Chapitre 1
Le modèle de Malthus

Ce que nous verrons dans ce chapitre :

- la notion de *modèle*,
- l'*équation de Malthus* : un modèle de population se développant à taux constant,
- un modèle de *population raréfiée*,
- des *méthodes numériques pour simuler et tester* le modèle de Malthus,
- l'équation de Malthus intervient aussi dans les *datations au carbone-14*.

1 Du monde réel à l'équation différentielle :

1.1 Qu'est-ce qu'un modèle ? Quelques définitions

Soit à étudier une situation concrète : une ou plusieurs populations dans un milieu donné, des réactifs chimiques dans une éprouvette, l'expansion d'une tumeur ou d'une épidémie, des atomes radio-actifs dans un morceau de bois, un gaz dans un récipient, etc... Modéliser la situation consiste à :

- sélectionner un certain nombre d'observables : nombre d'individus, nombre de molécules ou concentration en réactifs, nombre de cellules ou d'individus malades, nombre d'atomes radio-actifs, pression et température (par exemple). Ici on fait une abstraction car en choisissant certains observables on en ignore d'autres. Par exemple on peut dans un premier temps négliger l'âge des individus d'une population et les considérer tous identiques. Ces observables sont représentées par des fonctions, qui peuvent dépendre du temps.

- postuler une structure sous-jacente au phénomène et les relations satisfaites par le ou les observables. Dans le cas d'un gaz parfait, on obtient l'équation bien connue :

  $$PV = nRT$$

  où P est la pression, V le volume du récipient, n le nombre de molécules, R une constante et T la température. Ici l'équation ne dépend pas du temps. Un exemple d'équation faisant intervenir le temps est $y'(t) = ay(t)$ où a est une constante et $y(t)$ peut représenter un nombre d'individus ou d'atomes (équation de Malthus)

- Faire l'étude mathématique (ou numérique) de l'équation : si c'est une équation différentielle, la résoudre (si possible !). L'exploration du modèle mathématique peut faire paraître des phénomènes inconnus *a priori*, susceptibles de tests.

- Etude *à posteriori* : comparer les valeurs théoriques (prévues par le modèle) aux données expérimentales et estimer la validité du modèle. Cette partie est cruciale si on veut utiliser le modèle pour faire des prévisions dans le futur ou pour une situation légèrement différente. Notons qu'il n'y a pas ici d'absolu : une série d'expériences positives peut donner une certaine validité au modèle mais ne démontre pas qu'il est "vrai". C'est un problème délicat de décider si un modèle est "bon". En revanche une expérience négative peut invalider un modèle. Ce cas conduit à re-examiner les observables choisies et/ou la relation postulée.

Faire des modèles est avant tout un outil pour penser. Les modèles (biologiques) sont rarement faits pour prédire, mais plutôt pour explorer les conséquences d'une hypothèse particulière à l'intérieur d'un système, ou pour identifier les éléments clés, des liens ou des paramètres importants.

Précisons quelles fonctions nous utiliserons et quelques définitions.

Définition 1.1 *Considérons une quantité qui évolue avec le temps (qui décrit* $\mathbb{R}$ *ou* $\mathbb{R}^+$*) et dont la valeur à l'instant* t *est notée* $y(t)$ *; l'évolution avec le temps de cette quantité est donc donnée par la fonction* $t \mapsto y(t)$ *, que nous supposerons dérivable.*

– l'accroissement de cette quantité entre les instants t *et* $t+\Delta t$ *est la quantité* $y(t+\Delta t)-y(t)$ *(elle peut être positive, négative ou nulle),*

– la vitesse moyenne d'évolution de cette quantité entre les instants t *et* $t+\Delta t$ *est la quantité* $\frac{y(t+\Delta t)-y(t)}{\Delta t}$ *.*

– la vitesse instantanée d'évolution de cette quantité à l'instant t *est la limite, quand* Δt *tend vers zéro, de* $\frac{y(t+\Delta t)-y(t)}{\Delta t}$ *, c'est donc la dérivée* $y'(t)$ *de la fonction* $y(t)$ *.*

– le taux instantané d'évolution de cette quantité à l'instant t *est la quantité* $\frac{y'(t)}{y(t)}$ *.*

Lorsque la quantité retenue est le nombre $N(t)$ d'individus à l'instant t dans une population donnée, on aurait envie de poser $y(t) = N(t)$; cependant, si on veut être parfaitement rigoureux, la dérivée $N'(t)$ de la fonction $N(t)$ définie ci-dessus n'existe pas car, une population ne pouvant avoir qu'un nombre entier d'individus, le nombre $N(t)$ d'individus progresse par sauts d'au moins une unité à la fois, la fonction $t \mapsto N(t)$ n'est donc pas continue, et encore moins dérivable.
C'est pourquoi $y(t)$ ne sera pas exactement le nombre $N(t)$ d'individus à l'instant t, mais une approximation de ce nombre à une unité-près (i. e. $|y(t)-N(t)| \leq 1$ pour tout t) par une fonction $y(t)$ dérivable par rapport à t. L'erreur ainsi faite est négligeable si on la rapporte à une population suffisamment nombreuse. On pourra donc considérer que (à une erreur négligeable près) le taux instantané d'évolution de cette population à l'instant t est la quantité $\frac{y'(t)}{y(t)}$.

