

BROUILLON - MODÉLISER, C'EST QUOI.

CHRISTOPHE BAL

*Document, avec son source L^AT_EX, disponible sur la page
<https://github.com/bc-writing/drafts>.*

Mentions « légales »

Ce document est mis à disposition selon les termes de la licence Creative Commons “Attribution - Pas d’utilisation commerciale - Partage dans les mêmes conditions 4.0 International”.



TABLE DES MATIÈRES

1. Le principe de « multi-proportionnalité »	2
1.1. Quelques modèles célèbres	2
1.2. Les modèles sont-ils bons ?	3
2. Coût marginal – Polynômes de degré 2 ou 3	5
2.1. Une modélisation	5
2.2. Le modèle est-il bon ?	7
3. Désintégration radioactive – Probabilités et lois continues, ou pas...	8
3.1. Une modélisation	8
3.2. Le modèle est-il bon ?	9
3.3. Annexe : le cas d’une probabilité quelconque	9
4. Daniel BERNOULLI et la variole – Équations différentielles non linéaires	10
4.1. Une modélisation	10
4.2. Le modèle est-il bon ?	12
4.3. Annexe : résoudre autrement le système d’équations différentielles	13
5. AFFAIRE À SUIVRE...	14

Dans ce modeste document, nous allons présenter quelques modèles classiques en cherchant surtout à nous interroger sur leur validité du point de vue mathématique. Ceci nous amènera à nous questionner sur la signification du mot « *modéliser* », un mot dont nous reportons la définition à la toute dernière section de ce document.

1. LE PRINCIPE DE « MULTI-PROPORTIONNALITÉ »

1.1. **Quelques modèles célèbres.** Les sciences du réel sont avant tout des sciences de la mesure. Il peut arriver que des phénomènes fassent apparaître, via de simples mesures, de la proportionnalité sur plusieurs variables étudiées indépendamment les unes des autres : l'auteur de ces lignes parle de « *multi-proportionnalité* ». Voici des exemples classiques de ce type fournis par les Sciences Physiques¹.

- En Électricité, en régime continu, la tension et l'intensité du courant aux bornes d'une résistance vérifient $U = RI$. Cette loi est due à Georg² Simon Ohm (1789 – 1854). On imagine bien un ensemble de mesures de U en fonction de I faisant apparaître une relation de quasi proportionnalité du fait de certaines imprécisions de mesure. Une fois cette observation faite, il devient « *naturel* » de poser la loi $U = RI$.
- Amedeo Avogadro (1776 – 1846) a énoncé la loi des gaz parfaits : $PV = nRT$. Dans cette relation, V est un volume de gaz en m^3 , P une pression en pascal, T une température en K , n la quantité de matière en mol , et $R \approx 8,314$ la constante des gaz parfait en $J \cdot mol^{-1} \cdot K^{-1}$. On imagine bien ici aussi une étude, à quantité de matière fixée, d'une relation entre P , V et T en s'intéressant à deux variables, la troisième restant constante. On observerait alors les phénomènes suivants, sous des conditions physiques acceptables correspondant en fait au domaine de validité de la loi.
 - (a) P étant fixé, on observe que la température est proportionnelle à V d'où l'on pose $T = a(P)V$ avec $a(P)$ une constante dépendant a priori de P .
 - (b) V étant fixé, on observe que la température est proportionnelle à P d'où $T = b(V)P$ avec $b(V)$ une constante dépendant a priori de V .

À ce stade, le physicien aguerri propose une formule du type $T = kPV$ avec k une constante ne dépendant pas de T , P et V . Intuitivement c'est facile à comprendre mais est-ce mathématiquement correct ? Nous verrons dans la section suivante que oui !

Le passage de $T = kPV$ à la loi des gaz parfaits s'obtiendrait de façon analogue via la prise en compte en plus du paramètre n . Le choix et la signification de la constante R sont motivés par le physicien qui veut rendre sa formule la plus expressive possible.

- Nous devons à Isaac Newton (1642 – 1727) la loi de l'attraction universelle qui dit que l'intensité en N de la force exercée entre deux corps A et B est $F_{A \leftrightarrow B} = G \frac{m_A m_B}{d^2}$ avec m_A et m_B les masses en kg de chacun des deux corps, d la distance en m les séparant et $G \approx 6,674$ la constante gravitationnelle en $N \cdot m^2 \cdot kg^{-2}$. Tout comme pour la loi des gaz parfaits, on pourrait imaginer étudier l'intensité relativement à m_A , m_B et l'inverse de d^2 respectivement, les deux autres paramètres étant fixés à chaque fois.

Pourquoi l'inverse de d^2 ? On peut imaginer une intuition relativement à l'inverse de d , puis lors de mesures on s'aperçoit qu'il y a en fait une relation quadratique relativement $\frac{1}{d}$. Dès lors, il suffit de tenter sa chance avec l'inverse de d^2 .

Dans chacun des cas proposés, nous avons rencontré une fonction f à n variables x_1, \dots, x_n qui est proportionnelle relativement à chaque variable x_i dès que l'on fixe les autres, puis nous

1. L'auteur n'affirme pas que ce qui suit est effectivement ce qui a guidé l'établissement des lois présentées mais cette approche de modélisation reste concrètement très acceptable. Afin de préciser le propos, une étude historique serait nécessaire mais le temps manque au papa qui écrit ces lignes... Toute contribution est bienvenue !

2. L'absence d'un « e » final n'est pas une faute de frappe.

sommes arrivés à une relation du type $f(x_1; \dots; x_n) = kx_1 \cdots x_n$ avec k une constante. Bien que cela semble intuitivement clair, il serait bien de prouver la validité de cette affirmation. Ceci va être fait dans la section suivante.

