
Contrôle terminal, durée 2h

Exercice 1

Une société immobilière dispose de 600 appartements dont les surfaces sont données par le tableau suivant:

Surface	[25;50[	[50;60[	[60;80[	[80;100[	[100;120[	[120;145[
Fréquence	0,02	0,15	0,13	0,22	0,28	0,2

1. Précisez la population considérée, la variable étudiée et sa nature, la taille de l'échantillon.
2. Retrouver le tableau des effectifs à partir du tableau ci-dessus. Calculer la moyenne, la variance et l'écart-type associées à ces données.
3. Etablir les fréquences cumulées. Tracer le graphe des fréquences cumulées. Calculer la valeur des quartiles.
4. Tracer la boîte à moustache.
5. Donner les densités et tracer l'histogramme. Quelle est la classe modale?

Exercice 2

Le tableau suivant contient, pour treize sportifs, l'âge et la fréquence cardiaque maximale (FCM).

Age	40	36	51	49	47	51	32	55	55	23	49	52	35
FCM	187	195	180	190	185	183	195	185	189	201	189	185	195

1. On voudrait savoir si les deux variables sont liées par une relation linéaire. Quelle serait la première étape à franchir avant d'aborder tout calcul préliminaire?

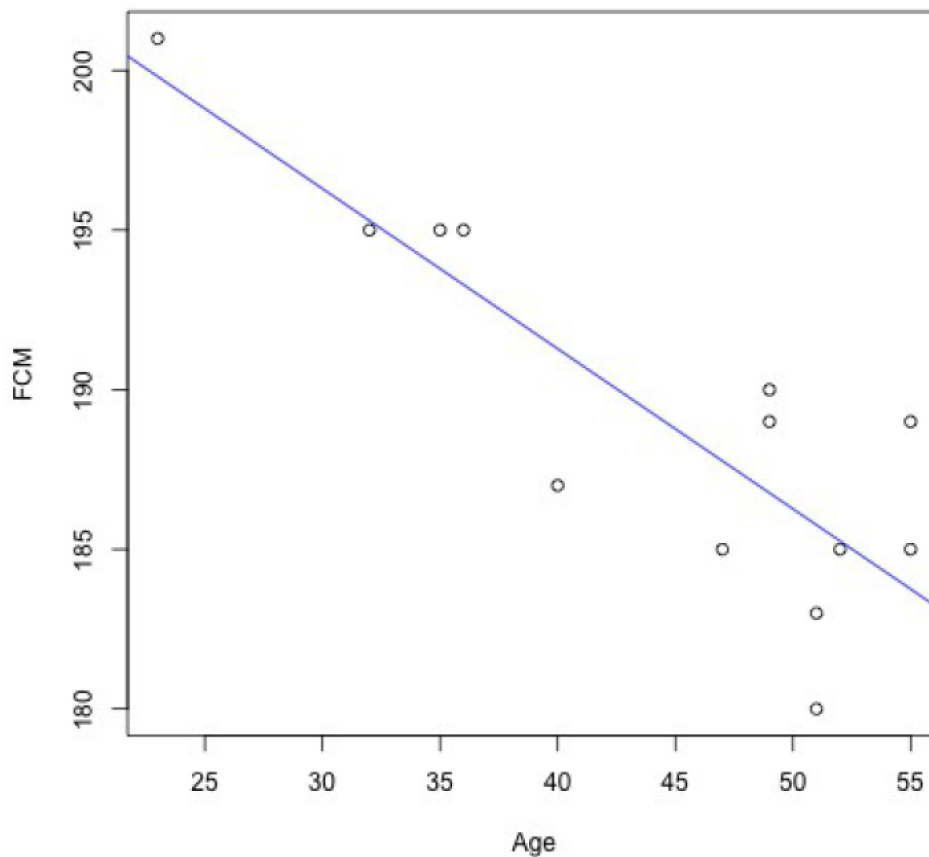
La personne chargée d'étudier ces données envisage d'utiliser le modèle linéaire simple comme modèle prévisionnel. Pour cela, elle utilise le logiciel R et définit d'abord les vecteurs age et fcm contenant ces observations. Puis, elle tape les commandes suivantes:

```
> reg=lm(fcm~age)
> plot(age, fcm, xlab="Age",ylab="FCM")
> abline(reg, col='blue')
> summary(reg)
```

Elle obtient la FIGURE 1 (page 3) ainsi que les résultats suivants:

```
Call:
lm(formula = fcm ~ age)
```

Figure 1:



Residuals:

Min	1Q	Median	3Q	Max
-5.756	-2.756	1.190	1.715	5.251

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	211.35371	4.25279	49.698	2.69e-14 ***
age	-0.50191	0.09394	-5.343	0.000236 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.266 on 11 degrees of freedom

Multiple R-squared: 0.7218, Adjusted R-squared: 0.6965

F-statistic: 28.54 on 1 and 11 DF, p-value: 0.0002365

2. Expliquer sommairement ce que fait chacune des quatre commandes précédemment données.
3. Donner l'équation de la droite de régression linéaire et interpréter les valeurs de ses coefficients.
4. Donner la valeur du coefficient de détermination. Le modèle linéaire vous semble-t-il approprié?
5. Donner une estimation moyenne de la FCM d'un sportif de 45 ans. Donner la commande R permettant de faire cela.

Exercice 3 On effectue un contrôle de fabrication sur des pièces dont une proportion p inconnue est défectueuse. On contrôle un lot de 200 pièces choisies au hasard et on trouve 20 pièces défectueuses. On désigne par X_i la variable aléatoire telle que $X_i = 1$ si la i -ème pièce est défectueuse, et $X_i = 0$ sinon. On note p la proportion de pièces défectueuses et on désigne par N la variable aléatoire donnant le nombre de pièces défectueuses dans l'échantillon aléatoire.

1. Quelle est la loi des X_i ? Quelle est la loi de N ? Justifier votre réponse.
2. Construire un estimateur de p en utilisant la loi des grands nombres et donner sa valeur numérique.
3. Construire un intervalle de confiance pour p , avec un niveau de confiance de 95%.

Valuers de quantiles pouvant vous être utiles:

```
> qnorm(.95)      > qnorm(.995)
[1] 1.644854      [1] 2.575829
> qnorm(.975)     > qnorm(.99)
[1] 1.959964      [1] 2.326348
> qnorm(.98)      > qt(.99,99)
[1] 2.053749      [1] 2.364606
> qnorm(.02)      > qt(.98,99)
[1] -2.053749     [1] 2.081162
```