1.2 Une disgression : Le modèle de Malthus discret

Soit N_n le nombre d'individus d'une population donnée (d'insectes par exemple) au temps n où maintenant n est un entier (n peut être mesuré en nombre d'années, de mois, d'heures, ... selon ce qui semble le plus perinent). On utilise plutôt de type de description lorsque l'évolution de la population est lente (et donc il est inutile de l'étudier à tout temps). On suppose que le nombre d'individus au temps $n+1$ ne dépend que du nombre d'individus au temps n. Soit d la probabilité pour un individu de mourrir pendant l'intervalle de temps $[n, n+1[$ et soit b le nombre moyen par individu de naissances pendant la même période. Ainsi, d est le taux de mortalité et b le taux de fertilité de la population étudiée. On a alors $N_{n+1} = (1+b-d)N_n$ ou encore $N_{n+1} = \lambda N_n$ si on pose $\lambda = 1+b-d$ (taux de croissance de la population). Supposons que $\lambda > 0$ ne dépend pas de n (ce qui est trés discutable !). On sait qu'alors $N_n = N_0\lambda^n$ où N_0 est le nombre initial d'individus. Il y donc trois situations différentes.

(i) $0 < \lambda < 1$: La population décroit et tend à disparaitre.

(ii) $\lambda = 1$: La population reste constante.

(iii) $\lambda > 1$: La population est croissante et tend à devenir infinie ! Ceci n'est pas raisonnable car la quantité de ressources (nourriture par exemple) est limitée. On pourrait supposer qu'il existe un nombre critique N_c d'individus au delà duquel les ressources vitales de la population deviennent insuffisantes. La population suivrait alors un modèle de Malthus avec un taux de croissance $\lambda' < 1$.

1.3 Analyse a priori de deux modèles :

Considérons une population donnée, dont nous appellerons $y(t)$ une approximation dérivable du nombre $N(t)$ d'individus à l'instant t. Nous supposerons que cette population se reproduit de manière sexuée. Nous allons considérer 2 cas de figure, selon que la population est dense ou très dispersée (raréfiée). Dans les deux cas, il n'y aura pas de contrainte environnementale au sens que les ressources seront supposées toujours suffisantes.

a) Population dense sans contrainte : Si la population est suffisamment dense, dans le processus de reproduction, on note p_i le nombre de rapports (par unité de temps) qu'a un individu i de la population avec un individu du sexe opposé de la même population. Nous supposerons le nombre p_i limité[1]. Nous noterons $\bar{p}$ la valeur moyenne des p_i. Nous supposerons que :

(i) $\bar{p}$ reste constant lorsque la population évolue.

(ii) le taux de fécondité moyen d'un rapport reste constant.

(iii) le taux de mortalité : $\left(\dfrac{\text{Nombre de morts par UT}}{y(t)}\right)$ reste constant dans le temps,

Si on admet ces trois présupposés, montrer que :

- une bonne estimation du nombre de rapports effectifs sur une unité de temps est $\frac{1}{2}\,\bar{p}\,.\,y(t)$,
- le rapport $\left(\dfrac{\text{Nombre de naissances par UT}}{\text{Nombre de rapports par UT}}\right)$ reste constant dans le temps,
- le rapport $\left(\dfrac{\text{Nombre de naissances par UT}}{y(t)}\right)$ reste constant dans le temps,
- le rapport $\left(\dfrac{y'(t)}{y(t)}\right)$ reste constant dans le temps,
- la fonction y obéit à une équation différentielle du type : $y' = a\,y$, où a est une constante réelle.

Discutez a priori[2] les présupposés (i), (ii) et (iii). Recherchez des exemples expérimentaux ou historiques de cas où les quantités supposées ici constantes ont en fait évolué dans le temps.

[1] En effet, il y a des limites biologiques au nombre de partenaires effectifs d'un même individu, et ce n'est pas parce que la population passe de 50 Millions d'individus à 100 Millions d'individus que le nombre de rapports d'un même individu passe du simple au double.

[2] Une discussion plus précise et plus approfondie sera faite en fin de section 2.2, lorsque nous saurons calculer $y(t)$ pour tout t et le comparer à un grand nombre de relevés statistiques pour différentes populations.

b) Population raréfiée sans contrainte : Si la population est très raréfiée (en particulier s'il s'agit d'une espèce en voie d'extinction), les occasions de rencontres entre individus de sexe opposé deviennent statistiquement rares, et le nombre de relations effectives devient inférieur à la limite biologique au nombre de relations que peut avoir chaque individu[3]. On suppose que le nombre d'individus occupant un territoire donné est $y(t)$ et qu'il y a autant de mâles que de femelles. On subdivise le territoire en N cases d'áire égale, et on suppose que $N >> y$. On suppose que les individus sont répartis aléatoirement sur le territoire, et qu'un mâle et une femelle se trouvant dans une même case ont une probabilité $p \in (0, 1]$ de se rencontrer (p est supposé indépendant de la case et du tempsu temps). S'ils ne sont pas dans la même case la probabilité est nulle. Montrer que, en moyenne :

- le nombre de rencontres (entre un mâle et une femelle) est

  $$\frac{p}{4N}\, y(t)^2.$$

  Si le territoire est agrandi, ce nombre augmente t'il ou diminue t'il ?

