

INSTITUT DE FORMATION ET DE RECHERCHE DÉMOGRAPHIQUES

MOUVEMENTS ET MODÈLES DE POPULATION

Nicolas BROUARD

I.Fo.R.D Yaoundé

Cameroun

Décembre 1985

Ce cours a pour but d'apporter aux étudiants une connaissance théorique des principaux modèles de dynamique des populations. Les démographes utilisent principalement les modèles de population stable ou quasi-stable, mais il est important de connaître les premiers modèles théoriques, de Lotka en particulier, sur les compétitions entre espèces (modèle proie-prédateur et modèle de pêche). Un chapitre est donc consacré aux modèles de population généraux où seules les populations totales, sans considération de classe d'âge, interviennent.

Pour traiter des populations à classes d'âge stables, il est nécessaire de revoir les deux composantes que sont mortalité et fécondité par âge. Un chapitre est alors consacré d'une part à une présentation de la table de mortalité et à son associé le modèle de population stationnaire, d'autre part à la fécondité par âge. Ces chapitres sont très descriptifs et montrent concrètement sur une très longue période (la France depuis la fin du 18<sup>e</sup> siècle) l'évolution des indices classiques dégagés par la théorie. Enfin il permet une confrontation des analyses longitudinales et transversales.

Avant d'aborder les modèles de population stable un chapitre est consacré à la modélisation de la mortalité et de la fécondité par âge.

Le chapitre sur les populations stables ne traite que la théorie de Lotka et des problèmes de convergence (forte et faible), laissant au cours de statistiques imparfaites le soin d'en étudier les applications pratiques.

# Chapitre 1

## Mesures des variations d'une population

### 1.1 Diagramme de Lexis <sup>point</sup>

Un diagramme de Lexis [31] est un système de coordonnées où l'on trouve en abscisse le temps et en ordonnée l'âge. Une *ligne de vie* propre à un individu est une portion de droite, parallèle à la première bissectrice (puisque lorsque le temps passe, l'âge augmente de ce même temps passé), dont l'origine commence à la date d'apparition de l'individu dans la population (généralement la naissance) et l'extrémité à la date de sortie (généralement la mort).

D'une manière plus générale, on peut noter sur ces lignes de vie d'autres événements démographiques comme les mariages, naissances des enfants, migrations, divorces, remariages, etc. Mais, dans le cas de la mortalité par exemple, on ne s'intéressera qu'aux décès et à la population survivante ; dans le cas de la fécondité, qu'aux naissances, âges des pères et mères, et population en âge de procréer, etc.

Une analyse démographique peut être restreinte à une période de l'histoire et à certains âges seulement ; plus généralement on peut parler d'un *domaine* du diagramme de Lexis (voir figure 1.1).

Si on ne s'intéresse, par exemple, qu'aux événements qui modifient la taille d'une population entre deux dates  $t$  et  $t + h$ , le domaine a la forme d'un  $\square$ , et les événements qui peuvent survenir sont tous regroupés sur la figure 1.2.

L'un des buts de la statistique et plus particulièrement de la démographie est de calculer des *risques* et d'analyser les variations de ces risques selon divers critères comme l'âge, le sexe, ou plus généralement toute variable qui pourrait contribuer à expliquer les comportements de l'homme. Avant d'aborder la notion de risque, nous allons définir un outil intermédiaire mais fondamental : un *taux*.

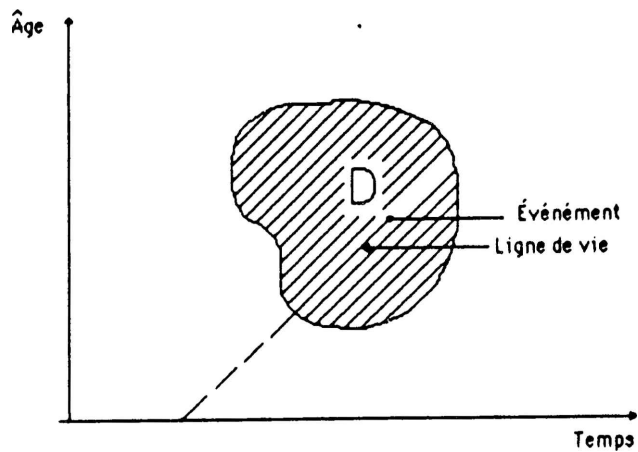


Figure 1.1: Définition d'un taux sur un domaine quelconque du diagramme de Lexis.

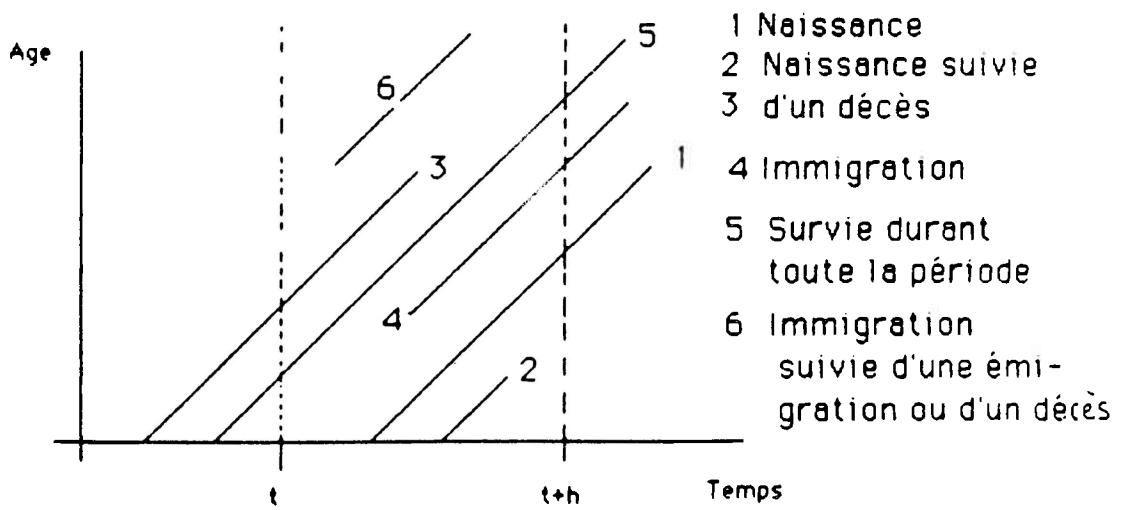


Figure 1.2: Diagramme de Lexis montrant les différents événements démographiques qui concourent à l'accroissement d'une population.



## 1.2 Le taux

**Définition 1** On définit un *taux*  $ME(D)$  relatif à un type d'événement démographique  $E$ , et se rapportant à un domaine  $D$ , comme le rapport du nombre d'événements du type  $E$  survenus dans le domaine  $D$ , au nombre d'unités de temps vécues par l'ensemble de la population concernée par ce même domaine.

Dans le cas de la mortalité, par exemple, les événements sont des décès, et les taux correspondants sont appelés *taux de mortalité*.

Comme en mécanique, où une *vitesse* est une quantité homogène au rapport d'une longueur (L) par un temps (T), un *taux* en démographie n'est pas sans dimension. Un *taux* est le rapport d'un *effectif* au produit d'un effectif et d'un temps, et est donc homogène à l'*inverse d'un temps* ( $\frac{1}{T}$ ). On doit donc exprimer l'unité de temps d'un *taux* ; lorsque celle-ci est sous-entendue, c'est que l'unité de temps est l'année.

densité de population.  
passage d'un d

## 1.3 Taux d'accroissement d'une population

Limitons nous dans ce paragraphe aux définitions du *taux d'accroissement* d'une population entre deux dates  $t$  et  $t + h$ . Si nous pouvions représenter chacun des individus présents à un moment quelconque de la période par une ligne de vie qui lui est propre, le nombre total d'années vécues par cette population serait la somme des longueurs des portions de ligne de vie propre à chaque individu. Ainsi, une personne présente dans la population durant toute la période compterait pour  $h$  années, une autre qui décéderait ou naîtrait en milieu de période pour  $\frac{h}{2}$ , etc.

La *taille* de la population à une date,  $\tau$ , quelconque, est égale au nombre des lignes de vie qui traversent la ligne verticale d'équation,  $t = \tau$ . Notons  $P(\tau)$  cette *taille* ; le nombre total d'années vécues par la population durant la période est aussi égal à :

$$\int_t^{t+h} P(\tau) d\tau,$$

c'est-à-dire à l'aire située sous la courbe  $P(\tau)$  comme l'indique la figure 1.3.

Suivant la définition exacte d'un *taux*, le numérateur<sup>1</sup> c'est à dire le nombre

<sup>1</sup>Notons que les événements du type *immigration* suivi d'*émigration* durant la période, concourent, en toute rigueur, à l'accroissement du dénominateur sans compensation au numérateur.

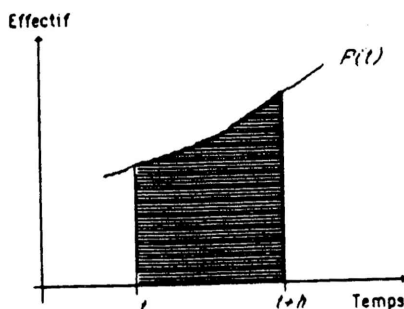


Figure 1.3: Le nombre d'années-vécues par l'ensemble de la population entre les deux dates est aussi égale à la surface hachurée.

des événements qui concourent, durant la période, à l'accroissement<sup>2</sup> de la population est égal à la variation de la population entre les deux dates,  $P(t+h) - P(t)$ .

Nous avons ainsi la définition suivante :

**Définition 2** Le taux d'accroissement,  $r$ , d'une population entre deux dates,  $t$  et  $t+h$ , est égal à :

$$r = \frac{P(t+h) - P(t)}{\int_t^{t+h} P(\tau) d\tau} \quad (1.1)$$

En introduisant la valeur moyenne de la population sur la période  $[t, t+h]$  qui vaut par définition (voir la figure 1.3) :

$$\bar{P}_{[t,t+h]} = \frac{1}{h} \int_t^{t+h} P(\tau) d\tau,$$

on peut aussi écrire :

$$r = \frac{P(t+h) - P(t)}{h \times \bar{P}_{[t,t+h]}} \quad (1.2)$$

**Calculs pratiques** Ces présentations mathématiques ont pour objet de justifier les pratiques courantes de calcul de taux d'accroissement, et en particulier celui d'une population entre deux recensements. On prend alors, comme approximation de la valeur moyenne de la population entre les deux dates, le produit de la durée de la période intercensitaire par la moyenne arithmétique des populations aux deux recensements.

