ANNEXE T.2. - ÉLISE & DAVID

APP0

Quand un biologiste et une ingénieur s'entraident

Élise & David 1

David, jeune étudiant en biologie et passionné par ses études, décide de reproduire des expériences classiques concernant la croissance d'une population d'organismes unicellulaires et demande l'aide d'Élise, jeune étudiante ingénieur, pour modéliser les résultats obtenus. Durant la première expérience, les conditions expérimentales mises en œuvre par David sont telles que l'on peut faire l'hypothèse que les organismes unicellulaires ont suffisamment d'espace et de nourriture pour pouvoir se développer sans contraintes. La croissance de la population dépend donc essentiellement de deux paramètres : le taux de "naissance [1]", noté a, et le taux de "décès", noté b. Pendant que David prépare ses instruments de mesure (son microscope notamment) pour compter ses organismes unicellulaires, Élise se préoccupe de trouver un modèle d'évolution de la population valable pour tout temps t. Mais qu'est-ce qu'Élise entend réellement par *modèle* et que pourrait-elle proposer, à votre avis ? Considérez que la population initiale est P_0 et discutez vos résultats en fonction de la valeur de $r = a - b$. Ce paramètre r est généralement appelé taux de reproduction.

1. Modélisation
   Il est important de réaliser que dans ce qui suit, deux "taux" différents vont apparaître. L'énoncé parle de *taux de naissance et de décès* qui sont des rapports entre le nombre de naissances (ou de décès) et la population, à un instant donné et pour un intervalle de temps donné. Si on modélise la population comme une fonction du nombre de jours ($P(t)$ où l'unité de t est "jour"), on considère alors que le taux de naissance est le rapport entre le nombre de naissances durant un jour et la population à cet instant (exprimé en jours) de sorte que l'accroissement de population durant une fraction de jour Δt sera $aP(t)\Delta t$. Ensuite, pour modéliser l'évolution de la population au cours du temps, on va introduire un *taux d'accroissement* qui est le rapport entre l'accroissement de la population et l'intervalle de temps pendant lequel cet accroissement est mesuré.
   Soit $P(t)$ la population à l'instant t. Alors, pour un petit intervalle de temps Δt, on peut écrire, d'après l'énoncé,
   $$P(t + \Delta t) - P(t) \approx aP(t)\Delta t - bP(t)\Delta t$$
   donc $\frac{P(t+\Delta t)-P(t)}{\Delta t} \approx rP(t)$ avec $r := a - b$.
   Par passage à la limite, quand $\Delta t \to 0$
   $P'(t) = \frac{dP}{dt} = \lim_{\Delta t \to 0} \frac{P(t+\Delta t)-P(t)}{\Delta t} = rP(t)$.
   Ceci est une équation différentielle du premier ordre, linéaire, homogène, à coefficients constants.
   — équation différentielle : l'inconnue est une fonction P (de la variable indépendante t) dont on connaît une relation avec sa (ou ses) dérivée(s)(et pas $P(t)$ elle-même) ;
   — du premier ordre : dans cette équation, seule la dérivée première intervient (et pas les dérivées d'ordre supérieur) ;
   — linéaire, homogène : le second membre est une fonction linéaire et homogène de l'inconnue $P(t)$;
   — à coefficients constants : le coefficient r qui intervient dans le second membre est une constante (ne dépend pas de t).

1. Dans le présent contexte, la "naissance" correspond au processus de division cellulaire, générant donc deux cellules "filles" à partir d'une cellule "mère".

2. Résolution de $P'(t) = rP(t)$
   Deux pistes de résolution sont proposées :
   — Quelle est la fonction dont la dérivée est elle-même multipliée par une constante ? C'est la fonction exponentielle et donc $P(t) = Ce^{rt}$.
     Alternativement,
   — $\frac{dP}{dt} = rP \Rightarrow \frac{dP}{P} = rdt$ (c'est la méthode de séparation des variables (dépendante (P) et indépendante (t))).
     On intègre de part et d'autre du signe = en n'oubliant pas une constante d'intégration :
     $\ln|P| = rt + C'$
     $\Rightarrow P(t) = e^{rt+C'} = e^{C'}e^{rt} = Ce^{rt}$.