- le rapport $\left(\dfrac{\text{Nombre de rencontres donnant lieu a un rapport par UT}}{\text{Nombre de rencontres effectives par UT}}\right)$ représente une probabilité pour qu'un événement ait lieu, lequel ? Nous considérerons cette probabilité comme constante.

- le rapport $\left(\dfrac{\text{Nombre de naissances par UT}}{y(t)^2}\right)$ est une constante[4], notée a.

- si on suppose (comme dans le cas a)) que le taux de mortalité : $\left(\dfrac{\text{Nombre de morts par UT}}{y(t)}\right)$ reste constant dans le temps, alors il existe une constante b telle que, pour tout t, $y'(t) = a\, y(t)^2 - b\, y(t)$.

  Discutez a priori[5] les présupposés sur lesquels est fondé ce modèle. Recherchez des exemples expérimentaux ou historiques de cas où les quantités supposées ici constantes ont en fait évolué dans le temps.

De manière humoristique, on pourrait appeler "paradis" et "désert" les deux modèles précédents. L' "enfer", lorsque les ressources sont limitées et que la compétition fait rage, sera abordée dans le chapitre 3 (loi logistique).

1.4 Que signifie l'équation de Malthus ?

Etant donnée une constante non nulle a, considérons l'équation différentielle :

$$x' = a\, x \tag{1}$$

ce qui signifie qu'une fonction x est solution de cette équation différentielle si et seulement si, pour toute valeur t pour laquelle $x(t)$ est définie, x est dérivable et vérifie $x'(t) = a\, x(t)$.

[3]Ceci signifie que le facteur limitant le nombre de relations effectives de chaque individu n'est plus, comme dans le cas a), la limite biologique au nombre de relations, mais le nombre de rencontres (supposées rares) entre deux individus de sexe opposé.

[4]Ici, le rapport $\left(\dfrac{\text{Nombre de naissances par UT}}{\text{Nombre de rencontres donnant lieu a un rapport par UT}}\right)$, qui représente le taux de fécondité moyen, est supposé constant.

[5]Une discussion plus précise et plus approfondie sera faite en fin de section 2.2, lorsque nous saurons calculer $y(t)$ pour tout t et le comparer à un grand nombre de relevés statistiques pour différentes populations.

a) Considérons une population donnée, dont nous appellerons $y(t)$ une approximation dérivable du nombre $N(t)$ d'individus à l'instant t.
Si nous supposons que la fonction y obéit à l'équation différentielle (1), les assertions suivantes sont-elles supposées vraies ou fausses ?

- La différence entre taux de natalité et taux de mortalité est supposée constante dans le temps.
- Si $a > 0$, l'accroissement de la population est supposé proportionnel au nombre de couples potentiellement possibles.
- La population est supposée croissante lorsque $a > 0$ et décroissante lorsque $a < 0$.
- Si $a > 0$, l'accroissement de la population sur une année est supposé proportionnel à la valeur moyenne de cette population sur la même année.

b) Supposons que $a \neq 0$. Rappeler (cf. cours mat111) quelles sont toutes les solutions de l'équation différentielle (1)[6].

c) Supposons toujours que $a \neq 0$, soit $t_0 \in \mathbb{R}$ un temps initial [7] et $y_0 \in \mathbb{R}$. Quelles sont les solutions de (1) qui satisfont $y(t_0) = y_0$?

d) Soient $y_1(t)$ et $y_2(t)$ deux solutions de l'equation (1), et $t_0 \in \mathbb{R}$. On suppose que $y_1(t_0) < y_2(t_0)$. Que peut-on en déduire pour $y_1(t)$ et $y_2(t)$?

2 Méthodes numériques pour la simulation et validité du modèle de Malthus

Les statistiques concernant la population du Canada de 1851 à 2006 donnent les chiffres suivants (exprimés en millions d'individus) :

dates t_i	1851	1861	1871	1881	1891	1901
population y_i	2,436297	3,229633	3,737257	4,381256	4,932206	5,418663

1911	1921	1931	1941	1951	1956	1961
7,221662	8,800429	10,376379	11,506655	14,009429	16,080791	18,238247

1966	1971	1976	1981	1986	1991	1996
20,014880	21,568305	22,992595	24,343177	25,309330	27,296856	28,846758

2001	2006
30,007094	31,612897

Si on suppose que cette population a suivi la loi de Malthus, la valeur théorique $y_{\text{theo}}(t)$ (exprimée en millions d'individus) du nombre d'habitants du Canada à l'instant t est donnée par une équation de la forme:

$$y_{\text{theo}}(t) = C.e^{a\,t} \qquad , \text{ soit } \qquad ln\,(\,y_{\text{theo}}(t)\,) = a\,t + b \quad , \tag{2}$$

[6] Vous souvenez-vous comment on le démontre ?