<sup>2</sup> celui-ci peut être négatif.

**Exemple**

$$P_0 = 9\,900 \quad \text{au 1er Janvier 1973}$$

$$P_1 = 9\,950 \quad \text{au 31 Mars 1983}$$

$$\text{d'où } r = \frac{9\,950 - 9\,900}{\frac{(9\,900 + 9\,950)}{2} \times \left(\frac{3}{12} + 10\right)} = 0,049\% \quad \text{par an.}$$

**1.3.1 Taux d'accroissement instantané**

Soit  $P(t)$  la taille d'une population à l'instant  $t$ , et  $P(t + \Delta t)$  celle à l'instant  $t + \Delta t$ . Selon la définition 2, le taux d'accroissement durant la période  $[t, t + \Delta t]$  est :

$$\rho(t, t + \Delta t) = \frac{P(t + \Delta t) - P(t)}{\int_t^{t + \Delta t} P(\tau) d\tau}$$

En faisant tendre  $\Delta t$  vers 0, on obtient la définition d'un taux instantané :

**Définition 3** On définit le taux de croissance instantané  $\rho_t$  de la population à l'instant  $t$ , comme la limite, si elle existe, du taux de croissance  $\rho_{[t, t+h]}$  lorsque  $h$  tend vers zéro :

$$\rho_t = \lim_{\Delta t \rightarrow 0} \rho_{[t, t + \Delta t]} \quad (1.3)$$

Comme  $\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_t^{t + \Delta t} P(\tau) d\tau = P(t)$ , nous pouvons affirmer que si la fonction  $P(t)$  est dérivable, de dérivée  $P'(t)$ , le taux instantané existe et vaut :

$$\rho_t = \lim_{\Delta t \rightarrow 0} \frac{P(t + \Delta t) - P(t)}{\Delta t \times \frac{1}{\Delta t} \int_t^{t + \Delta t} P(\tau) d\tau} = \frac{P'(t)}{P(t)} \quad (1.4)$$

Rappelons que la fonction  $P(t)$ , qui représente un effectif entier d'hommes ou de femmes, est une fonction *en escalier* et n'est donc pas dérivable au sens strict. Pourtant tous les modèles de démographie classique utilisent les notions de taux de croissance instantané ou de dérivée d'effectif. C'est qu'en effet, le démographe s'intéresse le plus souvent aux grands effectifs<sup>3</sup> de telle sorte qu'il peut, dans la plupart des cas, raisonner sur des effectifs *moyens* ou plus exactement l'*espérance mathématique* des effectifs. La figure 1.4 montre le graphe de la fonction  $P(t)$  dont une partie serait agrandie à l'aide d'une loupe jusqu'à ce que le processus élémentaire apparaisse. Dans le cas simple où seules des naissances et des décès interfèrent, ce graphe correspond en statistique au processus dit *processus de vie ou de mort*. La courbe *lissée* est alors l'espérance mathématique de

<sup>3</sup>et des périodes de temps assez longues.

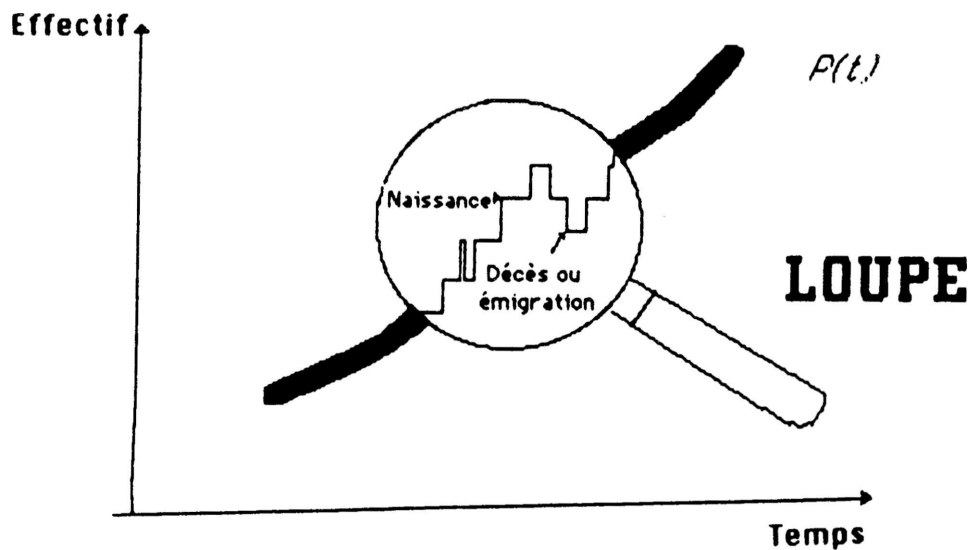


Figure 1.4: Graphe de la taille d'une population en fonction du temps, agrandie "à la loupe" jusqu'à ce que le processus élémentaire de vie ou de mort apparaisse.

la variable aléatoire effectif<sup>4</sup> dont on prouve le plus souvent la dérivabilité.

Le processus élémentaire le plus simple est celui où les taux instantanés de naissance, mort, émigration ou immigration sont constants dans le temps, ou se combinent pour donner un taux de croissance constant.

Étudions plus en détail ce cas particulièrement important.

### 1.3.2 Taux de croissance constant

Si nous supposons que le taux de croissance est indépendant du temps, il devient aisé d'intégrer l'équation 1.4 :

$$\frac{dP}{P} = \rho dt \iff \int_0^t \frac{dP}{P} = \int_0^t \rho dt$$

soit encore :

$$P(t) = P(0)e^{\rho t} \quad (1.5)$$

Le graphe de la fonction est donc celui d'une exponentielle.

<sup>4</sup>définie comme le solde des naissances et des décès.

**Temps de doublement**

Calculons la durée,  $\tau$ , au bout de laquelle la taille de la population,  $P(t + \tau)$  sera le double de ce qu'elle est à la date  $t$ .

$$P(t + \tau) = 2P(t) \iff \begin{aligned} P(0)e^{\rho(t+\tau)} &= 2 \times P(0)e^{\rho t} \\ P(0)e^{\rho t} \times e^{\rho \tau} &= 2P(0)e^{\rho t} \end{aligned}$$

Comme on peut simplifier les deux membres par  $e^{\rho t}$ , le *temps de doublement est donc indépendant de la date de référence  $t$* , ce qui justifie son appellation. Il vient alors  $e^{\rho \tau} = 2$  ou

$$\tau = \frac{\log 2}{\rho} \quad (1.6)$$

Cette propriété d'indépendance par rapport à la date de référence est la base d'une progression géométrique.

**Règle pratique de calcul d'un temps de doublement** Comme les taux d'accroissement sont généralement exprimés en pourcentage, on peut écrire :

$$\begin{aligned} \tau &= \frac{\log 2}{100 \times \frac{\rho}{100}} = \frac{0,693}{100 \times \rho(\%)} \\ \tau &= \frac{70}{\rho(\%)} \end{aligned} \quad (1.7)$$

**Exercice 1** La ville de Yaoundé a un taux de croissance de 9 % annuel, calculer son temps de doublement.

**Solution 1** Le temps de doublement de la ville de Yaoundé est proche de  $\frac{70}{9} = 7,7$  ans.

**Calcul d'un taux de croissance sur une période finie**

Revenons maintenant au calcul du taux de croissance d'une population entre deux dates  $t_0$  et  $t_0 + T$ . Nous pouvons faire l'hypothèse que la loi de la population est du type exponentiel durant cette période et calculer l'unique exponentielle qui passe par les deux points. Les calculs sont alors les suivants :

$$\text{pour } t = t_0 \quad P(t_0) = P_0 e^{\rho t_0}, \quad \text{et } t = t_0 + T, \quad P(t_0 + T) = P_0 e^{\rho(t_0 + T)}$$

d'où :

$$\rho = \frac{1}{T} \log \frac{P_{t_0+T}}{P_{t_0}} \quad (1.8)$$

**Exercice 2** Vérifier que dans le calcul d'une population de type exponentiel, le taux d'accroissement  $r$  de la population entre deux dates est bien le taux de croissance instantanée  $\rho$ .

**Solution 2** Par définition du taux nous avons en effet :

$$r = \frac{P_{t_0+T} - P_{t_0}}{\int_{t_0}^{t_0+T} P_0 e^{\rho t} dt} = \frac{P_{t_0+T} - P_{t_0}}{\frac{P_0}{\rho} [e^{\rho(t_0+T)} - e^{\rho t_0}]} = \rho$$

Si on se contente de calculer l'intégrale par l'aire du trapèze (ce qui revient à calculer la population moyenne comme la moyenne arithmétique des deux populations extrêmes) :

$$r_m = \frac{P_{t_0+T} - P_{t_0}}{T \times \left( \frac{P_{t_0+T} + P_{t_0}}{2} \right)} \quad (1.9)$$

l'erreur relative commise est du second ordre en  $\rho T$ , comme le prouve le calcul suivant : (on pose  $t = T$  et  $r_m$  est le taux déduit de la *moyenne arithmétique*)

$$r_m = \frac{2(e^{\rho t} - 1)}{t(1 + e^{\rho t})} = \frac{2}{t} \tanh(\rho t) = \frac{2}{t} \frac{(\rho t + \rho^2 \frac{t^2}{2!} + \rho^3 \frac{t^3}{3!} + \dots)}{(2 + \rho t + \rho^2 \frac{t^2}{2!} + \rho^3 \frac{t^3}{3!} + \dots)}$$

*du par Taylor*

La division suivant les puissances croissantes<sup>5</sup> de  $\rho$  montre :

$$\begin{array}{r|l} 2\rho + \rho^2 t + \rho^3 \frac{t^2}{3} + \rho^4 \frac{t^3}{12} + \rho^5 \frac{t^4}{60} + \dots & 2 + \rho t + \rho^2 \frac{t^2}{2} + \rho^3 \frac{t^3}{6} + \dots \\ 2\rho + \rho^2 t + \rho^3 \frac{t^2}{2} + \rho^4 \frac{t^3}{6} + \dots + & \rho - \frac{1}{12} \rho^3 t^2 + \dots \\ \quad - \frac{1}{6} \rho^3 t^2 - \frac{1}{12} \rho^4 t^3 + \dots & \\ \quad - \frac{1}{6} \rho^3 t^2 - \frac{1}{12} \rho^4 t^3 + \dots & \\ \quad 0 \quad \quad \quad 0 & \end{array}$$

soit  $r_m = \rho - \frac{1}{12} \rho^3 t^2 + O(\rho^5 t^4)$ . L'erreur relative vaut :

$$\frac{r_m - \rho}{\rho} = -\frac{1}{12} (\rho^2 t^2) + O(\rho^4 t^4). \quad (1.10)$$

Nous vérifions en effet que le terme du premier ordre est nul. Il n'en serait pas de même si nous avions pris comme dénominateur la valeur de la population en début de période<sup>6</sup> :

$$r_d = \frac{P_{t_0+T} - P_{t_0}}{t \times P_{t_0}} = \rho + \frac{\rho^2}{2} t + \rho^3 \frac{t^2}{6} + O(\rho^4 t^3),$$

car l'erreur relative serait :

$$\frac{r_d - \rho}{\rho} = \frac{\rho t}{2} + \frac{\rho^2 t^2}{6} + O(\rho^3 t^3). \quad (1.11)$$

c'est à dire beaucoup plus forte, puisque généralement  $\rho^2 t^2 \ll \rho t$ .

