



## Lettre n° 2

30 avril 2020

*Les modèles probabilistes, S I R, S E I R et son extension S E I R D I M.*

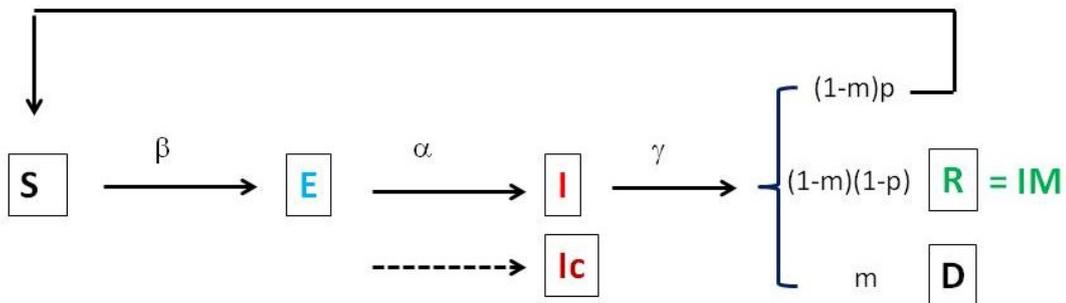
*Equations dynamiques et effet des différents paramètres.*

*Exemple de résolution au tableur EXCEL. Résultats pour les pays : France, Allemagne, Italie, Espagne, Suède, Royaume-Uni, USA, Brésil.*

Bonjour, ou bonsoir, si vous êtes à l'autre bout du monde.

Nous entamons les choses sérieuses avec le modèle probabiliste découlant de l'hypothèse simplificatrice d'une évolution homogène de la maladie dans un territoire fermé, par exemple un territoire national dont les frontières sont fermées. Une telle hypothèse est fréquente et efficace pour des phénomènes collectifs tels que les transitions de phase (bien que les interactions soient le plus souvent à courte portée). Dans cette description homogène, on décrit les populations en utilisant des quantités fractionnaires. Le premier modèle appelé S.I.R., datant des années 1930, considérait trois fractions : la fraction saine (S), la fraction infectée (I) et la fraction guérie (R pouvant signifier « retirés » aussi bien que « guéris » (recovered). Nous avons introduit une quatrième fraction (D), « décédés » non par attrait morbide mais parce que les chiffres des décès sont bien connus et, comme on le verra, contiennent une information utile pour bien conduire l'analyse.

D'autre part l'existence de « porteurs sains » (asymptomatiques) avait déjà conduit les épidémiologistes à introduire une fraction supplémentaire (E) « exposés ». Le parcours de la maladie est alors  $S \rightarrow E \rightarrow I \rightarrow R$  ou D. Plus récemment, s'est posée la question de l'immunité acquise, avec des récives possibles, qui ont pour effet de renvoyer une fraction de la population guérie vers le réservoir de personnes saines. Nous avons fait le schéma de ce parcours complet, avec les coefficients qui définissent la cinétique de chaque passage, et des coefficients (probabilités !) pour l'embranchement  $I \rightarrow R$  ou D ou S. Pour mémoire le nom du modèle devient alors SEIRDIm.



S Susceptible (sain)  
 E Exposed  
 I Infectious  
 Ic cumul infectés  
 R Retired (immunisé)  
 D Dead

$\beta$  = taux de contamination  
 $\alpha$  = 1/ temps d'incubation  
 $\gamma$  = 1/ temps de guérison  
 m = taux de mortalité  
 p = taux d'immunisation