Remarque 1.1. *Bien entendu une fois ces lois établies, il va falloir les valider par des mesures et étudier leur domaine d'applicabilité. Ceci se fait de façon directe et indirecte : par exemple, la loi d'attraction de Newton permet de valider les observations empiriques de Johannes Kepler (1571 – 1630) sur les orbites coniques des astres³. N'est-ce pas joli ? C'est là tout le beau travail du scientifique du réel.*

1.2. Les modèles sont-ils bons ? Expliquons pourquoi ce qui précède est mathématiquement acceptable. Nous examinerons d'abord le cas de trois variables afin de nous permettre d'envisager sereinement le cas général via une récurrence sur le nombre de variables.

Cas avec trois variables. Considérons une fonction réelle $f(x; y; z)$ définie sur $(\mathbb{R}_+)^3$ qui vérifie les conditions suivantes.

- (1) $\forall (y; z) \in (\mathbb{R}_+)^2$, la fonction $x \rightarrow f(x; y; z)$ est proportionnelle à x .
- (2) $\forall (x; z) \in (\mathbb{R}_+)^2$, la fonction $y \rightarrow f(x; y; z)$ est proportionnelle à y .
- (3) $\forall (x; y) \in (\mathbb{R}_+)^2$, la fonction $z \rightarrow f(x; y; z)$ est proportionnelle à z .

Autrement dit, nous avons :

- (1) $\forall (y; z) \in (\mathbb{R}_+)^2$, il existe $a(y; z)$ telle que $f(x; y; z) = a(y; z)x$.
- (2) $\forall (x; z) \in (\mathbb{R}_+)^2$, il existe $b(x; z)$ telle que $f(x; y; z) = b(x; z)y$.
- (3) $\forall (x; y) \in (\mathbb{R}_+)^2$, il existe $c(x; y)$ telle que $f(x; y; z) = c(x; y)z$.

Nous avons alors :

$a(y; z)x = b(x; z)y$ sur $(\mathbb{R}_+)^3$ puis $\frac{a(y; z)}{y} = \frac{b(x; z)}{x}$ sur $(\mathbb{R}_+^*)^3$. Comme à droite et à gauche du signe égal, les expressions ne dépendent pas de x et y respectivement, mais peuvent dépendre de z , nous avons l'existence de $d(z)$ telle que $\frac{a(y; z)}{y} = \frac{b(x; z)}{x} = d(z)$ d'où $f(x; y; z) = d(z)xy$ sur $(\mathbb{R}_+^*)^3$.

Ensuite nous obtenons :

$d(z)xy = c(x; y)z$ puis $\frac{d(z)}{z} = \frac{c(x; y)}{xy}$ sur $(\mathbb{R}_+^*)^3$. Ceci prouve l'existence d'une constante k telle que $\frac{d(z)}{z} = \frac{c(x; y)}{xy} = k$ d'où $f(x; y; z) = kxyz$ sur $(\mathbb{R}_+^*)^3$.

La formule $f(x; y; z) = kxyz$ reste vraie sur $(\mathbb{R}_+)^3$ puisque $f(x; y; z)$ et $kxyz$ s'annulent dès que l'une au moins des variables est nulle.

Cas général. Faisons une preuve par récurrence sur $n \in \mathbb{N}^*$ pour démontrer la validité de la propriété $\mathcal{P}(n)$ définie comme suit : « Pour toute fonction f à n variables x_1, \dots, x_n définie sur \mathbb{R}_+ telle que chacune des fonctions $f_i : t \rightarrow f(x_1; \dots; x_{i-1}; t; x_{i+1}; \dots; x_n)$ ⁴, pour $i \in \llbracket 1; n \rrbracket$, soit linéaire en t , il existe une constante k telle que $f(x_1; \dots; x_n) = kx_1 \cdots x_n$ sur $(\mathbb{R}_+)^n$ ».

- *Cas de base.*

$\mathcal{P}(1)$ est clairement vraie.

3. L'auteur conseille vivement au lecteur que les questions de géométrie passionnent la lecture de l'ouvrage « Newton implique Kepler : méthodes géométriques élémentaires pour l'enseignement supérieur en mathématiques » de Benoît Rittaud aux éditions ellipses.

4. Désolé pour les abus évidents de notations pour $i = 1$ et $i = n$.

- *Hérédité.*

Supposons $\mathcal{P}(n)$ valide pour un naturel n fixé, mais quelconque, puis considérons une fonction f de $(n + 1)$ variables qui vérifie les conditions de la propriété $\mathcal{P}(n + 1)$.

Fixons $x \in \mathbb{R}_+$ et considérons la fonction $f_x : (x_1; \dots; x_n) \rightarrow f(x_1; \dots; x_n; x)$. Comme f_x vérifie les conditions de la propriété $\mathcal{P}(n)$, nous avons par hypothèse de récurrence $f_x(x_1; \dots; x_n) = a(x)x_1 \cdots x_n$, soit $f(x_1; \dots; x_n; x) = a(x)x_1 \cdots x_n$ où $a(x)$ est une constante dépendant du réel x .

D'autre part, par hypothèse $f(x_1; \dots; x_n; x) = b(x_1; \dots; x_n)x$ où $b(x_1; \dots; x_n)$ est une constante dépendant des réels x_1, \dots, x_n .