<sup>5</sup>Le développement de  $\tanh z$  par la formule de Taylor fournit un résultat équivalent.

<sup>6</sup>multiplié par la durée.

**Taux d'accroissement constant — Formulation discrète**

Si on observe une population exponentielle à des intervalles de temps réguliers  $0, 1, 2, \dots, n$ , nous pouvons écrire :

$$P_n = P_0 e^{rn} = P_0 (e^r)^n.$$

Si nous posons :

$$e^r = 1 + r,$$

soit encore

$$\log(1 + r) = \rho, \quad (1.12)$$

il vient  $P_n = P_0(1 + r)^n$ . La suite  $P_0, P_1, \dots, P_n$  est alors une suite géométrique de raison  $1 + r$ .

1. Il ne faut pas confondre  $r$  et  $\rho$ , même si, quantitativement leurs valeurs sont proches lorsqu'elles sont petites :

$$\rho = \log(1 + r) \simeq r \quad \text{pour } r \text{ petit.}$$

2. Même si les intervalles de temps ne sont pas réguliers, certains auteurs gardent la notation  $(1 + r)$  et écrivent pour toute date  $t$  :  $P_t = P_0(1 + r)^t$ , la notation  $P_t = P_0 e^{rt}$  semble néanmoins préférable.

**1.3.3 Taux d'intérêt, calcul d'un intérêt**

Lorsqu'on place une somme d'argent  $S$  chez un intermédiaire financier, il est important de connaître, non seulement l'unité de temps dans laquelle est exprimée le taux d'intérêt, mais aussi son mode de calcul. Généralement, l'unité est l'année et le mode de calcul annuel, mais il peut en être autrement et il est nécessaire de connaître les alternatives et leurs conséquences. Nous allons raisonner sur l'exemple d'un taux  $r = 6,5\%$  l'an, en faisant varier le mode de calcul.

- Si le mode de calcul est annuel, la somme doit être évaluée de nouveau tous les ans d'un facteur  $(1 + 0,065)$  ; autrement dit la valeur de la somme est sur la courbe exponentielle (voir la figure 1.5) dont l'équation est :

$$S \times (1,065)^t = S \times e^{t \log(1,065)}.$$

Le taux de croissance instantané correspondant à cette exponentielle est alors :

$$\rho_1 = \log(1,065) = 0,0629 \quad \text{par an.}$$

Au bout de 10 ans la valeur à rembourser au prêteur est  $S \times (1,065)^{10} = S \times 1,877$ .

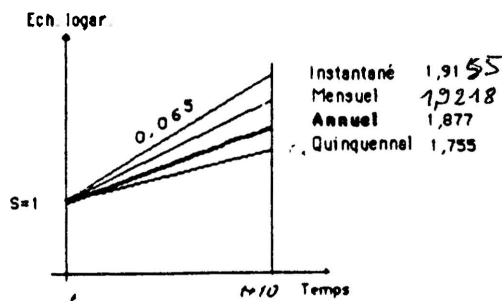


Figure 1.5: Différents modes de calcul d'un intérêt pour une même valeur du taux, ici de 6,5% annuel.

- Si le mode de calcul est quinquennal, la somme n'est réévaluée que tous les cinq ans d'un facteur  $(1 + 5r) = 1 + 0,325$ . Autrement dit l'exponentielle est différente, moins avantageuse pour le prêteur, et le taux de croissance instantané correspondant est :

$$\rho_5 = \frac{1}{5} \log(1 + 5r) = \frac{1}{5} \log(1,325) = 0,05628 \quad \text{par an.}$$

Au bout de 10 ans la valeur du capital placé est  $S \times (1 + 5 \times 0,065)^2 = S \times 1,755$ .

- Inversement, si le capital est réévalué tous les mois, il vaudra au bout d'un an,  $S \times (1 + \frac{r}{12})^{12}$  et, s'il est réévalué tous les jours,  $S \times (1 + \frac{r}{365})^{365}$ .
- A la limite, si le capital était réévalué à chaque instant le capital vaudrait en fin d'année<sup>7</sup> :

$$\lim_{n \rightarrow \infty} S(1 + \frac{r}{n})^n = S \times e^r,$$

le taux de croissance instantané,  $\rho$ , serait alors égal au taux d'intérêt lui-même. Dans l'exemple ci-dessus, un capital placé avec un taux annuel de 6,5 %, mais dont le mode de calcul de l'intérêt serait du type instantané, vaudrait au bout de 10 ans :  $e^{10 \times 0,065} = 1,91$  contre 1,877 si le mode de calcul annuel, classique, était appliqué.

Quand on parle d'un taux d'accroissement annuel,  $r$ , il est généralement sous entendu que le mode de calcul est annuel, et le taux instantané,  $\rho$ , qui

<sup>7</sup>Nous avons en effet :

$$\begin{aligned} \frac{1}{1+z} &= 1 - z + \frac{z^2}{2} - \frac{z^3}{3} + \dots \\ \log\left(\frac{1}{1+z}\right) &= z - \frac{z^2}{2} + \frac{z^3}{3} + \dots \quad \text{d'où} \\ n \log\left(1 + \frac{r}{n}\right) &= r - \frac{r^2}{2n} + \frac{r^3}{3n^2} + \dots \end{aligned}$$



lui correspond vaut :  $\rho = \log(1 + r)$ . La valeur de  $r$  est calculée comme le pourcentage d'accroissement en un an du capital,  $r = \frac{S_{t+1} - S_t}{S_t}$ . Cette valeur est un nombre sans dimension<sup>8</sup> et est souvent exprimée en pourcentage. Quitte à exprimer l'accroissement en pourcentage, les économistes ont l'usage des *indices*.

### 1.3.4 Indice

**Définition 4** On définit l'indice de l'année  $j$ , en référence à l'année de base 0, et on note  $i_{j/0}$  le rapport  $\frac{P_j}{P_0} \times 100$  :

$$i_{j/0} = \frac{P_j}{P_0} \times 100 \quad (1.13)$$

La relation entre  $r$  et  $i_{j/0}$  est alors une relation comptable simple,  $i = 100 \times (1 + r)$  (un accroissement de 50% correspond à un indice de 150). Il n'en est pas de même entre un indice (ou un pourcentage d'accroissement) et un taux (instantané ou calculé par l'équation 1.9) car les différences peuvent être très importantes comme le montre le tableau 1.1

Les indices sont fort utiles en économie, et donnent d'une manière simple l'ordre de grandeur des évolutions comme le montre, par exemple, le tableau 1.2. Néanmoins, en démographie, on raisonne très souvent sur des effectifs, c'est à dire des quantités cumulables, et le taux est souvent le premier outil d'analyse que l'on peut utiliser ; le second étant le *quotient* que nous allons voir maintenant.

### 1.3.5 Quotients

Les démographes ont introduit la notion de *quotient* pour étudier et décrire plus facilement la *décroissance* d'un effectif, comme par exemple la survie à divers âges d'une cohorte. Un quotient est une simple probabilité, et les anglosaxons n'ont pas de terme approprié et parlé de *probability of deaths* par exemple. Le quotient est un outil très utile dans les cas simples, comme dans le calcul d'une table de mortalité, et en l'absence de calculateur électronique. Il est moins et son calcul est plus délicat (voir Pressat [38]) dès que les phénomènes sont plus complexes, c'est à dire quand la décroissance d'un effectif est due à plusieurs phénomènes concurrents, mortalité et émigration par exemple, ou même quand certains phénomènes antagonistes font croître les effectifs, (entrées et sorties multiples de la vie active, migrations<sup>9</sup> etc. Les notions de taux et plus particulièrement celle de taux instantané sont les seules notions indiscutables

<sup>8</sup> alors qu'un taux est homogène à l'inverse d'un temps.

<sup>9</sup> Le terme anglais qui décrit ces processus est *increment-decrement life table*.

Indice $i = 100 \times (1 + r)$	Accroissement $r$	Croissance instantanée $\rho = \log(1 + r)$
0	-1	$-\infty$
20	-0,8	-1,609
50	-0,5	-0,693 <sup>(a)</sup>
90	-0,1	-0,105
95	-0,05	-0,0512
97	-0,03	-0,0304
99	-0,01	-0,01005
99,9	-0,001	-0,0010005
100	0	0
+100,1	0,001	0,0009995
+101	0,01	0,00995
+103	0,03	0,0295
105	0,05	0,0488
110	0,1	0,0953
150	0,5	0,4054
180	0,8	0,587
200	1	0,693 <sup>(a)</sup>

<sup>(a)</sup> = log 2Tableau 1.1: Différence entre un *taux* instantané et un *pourcentage* d'accroissement.

Tableau 1.2: Production d'un travailleur dans l'agriculture en France de 1750 à 1980.

base 100 = 1934-1938	Indice
1750-1800	20
1894-1898	68
1909-1913	75
1934-1938	100
1980	200

Source : "Le grand espoir du XX<sup>e</sup> siècle", J. Fourastié, Presses Universitaires de France, 1949, (l'estimation pour l'année 1980 provient de M. Alfred Sauvy).

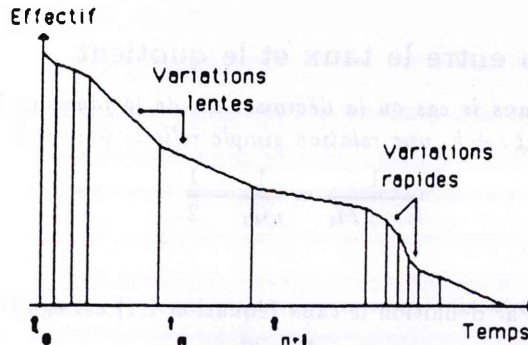


Figure 1.6: Temps intermédiaires.

dans les cas complexes. Même dans le cas simple d'une table de mortalité, il est possible d'introduire la notion d'espérance de vie, sans mentionner le mot quotient ~~mais~~ en raisonnant uniquement sur les taux infinitésimaux, (comme le fait M. Bourgeois Pichat [2].