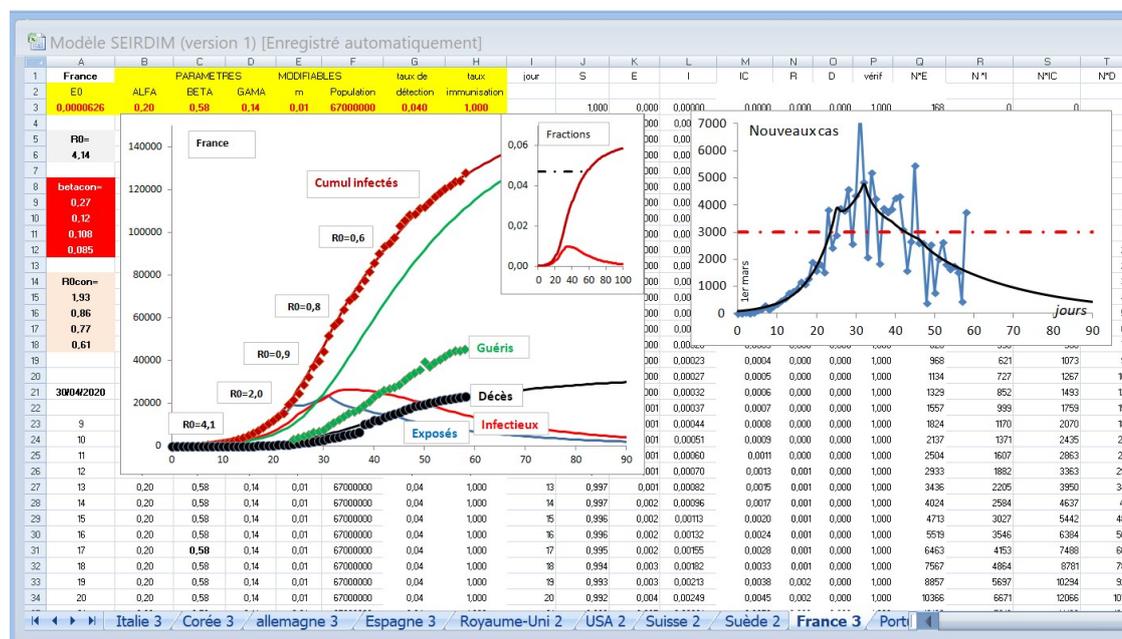
$R_0 = \beta/\gamma$  est le taux de reproduction = le nombre de personnes contaminées par personne infectée au départ de l'épidémie (quand la croissance est exponentielle).

Dans ce schéma nous avons introduit une catégorie « cumul infectés » (Ic), qui n'est pas impliquée dans le calcul, mais dérivée de ses résultats. En effet, les chiffres des sites officiels comptabilisent sous l'appellation « total confirmed cases » ou « total cases » le nombre des personnes qui sont entrées dans la catégorie I. Mais il leur est plus difficile de comptabiliser celles qui en sortent, parce que le décompte des « guéris » pose des problèmes (notamment de délai pour la certification). Notons enfin que la fraction Ic ne peut qu'augmenter avec le temps, comme celle des décès et des immunisés, alors que la fraction I décroît dès que le « pic » est passé. L'évolution dans le temps de ces différentes quantités fractionnaires est régie par les équations différentielles couplées :

$$dS/dt = -\beta S I + (1-m)(1-p) \gamma I ; dE/dt = \beta S I - \alpha E ; dI/dt = \alpha E - \gamma I ; dR/dt = (1-m) p \gamma I ; dD/dt = m \gamma I$$

Ajoutons-y  $dIc/dt = \alpha E$  ; la somme de ces dérivées est nulle, avec  $S + E + I + R + D = 1$ .

*Remarque :* Au début de l'épidémie, la croissance de  $E$ ,  $I$ ,  $Ic$  est quasi-exponentielle, et  $R_0$  prend des valeurs élevées, 3, 4 ou plus dans la plupart des pays. Pour la France, la valeur observée, voisine de 4, correspond à un doublement du nombre d'infectés en presque 4 jours ; la croissance est d'un facteur 3 par semaine. On évalue donc la conséquence d'un retard d'une semaine dans la décision de confinement dû au maintien du premier tour des élections municipales le 15 mars dernier entre autres.