Donc sur $(\mathbb{R}_+^*)^{n+1}$, nous avons $a(x_{n+1})x_1 \cdots x_n = b(x_1; \dots; x_n)x_{n+1}$, puis ensuite $\frac{a(x_{n+1})}{x_{n+1}} = \frac{b(x_1; \dots; x_n)}{x_1 \cdots x_n} = k$ avec k une constante d'où $f(x_1; \dots; x_n; x_{n+1}) = kx_1 \cdots x_n x_{n+1}$ qui est valable sur $(\mathbb{R}_+^*)^{n+1}$. Cette formule s'étend à $(\mathbb{R}_+)^{n+1}$ puisque $kx_1 \cdots x_n x_{n+1}$ et $f(x_1; \dots; x_n; x_{n+1})$ s'annulent dès que l'une au moins des variables est nulle.

Nous avons bien déduit la validité de $\mathcal{P}(n + 1)$ à partir de celle de $\mathcal{P}(n)$.

- *Hérédité.*

Par récurrence sur $n \in \mathbb{N}^*$, la propriété $\mathcal{P}(n)$ est vraie pour tout naturel non nul n .

2. COÛT MARGINAL – POLYNÔMES DE DEGRÉ 2 OU 3

2.1. Une modélisation. À des fins prospectives, on souhaite modéliser le capital d'une entreprise qui fabrique et vend un produit. On utilise les notations suivantes.

- $C(x)$ est le coût de production en euros pour x unités produites.
- $C_m(x) = C(x) - C(x-1)$ est le coût marginal de production pour $x \in \mathbb{N}^*$ (*c'est le coût spécifique à la x^e unité produite*).
- $CM(x) = \frac{C(x)}{x}$ est le coût moyen de production pour $x \in \mathbb{N}^*$.

Cas 1. Supposons que des études statistiques aient montré que lorsque le nombre d'unités vendues augmente le coût marginal évolue de façon quasi linéaire.

Dans ce cas, on peut faire l'approximation $C_m(x) = ax + b$ avec $(a; b) \in (\mathbb{R}_+^*)^2$ (*cette fonction affine peut être celle obtenue via une régression linéaire par la méthode des moindres carrés*). Nous avons alors :

$$\begin{aligned} C(x) &= C(0) + \sum_{k=1}^x C_m(k) \\ &= C(0) + \sum_{k=1}^x (ak + b) \\ &= mx^2 + px + q \end{aligned}$$

Pour la dernière égalité, nous avons utilisé le fait que $\forall (e, x) \in \mathbb{N} \times \mathbb{N}^*$, $\sum_{k=1}^x k^e$ est un polynôme en x de degré $(e+1)$. En utilisant les formules exactes de ces polynômes, il est facile de vérifier que nos coefficients m , p et q sont tous positifs et même que $m > 0$.

Avec ce modèle, nous avons $CM(x) = mx + p + \frac{q}{x}$ pour $x \in \mathbb{N}^*$. Nous pouvons donner l'interprétation concrète suivante des termes mx , p et $\frac{q}{x}$.

- (1) p correspond à un coût fixe de fabrication : quelque soit le nombre d'unités vendues, l'entreprise doit déboursier en moyenne q euros par unité fabriquée (*penser à l'usage de bâtiments, aux salaires des employés en CDI...*).
- (2) mx augmente strictement avec x car $m > 0$. Ce terme correspond à un coût d'usage (*penser par exemple à l'entretien d'une chaîne de production, à l'énergie dépensée en plus pour produire plus...*).
- (3) $\frac{q}{x}$ indique un coût amorti en ce sens où plus x augmente moins $\frac{q}{x}$ influence $CM(x)$ (*penser par exemple à des investissement faits pour une nouvelle chaîne de production*).

Cas 2. Supposons que des études statistiques aient montré que lorsque le nombre d'unités vendues augmente le coût marginal diminue strictement puis qu'ensuite il atteint un minimum pour enfin ne cesser d'augmenter strictement. Notons que le minimum est forcément positif, et la valeur où il est atteint strictement positive. Nous supposons de plus que les données sont telles que la courbe de C_m soit quasi parabolique.

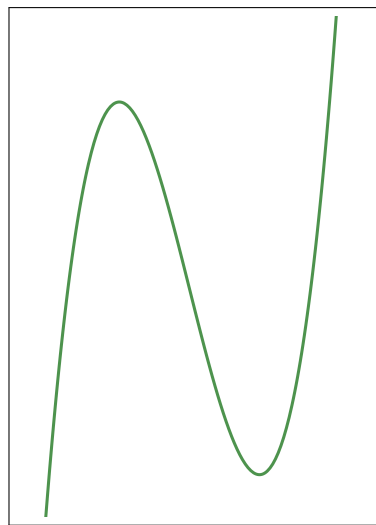
Dans ce cas, on peut faire l'approximation $C_m(x) = ax^2 + bx + c$ avec les contraintes $a > 0$, $\Delta \stackrel{\text{def}}{=} b^2 - 4ac \leq 0$ ainsi que $\frac{-b}{2a} > 0$ (*ce polynôme du 2^e degré peut être celui obtenu via une régression polynomiale par la méthode des moindres carrés*). Pourquoi ces conditions ?

- (1) $a > 0$ permet de vérifier les contraintes de variation.
- (2) $\Delta \leq 0$ et $\frac{-b}{2a} > 0$ servent à valider la contrainte du minimum.

De nouveau, nous avons :

$$\begin{aligned}
 C(x) &= C(0) + \sum_{k=1}^x C_m(k) \\
 &= C(0) + \sum_{k=1}^x (ak^2 + bk + c) \\
 &= mx^3 + px^2 + qx + r
 \end{aligned}$$

Si une étude statistique rapide sur le coût de fabrication C fournit une courbe similaire à celle ci-contre, on pourra espérer modéliser correctement C avec un polynôme de degré 3 comme ci-dessus. Dans ce cas, il faudra bien choisir certaines données concrètes afin de trouver de « bonnes » valeurs des coefficients m , p , q et r . C'est à l'usage que l'on pourra voir la pertinence ou non de ces choix.