[7] C'est en général l'utilisateur qui choisit cet instant initial à sa convenance, mais l'instant initial peut parfois nous être imposé, il peut même nous être inconnu, comme c'est le cas dans la datation au Carbone-14 (cf. la section 3.2).

où a , b et C sont des constantes à déterminer et où $b = ln(C)$. Les valeurs expérimentales (t_i, y_i) sont bien approchées par les valeurs théoriques $(t_i, y_{\text{theo}}(t_i))$ si les points $(t_i, \ln(y_i)$ sont bien approchés par les $(t_i, \ln(y_{\text{theo}}(t_i) = (t_i, at_i + b)$. On s'attend donc à ce que les points $(t_i, \ln(y_i)$ soient presque alignées, et on cherche la droite $t \mapsto at + b$ qui passe "au plus près" de ce nuage de points. Cette idée est formalisée dans la *méthode des moindre carrés*, qu'on appelle aussi méthode de *régression linéaire*, que nous détaillons maintenant.

Premier problème : Calculer les valeurs de a et b telles que les écarts entre les valeurs mesurées $z_i = ln\,(y_i)$ à l'instant t_i et les valeurs théoriques $ln\,(\,y_{\text{theo}}(t_i)\,) = at_i + b$, au même instant soient les plus petits possibles.

Le critère[8] que nous retiendrons est de choisir les valeurs de a et b qui rendent la fonction

$$E(a,b) = \sum_{i=1}^{n} (\,ln\,(\,y_{\text{theo}}(t_i)\,) - ln(y_i)\,)^2 = \sum_{i=1}^{n} (\,a\,t_i + b - z_i\,)^2 \qquad (3)$$

la plus petite possible (ici, on a $n = 22$). Il s'agit de minimiser, lorsque a et b varient, la somme des carrés des écarts entre valeurs expérimentales et valeurs théoriques. Il s'agit donc d'un problème de recherche de minimum. Sa résolution mélange des arguments d'algèbre linéaire, de géométrie euclidienne et de fonctions de plusieurs variables. La théorie nous fournit un unique minimum pour la fonction $E(a, b)$, qui est atteint lorsque

$$a = \frac{\sum_{i=1}^{n} z_i\,(\,t_i - m_t\,)}{\sum_{i=1}^{n} (\,t_i - m_t\,)^2} \qquad , \qquad b = m_z - a\,m_t \ , \qquad (4)$$

où m_t et m_z désignent les moyennes des t_i et z_i , i. e. $m_t = \frac{1}{n}\sum_{i=1}^{n} t_i$ et $m_z = \frac{1}{n}\sum_{i=1}^{n} z_i$ Autrement dit, la droite qui passe "au plus près" du nuage de points (t_i, z_i) est bien unique, et ces coefficients sont donnés par (4).

Exercice 1 : On veut calculer a et b à l'aide d'un tableur Open Office sous Linux (équivalent de Excell, NeoOffice ou LibreOffice). Ouvrir une feuille de calcul et enregistrer là sous le nom "canada.xls".

1) Rentrer les données t_i et y_i dans deux colonnes adjacentes, par exemple A et B.

   Conserver la ligne 1 pour du commentaire (exemple : A1= "Années t_i", B1="Population y_i") Utiliser des formules automatiques quand c'est possible. Exemple : si A2="1861", on peut rentrer dans A3 la formule " $= A2 + 10$"*, copier cette formule et la coller dans les cellules A4 à A12 (ensuite il faut incrémenter par 5 années).*

2) Faire les moyennes des t_i et des y_i, par exemple en A26 et B26, en utilisant les formules adéquates (voir dans le menu des formules).

   Garder la ligne 25 pour du commentaire ("m_t", "m_y")

3) Utiliser les colonnes C, D, E, etc... pour faire des calculs intermédiaires : $z_i = \ln(y_i$ dans la colonne C, ensuite les $t_i - m_t$, les $(t_i - m_t)^2$, les $z_i(t_i - m_t)$ et toute expression utile.

   Utile à savoir : le numéro d'une cellule apparaissant dans une formule n'est pas incrémentée lors des copier-coller de la formule, si la lettre désignant la cellule est entourée par des $ $ (par exemple utiliser A26 dans les formules réfé"rant à la moyenne m_t

[8]La validité de ce critère est fondée sur des arguments et hypothèses probabilistes et statistiques, voir votre futur cours de statistiques l'an prochain.

4) Utiliser la ligne 26 pour faire la somme des $(t_i - m_t)^2$, la somme des $z_i(t_i - m_t)$

5) Conclure.

Deuxième problème : Comparer les résultats de la simulation et les données statistiques.

Considérons les valeurs a et b calculées par le programme précédent, on a donc, pour tout t :

$$y_{\text{theo}}(t) = C.e^{a\,t} \qquad \text{et} \qquad ln\,(\,y_{\text{theo}}(t)\,) = a\,t + b \quad , \tag{5}$$

où $b = ln(C)$.

Définition 2.1 *On appelle* "résidus" *l'écart* r_i *entre données statistiques*[9] *et valeurs calculées par la simulation*[10]*, i. e.* $r_i = ln(y_i) - ln\,(\,y_{\text{theo}}(t_i)\,)$.

Question : lorsque $r_i = +0,05$, quelle est l'erreur relative faite en remplaçant y_i par $y_{\text{theo}}(t_i)$? même question lorsque $r_i = -0,05$.

Le modèle utilisé pour la simulation est considéré comme idéal lorsque les deux critères suivants sont remplis :

(i) les valeurs absolues (ou relatives) des r_i sont faibles,

(ii) chaque r_i peut être considéré comme le résultat du tirage d'une variable aléatoire R_i d'espérance mathématique nulle, les variables aléatoires R_1 , R_2 , ... , R_n pouvant être considérées comme indépendantes entre elles[11] et de même variance[12].