   Pour déterminer la constante C, en $t = 0$, on a : $P(0)P_0 = Ce^0 = C$.
   L'expression finale de la solution est donc $P(t) = P_0e^{rt}$.
3. Discussion
   Si $r > 0$, il s'agit d'une exponentielle croissante et la population "explose".
   Si $r < 0$, il s'agit d'une exponentielle décroissante et la population s'éteint.
   $r > 0 \Leftrightarrow$ taux de naissance > taux de décès
   $r < 0 \Leftrightarrow$ taux de naissance < taux de décès.

Élise & David 2

David retrouve Élise quelques jours plus tard et lui fait part de ses mesures. Il avait 10 organismes au départ, a compté 25 cellules après 24 heures, 45 après 36 h, 144 après 3 jours et 360 à la fin du 4^e jour. Élise lui explique sa démarche de modélisation et tous deux essaient de déterminer quelle sera la taille de la population après 10 jours. Quel pourrait être le résultat trouvé, en essayant de prendre en compte l'ensemble des mesures patiemment effectuées par David ?

Élise et David sont impressionnés par le résultat de cette croissance après 10 jours, et afin de garantir, si possible, que l'hypothèse initiale reste satisfaite (suffisamment d'espace et de nourriture) David envisage de retirer de son milieu de culture R cellules par unité de temps. Il se fixe comme objectif d'enlever 20 cellules par jour, à partir du deuxième jour. Cela aura-t-il un effet déterminant sur l'évolution de la population ?

1. Ajustement des paramètres aux données.
   L'équation $\ln|P| = rt + C'$ est l'équation d'une droite dans le plan $(\ln|P|, t)$, dépendant de deux paramètres, la pente r et l'ordonnée à l'origine C'. Il s'agit donc de faire passer une droite au mieux parmi 5 points donnés qui ne sont pas parfaitement alignés (erreurs de mesures ou erreurs de modélisation), et pour cela trouver les meilleures valeurs des deux paramètres. C'est un problème de régression linéaire. Si les étudiants sont un peu débrouillards, ils peuvent trouver la solution dans un tableur Excel (la fonction "droitereg" calcule une droite de régression). Sinon, ils peuvent aussi essayer de "résoudre" ce problème de régression graphiquement. Il ne faut pas oublier de modifier les données en remplaçant les valeurs des populations mesurées par leur logarithme. Si on encode le temps en heures, les calculs avec excel donnent $C' = 2.354$ soit $C = P(0) = 10,527$ et $r = 0,0369$ (notez que si on encode le temps en jours, on obtient la même valeur pour C' (évidemment !) mais une valeur différente pour $r = 0.886 = 24 \cdot 0.0369$). Après 10 jours, la population vérifiera l'équation $\ln|P| = 0.0369 \cdot 240 + 2.354$ ou $P = e^{0.0369 \cdot 240 + 2.354} = 73874$. Chiffre impressionnant sachant que la population initiale était constituée de 10 cellules !
2. Nouvelle modélisation
   $P(t + \Delta t) - P(t) \approx aP(t)\Delta t - bP(t)\Delta t - R\Delta t$
   donc $\frac{P(t+\Delta t)-P(t)}{\Delta t} \approx (a - b)P(t) - R = rP(t) - R$
   Par passage à la limite, quand $\Delta t \to 0$
   $P'(t) = \frac{dP}{dt} = \lim_{\Delta t \to 0} \frac{P(t+\Delta t)-P(t)}{\Delta t} = rP(t) - R$
   Ceci est aussi une équation différentielle du premier ordre, linéaire, à coefficients constants. Par contraste avec la première, celle-ci est "non-homogène" (la première était homogène) car le second membre n'est pas "homogène" en l'inconnue $P(t)$.
3. Différentes méthodes de résolution.
   — Solution générale de l'équation linéaire non-homogène = solution générale de l'équation linéaire homogène + solution particulière de l'équation non-homogène. Ici, comme le terme non-homogène est une constante, on peut essayer de trouver une solution particulière constante (plus généralement, si le terme non-homogène est une fonction polynomiale de t, on peut essayer de trouver une solution particulière de la même forme. Il en est de même pour des termes non-homogènes trigonométriques ou exponentiels).
   Cherchons donc une solution particulière constante : $P(t) = P_c$. Insérant cette solution dans l'équation, on obtient $0 = rP_c - R$, d'où $P_c = \frac{R}{r}$.
   La solution générale de l'équation non-homogène est donc : $P(t) = Ce^{rt} + \frac{R}{r}$.
   — Variation de la constante. On postule une solution de la forme $P(t) = K(t)e^{rt}$ (on remplace dans la solution de l'équation homogène la constante C par une fonction $K(t)$). En injectant cette solution dans l'équation différentielle, on obtient :