Avec ce modèle, nous avons $CM(x) = mx^2 + px + q + \frac{r}{x}$ pour $x \in \mathbb{N}^*$. Bien que les significations concrètes de px , q et $\frac{r}{x}$ peuvent être similaires à celles données dans le cas 1, il devient difficile de donner un sens concret au terme mx^2 .

Remarque 2.1. Si le coût marginal C_m admet une « courbe en U », on pourra faire l'approximation $C_m(x) = c(x - \alpha)^{2p} + \beta$ avec $(c; \alpha; \beta) \in (\mathbb{R}_+)^3$ et $p \in \mathbb{N}^*$. Dans ce cas, $C(x)$ sera un polynôme de degré $(2p + 1)$. Est-ce plus efficace au final que de considérer une « courbe en U » comme quasi parabolique ? Ne pas oublier qu'un modèle doit être aisément utilisable quitte à être un peu approximatif. Il est donc parfois inutile de vouloir être trop précis.

Remarque 2.2. Plus généralement, on pourrait considérer d'autres types de fonction avec des variations collant à une étude statistique pour peu que le comportement du coût marginal C_m soit relativement régulier. Le problème que l'on aura à résoudre sera soit d'évaluer simplement $\sum_{k=1}^x C_m(k)$ pour une utilisation « humaine », soit de le faire efficacement de façon informatique.

Remarque 2.3. On justifie très souvent les deux modèles précédents via l'utilisation de l'approximation $C_m(x) \approx C'(x)$, puis ensuite on utilise les polynômes ainsi trouvés pour des applications numériques avec de faibles valeurs pour x . On voit que cette démarche est fragile⁵. Le piège est que cette approche fournit tout de même un polynôme de degré 2 ou 3 suivant les hypothèses de départ. Il pourrait être intéressant d'évaluer l'erreur commise par rapport à la méthodologie présentée dans ce document.

5. L'auteur, qui enseigne les mathématiques au lycée, doit admettre qu'il utilise aussi ce type de raisonnement dans des exercices, ce qui est pédagogiquement intéressant dans le cadre du calcul de dérivées. Par contre, les élèves sont avertis de l'arnaque faite en utilisant l'expression polynomiale pour de petites valeurs de x .

2.2. Le modèle est-il bon ? Il se trouve que la modélisation du coût de production par un polynôme de degré 2 ou 3 apparaît souvent dans des livres présentant des mathématiques élémentaires pour l'Économie. Cela montre que ces modèles ont leur utilité et ceci même si les raisonnements précédents sont fragiles car ils reposent sur des « *approximations de courbes* ». La technique d'approximation de formules, ou plus précisément d'interpolation, reste tout de même intéressante.

3. DÉSINTÉGRATION RADIOACTIVE – PROBABILITÉS ET LOIS CONTINUES, OU PAS...

3.1. Une modélisation. Les Sciences Physiques nous donnent la loi **[D]** suivante : « *la probabilité qu'à un instant t un noyau radioactif se désintègre dans l'intervalle $[t; t + s]$ ne dépend pas de son âge t et n'est jamais égale à un* ». Concrètement, ceci signifie que l'on suppose que les noyaux radioactifs ne vieillissent pas avant leur désintégration, et que de plus il est impossible qu'ils disparaissent majoritairement d'un seul coup. Il est aussi implicite qu'il est impossible qu'aucun noyau ne disparaisse sur l'intervalle $[0; t]$ pour $t > 0$, ceci revient à n'étudier que le phénomène de radioactivité.

Nous allons supposer de plus que la probabilité P de désintégration suit sur \mathbb{R}_+ une loi de densité continue f de sorte que $P([0; t]) = \int_0^t f(x) dx$ pour $t \geq 0$. Posant $F(t) = P([0; t])$ pour $t \geq 0$, nous avons alors les faits suivants.

- F est une primitive de f sur \mathbb{R}_+ puisque f est continue sur \mathbb{R}_+ par hypothèse.
- La loi **[D]** signifie que la probabilité que le noyau se désintègre dans l'intervalle $[t; t + s]$ sachant qu'il ne s'est pas désintégré dans l'intervalle $[0; t]$ est égale à $P([0; s]) = F(s)$ pour tout couple $(t; s) \in (\mathbb{R}_+)^2$.
- Pour $(t; s) \in (\mathbb{R}_+)^2$, la probabilité que le noyau se désintègre dans l'intervalle $[t; t + s]$ sachant qu'il ne s'est pas désintégré dans l'intervalle $[0; t]$ est $\frac{P([t; t+s])}{1-P([0; t])}$, c'est à dire $\frac{F(t+s)-F(t)}{1-F(t)}$ (notons qu'il n'y a aucun souci car par hypothèse $F(t) \neq 1$).
- Nous obtenons donc $\frac{F(t+s)-F(t)}{1-F(t)} = F(s)$ d'où $F(t+s) - F(t) = F(s)(1 - F(t))$ pour tout couple $(t; s) \in (\mathbb{R}_+)^2$.
- L'équation fonctionnelle vérifiée par F n'est pas très sympathique. Nous allons essayer d'en déduire une autre qui nous éclairera sur la marche à suivre. Il est naturel d'introduire $G(t) = 1 - F(t)$ la probabilité que le noyau ne se soit pas désintégré dans l'intervalle $[0; t]$. Nous avons alors :

$$\begin{aligned} F(t+s) - F(t) &= F(s)(1 - F(t)) \iff 1 - G(t+s) - (1 - G(t)) = (1 - G(s))G(t) \\ &\iff -G(t+s) + G(t) = G(t) - G(s)G(t) \\ &\iff G(t+s) = G(s)G(t) \end{aligned}$$

Nous obtenons donc que la fonction G dérivable sur \mathbb{R}_+ vérifie l'équation fonctionnelle $G(t+s) = G(s)G(t)$. Ceci nous donne alors $G(x) = e^{kx}$ avec $(k; x) \in \mathbb{R}^* \times \mathbb{R}_+$.