L'idéal étant rarement atteint, on mesurera la plus ou moins bonne adéquation d'un modèle à la réalité au fait que les critères (i) et (ii) ci-dessus sont plus ou moins bien vérifiés.

Le but de l'exercice suivant est de calculer et de visualiser les résidus r_i dans le cas de l'étude de la population du Canada par le modèle de Malthus.

Exercice 2 On utilise la même feuille de calcul "canada.xls" que précédement.

1) Calculer les valeurs théoriques $x_i = at_i + b$ dans une colonne.

2) Calculer les résidus $r_i = z_i - x_i$ dans une autre colonne.

3) Visualiser les résidus en appliquant l'outil "graphique" aux colonnes des t_i et des r_i.
Choisir "colonne" puis "première colonne comme label" dans les options d'affichage. On peut peut ajouter un titre "Graphique des résidus", X-Axis: "Années" , Y-axis : "Résidus"

---

[9]Dans un contexte différent, lorsque les y_i seront les résultats de mesures expérimentales, on parlera de "données expérimentales".

[10]Ici, puisque le calcul de a et b passe par la tentative de restituer $ln(y_i)$ sous la forme $a\,t_i+b$, les données statistiques seront les $ln(y_i)$ et les valeurs simulées au même instant seront les $ln\,(\,y_{\text{theo}}(t_i)\,) = a\,t_i + b$.

[11]En particulier, la probabilité pour que R_{i+1} prenne le même signe que R_i est la même que la probabilité pour que R_{i+1} prenne le signe contraire de R_i .

[12]En particulier, l'écart entre deux résidus consécutifs R_{i+1} et R_i n'a aucune raison d'être plus faible que l'écart entre deux résidus non consécutifs R_j et R_i : en termes plus précis, l'espérance mathématique de $(R_{i+1} - R_i)^2$ devrait être a priori la même que celle de $(R_j - R_i)^2$.

4) Faire les mêmes opérations avec les résidus relatifs r_i/z_i.

Interprétation des résultats Au regard des critères (i) et (ii) ci-dessus, répondre aux questions suivantes :

a) Le modèle de Malthus est-il bien adapté pour interpréter l'évolution de la population du Canada de 1851 à 2006 inclus ? Observez-vous une distorsion par rapport au modèle ? entre quelles dates ?

b) Le modèle de Malthus est-il mieux ou moins bien adapté[13] pour interpréter l'évolution de la population du Canada de 1851 à 1956 inclus ?

Exercice 3 : Les estimations et statistiques concernant la population des USA de 1790 à 1950 donnent les chiffres suivants (exprimés en millions d'individus) :

dates t_i	1790	1800	1810	1820	1830	1840	1850	1860
population y_i	$3,929$	$5,308$	$7,240$	$9,638$	$12,866$	$17,069$	$23,192$	$31,443$

1870	1880	1890	1900	1910	1920	1930	1940	1950
$38,558$	$50,156$	$62,948$	$75,995$	$91,972$	$105,711$	$122,775$	$131,669$	$150,697$

Si on suppose que cette population a suivi la loi de Malthus, la valeur théorique $y_{\text{theo}}(t)$ (exprimée en millions d'individus) du nombre d'habitants des USA à l'instant t est donnée par une équation de la forme:

$$y_{\text{theo}}(t) = C.e^{a\,t} \quad , \text{ soit } \quad ln\,(\,y_{\text{theo}}(t)\,) = a\,t + b \ ,$$

où a, b et C sont des constantes à déterminer et où $b = ln(C)$. En adaptant les programmes ci-dessus[14], répondez aux questions suivantes :

a) Calculer les valeurs de a et b minimisant l'expression (3) (ici, on a $n = 17$), par la méthode des moindres carrés.

b) Afficher les résidus dans le cas où on cherche à simuler la population des USA par le modèle de Malthus (en optimisant le choix des paramètres a et b). Au regard des critères (i) et (ii) ci-dessus, répondre aux questions suivantes :

(1) Le modèle de Malthus est-il bien adapté pour interpréter l'évolution de la population des USA de 1790 à 1950 inclus ?

(2) Le modèle de Malthus est-il bien adapté pour interpréter l'évolution de la population des USA de 1870 à 1950 inclus ?

(3) Le modèle de Malthus est-il mieux ou moins bien adapté pour interpréter l'évolution de la population des USA de 1790 à 1860 inclus ?

[13]Le programme précédent est utilisé pour étudier la bonne ou mauvaise adéquation du modèle aux données statistiques prises sur un intervalle de temps donné. Si on considère la même population sur un intervalle de temps différent, on devra recalculer les nouvelles valeurs de a et b (donc recalculer la nouvelle fonction $t \mapsto y_{\text{theo}}(t)$) en fonction des données statistiques prises sur le nouvel intervalle de temps. à chaque fois qu'on change d'intervalle. Pour ce faire, enregistrer la feuille de calcul sous un nom différent, par exemple canada1851-1956.xls, avant de supprimer les lignes superflues. C'est pourquoi le programme recalcule a et b

[14]Enregistrer la feuille "canada.xls" sous le nom "usa.xls", avant de faire les modifications d'années et de populations nécéssaires

c) Selon les observations faites sur les populations des USA et du Canada, est-il possible de deviner les limites de validité du modèle de Malthus?

