de Q_1 à Q_3 s'étendent de la plus petites à la plus grande des valeurs non-exceptionnelles.

A noter que la commande `boxplot` (tout comme la commande `hist`) ne produit pas seulement un dessin mais aussi une valeur, qui peut être récupérée dans une variable (`x.boxplot`, par exemple) qui est en fait un tableau. Pour cela, il suffit de remplacer `boxplot(x)` par `x.boxplot<-boxplot(x)`. Il suffit ensuite de demander la valeur de cette variable pour trouver les valeurs caractérisant la boîte à moustache : extrémité de la moustache inférieure, les trois quartiles Q_1 , Q_2 , Q_3 , extrémité de la moustache supérieure.

3 Quartiles d'une densité gaussienne