En effet, il suffit de fixer $t \in \mathbb{R}_+$ puis de dériver par rapport à s ce qui nous donne $G'(t+s) = G'(s)G(t)$. Comme G ne peut pas être constante⁶, on a : $k \stackrel{\text{déf}}{=} G'(0) \neq 0$. Nous avons alors pour $t \in \mathbb{R}_+$ quelconque $G'(t) = kG(t)$ soit une équation différentielle classique (il n'est pas gênant que nous soyons juste sur \mathbb{R}_+ au lieu de \mathbb{R} tout entier).

6. Sur \mathbb{R}_+^* , nous avons toujours $G(t) \neq 0$ et $G(t) \neq 1$, i.e. $F(t) \neq 1$ et $F(t) \neq 0$ par hypothèse car on n'exclut les cas de désintégrations probablement certaine et impossible.

- Sur \mathbb{R}_+ , $f(x) = F'(x) = -G'(x) = -ke^{kx}$. En fait nécessairement $k < 0$ car nous devons avoir $\int_0^{+\infty} f(x) dx = 1$. Faisons comme en Physique en n'utilisant que des paramètres positifs : notant $\lambda = -k > 0$, nous avons : $f(x) = \lambda e^{-\lambda x}$.
- Réciproquement si $f(x) = \lambda e^{-\lambda x}$ avec $\lambda > 0$ alors la probabilité P de désintégration de loi de densité continue f vérifie bien la loi **[D]** (*facile à vérifier en « remontant » certains des calculs faits ci-dessus*).

En résumé, si la probabilité P de désintégration vérifie la loi **[D]**, alors elle suit une loi de densité continue f si et seulement si sur \mathbb{R}_+ , $f(x) = \lambda e^{-\lambda x}$ avec $\lambda > 0$.

Remarque 3.1. Les physiciens appellent $\tau \stackrel{\text{déf}}{=} \frac{1}{\lambda}$ la constante de temps radioactive.

3.2. Le modèle est-il bon ? Vous aurez noté qu'en ajoutant juste une petite hypothèse, à savoir que la loi de probabilité de désintégration suit une loi de densité continue, nous arrivons à une modélisation robuste vis à vis de la loi physique **[D]**. Ceci étant dit, il serait bien de pouvoir se passer de cette hypothèse additionnelle dont la loi **[D]** ne parle pas. Est-ce possible ? L'annexe suivante répond positivement à cette question. Plus précisément, nous allons voir que l'hypothèse sur la loi de densité continue peut-être omise même si au final nous allons démontrer que la probabilité suit forcément une loi de densité continue. Nous avons donc pu construire un modèle très robuste au regard de la loi physique **[D]**. Ceci est très joli !

3.3. Annexe : le cas d'une probabilité quelconque. Posons $F(t) = P([0; t])$ ainsi que $G(t) = 1 - F(t)$ pour $t \geq 0$ comme précédemment. On sait juste que $0 < G \leq 1$ sur \mathbb{R}_+ , que $0 < G < 1$ sur \mathbb{R}_+^* , et que G est décroissante sur \mathbb{R}_+ (la monotonie découle de la croissance des probabilités).

Reprenant le raisonnement de la section sur la modélisation, sans passer par aucun calcul intégral, nous avons de nouveau que $G(t+s) = G(s)G(t)$ pour $(t; s) \in (\mathbb{R}_+)^2$. Démontrons alors l'existence de $\lambda \in \mathbb{R}_+^*$ tel que $G(x) = e^{-\lambda x}$.

- $G(0+0) = G(0)G(0)$ donne $G(0)(1 - G(0)) = 0$ soit $G(0) = 0$ ou $G(0) = 1$. Seul le second cas est possible.
- Pour tout réel $x \geq 0$, une récurrence facile montre que $G(nx) = G(x)^n$ pour tout naturel n .
- Dans la suite, nous poserons $\alpha = G(1) \in]0; 1[$. En particulier, $\forall n \in \mathbb{N}$, $G(n) = \alpha^n$.
- $\forall (p; q) \in \mathbb{N} \times \mathbb{N}^*$, $G(p) = G\left(q \times \frac{p}{q}\right)$, soit $\alpha^p = G\left(\frac{p}{q}\right)^q$. Comme de plus $G > 0$ sur \mathbb{R}_+ , on a : $G\left(\frac{p}{q}\right) = \sqrt[q]{\alpha^p} = \alpha^{\frac{p}{q}}$. Autrement dit, $\forall r \in \mathbb{Q}_+$, $G(r) = \alpha^r$.
- Soit enfin $x \in \mathbb{R}_+$. Par densité de \mathbb{Q} dans \mathbb{R} , il existe deux suites de rationnels positifs (r_n) et (R_n) qui convergent vers x et vérifient $\forall n \in \mathbb{N}$, $r_n \leq x \leq R_n$. Comme G décroît, $\forall n \in \mathbb{N}$, $G(r_n) \geq G(x) \geq G(R_n)$, soit $\alpha^{r_n} \geq G(x) \geq \alpha^{R_n}$. Par continuité de la fonction $t \rightarrow \alpha^t$ sur \mathbb{R}_+ , un passage à la limite donne $G(x) = \alpha^x$.
- Il ne reste plus qu'à poser $\lambda = -\ln \alpha > 0$ de sorte que $G(t) = e^{-\lambda t}$ sur \mathbb{R}_+ .