Nous allons donc définir un quotient dans le cas où l'effectif décroît en raison d'une seule cause et que cette cause (ou événement) est non renouvelable (décès) ou supposé non renouvelable (mariage).

**Définition 5** Soit  $P(t)$  un effectif décroissant en fonction du temps  $t$ , le quotient,  ${}_h Q_t$ , à la date  $t$ , et pour une durée  $h$ , est défini comme le rapport :

$${}_h Q_t = \frac{P(t) - P(t+h)}{P(t)} \quad (1.14)$$

Ce nombre est sans dimension et compris entre 0 et 1 ; c'est pourquoi on l'assimile à une probabilité : probabilité pour qu'un individu n'ayant pas encore subi l'événement à la date  $t$  le subisse entre les instants  $t$  et  $t+h$ . La valeur  ${}_h P_t = 1 - {}_h Q_t$  est alors la probabilité de ne pas le subir.

L'introduction du quotient résulte du fait qu'en démographie on ne peut observer un processus continu qu'à des intervalles de temps discrets. Ainsi, partant d'un effectif donné, à une date  $t_0$ , on ne peut mesurer l'effectif restant, qu'à des dates  $t_1, t_2, \dots, t_n$ . En l'absence d'information intermédiaire entre ces deux dates, il paraît justifié de faire l'hypothèse que la décroissance des effectifs est linéaire entre deux dates. En effet, plus la décroissance est mouvementée, plus il paraît important d'introduire des temps intermédiaires pour mieux étudier ces variations (voir la figure 1.6).

Dans le cas de la mortalité, par exemple, il est d'usage de découper le cycle de vie en : (1) première année, (2) 1 an à 5 ans, (3) à (18) de 5 ans en 5 ans<sup>10</sup>.

Si on considère que les temps, ou âges intermédiaires, sont en nombre suffisant pour décrire correctement la fonction  $P$  (généralement la fonction de survie), il est possible de faire l'hypothèse que la fonction  $P$  est linéaire par morceau, et l'on obtient une relation très simple entre le quotient et le taux.

<sup>10</sup>Ce découpage n'est d'ailleurs pas suffisant durant la première année, et il est nécessaire d'étudier la mortalité durant les premiers jours, la première semaine, le premier mois, etc.

*de*  
*un l'approche*  
*à venir*  
*prendre des données pour l'analyse de la mortalité de la population*  
*1050*

## 1.3.6 Relation entre le taux et le quotient

**Propriété 1.3.1** Dans le cas où la décroissance de la fonction  $P$  est linéaire entre les instants  $t$  et  $t + h$ , une relation simple relie le taux et le quotient :

$$\frac{1}{h \times {}_hM_t} = \frac{1}{hQ_t} - \frac{1}{2} \quad (1.15)$$

**Démonstration** Par définition le taux (équation 1.1) est égal<sup>11</sup> à :

$${}_hM_t = \frac{P_t - P_{t+h}}{h \times \left(\frac{P_t + P_{t+h}}{2}\right)} \quad (1.16)$$

Or, par définition du quotient (équation 1.14), nous avons :  $P_{t+h} = P_t(1 - hQ_t)$  ; ainsi :

$${}_hM_t = \frac{{}_hQ_t P_t}{h \times P_t \left(\frac{2 - hQ_t}{2}\right)},$$

soit finalement :

$${}_hM_t = \frac{2 \times {}_hQ_t}{h \times (2 - hQ_t)} \quad (1.17)$$

$${}_hQ_t = \frac{2h \times {}_hM_t}{2 + h {}_hM_t} \quad (1.18)$$

On en déduit aussi la relation 1.15, ce qui achève la démonstration. La relation 1.15 est sans doute la plus simple à retenir, car (1) c'est une relation entre nombre sans dimension (le produit du taux  ${}_hM_t$  par une durée  $h$  est un nombre sans dimension) (2) l'inverse d'un taux annuel ( $h = 1$ ) est inférieur à l'inverse d'un quotient, puisque le numérateur est dans le cas du taux (population moyenne entre les deux dates) inférieur à celui du quotient (population en début de période) ; il faut donc lui ôter une quantité, en l'occurrence  $\frac{1}{2}$ .

Dans le cas annuel, les valeurs numériques du taux et du quotient ne diffèrent pas beaucoup, ce qui est parfois une source de confusion entre deux quantités qui ne résultent pas du tout de la même démarche.

Dans le cas de données quinquennales, le taux  ${}_5M_t$  sera légèrement supérieur au cinquième du quotient  ${}_5Q_t$ .

**Cas où la décroissance n'est pas linéaire** Dans le cas où la décroissance ne peut être assimilée à une droite, la formule 1.15 sera une approximation d'autant plus inexacte, que l'arc s'éloignera du segment. Si l'arc est une exponentielle décroissante, nous avons vu que l'erreur relative commise sur le taux en assimilant l'arc à un segment de droite est  $\frac{\sigma^2 h^2}{12}$ . En vertu de l'équation 1.15, l'erreur relative sur le quotient est identique.

<sup>11</sup>On considère ici un taux de décroissance qui est donc l'opposé de celui donné par l'équation 1.1.

le 2<sup>o</sup> fait  
fait  
à rebours

## Chapitre 2

# Modèles de population sans classe d'âge

Même si l'espèce humaine domine aujourd'hui les autres espèces il est important de connaître les mécanismes qui régulent les autres espèces de la planète ou ceux qui peuvent amener à leur extinction. D'autre part, on ne peut pas supposer que l'homme domine entièrement l'environnement dans lequel il vit, mais qu'il évolue, se multiplie, selon des modèles certes complexes, mais qui ne sont peut-être pas si différents des modèles *écologiques* d'interaction entre les espèces. L'échec relatif de la théorie des populations stables (croissance exponentielle) dans son explication de l'évolution démographique des pays développés nous conduit en effet à envisager d'autres modèles (Easterlin [12], Samuelson [40]), cycliques cette fois, qui s'apparentent du point de vue mathématique à certains modèles écologiques. *de la population*

L'objet de ce chapitre est d'étudier la façon dont on peut *modéliser* l'évolution quantitative des espèces lorsqu'elles sont en interaction. Nous suivons l'approche du célèbre démographe A. Lotka [32] qui a consacré une partie de ses recherches à la compétition entre les espèces avant d'aborder l'étude de son œuvre maîtresse sur les populations stables.

Les modèles les plus simples ne prennent en compte que les tailles des différentes espèces en interaction, sans s'intéresser à leur composition par âge, si importante en démographie humaine, et c'est pourquoi le chapitre s'intitule "modèles de population sans classe d'âge", mais il va de soit que les modèles de compétition entre espèces peuvent comporter des classes d'âge. Lorsque nous aborderons au chapitre 5 la théorie des populations stables, nous introduirons bien évidemment les classes d'âge, mais le modèle sans classe d'âge qui lui correspondra sera plus pauvre puisqu'il s'agira du modèle exponentiel étudié au paragraphe 1.3.2 du chapitre 1. En effet le modèle démographique le plus simple est celui qui considère que la population a dominé le milieu dans lequel

elle vit, qu'il n'y a plus de compétition directe avec d'autres espèces, ni même avec elle-même. Dans ce cas le taux de croissance est supposé constant et le modèle de croissance est exponentiel.

Malthus [33] a, le premier, décrit cette croissance exponentielle naturelle, et a rappelé la contrainte de ressources<sup>1</sup>, Verhulst (1838) [44] et Pearl (1922) [35] ont modélisé l'évolution d'une telle population contrainte, en postulant que le taux de croissance ne pouvait rester constant mais devait être une fonction décroissante de la taille de la population. Le cas le plus simple envisagé est celui où la décroissance est *linéaire* ; la courbe de l'évolution d'une telle population est appelée *courbe logistique*. La réalité d'une telle loi a tout d'abord été vérifiée sur des populations animales comme les drosophiles de Pearl ou les paramécies de Gause (1935)[14], avant d'être proposée comme modèle d'évolution d'une population humaine.

Avant d'étudier en détail les différents modèles, nous allons, à la manière de A. Lotka, exposer les divers types de dépendance mutuelle possible entre les espèces afin de comprendre à quel niveau se situent les modèles que nous étudierons. A. J. Lotka distingue deux types de dépendance mutuelle, le type coordonné et le type subordonné.

**Dépendance coordonnée** Dans la dépendance coordonnée, la compétition peut être *passive* : chaque espèce influence l'autre indirectement par une diminution de la nourriture par exemple, ou *active* : les espèces se disputent les biens dont elles cherchent toutes deux à profiter dans un combat réel.

**Dépendance subordonnée** La dépendance du type subordonnée est celle dans laquelle l'une des deux espèces est utile à l'autre. On retient en particulier le cas le plus commun où une espèce victime (ou un produit de cette espèce) sert de nourriture à une seconde espèce. Le miel et le lait en sont des exemples.

Si les individus de l'espèce victime sont eux-mêmes exploités par l'espèce prédatrice, on distingue le cas où la victime est dévorée sur le champs par le prédateur (carnivores par exemple) et le cas où la victime reste en vie (parasitisme).

Plus généralement, dans un système écologique réel plusieurs espèces sont en présence et leurs règles de cohabitation sont très complexes (voir le tableau 2.1).

On a pu évaluer quelques chaînes de nourriture et par exemple celle d'un poisson, le hareng, dont le type de nourriture et la taille de ses proies dépendent de son âge (voir la figure 2.2).

Du point de vue mathématique, il est possible d'écrire les équations qui régissent un tel système, aussi complexe soit-il. Ainsi, si  $n$  espèces, en quantité

---

<sup>1</sup>Rappelons que l'auteur suppose que les ressources ne peuvent avoir qu'une croissance arithmétique.