Troisième Problème : Calculer la valeur simulée $y_{\text{theo}}(t)$ de la population à tout instant t.

Ce calcul n'a de sens que si le modèle a déjà été dûment validé selon la démarche exposée ci-dessus en section 2, intitulée "deuxième problème". Dans le cas du modèle de Malthus, comme on a préalablement calculé a et b (cf. la section 2, intitulée "premier problème"), ce calcul est facile : on a

$$y_{\text{theo}}(t) = e^b . e^{a\,t} \ .$$

Le programme qui réalise ce calcul est simple.

Exercice 4 : reprendre la feuille "canada.xls".

1) Calculer dans une colonne les valeurs théoriques $exp(x_i)$ entre 1851 et 2006.

2) Appliquer l'outil "graphique" sur les colonnes des t_i des y_i et des $exp(x_i)$.

   Choisir "XY" puis "Points et lignes" dans les options de l'outil graphique. Mettre un titre "Comparaison données/modèle", X-Axis: "Années", Y-Axis: population .

3) Ajouter les années 2006 à 2030, et calculer $exp(at_i + b)$ pour les années correspondantes.

   (a) Quelle population le modèle de Malthus prévoit-il pour le Canada en 2030 ? Cette prévision est-elle fiable ?

   (b) Adapter et exécuter le programme ci-dessus de manière à obtenir l'affichage graphique annoncé.

Remarque : Le même programme peut être utilisé pour tester et simuler un modèle du type

$$y_{\text{theo}}(t) = \phi(a\,t + b) \ ,$$

où a et b sont des paramètres à déterminer et où ϕ^{-1} est une fonction connue: il suffit de modifier la adéquate en posant $z(i) = \phi^{-1}(y(i))$.

Exercice difficile[15] *:* Comment adapter le programme ci-dessus pour tester et simuler un modèle du type $y_{\text{theo}}(t) = \phi(a\,f_1(t) + b\,f_2(t))$, où a et b sont des paramètres à déterminer et où ϕ, f_1 et f_2 sont des fonctions connues ?

3 Un phénomène totalement différent relevant de la même équation différentielle :

3.1 Désintégration :

Dans le processus de désintégration radioactive d'un corps A (le cas du Carbone 14 est un exemple célèbre), la quantité de corps A qui est désintégrée par Unité de Temps est supposée proportionnelle

[15] Cet exercice étant pour l'instant trop difficile, sa correction est différée à une feuille d'exercice ultérieure, quand nous en aurons vraiment besoin.

à la quantité $a(t)$ de corps A présente à cet instant $a(t)$ (cette quantité est généralement une concentration, mesurée par exemple en nombre de molécules par unité de volume). En d'autre termes, le taux de désintégration $\frac{a'(t)}{a(t)}$ est une constante.

a) Montrer que cette constante est strictement négative, nous la noterons $-k$.

b) Quelle est l'équation différentielle satisfaite par la fonction $t \mapsto a(t)$.

c) Connaissant l'instant initial t_0 et la valeur initiale $a(t_0)$, calculer $a(t)$ pour tout $t > t_0$. Donner l'allure du graphe de la fonction a.

d) Calculer (en fonction de k) le nombre $T_{1/2}$ d'années (à partir de l'instant initial) nécessaires pour que la concentration diminue de moitié par rapport à la concentration initiale.

e) Pour toute valeur de t, calculer $\frac{a(t+T_{1/2})}{a(t)}$. Cette quantité dépend-t-elle de t ?

f) Pour toute valeur de t, calculer $\frac{a'(t+T_{1/2})}{a'(t)}$. Cette quantité dépend-t-elle de t ?
La quantité $T_{1/2}$ est appelée temps de demi-vie du processus.

3.2 Datation au Carbone-14 :

Un des procédé utilisé[16] pour dater certains objets trouvés lors de fouilles archéologiques (restes animaux ou végétaux, bois et sutout charbon de bois) est de mesurer le nombre de molécules de Carbone-14 qui se désintègrent par unité de temps. Le principe est le suivant : l'atmosphère est bombardée par les rayons cosmiques ce qui produit des neutrons libres qui, en réagissant avec l'azote de l'atmosphère, produisent un isotope radioactif du carbone : le Carbone-14. En réagissant avec l'oxygène, ce Carbone-14 produit un dioxyde de carbone qui, mélangé au dioxyde de carbone usuel, se répand de manière uniforme dans l'atmosphère et est absorbé, au même titre que le dioxyde de carbone usuel, par les plantes et, à travers les plantes, par les animaux.