$$\begin{aligned} K'(t)e^{rt} + rK(t)e^{rt} &= rK(t)e^{rt} - R \\ K'(t) &= -Re^{-rt} \\ K(t) &= \frac{R}{r}e^{-rt} + C \\ P(t) &= \frac{R}{r} + Ce^{rt} \end{aligned}$$

   Notons que la deuxième équation ci-dessus est une équation différentielle directement intégrable car la fonction inconnue $K(t)$ n'apparait pas dans le second membre de l'équation.

— Le livre Calculus propose également une autre méthode, dite du facteur d'intégration. Si les étudiants veulent l'utiliser, il faut exiger qu'ils comprennent réellement ce qu'ils font et qu'ils ne se contentent pas d'appliquer une formule.

4. Prévision de la valeur de la population après 10 jours.
   Supposons que le taux de reproduction est le même que celui calculé précédemment ($r = 0.0369$) car il s'agit de la même population d'organismes unicellulaires, dans des conditions de culture pratiquement identiques. Le retrait s'effectuant à partir du deuxième jour, prenons comme condition initiale $P(0) = 25$, ce que David avait mesuré à la fin du premier jour. Le taux de reproduction étant un taux horaire (vu la manière dont il a été calculé), il convient de prendre pour R une valeur de retrait (constant) horaire également, soit $20/24$ cellules par heure d'après l'énoncé. Ceci nous permet de calculer la constante C de l'équation : $P(0) = 25 = \frac{20/24}{0.0369} + C$ qui nous donne $C = 2.416$. L'équation devient donc $P(t) = 22.584 + 2.416e^{0.0369t}$ et après 9 jours d'évolution suivant cette équation (soit 10 jours après le début de l'expérience de David), on obtient $P(216) \approx 7015$. La croissance est effectivement ralentie mais le jour suivant, les organismes seront déjà près de 17000 et plus de 41000 après 12 jours. L'explosion n'est donc que retardée...

Élise & David 3

David et Élise se rendent bien compte que le modèle de croissance qu'ils ont déterminé n'est pas soutenable et que l'hypothèse initiale ne pourra pas être satisfaite très longtemps. Expliquez le défaut majeur du premier modèle proposé par Élise.

Une manière plus réaliste de tenir compte des contraintes d'espace et de nourriture (notamment) sur l'évolution de la population de ces organismes monocellulaires étudiés par David avec l'aide d'Élise est de supposer que le taux de décès et le taux de division cellulaire ne sont pas constants au cours du temps. Élise propose alors, après discussion avec David, de supposer que ces taux sont deux fonctions affines de la population, croissante pour les décès et décroissante pour les naissances. Commentez cette hypothèse de modélisation. Vous parait-elle raisonnable ?

Élise est bien décidée à explorer ce nouveau modèle et se donne d'abord comme objectif de le mettre sous une forme semblable au premier modèle, en faisant apparaitre un taux de reproduction, qui n'est, cette fois-ci, plus constant. Et en analysant ce modèle plus attentivement, elle observe qu'il prédit une stabilisation de la population. Pouvez-vous confirmer à David cette bonne nouvelle ?

David relance une nouvelle expérience et obtient les résultats suivants :
- 10 cellules comme population initiale,
- 20 cellules après un jour,
- une population d'équilibre (constaté après une vingtaine de jours) évaluée à 1000 cellules.

Pouvez-vous produire un graphique illustrant l'évolution de la population au cours du temps ? Ce graphique permettrait de clôturer le rapport de David de la plus belle manière !