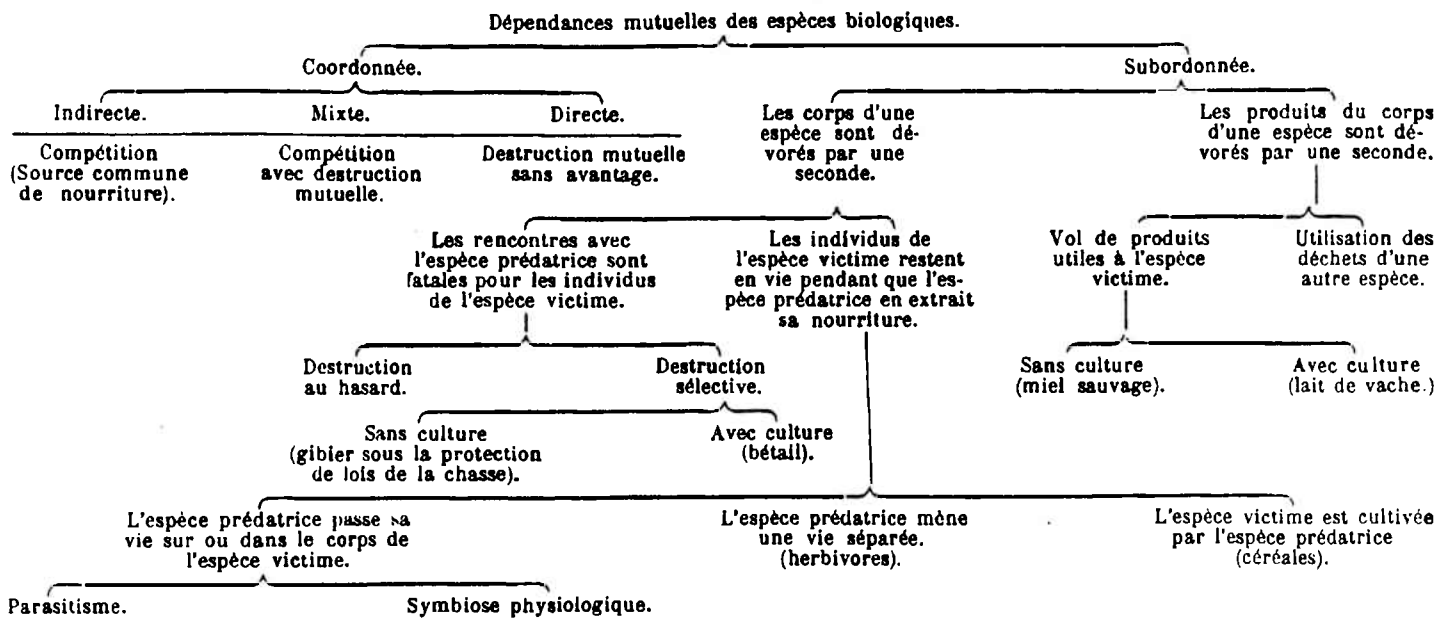


Figure 2.1: Dépendances mutuelles des espèces biologiques. Source : A. Lotka.

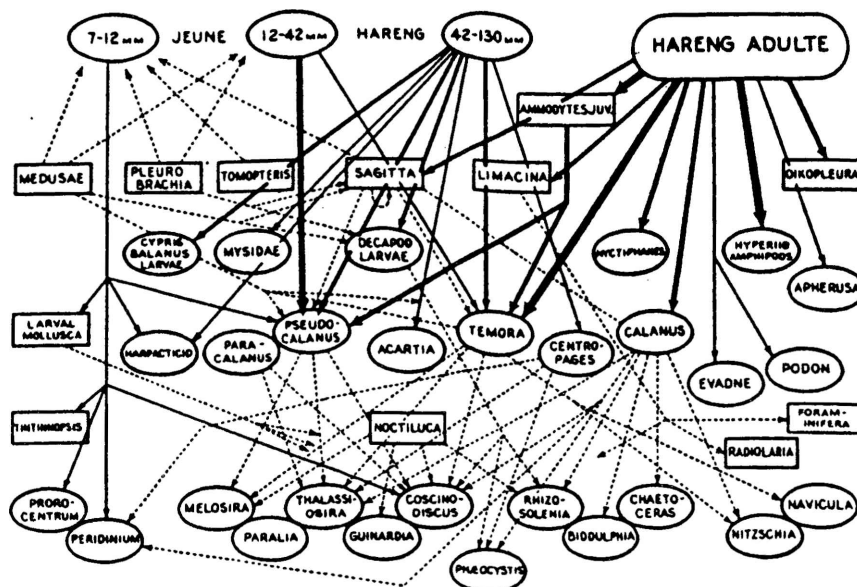


TABLEAU II. Le hareng et les espèces qui lui servent directement ou indirectement de nourriture. On observe l'influence de la taille du poisson jeune et adulte sur le choix de sa nourriture. (D'après A. C. Hardy, The herring in relation to its animate environment, Part I. Ministry of Agriculture and Fishery Investigations, 1924, Series 2, Vol. 7, N° 3. Reproduced by permission of the Controller of His Britannic Majesty's Stationery Office.)

Figure 2.2: Le hareng et les espèces qui lui servent directement ou indirectement de nourriture.

Source : A. Lotka [32].



$X_1, X_2, \dots, X_n$ , à un instant  $t$  donné sont en dépendance mutuelle, on peut caractériser l'évolution de chacune d'entre elles par une équation différentielle du type :

$$\frac{dX_i}{dt} = f_i(X_1, X_2, \dots, X_n) \quad \forall i \in [1, n]. \quad (2.19)$$

Le système général n'est souvent pas soluble analytiquement, par contre, si on ne s'intéresse qu'aux situations d'équilibre éventuel et à leur stabilité, il est possible par un changement de variable et une linéarisation ~~par un développement en série de Taylor~~, d'obtenir un système linéaire simple dont on peut décrire les solutions. à manier

Dans le cas de *deux espèces*, les formes des *courbes intégrales*, solutions de ce système différentiel ont été répertoriées par A. J. Lotka et reproduite sur la figure 2.3. Nous ne nous intéresserons dans ce chapitre qu'aux cas les plus simples de dépendance coordonnée (compétition) et subordonnée (modèles proie-prédateur ou hôte-parasite). Dans le cas de la dépendance subordonnée nous verrons le cas particulier où la population est soumise à un prélèvement régulier, comme dans le cas d'une population soumise à une émigration constante, qui correspond en écologie au cas de la pêche ou de la chasse.

## 2.1 Population logistique

Le modèle de croissance exponentielle vu précédemment suppose que la population n'est soumise à aucune contrainte qui puisse modifier sa fécondité et sa mortalité. Après Malthus, l'idée d'une régulation possible de la population a vu le jour, et le premier modèle régulateur, où le taux de croissance diminue lorsque la population augmente, nous vient de Verhulst et de Pearl<sup>2</sup> avec leurs célèbres expériences sur les drosophiles. La contrainte imposée par le milieu (surpopulation, raréfaction des ressources) est supposée créer une hausse de la mortalité et/ou induire une baisse de la fécondité.

On peut donc faire l'hypothèse que l'équation qui régit la taille de la population  $P$  en fonction du temps  $t$  est de la forme :

$$\frac{dP}{P dt} = \rho(P) \quad \text{avec} \quad \rho'(P) < 0, \quad (2.20)$$

ou encore :

$$\int_{P_0}^{P_t} \frac{1}{P \rho(P)} dP = t - t_0. \quad (2.21)$$

Dans le cas où la fonction  $\rho(P)$  est linéaire décroissante en fonction de  $P$ , nous avons la définition d'une population logistique.

**Définition 6** *Une population est dite population logistique si son taux de croissance est une fonction linéaire décroissante de sa taille, c'est à dire, si le taux*

<sup>2</sup>Verhulst et Pearl op.cit.

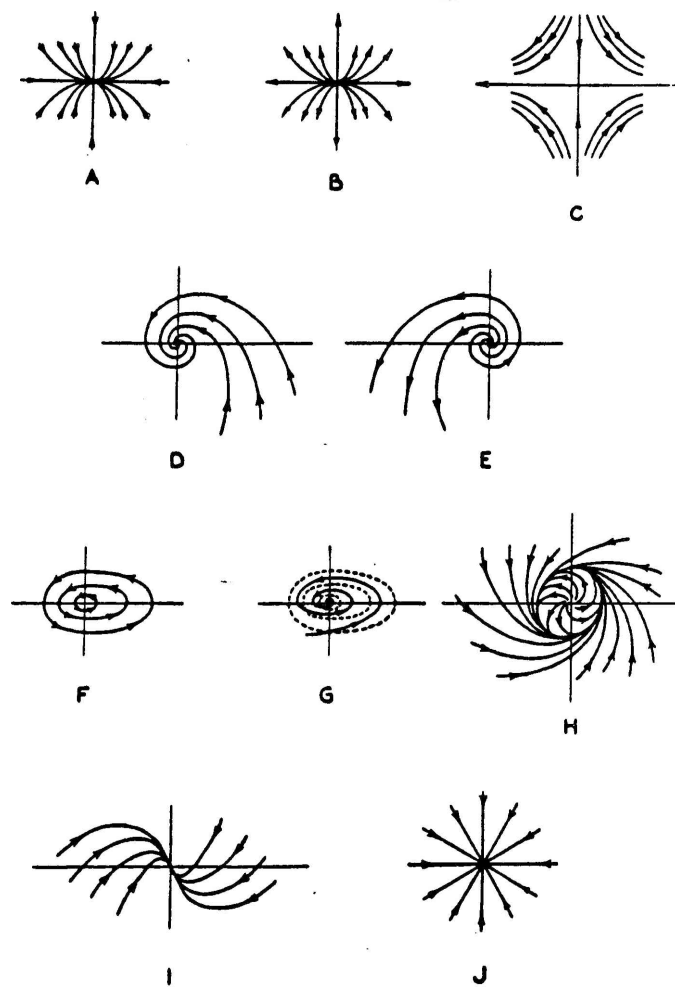


Figure 2.3: Diagrammes représentant plusieurs types d'états stationnaires dans un système comprenant deux espèces biologiques.  
Source : A. Lotka [32].

de croissance instantané  $\rho$  peut s'exprimer en fonction de la taille  $P$  de la population à l'instant  $t$  sous la forme :

$$\rho \stackrel{\text{def}}{=} \frac{dP}{P dt} = a - bP \quad (2.22)$$

où  $a$  et  $b$  sont des constantes positives.