Finalement, $F(t) = 1 - G(t) = 1 - e^{-\lambda t} = \int_0^t \lambda e^{-\lambda x} dx$ prouve que la probabilité P suit forcément la loi de densité continue $f(x) = \lambda e^{-\lambda x}$.

4. DANIEL BERNOULLI ET LA VARIOLE – ÉQUATIONS DIFFÉRENTIELLES NON LINÉAIRES

4.1. **Une modélisation.** Daniel Bernoulli⁷ (1700 – 1782) était un médecin, physicien et mathématicien suisse. Il présente en 1760 à l'Académie des sciences de Paris un mémoire intitulé « *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole⁸ et des avantages de l'inoculation pour la prévenir* »⁹. Nous allons présenter son raisonnement.

Propagation de la variole. Pour modéliser ceci, nous aurons besoin des notations suivantes où l'unité de temps sera l'année.

- q_I désigne la probabilité d'être infecté par la variole pendant une année.
- q_V désigne la probabilité de mourir lorsque l'on attrape la variole pour la 1^{re} fois. Une personne ne mourant pas pendant une année sera considérée comme immunisée contre la variole.
- $m(t)$ désigne, à la fin de l'année t , le taux de mortalité « normale » sans tenir compte d'une épidémie de variole (*l'année 0 est celle du début de l'étude réelle*).

Les notations précédentes sous-entendent que les quantités nommées sont constantes. C'est une hypothèse de modélisation ! Continuons avec les fonctions suivantes où t désigne une année.

- $S(t)$ est le nombre de personnes saines n'ayant pas la variole à la fin de l'année t .
- $I(t)$ est le nombre de personnes qui se sont immunisées contre la variole au cours de l'année t . Ces personnes sont celles qui ont attrapé la variole pendant l'année t sans mourir.
- $P(t) = S(t) + I(t)$ est le nombre personnes toujours vivantes à la fin de l'année t .

Faisons le bilan entre l'année t et l'année $(t + 1)$.

- $S(t + 1) = S(t) - q_I S(t) - m(t) S(t)$
 - $q_I S(t)$ est le nombre de personnes mortes de la variole.
 - $m(t) S(t)$ est le nombre des personnes saines mortes de cause « normale ».
- $I(t + 1) = I(t) + (1 - q_V) q_I S(t) - m(t) I(t)$
 - Parmi les $q_I S(t)$ nouvelles personnes infectées, il y en a juste $(1 - q_V) q_I S(t)$ qui ne meurent pas de la variole.
 - $m(t) I(t)$ est le nombre d'infectés morts d'autre chose que de la variole.
- Nous en déduisons :

$$P(t + 1) = P(t) - q_V q_I S(t) - m(t) P(t)$$

Notant $\Delta f_b^a = \frac{f(a) - f(b)}{a - b}$ le taux de variation de la fonction f entre a et b distincts, nous avons :

- $\Delta S_{t+1}^t = -q_I S(t) - m(t) S(t)$
- $\Delta P_{t+1}^t = -q_V q_I S(t) - m(t) P(t)$

7. Il était le fils de Jean Bernoulli (1667-1748) et le neveu de Jacques Bernoulli (1654-1705) qui étaient tous les deux des mathématiciens et physiciens suisses.

8. La petite vérole était le nom que l'on donnait à la variole.

9. Pour les aspects historiques, se reporter au livre « *Histoires de mathématiques et de populations* » de Nicolas Bacaër aux éditions CASSINI.

Il n'est pas aisé d'étudier des systèmes d'équations sur des suites comme celui ci-dessus. Nous allons donc émettre de nouvelles hypothèses de modélisation pour aboutir à un modèle utilisant des équations différentielles. L'idée simple, mais peut-être un peu trop, est de dire que le raisonnement fait sur une année pourrait être fait de la même façon sur une portion plus petite d'années, une portion aussi petite que nécessaire. Dès lors, en considérant S et P définies sur \mathbb{R}_+ , et non juste sur \mathbb{N} , pour tout couple $(t; \delta t) \in \mathbb{R}_+ \times \mathbb{R}$ tel que $t + \delta t \geq 0$, nous avons :

- $\Delta S_{t+\delta t}^t = -q_I S(t) - m(t)S(t)$
- $\Delta P_{t+\delta t}^t = -q_V q_I S(t) - m(t)P(t)$

Il ne reste plus qu'à passer un cap supplémentaire en supposant S et P dérivables sur \mathbb{R}_+ . Dès lors par passage à la limite via $\delta t \rightarrow 0$, ou juste $\delta t \rightarrow 0^+$ pour $t = 0$, on obtient¹⁰ :

- $S'(t) = -q_I S(t) - m(t)S(t)$
- $P'(t) = -q_V q_I S(t) - m(t)P(t)$

Avec ces nouveaux choix de modélisation, notre sac d'hypothèses s'alourdit !

