Exercice 1. Régression linéaire simple.
Charger sur `R` le jeu de données `rmr` en tapant les lignes de code suivantes : `library(ISwR), data(rmr)` et `attach(rmr)`. Ce jeu de données contient une colonne `body.weight` et une colonne `metabolic.rate`. On suppose que `metabolic.rate` dépend linéairement de `body.weight`. Donner l'estimateur des moindres carrés de la pente et de l'intersection avec l'axe OY de la droite de régression, estimer le taux métabolique pour une personne pesant 70 kg et représenter graphiquement la situation

(a) en programmant selon la théorie.
Aide `R` : `abline` ajoute une droite d'ordonnée et de pente de données à un graphique existant, `points` y ajoute des points.

(b) en utilisant les routines `R`.
Aide `R` : consulter la section du manuel en ligne consacrée aux modèles linéaires, disponible à l'adresse `http://cran.r-project.org/doc/manuals/R-intro.html#Linear-models`.

Exercice 2. Fit d'un polynôme de degré 2.
Selon la 2^e loi de Newton, un corps proche de la surface de la Terre tombe verticalement selon l'équation $s(t) = s_0 + v_0 t + \frac{1}{2} g t^2$, où

$$\begin{array}{ccl} s & : & \text{position verticale} \\ s_0 & : & \text{position initiale (inconnue)} \\ v_0 & : & \text{vitesse initiale (inconnue)} \\ g & : & \text{accélération} \end{array}$$

On veut estimer g.
On mesure, aux temps $t = 0.1, 0.2, 0.3, 0.4, 0.5$ les positions $s = -0.0549, 0.0945, 0.3139, 0.7559, 1.1369$.
Donnez l'estimateur des moindre carrés $\hat{g}$ de g, puis programmez-le en `R` et représentez graphiquement la situation.

Exercice 3. Une loi de physiologie dit que, parmi les animaux à sang chaud, le métabolisme au repos (en kcal/jour) M est proportionnel au poids (en kg) P à la puissance 3/4, i.e. $M \approx cP^{3/4}$ d'où $Y \approx \log_{10}(c) + \frac{3}{4}\log_{10}(P)$.
On aimerait vérifier cette loi et on propose le modèle statistique suivant qui semble adéquat

$$Y = \theta_1 + \theta_2 X + Z \quad \text{où} \quad Z \sim \mathcal{N}(0, \sigma^2),$$

Z représente la fluctuation biologique.
Des observations ont donné les résultats suivants :

	P (kg)	M (kcal/jour)
souris	0.021	3.6
rat	0.282	28.1
chat	3	152
chien	24.8	875
chèvre	36	800
mouton	46.8	1330
vache	300	4221
taureau	600	7877

(a) Calculez le coefficient de corrélation empirique, pour cela, faites une fonction `R`.

(b) Donnez les estimateurs des moindres carrés pour θ_1 et θ_2. Programmez-le en `R` et représentez graphiquement la situation.

Exercice 4. On considère un échantillon aléatoire $(X_1, \ldots, X_n)$ de densité de Pareto avec k fixé, i.e. :

$$f(x; \theta, k) = \theta k^{\theta} \left(\frac{1}{x}\right)^{\theta+1}, \quad x \geq k \geq 0, \quad \theta > 1.$$

(a) Calculer l'estimateur pour θ grâce à la méthode des moments.

(b) Calculer l'estimateur pour θ grâce à la méthode du maximum de vraisemblance.

(c) Avec `R`, simuler 1000 réalisations d'une variable suivant une loi de Pareto de paramètre θ et k en utilisant la méthode directe (cf. série 3 ex. 1, calculer la fonction de répartition) et comparer les deux estimateurs obtenus précédemment.

À rendre jusqu'au **jeudi 28 mars 2013, à 18h**