A) **Les hypothèses sur lesquelles est fondée la méthode :**

La possibilité d'utiliser la désintégration du Carbone-14 pour calculer effectivement la date de mort d'un organisme est fondée sur les hypothèses suivantes :

(i) Le rapport $\left(\frac{\text{Nombre de molecules de Carbone}-14}{\text{Nombre total de molecules de Carbone}}\right)$ par unité de masse, que nous noterons dans la suite $\left(\frac{^{14}C}{C \text{ total}}\right)$ est considéré comme uniforme dans l'atmosphère (où ces deux types de carbone sont diffusés sous forme de dioxyde de carbone) et est, dans les organismes vivants, égal à ce qu'il est dans l'atmosphère, en raison du fait que les organismes vivants absorbent et fixent le Carbone-14 et le Carbone normal sans faire de différence entre l'un et l'autre, donc dans les mêmes proportions que dans le dioxyde de carbone atmosphérique. Bien sûr, à peine absorbé, le Carbone-14 commence à se désintégrer, donc sa proportion tend à diminuer, mais ce phénomène (très lent) est, chez les organismes vivants, compensé par les échanges permanents entre les organismes vivants et leur milieu, qui rétablissent en permanence la proportion $\left(\frac{^{14}C}{C \text{ total}}\right)$ présente dans l'atmosphère.

[16] Le procédé a été développé par Willard Frank Libby et son équipe de 1945 à 1949, et lui a valu le Prix Nobel de Chimie en 1960. La date 1949 est celle du premier test par le Carbone-14, effectué par cette équipe, sur deux échantillons de bois provenant d'une tombe égyptienne : il s'agissait d'un premier test de validité de la méthode, puisque les archéologues savaient dater cette tombe comme vieille de 4600 ans, voir l'exercice à ce sujet plus loin.

(ii) Le bombardement de l'atmosphère par les rayons cosmiques est l'unique source[17] de Carbone-14 et ce rayonnement cosmique est resté à peu près constant dans les 500 derniers siècles[18], ce qui implique que le rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ dans l'atmosphère est resté constant au cours des siècles.

B) **Traduction de ces hypothèses en termes mathématiques :**

– Une première conséquence de ces hypothèses est que le rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ se maintient à un niveau constant (noté a_0) pendant toute la vie de l'organisme considéré. Notons t_0 l'instant où cet organisme meurt; à partir de cet instant, il n'y a plus ingestion de Carbone-14 et, la désintégration radioactive n'étant plus compensée, la concentration en Carbone-14 commence à diminuer selon une loi décrite dans la section 3.1. Notons $a(t)$ la valeur du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ à un instant t postérieur à t_0, la fonction a obéit donc, sur l'intervalle $[\,t_0\,,\,+\infty\,[$, à l'équation différentielle

$$a' = -k\,a \tag{6}$$

où k est une constante.

– Une autre conséquence des hypothèses (i) et (ii) est que la valeur $a_0 = a(t_0)$ du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ à l'instant de la mort de l'organisme est la même pour un organisme qui est mort à un instant t_0 situé il y a plusieurs siècles et pour un organisme mort en 1950. Par exemple, si on appelle $a(t)$ la valeur à l'instant t du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ dans un charbon de bois $\mathcal{A}$ produit à une date t_0 (située, par exemple, pendant la préhistoire), si on appelle $\tilde{a}(t)$ la valeur à l'instant t du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ dans un charbon de bois $\widetilde{\mathcal{A}}$ produit en 1949, on a

$$a(t_0) = \tilde{a}(1949) \quad . \tag{7}$$

Question : Justifier l'égalité (7) et prouvez que ceci implique que

$$a'(t_0) = \tilde{a}'(1949) \quad , \tag{8}$$

donc qu'il a suffi de mesurer le nombre de molécules (par gramme) de $\widetilde{\mathcal{A}}$ qui se désintégraient par unité de temps en 1950 pour en déduire la valeur de $a'(t_0)$.

C) **Quelles sont les données qui nous sont accessibles pour faire le calcul ?**

Evidemment, la date t_0 à laquelle commence le processus de désintégration nous est ici inconnue, puisque c'est précisément cette date que la méthode de datation cherche à déterminer!

[17]Ceci est à peu près vrai, mais seulement pour les organismes qui sont morts avant 1950. Depuis 1950 en effet, on observe une augmentation du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ dans l'atmosphère (et donc dans la biosphère), que l'on suppose due aux expériences atomiques. C'est pourquoi 1950 sert souvent de date de référence : pour les organismes morts après cette date, il faudrait revoir les paramètres des calculs. Ce phénomène récent n'a évidemment aucun impact sur la datation d'objets anciens, pour lesquels l'hypothèse faite reste donc valable.

[18]Ceci n'est vrai qu'en première approximation, les variations du rayonnement cosmique au cours des siècles sont donc sources d'erreurs dans cette méthode de datation.

a) *Quelles sont les données qui étaient accessibles lors de l'invention de la méthode ?*

A cette époque, la valeur $a(t)$ à l'instant t du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ pour l'objet à dater était trop petite (de l'ordre de 10^{-12}) pour être directement accessible à la mesure, on mesurait donc la radioactivité d'un échantillon de l'objet à dater[19], i. e. le nombre $R(t)$ de molécules de Carbone-14 qui disparaissaient par Unité de Temps et par gramme, mesuré à l'instant t.

Question 1 : Si N_0 est le nombre de molécules de Carbone total par gramme, montrer que $a'(t) = -\frac{R(t)}{N_0}$.

Pour la même raison, la valeur $\tilde{a}(1949)$ du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ pour un échantillon "neuf" (i. e. mort en 1949) n'était pas accessible à la mesure, on mesurait la radioactivité de l'échantillon "neuf", i. e. le nombre $\tilde{R}$ de molécules de Carbone-14 de l'échantillon "neuf" qui disparaissaient par Unité de Temps et par gramme, mesuré en 1949.