### 2.1.1 Expression de la taille de la population en fonction du temps

L'équation 2.22 peut aussi s'écrire sous la forme :

$$\frac{dP}{P \times (a - bP)} = dt,$$

ce qui montre qu'on peut décomposer en éléments simples la partie gauche de cette équation :

$$\frac{dP}{P \times (a - bP)} = \frac{A}{P} + \frac{B}{a - bP}$$

d'où l'on déduit :

$$A = \frac{1}{a} \quad \text{et} \quad B = \frac{b}{a}.$$

On peut alors intégrer chaque élément simple :

$$\frac{1}{a} \left[ \int_{P_0}^{P_t} \frac{dP}{P} + b \int_{P_0}^{P_t} \frac{dP}{a - bP} \right] = t - t_0$$

soit encore en intégrant :

$$\log \left( \frac{P_t}{P_0} \times \frac{|a - bP_0|}{|a - bP_t|} \right) = a(t - t_0).$$

Par continuité<sup>3</sup> à partir de  $P_0$ ,  $P$  et  $P_0$  sont du même côté par rapport à la limite  $\frac{a}{b}$  et nous obtenons :

$$P(a - bP_0) = e^{at} (a - bP) P_0, \quad (2.23)$$

et en posant  $\frac{a}{b} = L$ , où  $L$  est la valeur, autre que 0 pour laquelle le taux de croissance est nul :

$$e^{-at} P \left( 1 - \frac{P_0}{L} \right) = P_0 \left( 1 - \frac{P}{L} \right) \quad (2.24)$$

ce qui implique :

$$P(t) = L \left( \frac{1}{1 + \frac{L - P_0}{P_0} e^{-at}} \right) \quad (2.25)$$

<sup>3</sup>La fonction  $P$  étant dérivable, elle est continue.

- Dans le cas où  $P_0$  est inférieur à  $L$ ,  $P(t)$  est aussi inférieur à  $L$  pour tout  $t$ . Ce cas est le cas le plus fréquent dans les populations humaines.
- Dans le cas où  $P_0$  est supérieur à  $L$ , l'équation 2.25 s'écrit plutôt sous la forme :

$$P(t) = L \left( \frac{1}{1 - \frac{P_0 - L}{P_0} e^{-at}} \right), \quad (2.26)$$

ce qui implique  $P(t) > L \quad \forall t_0$ , et une asymptote verticale pour :

$$t = -\frac{1}{a} \log \frac{P_0}{P_0 - L} \quad (2.27)$$

### 2.1.2 Étude des variations de la fonction logistique

Par définition, la dérivée vaut :

$$P' = \frac{dP}{dt} = aP \left( 1 - \frac{P}{L} \right). \quad (2.28)$$

On déduit alors la dérivée seconde :

$$P'' = \frac{d^2P}{dt^2} \quad (2.29)$$

$$= \frac{dP'}{dt} = \frac{d(aP(1 - \frac{P}{L}))}{dP} \times \frac{dP}{dt} \quad (2.30)$$

$$= \left( a(1 - \frac{P}{L} - a\frac{P}{L}) \right) \times aP(1 - \frac{P}{L}) \quad (2.31)$$

$$= a^2(1 - \frac{2P}{L})P(1 - \frac{P}{L}). \quad (2.32)$$

**Cas où  $P_0 < L$ .** La dérivée de  $P$  par rapport au temps est toujours positive, la dérivée seconde s'annule pour  $P = \frac{L}{2}$ . Nous avons le tableau de variation suivant :

$t$	$-\infty$	$0$	$+\infty$
$P''$	0	+	+
$P'$	0	+	+
$P$	0	$\nearrow P_0$	$\nearrow L$

**Cas où  $P_0 > L$ .** La dérivée  $P'$  est toujours négative, et  $P''$  toujours positive. Le tableau de variation est alors le suivant :

$t$	$-\frac{1}{a} \log \frac{P_0}{P_0 - L}$	$0$	$+\infty$
$P''$	$+\infty$	+	+
$P'$	$-\infty$	-	-
$P$	$+\infty$	$\searrow P_0$	$\searrow L$

Le cas général se réfère toujours au cas où  $P_0 < L$ . Le point d'inflexion, s'il existe, est obtenu pour une population égale à la moitié de la population limite, et au bout d'un temps  $T$  qui vérifie :

$$\frac{L}{2} = \frac{L \times P_0}{P_0 + (L - P_0)e^{-aT}},$$

soit

$$T = \frac{1}{a} \log \frac{L - P_0}{P_0} \quad (2.33)$$

Si l'on prend  $T$  au lieu de  $P_0$  comme paramètre, nous obtenons l'équation sans dimension et simplifiée suivante :

$$\frac{P}{L} = \frac{1}{1 + e^{-a(t-T)}} \quad (2.34)$$

**Exercice 3** Montrer que le point d'inflexion est centre de symétrie.

**Solution 3** Le point d'inflexion est obtenu pour  $P_i = \frac{L}{2}$  et  $t_i = T$ . Pour que ce dernier soit centre de symétrie, il faut et il suffit de prouver que :  $P(t_i + h) + P(t_i - h) = 2 \times P_i$ , soit :

$$\frac{L}{1 + e^{ah}} + \frac{L}{1 + e^{-ah}} = L$$

ce qui se vérifie aisément.

**Exercice 4** Montrer que le taux de croissance  $\rho(t)$  suit lui aussi une loi logistique.

**Solution 4** On peut en effet écrire :

$$\rho = \frac{dP}{P dt} = a \left(1 - \frac{P}{L}\right),$$

et en dérivant :

$$\begin{aligned} \rho' &= \frac{d\rho}{dP} \frac{dP}{dt} \\ &= -\frac{a}{L} \times \rho P \end{aligned}$$

or  $P = L(1 - \rho/a)$  d'où :

$$\rho' = -a\rho \left(1 - \frac{\rho}{a}\right),$$

ce qui est bien l'équation d'une loi logistique.

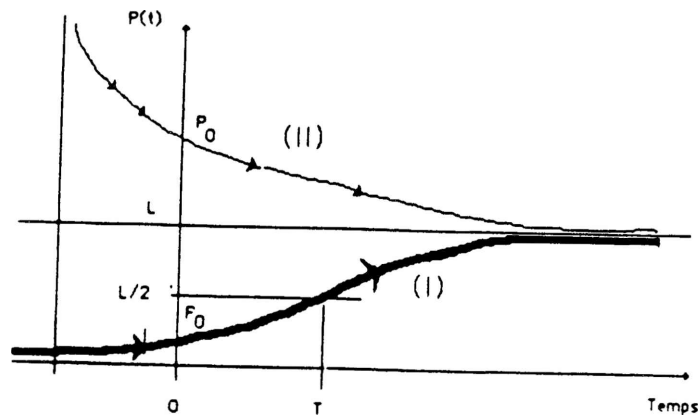


Figure 2.4: Population logistique : représentation de  $P$  en fonction du temps. (Partie A).

**Remarque** Prenons le cas particulier où  $a = 1$  et  $T = 0$ , et posons :  $x = \frac{P}{L}$  ; l'équation 2.34 devient :

$$t = \log \left( \frac{x}{1-x} \right) = \text{logit}(x).$$

Nous comprenons alors l'analogie avec la fonction "logit" souvent utilisée en démographie : elle permet de transformer une variable variant sur tout le domaine réel  $]-\infty, +\infty[$  en une variable comprise entre 0 et 1. On utilise aussi la transformation *logit* en statistique pour analyser des données binaires (réponses par oui ou par non dans les enquêtes par exemple) avec les méthodes économétriques classiques.

### Graphiques

Nous avons représenté deux types de graphiques l'un (figure 2.4) où les fonctions  $P$  (effectif de population) et  $\rho$  (taux de croissance) sont exprimées en fonction du temps, l'autre (figure 2.5) où  $P'$  et  $P''$  le sont en fonction de la population  $P$  elle-même. Les courbes notées (I) en trait gras correspondent au cas où l'effectif initial est inférieur à l'effectif limite, celles notées (II), en trait fin, au cas où il lui est supérieur.

### Interprétation démographique

Comme le taux de croissance peut être décomposé (dans une population fermée) entre taux brut de natalité et taux brut de mortalité, il suffit pour obtenir une population logistique que le taux de mortalité,  $m$ , croisse linéairement avec la taille de la population :

$$m = m_0(1 + MP)$$

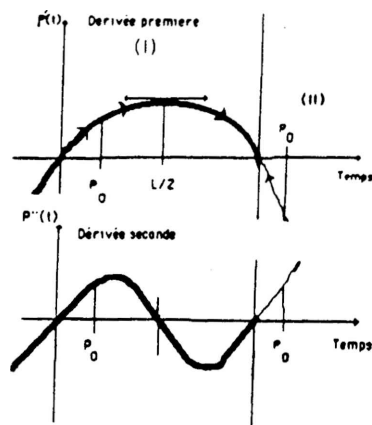


Figure 2.5: Population logistique : représentation de  $P'$  et  $P''$  en fonction de  $P$ . (Partie B).

et que le taux de natalité  $n$  décroisse :

$$n = n_0(1 - NP).$$

On trouve ainsi comme taux de croissance,  $\rho$  :

$$\rho = (n_0 - m_0) - (n_0N + m_0M)P$$

$$\rho = \rho_0\left(1 - \frac{P}{L}\right) \quad \text{avec } \rho \text{ quelconque et } L > 0.$$

### Ajustement d'une population logistique

Il peut être utile d'ajuster une population logistique à une série statistique observée. Les méthodes les plus modernes utilisent l'ordinateur en minimisant une *somme de carrés* [26] mais il existe des méthodes graphiques pratiques :

- On choisit sur une large étendue du graphique, deux points  $P_0$  et  $P_2$  équidistants d'un troisième point médian  $P_1$  qu'on prend comme nouvelle origine. Si  $n$  désigne la distance entre  $P_0$  et  $P_1$  (= distance entre  $P_1$  et  $P_2$ ), les 3 paramètres  $L$ ,  $a$  et  $T$  de l'équation 2.34 valent alors :

$$\begin{aligned} L &= \frac{2P_1P_2P_3 - P_2^2(P_1 + P_3)}{P_1P_3 - P_2^2} \\ a &= \frac{1}{n} \log \frac{P_2(L - P_1)}{P_1(L - P_2)} \\ T &= \frac{1}{a} \log \frac{L - P_2}{P_2} \end{aligned} \quad (2.35)$$

- On peut aussi trouver ces coefficients à l'aide de *nomogrammes* (voir [41] et [39]).

Tableau 2.1: Ajustement de la population des Etats Unis par une loi logistique.

à faire.

La méthode de Keyfitz consiste à développer par la formule de Taylor l'expression de la somme des carrés des écarts entre les valeurs ajustées et observées à partir d'un triplet quelconque  $(L_0, a_0, T_0)$  et à minimiser cette somme en annulant ses trois dérivées partielles par rapport à  $L$ ,  $a$  et  $T$ . D'où un nouveau triplet  $(L_1, a_1, T_1)$  à partir duquel on peut recommencer l'opération. On montre que le processus converge et peut être arrêté lorsqu'un seuil de précision est atteint.