1. Nouvelle modélisation
   $P(t+\Delta t) - P(t) \approx (a - cP(t))P(t)\Delta t - (b + dP(t))P(t)\Delta t = (a-b)P(t)\Delta t - (c+d)P^2(t)\Delta t$
   donc $\frac{P(t+\Delta t)-P(t)}{\Delta t} \approx (a-b)P(t) - (c+d)P^2(t)$
   Par passage à la limite, quand $\Delta t \to 0$
   $P'(t) = \frac{dP}{dt} = \lim_{\Delta t \to 0} \frac{P(t+\Delta t)-P(t)}{\Delta t} = (a-b)P(t) - (c+d)P^2(t)$
2. Reformulation
   Essayons de mettre cette équation sous une forme semblable à celle de la première équation. Pour cela, posons $r_0 = a-b$ et $c+d = r_0/K$, où $K = (a-b)/(c+d)$ est une nouvelle constante. L'équation devient alors :
   $$P'(t) = \frac{dP}{dt} = r_0(1 - P(t)/K)P(t) = rP(t).$$
   Ici, le taux de reproduction $r = r_0(1 - P(t)/K)$ n'est plus une constante mais, si P est très petit par rapport à K, ce taux est proche de la constante r_0 alors que si P s'approche de K, le taux de reproduction tend vers 0. Remarquons encore que si P est plus grand que K, le taux de reproduction est négatif et la population décroît. Cette équation porte le nom d'équation logistique et ce modèle logistique a été construit pour la première fois par le mathématicien belge Pierre-François Verhulst (1804-1849).
3. Analyse de l'équation
   Les commentaires un peu plus haut indiquent déjà comment la population peut évoluer. Partant d'une valeur initiale P petite, la population va croître selon une exponentielle, caractérisée par un taux de reproduction proche de r_0 mais qui va diminuer au fur et à mesure que P croît. Lorsque P se rapproche de K, le taux de reproduction tend vers 0, ce qui indique que la croissance ralentit mais P continue de se rapprocher de K et, à la limite, lorsque $P = K$, la dérivée P' s'annule et donc P reste égal à K. K est donc la valeur d'une population **d'équilibre**. $P = 0$ est aussi une valeur d'équilibre (la dérivée est nulle) mais une petite perturbation de cette valeur (positive) verra la population s'éloigner de cet équilibre et rejoindre la valeur K. Au contraire, si la population est à l'équilibre K, une petite perturbation (provoquant $P > K$ ou $P < K$) sera telle que la population rejoindra à nouveau cette valeur K. On dit que l'équilibre en K est (asymptotiquement) stable.
4. Résolution de l'équation logistique
   On utilise la méthode de séparation des variables (dépendante et indépendante). L'équation devient :
   $$\frac{dP}{P(1-P/K)} = r_0 dt \text{ ou encore } \frac{KdP}{P(K-P)} = r_0 dt$$

Comme $\frac{K}{P(K-P)} = \frac{1}{P} + \frac{1}{K-P}$ l'équation à résoudre est :

$$\left(\frac{1}{P} + \frac{1}{K-P}\right) dP = r_0 dt$$

et on peut intégrer chacun des membres pour obtenir

$$\ln|P| - \ln|K-P| = \ln\left|\frac{P}{K-P}\right| = r_0 t + C'$$

ou

$$\left|\frac{P}{K-P}\right| = e^{r_0 t + C'} = Ce^{r_0 t}$$

que l'on peut récrire

$$P(t) = \frac{KCe^{r_0 t}}{1 + Ce^{r_0 t}}$$

5. Ajustement aux données
   Cette fois-ci, trois données sont proposées et la fonction décrivant P dépend de trois paramètres : K, C et r_0. L'énoncé fournit la valeur de la population d'équilibre et on a donc $K = 1000$. L'énoncé fournit aussi la valeur de la population initiale :

   $$P(0) = \frac{KC}{1+C} = 10$$

   soit

   $$C = \frac{10}{K - 10} = 0.010$$

   Enfin, l'énoncé nous donne $P(24)$, la valeur de la population après un jour. De l'équation décrivant P on tire

   $$e^{r_0 t} = \frac{P(t)}{C(K - P(t))}$$

   et donc

   $$r_0 = \frac{1}{24} \ln \frac{P(24)}{C(K - P(24))} = \frac{1}{24} \ln \frac{20}{0.010 \cdot 980} = 0,0297$$

   Il reste à encourager les étudiants à représenter tout cela graphiquement...