Question 2 : Si N_0 est le nombre de molécules de Carbone total par gramme, montrer que $a'(t_0) = -\frac{\tilde{R}}{N_0}$.

Exercice 3 :

(1) Si on dispose d'un objet de bois dont on connaît de manière sûre la date de fabrication t_0 et si, pour cet objet, on connaît $R(1949)$ et $\tilde{R}$, donnez une formule pour le calcul du temps de demi-vie du carbone 14.

(2) En 1949, on a fait un test sur un échantillon de bois provenant d'une tombe égyptienne que les archéologues savaient dater de manière sûre comme vieille de 4600 ans. Supposons que les mesures faites en 1949 donnent un nombre de désintégrations de $3,7678$ par minute et par gramme pour cet échantillon et un nombre de désintégrations de $6,68$ par minute et par gramme pour un échantillon neuf de référence. Calculez le temps de demi-vie du Carbone-14.

(3) Reprenez cette question en supposant maintenant (de manière plus réaliste) que le nombre de désintégrations pour l'échantillon égyptien est compris entre $3,756$ et $3,78$ par minute et par gramme.

(4) Un charbon de bois provenant de la grotte de Lascaux donnait, en 1950, un nombre de de désintégrations de $0,97$ par minute et par gramme, alors qu'à la même date, le nombre de désintégrations par minute et par gramme pour un échantillon neuf de référence était de $6,68$.

- On suppose dans un premier temps que le temps de demi-vie du Carbone-14 est de 5568 années (estimation due à W. F. Libby). Estimer dans ce cas la date[20] à laquelle ce charbon de bois a été utilisé dans la grotte de Lascaux et, par conséquent, la date probable des peintures de cette grotte.
- On suppose dans un second temps que le temps de demi-vie du Carbone-14 est de 5734 années (estimation issue d'expériences plus récentes). Estimer dans ce cas la date à laquelle ce charbon de bois a été utilisé dans la grotte de Lascaux.

[19] Bien entendu, la faiblesse du signal induisait des erreurs assez importantes.

[20] Les mesures, que nous essayons de reconstituer ici, ont été faites à l'époque par Willard Frank Libby et son équipe. La date trouvée alors par datation au Carbone-14 a été mise en doute, car nettement postérieure à ce qui avait été suggéré par l'abbé Breuil, qui avait fouillé le site le premier. D'autres estimations faites depuis (y compris par d'autres méthodes) ont confirmé la première datation au Carbone-14 (en la situant à une date légèrement antérieure, voir la question suivante). Des études sur le style des peintures tendraient à situer la grotte à une date encore un peu antérieure. Le débat reste donc ouvert à cause de l'imprécision de toutes ces méthodes.

b) *Quelles sont les mesures accessibles actuellement ?*

L'utilisation du spectromètre de masse,permet actuellement de mesurer directement la valeur $a(t)$ à l'instant t du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ pour l'objet à dater.

Exercice[21] : On étudie un échantillon de bois trouvé dans une tombe datant de l'antiquité en vue de trouver la date t_0 de la taille de ce bois. Notons $a(t)$ la valeur à l'instant t du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ pour cet échantillon.

Pour cet échantillon, on dispose des valeurs $a(t_1)$ et $a(t_2)$ du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ aux dates $t_1 =$ 1er Janvier 2000 et $t_2 =$ 1er Janvier 2010
On dispose aussi de la mesure du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ pour un échantillon de référence taillé en 1950 : ce rapport prenait, en 1950, la valeur $\tilde{a}_0$.

* Calculer $a(t_0)$ en fonction des données ci-dessus.
* L'unité de temps étant l'année, calculer[22] la constante k de l'équation (6) en fonction des données $a(t_1)$ et $a(t_2)$.
* Calculer (en fonction de k , puis de a_1 et a_2) le temps $T_{1/2}$ de demi-vie du Carbone-14.
* Calculer la date t_0 de taille du bois trouvé dans la tombe en fonction de $\tilde{a}_0$, $a(t_1)$ et $a(t_2)$.

[21] Le but de cet exercice est de calculer le temps de demi-vie du Carbone 14 et de dater un objet utilisé à une date t_0 inconnue lorsqu'on ne dispose pas d'un échantillon de référence dont la date est certifiée.

[22] Comme $a(t_1)$ et $a(t_2)$ sont très voisins, le calcul de k par cette méthode nécessite que $a(t_1)$ et $a(t_2)$ soient calculés avec précision (au moins 3 chiffres significatifs sûrs), sinon l'erreur induite sur le calcul de k peut-être très grande. Ceci laisse peu d'espoir si on utilise la méthode de mesure de la radioactivité; en revanche la méthode de mesure directe du rapport $\left(\frac{^{14}C}{C \text{ total}}\right)$ par utilisation du spectromètre de masse laisse plus d'espoir, surtout si on fait des relevés à des instants t_1 , t_2 , ... , t_n suffisamment nombreux, sur un intervalle de temps suffisamment large; la valeur simulée de $a(t)$ pour tout t est alors donnée par $ln\,(a(t)) = -k\,t + b$, où on calcule $-k$ et b par la méthode des moindres carrés (cf. la section 2).