En résumé, nous devons trouver deux fonctions S et P dérivables sur \mathbb{R}_+ vérifiant les deux équations précédentes. A priori, ce n'est pas simple comme problème : ce qui crée une petite difficulté c'est la présence de $m(t)$ dont on ne sait rien. Nous décidons de chercher à obtenir des informations sur $R(t) = \frac{S(t)}{P(t)}$ le taux des personnes saines (*en annexe de cette section sont données deux autres méthodes de résolution du système ci-dessus dont celle utilisée par Daniel Bernoulli*). Au passage, nous avons besoin de supposer, de façon non abusive, que S et P ne s'annulent jamais.

$$\begin{aligned} R'(t) &= \frac{1}{P^2(t)} (S'(t)P(t) - S(t)P'(t)) \\ &= \frac{1}{P^2(t)} ([-q_I S(t) - m(t)S(t)] P(t) - S(t) [-q_V q_I S(t) - m(t)P(t)]) \\ &= \frac{1}{P^2(t)} (-q_I S(t)P(t) + q_V q_I S^2(t)) \\ &= -q_I R(t) + q_V q_I R^2(t) \end{aligned}$$

Ce qui est beau, c'est que l'on tombe sur une équation différentielle peu agressive a priori : $R'(t) = -q_I R(t) + q_V q_I R^2(t)$. Quiconque connaissant les formules générales de dérivation trouvent qu'il suffit de raisonner comme suit en se souvenant que S et P ne s'annulent jamais et donc R non plus.

$$\begin{aligned} R'(t) = -q_I R(t) + q_V q_I R^2(t) &\iff -\frac{R'(t)}{R^2(t)} = \frac{q_I}{R(t)} - q_V q_I \\ &\iff f'(t) = q_I f(t) - q_V q_I \text{ où on a posé } f(t) = \frac{1}{R(t)}. \end{aligned}$$

Nous aboutissons à une simple équation différentielle linéaire du 1^{er} ordre et il est connu que nécessairement $f(t) = q_V + k e^{q_I t}$ où $k \in \mathbb{R}$ est une constante dépendant des conditions réelles¹¹. Nous avons finalement : $\frac{S(t)}{P(t)} = \frac{1}{q_V + k e^{q_I t}}$. Tout ceci a été obtenu via de bien jolis calculs...

10. Peut-être que le lecteur aura noté une grosse arnaque ici. Nous en discuterons dans la section suivante.

11. Notre équation différentielle est juste sur \mathbb{R}_+ mais cela n'est pas gênant ici.

Remarque 4.1. L'équation différentielle $R'(t) = -q_I R(t) + q_V q_I R^2(t)$ est une équation du type $y' = ay(1 - by)$ avec $a \neq 0$ et $b \neq 0$. De telles équations sont dites « logistiques »¹². Elles se résolvent, pour les solutions ne s'annulant pas, via le changement de fonctions $f = \frac{1}{y}$.

Utilité de l'inoculation pour vacciner. L'idée première de Daniel Bernoulli est de supposer que l'inoculation de la variole à tous les jeunes enfants les rend tous immunisés sans les tuer¹³. Sous cette hypothèse, il faut étudier $Q(t) = \frac{P(t)}{N(t)}$ où $N(t)$ désigne le nombre, non nul par hypothèse, de personnes à la fin de l'année t s'il n'y a aucun cas de variole. En raisonnant comme ci-dessus, nous avons $\Delta N_{t+1}^t = -m(t)N(t)$ puis $N'(t) = -m(t)N(t)$ en supposant N dérivable sur \mathbb{R}_+ .

$$\begin{aligned} Q'(t) &= \frac{1}{N^2(t)} (P'(t)N(t) - P(t)N'(t)) \\ &= \frac{1}{N^2(t)} ([-q_V q_I S(t) - m(t)P(t)]N(t) + m(t)P(t)N(t)) \\ &= -q_V q_I \frac{P(t)}{N(t)} \cdot \frac{S(t)}{P(t)} \\ &= -q_V q_I Q(t) \frac{e^{-q_I t}}{q_V e^{-q_I t} + k} \text{ car } \frac{S(t)}{P(t)} = \frac{1}{q_V + k e^{q_I t}} = \frac{e^{-q_I t}}{q_V e^{-q_I t} + k} \end{aligned}$$

Comme $\frac{Q'(t)}{Q(t)} = \frac{-q_I q_V e^{-q_I t}}{q_V e^{-q_I t} + k}$, la dérivation logarithmique nous donne $Q(t) = K(q_V e^{-q_I t} + k)$ avec $K \in \mathbb{R}^*$ une constante. Ceci permet d'évaluer l'efficacité de l'inoculation de la variole en utilisant les données du terrain pour déterminer les différents paramètres.

4.2. Le modèle est-il bon ? Revenons sur la toute « petite » liste d'hypothèses de modélisation que nous avons faites.

- (1) La probabilité q_I d'être infecté dans l'année, et celle q_V de mourir dans l'année une fois infecté sont supposées indépendantes de l'année t considérée.
- (2) On a supposé S , P et N définies et dérivables sur \mathbb{R}_+ tout entier.
- (3) Enfin nous avons besoin que S , P et N ne s'annulent jamais.

Hypothèse 1. Ceci sous-entend un type particulier de propagation. Par exemple, cette première hypothèse reste-t-elle valable si les personnes malades sont mises en quarantaine ? Non. C'est ce qui permet d'éradiquer des maladies très virulentes. Une autre critique : plus il y a de malades contagieux, plus grande devient la probabilité de tomber malade. Du point de vue de la modélisation, pour une maladie « pas trop violente » et en début d'épidémie, on peut tout de même accepter la première hypothèse¹⁴.

12. Pierre François Verhulst (1804 – 1849) a proposé l'équation différentielle $y' = ay(1 - by)$ pour modéliser une évolution de population à croissance non exponentielle. C'est à lui que l'on doit l'appellation « logistique » sans en expliquer la raison.