**Application à la population des Etats-Unis** La population des Etats-Unis<sup>4</sup> semble suivre, depuis le XIX<sup>e</sup> siècle à peu près une loi logistique. Le tableau 2.1 et La figure 2.6 montre que les deux ajustements proposés donnent des résultats assez voisins sauf pour les populations limites.

La fonction logistique est très utilisée dans les *études de marché* car il a été remarqué que lorsque qu'un nouveau produit fait son apparition sur le marché (téléphone, voiture, télévision etc.) la courbe qui donne la proportion des détenteurs de ce nouveau produit en fonction du temps est une logistique.

**Exercice 5** Dans un pays  $A$ , le marché des télévisions est récent mais plusieurs enquêtes annuelles auprès des ménages ont permis de mesurer le nombre de ménages déjà équipés. Proposer une méthode pour pouvoir prévoir l'évolution du marché, et surtout prévoir la date où le nombre des importations ou bien la vitesse de fabrication devra être maximum afin de minimiser les stocks.

**Solution 5** Nous proposerons d'ajuster une loi logistique. Il faut alors au moins trois points pour pouvoir estimer  $L$ ,  $a$  et  $T$  selon la formule de l'équation 2.35. Le taux de fabrication  $\frac{dP}{dt}$  est maximum lorsque sa dérivée s'annule, c'est à dire au point d'inflexion de la logistique qui est atteint à la date  $T$ .

<sup>4</sup>Cet exemple est tiré de Gendreau [15] et de [35].



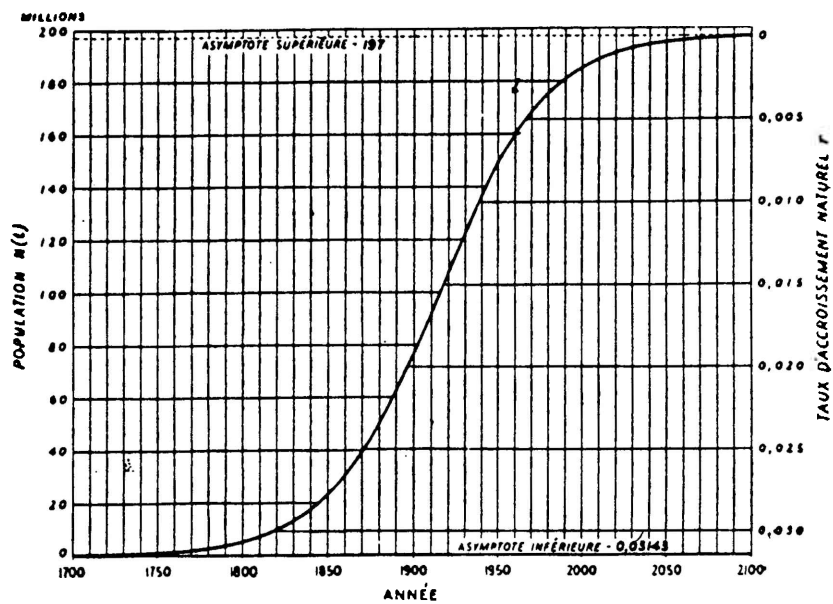


Figure 2.6: Ajustement de la population des Etats-Unis par une loi logistique.

**Exercice 6** Une des expériences de Gause [14] a consisté à disposer des tubes de laboratoire avec 5 protozoaires chacun et une solution nutritive. Il compta chaque jour le nombre de protozoaires dans les tubes. Dans les tous premiers instants, Gause estima la croissance à 230,9 % par jour, puis constata un nombre moyen maximum de 375 protozoaires par tube au bout du quatrième jour. Montrer que l'équation de la courbe logistique est :

$$P(t) = \frac{375}{1 + 74 \times e^{-2.309t}}$$

**Solution 6** Prenons la définition d'une loi logistique (équation 2.22) :

$$\rho = \frac{dP}{P dt} = a - bP.$$

Dans les premiers instants de l'expérience, on peut considérer que l'effectif de la population est pratiquement nul ; le taux de croissance initial vaut alors  $\rho = a = 2,309$  par jour. La population limite  $L = b/a$  vaut 375. L'équation de la courbe logistique est alors (voir équation 2.25) :

$$P(t) = 375 \times \frac{1}{1 + \frac{375-5}{5} e^{-2.309t}}$$

### Stabilité des points d'équilibre

**Définition 7** Une valeur particulière de l'effectif  $P$  sera appelé point d'équilibre du système si la valeur de la dérivée  $\frac{dP}{dt}$  est nulle en ce point.

Dans le cas de la population logistique, nous trouvons deux points d'équilibre, pour  $P = 0$  et  $P = L = \frac{a}{b}$ . Venons en maintenant à discuter brièvement de la stabilité de ces points d'équilibre. Selon la théorie vue précédemment et les graphiques 2.4 et 2.5, lorsque l'effectif initial  $P_0$  est supérieur à  $L$ , la population décroît jusqu'à atteindre la limite  $L$  ; de même lorsque l'effectif initial est inférieur à  $L$ , la population croît jusqu'à cette valeur limite.

**Présentation mathématique des problèmes de stabilité** La forme générale d'un modèle de population sans classe d'âge est régie par l'équation suivante :

$$\frac{dP}{dt} = f[P(t)] \quad f: \mathcal{R} \rightarrow \mathcal{R} \in C^1 \quad (2.36)$$

Exemple :

$$f[P(t)] = aP(t) \left[ 1 - \frac{P(t)}{L} \right]$$

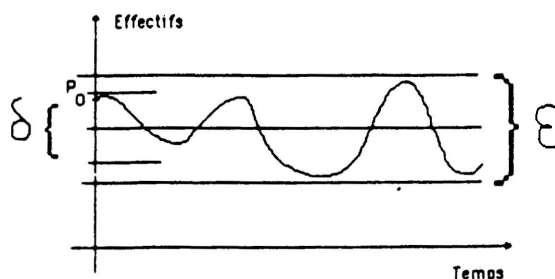


Figure 2.7: Définition d'un équilibre stable et asymptotiquement stable.

1. Une solution, notée  $\varphi(t, P_0)$ , ou *trajectoire* du système 2.36, prenant la valeur  $P_0$  à l'instant  $t = 0$  initial est une fonction  $\varphi$  continument dérivable telle que :

$$\varphi'(t) = f[\varphi(t)] \quad \text{avec} \quad \varphi(0) = P_0 \quad (2.37)$$

On montre qu'une solution existe toujours et est unique.

2. Un point d'équilibre pour le système 2.36 est un point  $P_e$  de  $\mathbb{R}$  vérifiant la relation  $f(P_e) = 0$ .
3. Le point  $P_e$  sera dit *stable* si pour tout intervalle  $\epsilon$  autour du point  $P_e$  il existe un intervalle d'épaisseur  $\delta$  suffisamment fin de telle sorte que toute trajectoire partant de cet intervalle origine( $\delta$ ) soit contenue dans cet intervalle final  $\epsilon$  (voir figure 2.7).

Autrement dit,  $P_e$  est stable si :

$$\forall \epsilon > 0, \quad \exists \delta > 0 \quad \text{tel que} \quad |P_0 - P_e| < \delta \Rightarrow |\varphi(t, P_0) - P_e| < \epsilon, \quad \forall t \geq 0. \quad (2.38)$$

4. Un point  $P_e$  sera dit *asymptotiquement stable*, s'il est *stable* et si :

$$\exists \delta > 0 \quad \text{tel que} \quad |P_0 - P_e| < \delta \Rightarrow \lim_{t \rightarrow \infty} |\varphi(t, P_0) - P_e| = 0.$$

On montre ainsi que si les trajectoires tendent asymptotiquement vers ce point d'équilibre, celui-ci est asymptotiquement stable.

### Exemple

Si l'équation de la population est la suivante :

$$\frac{dP(t)}{dt} = P(t)$$

nous avons les équilibres suivants :

$$\left\{ \begin{array}{l} \lambda > 0 \quad P_e = 0 \quad \text{est instable} \\ \lambda = 0 \quad \text{tous les points sont stables mais} \\ \quad \quad \quad \text{non asymptotiquement stables} \\ \lambda < 0 \quad P_e = 0 \quad \text{est asymptotiquement stable.} \end{array} \right.$$

**Exercice 7** Montrer que dans une population logistique :

$$dP/dt = aP(1 - P(t)/L) \quad (P > 0),$$

le point  $P = 0$  est un point d'équilibre instable et  $L$  un point d'équilibre asymptotiquement stable.

**Solution 7** Le point  $P = 0$  est évidemment un point d'équilibre instable, car dès qu'une trajectoire part d'un effectif  $P_0 = \epsilon > 0$  la trajectoire s'écarte de la ligne  $P = 0$ .

Etant donné que toutes les trajectoires tendent vers la limite  $P = L$ , ce point est bien évidemment instable.

**Exercice 8** Soit le système suivant : (extension du modèle logistique)

$$\frac{1}{P} \frac{dP}{dt} = a(P - \alpha)(P - \beta) \quad \begin{cases} P > 0 \\ a > 0 \\ 0 < \alpha < \beta \end{cases}$$

Trouver les points d'équilibre et discuter leur stabilité.

**Solution 8** Les 3 points d'équilibre ( $P' = 0$ ) sont obtenus pour  $P = 0$ ,  $P = \alpha$  et  $P = \beta$ . Pour  $P = \epsilon > 0$  la dérivée  $P' \simeq a\alpha\beta$  est positive. Ainsi au voisinage de  $P = 0$  la fonction  $P$  est croissante ; le point d'équilibre est donc instable. Par continuité, nous déduisons que la dérivée est positive dans l'intervalle  $]0, \alpha[$ , négative dans l'intervalle  $]\alpha, \beta[$  puis de nouveau positive au delà de  $\beta$ . Pour tout point origine proche de  $P = \alpha$  par valeur inférieure, l'effectif  $P$  croît puisque  $P' > 0$  ; s'il dépassait la valeur  $\alpha$ , la dérivée serait négative et l'effectif devrait de nouveau décroître. Le point  $P = \alpha$  est donc un point d'équilibre stable. Est-il asymptotiquement stable ? Il suffit pour cela de calculer un équivalent au voisinage de  $P = \alpha$  lorsque  $t$  tend vers l'infini. Nous avons dans ce cas en posant :  $P = p + \alpha$ ,

$$\frac{dp}{dt} = (p + \alpha)p(p - (\beta - \alpha))$$

soit pour  $p$  petit :

$$p' \simeq -\alpha(\beta - \alpha)p \simeq -kp \quad \text{avec } k > 0$$

soit encore :

$$p = p_0 e^{-kt}.$$

Chacune des trajectoires  $P(t)$  partant de l'intervalle  $]0, \beta[$ , converge vers le point d'équilibre  $P = \alpha$  qui est donc asymptotiquement stable. On déduit de même que  $\beta$  est instable.