Quelques mots maintenant sur le travail avec le tableur EXCEL : la capture d'écran jointe à ce texte vous donne un aperçu du travail de modélisation. Vous en trouverez d'autres dans les pages suivantes. Les cases colorées contiennent les valeurs des paramètres, dont certaines peuvent être ajustées de manière à obtenir le meilleur accord entre les courbes calculées et les données chiffrées de l'épidémie. Ces courbes se mettent à jour automatiquement à chaque modification de paramètres, c'est donc un ajustement « en temps réel ». Les paramètres du calcul sont :

$E_0$  = valeur de la fraction exposée, à l'origine des temps. Il faut la prendre non nulle pour que l'épidémie démarre. Plus cette valeur est élevée, plus tôt se fait le démarrage. Modifier la valeur de ce paramètre équivaut à changer la date de début de l'épidémie.

$\beta$ , « taux de contamination » est la probabilité par unité de temps (ici, le jour) qu'une personne saine soit contaminée. Ce paramètre traduit l'intensité (ou l'efficacité) des contacts sociaux et peut varier au cours du temps en fonction de circonstances (confinement, déconfinement, restrictions diverses). Pour le meilleur ajustement aux données nous l'avons fait varier par paliers, à intervalles d'une semaine. Ces changements sont responsables des

points anguleux sur la courbe de **E**. Les effets sur **I** et **Ic** sont nettement amortis en raison du délai d'incubation, et il ne nous a donc pas paru nécessaire de prendre des paliers plus courts.

$\alpha = 1/\text{temps d'incubation que nous avons pris égal à 5 jours}$

$\gamma = 1/\text{temps de guérison que nous avons pris égal à 7 jours}$

$m$  = taux de mortalité est la probabilité qu'une personne infectieuse (donc contaminée) décède. Nous l'avons fixé à 0,01, valeur qui semble généralement admise. Les données de mortalité posent le problème du taux de détection des personnes contaminées, qui est très loin de 1, et que nous allons détailler plus loin. Une estimation récente (20 avril) de l'Institut Pasteur prévoit pour début mai dans la population française près de 6% de personnes contaminées. Cela fait près de 4 millions de personnes ! Avec un nombre des décès à ce jour d'environ 20 000 (mais sont-ils tous recensés ?), cela ramènerait le taux de mortalité à 0,5 %.

« Population » est le nombre total d'habitants. Il est donc fixé, et vient multiplier toutes les fractions afin d'aboutir aux populations calculées de chaque catégorie.

« Taux de détection » est le facteur d'échelle qu'il a été nécessaire d'introduire pour obtenir l'accord final entre les données expérimentales et les courbes calculées. Il vient multiplier toutes les populations calculées (**E**, **I**, **Ic**, **R**) mais pas les décès qui sont supposés tous recensés. Une valeur inférieure à 1 traduit une insuffisance de la détection des personnes contaminées. Sa valeur remarquablement faible dans le cas de la France est attribuée à l'existence de nombreux porteurs asymptomatiques et/ou de malades pas assez atteints pour trouver utile de consulter le médecin. Notons que, dans le cas où le taux de détection des décès serait notablement inférieur à 1, il faudrait réduire proportionnellement la valeur du taux de mortalité, sans rien changer d'autre.

Rappelons, enfin, que les chiffres utilisés sont extraits de la base de données de l'Université Johns Hopkins, accessible sur le site <https://www.eficiens.com/coronavirus-statistiques/>.

Sur les pages suivantes, nous vous donnons un aperçu de la situation dans divers pays. Vous remarquerez que les données sur le nombre de guéris – quand elles existent – sont systématiquement en retard par rapport aux courbes calculées. En encart, nous avons repris les données essentielles **I**, **Ic**, exprimées sous leur forme fractionnaire. C'est là que l'on peut comparer le degré de contamination de chaque pays.

Le cas de la Suède, où le confinement n'a pas été obligatoire, est particulièrement intéressant, avec une épidémie qui progresse vers l'état d'immunité de groupes, apparemment sans risquer de saturer le système hospitalier.

La question de l'immunité de groupe fera l'objet de notre prochaine lettre. Nous aborderons l'étude de données régionales le rendu des tests massifs dès que les chiffres seront disponibles.

A bientôt, protégez-vous, et continuez à vous entraîner au port du masque et aux gestes de prévention !

François Varret, professeur émérite Université de Versailles Saint Quentin

Mathide Varret, chargée de recherche INSERM Hôpital Bichat