13. Dans le livre « *Histoires de mathématiques et de populations* », Nicolas Bacaër explique la suite du raisonnement consistant à prendre en compte les effets délétères de l'inoculation dans certains cas.

14. Des règles de quarantaines ne sont pas justifiables politiquement pour une maladie « pas trop violente » même si celle-ci peut être mortelle dans certains cas.

Hypothèse 2. Le passage de fonctions définies sur \mathbb{N} à des fonctions définies sur \mathbb{R}_+ ne pose pas de difficultés conceptuelle et concrète.

Ensuite arrive un grand classique des sciences du réel : les fonctions concrètes étudiées sont supposées suffisamment régulières. Pourquoi fait-on ceci ? Comme nous l'avons déjà indiqué, ceci permet de faire appel à des outils mathématiques puissants du calcul différentiel car l'on dispose de moins d'outils efficaces pour étudier les suites.

Ceci étant dit, supposer la dérivabilité sur \mathbb{R}_+ tout entier nous amène à étudier des fonctions au comportement très lisse. Ainsi les solutions des équations différentielles, dites logistiques, $y' = ay(1-by)$ sont très régulières, tandis que l'équation logistique discrète $u_{n+1} = au_n(1-bu_n)$ peut produire des suites au comportement chaotique.

De plus, comment passer d'une étude entre t et $t + \delta t$ avec δt aussi petit que nécessaire, au cas où δt tend vers 0 ? Rien ne justifie ceci sérieusement d'un point de vue concret. Nous sommes là face à un choix très fort de modélisation.

L'hypothèse 2 est donc lourde de conséquences... Ceci étant dit, cela reste un classique de la modélisation et l'histoire des sciences du réel prouve que ce type d'hypothèse est féconde à produire des modèles utiles.

Hypothèse 3. Ceci ne pose aucun souci sauf à vouloir étudier une population sans aucune personne saine, ou bien sans personne vivante.

Conclusion. Bien que nous ayons fait un très joli raisonnement mathématique, nous venons de voir que concrètement il y a de réelles failles de modélisation mais comme toujours en modélisation, c'est l'affrontement du modèle et des vraies données qui servira de juge d'utilité.

4.3. Annexe : résoudre autrement le système d'équations différentielles. Nous allons donner deux autres méthodes de résolution¹⁵ du système [EDV] suivant où les fonctions S et P sont dérivables sur \mathbb{R}_+ et ne s'annulent jamais.

$$[\text{EDV}] : \begin{cases} S'(t) = -q_I S(t) - m(t)S(t) \\ P'(t) = -q_V q_I S(t) - m(t)P(t) \end{cases}$$

Méthode de Daniel Bernoulli. Le médecin suisse raisonne comme suit.

$$\begin{cases} -m(t) = \frac{S'(t) + q_I S(t)}{S(t)} \\ -m(t) = \frac{P'(t) + q_V q_I S(t)}{P(t)} \end{cases}$$

$$(S'(t) + q_I S(t))P(t) = (P'(t) + q_V q_I S(t))S(t)$$

$$S'(t)P(t) - P'(t)S(t) = -q_I S(t)P(t) + q_V q_I S^2(t)$$

$$R'(t) = -q_I R(t) + q_V q_I R^2(t) \text{ en notant } R(t) = \frac{S(t)}{P(t)} \dots$$

15. Ces méthodes sont présentes dans le livre « *Histoires de mathématiques et de populations* » de Nicolas Bacaër.

Méthode via la dérivation logarithmique. Supposons la fonction m intégrable sur \mathbb{R}_+ ¹⁶. Les solutions de $S'(t) = -q_I S(t) - m(t)S(t)$ sont du type $S(t) = c_1 \exp\left(-q_I t - \int_0^t m(x) dx\right)$ où $c_1 \in \mathbb{R}$ est une constante (*ceci se démontre via la dérivée logarithmique de S sans aucune connaissance préalable sur les équations différentielles linéaires*). Concrètement, nous savons que $c_1 > 0$.

De même, les solutions de $P'(t) = -m(t)P(t)$ sont de la forme $P(t) = c_2 \exp\left(-\int_0^t m(x) dx\right)$ où $c_2 \in \mathbb{R}$ est une constante, tandis que $P'(t) = -q_V q_I S(t) - m(t)P(t)$ admet la solution particulière $P(t) = q_V S(t)$. Dès lors, les solutions de $P'(t) = -q_V q_I S(t) - m(t)P(t)$ sont du type $P(t) = c_2 \exp\left(-\int_0^t m(x) dx\right) + q_V S(t)$. Comme $\exp\left(-\int_0^t m(x) dx\right) = \frac{1}{c_1} e^{q_I t} S(t)$ nous obtenons :

$$P(t) = \frac{c_2}{c_1} e^{q_I t} S(t) + q_V S(t)$$

$$\frac{P(t)}{S(t)} = k e^{q_I t} + q_V \text{ avec } k = \frac{c_2}{c_1} \text{ une constante réelle}$$

$$R(t) = \frac{S(t)}{P(t)} = \frac{1}{q_V + k e^{q_I t}} \dots$$

5. AFFAIRE À SUIVRE...

16. L'ensemble des fonctions intégrables sur \mathbb{R}_+ ne se limite pas à celui des fonctions continues. On y trouve des fonctions plus ou moins complexes suivant le type de calcul intégral que l'on s'autorise. Le lecteur intéressé par le sujet pourra se reporter à l'excellent livre « *Intégration, de Riemann à Kurzweil et Henstock - La construction progressive des théories modernes de l'intégrale* » de Laurent Moonens chez ellipses.