## 2.2 Interaction entre les espèces

Nous avons vu au chapitre 2 que A. Lotka distinguait deux types de dépendance entre les espèces, à savoir la dépendance *coordonnée* et la dépendance *subordonnée*, ici, nous étudierons un modèle de chacun des deux types mais simplifierons en ne considérant qu'une interaction entre deux espèces seulement. Le modèle de dépendance coordonnée étudié ci-dessous s'appelle aussi *modèle de compétition entre deux espèces*.

### 2.2.1 Compétition entre deux espèces

Chacune des deux espèces, notées "1" et "2", est supposée croître suivant un modèle logistique en l'absence de l'autre. Ainsi si  $N_1$  désigne l'effectif de l'espèce 1 à une date  $t$ , l'équation qui donne la valeur de son taux de croissance est  $dN_1/N_1 dt = a_1 - b_{11}N_1$ . Mais la présence de la seconde espèce affecte la croissance de la première et cela d'autant plus que la seconde espèce est nombreuse. On suppose alors que le taux de croissance est diminué d'un facteur  $b_{12}N_2$ . En tenant un raisonnement analogue pour la seconde, on obtient le système différentiel suivant :

$$\begin{cases} \frac{dN_1}{N_1 dt} = a_1 - b_{11}N_1 - b_{12}N_2 \\ \frac{dN_2}{N_2 dt} = a_2 - b_{21}N_1 - b_{22}N_2 \end{cases} \quad (2.39)$$

Il n'est pas possible de résoudre analytiquement un tel système sauf dans un cas particulier que nous verrons plus loin : par contre on peut obtenir les courbes intégrales (c'est-à-dire les solutions  $N_1(t)$  et  $N_2(t)$ ) avec l'aide d'un micro-ordinateur. Ici, nous n'étudierons l'évolution du système que du point de vue *qualitatif*.

Nous avons vu au paragraphe 2.1.2 qu'une espèce est en équilibre, si la dérivée de son effectif par rapport au temps est nulle. Ainsi les points d'équilibre pour l'espèce 1 doivent vérifier :

$$N_1 = 0 \quad \text{et} \quad a_1 - b_{11}N_1 - b_{12}N_2 = 0.$$

Outre le cas trivial  $N_1 = 0$ , l'espèce 1 est en équilibre lorsque les effectifs  $N_1$  et  $N_2$  des deux espèces vérifient l'équation ci-dessus, c'est-à-dire lorsqu'ils appartiennent à une droite (notée  $D_1$  par la suite) du plan  $(N_1, N_2)$ . Plus précisément, nous pouvons dire que cette droite partage le plan en deux régions, l'une positive (l'origine y figure), l'autre négative. Dans la région positive, l'espèce 1 croît et ce d'autant moins que le point  $M(N_1, N_2)$  est proche de la droite, à tel point que si la trajectoire rencontre cette droite, la tangente à cette trajectoire est *verticale*. Un raisonnement symétrique nous amène à considérer une seconde droite  $D_2$  pour laquelle les tangentes aux trajectoires sont horizontales.

$$(D_1) \quad a_1 - b_{11}N_1 - b_{12}N_2 = 0.$$

$$(D_2) \quad a_2 - b_{21}N_1 - b_{22}N_2 = 0.$$

L'évolution de la cohabitation des deux espèces dépend de la position relative des deux droites.

Nous sommes amenés à distinguer quatre cas représentés sur les figures 2.8 à 2.11. Notre démarche empirique consiste, à partir d'un point quelconque du quadrant positif du plan, à considérer les signes des déplacements élémentaires  $dN_1$  et  $dN_2$  en fonction d'un laps de temps  $dt$ . Ce raisonnement nous conduit à nous placer au point de coordonnées  $(N_1 + dN_1, N_2 + dN_2)$ , puis de nouveau à itérer le processus en tenant compte de ce que l'approche d'une des droites  $D_1$  ou  $D_2$  amoindrit la valeur du déplacement d'une des espèces par rapport à l'autre et nous oblige à arrondir la trajectoire de sorte qu'au point d'intersection avec la droite, la tangente soit horizontale pour  $D_2$  et verticale pour  $D_1$ .

Nous trouvons ainsi dans le premier cas que toutes les trajectoires convergent vers le point d'équilibre  $N_1^{lim}$ . L'espèce 2 disparaît et l'espèce 1 se trouve à son point d'équilibre limite (équation logistique simple) :

$$N_1^{lim} = \frac{a_1}{b_{11}}.$$

Le deuxième cas correspond à l'extinction de l'espèce 1. Le troisième cas correspond au seul cas de coexistence entre les deux espèces, le point d'équilibre est le point d'intersection des deux droites. Ces coordonnées sont :

$$\begin{cases} N_1^{eq} = \frac{a_1 b_{22} - a_2 b_{12}}{b_{11} b_{22} - b_{12} b_{21}} \\ N_2^{eq} = \frac{a_2 b_{11} - a_1 b_{21}}{b_{11} b_{22} - b_{12} b_{21}} \end{cases} \quad (2.40)$$

Dans le quatrième cas, le point d'intersection des deux droites est *instable*. En effet toute trajectoire partant près de ce point origine s'en éloigne :

- pour un point situé en bas à droite du quadrant (point  $M_1$ ) la trajectoire va vers le point  $N_1^{lim}$ ,
- pour un point situé en haut à gauche, celle-ci se dirige vers le point  $N_2$ .

Ce raisonnement laisse entendre qu'il existe une ligne de démarcation qui sépare le quadrant en deux zones : l'une où l'espèce 1 gagne, l'autre où l'espèce 2 gagne.

**Remarque** Dans le cas particulier où les deux droites sont parallèles, la solution analytique du problème est simple. Posons en effet la relation :

$$\frac{b_{12}}{b_{11}} = \frac{b_{22}}{b_{21}} = \lambda;$$

et nous pouvons considérer une population moyenne,  $N = N_1 + \lambda N_2$ , puis écrire :

$$\begin{cases} \frac{dN_1}{N_1 dt} = a_1 - b_1 N & \text{avec } b_1 = b_{11} \\ \frac{dN_2}{N_2 dt} = a_2 - b_2 N & \text{et } b_2 = b_{22} \end{cases} \quad (2.41)$$

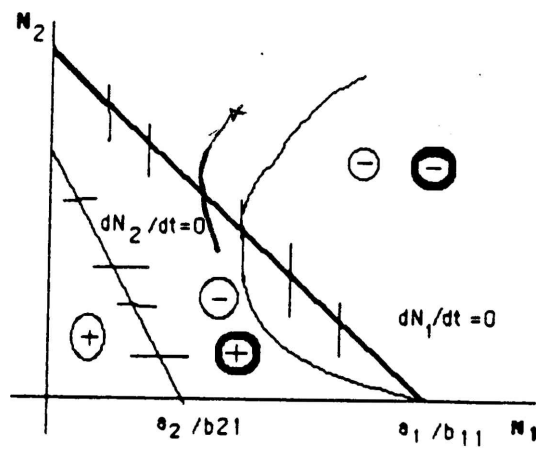
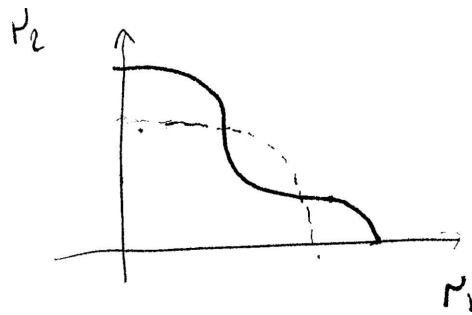


Figure 2.8: L'espèce 2 disparaît.

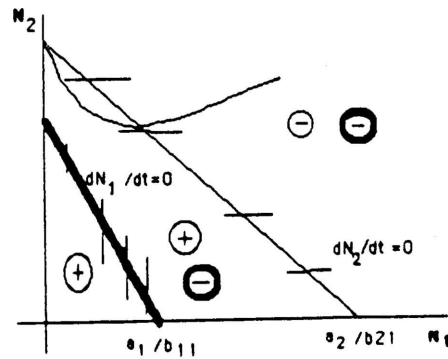


Figure 2.9: L'espèce 1 disparaît.

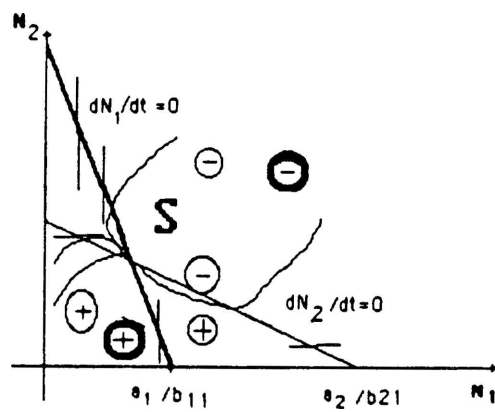


Figure 2.10: Équilibre stable entre les deux espèces.

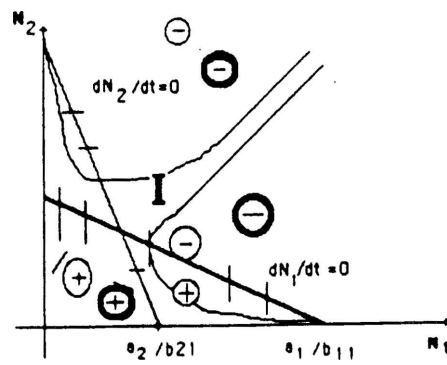


Figure 2.11: Équilibre instable.



Le taux de croissance de chacune des espèces est régi par la taille d'une population  $N$ , moyenne fixe des populations des deux espèces. En éliminant  $N$  entre les deux équations nous trouvons :

$$\frac{1}{b_1} \left( \frac{dN_1}{N_1 dt} - a_1 \right) = \frac{1}{b_2} \left( \frac{dN_2}{N_2 dt} - a_2 \right),$$

$$\frac{1}{b_1} \frac{dN_1}{N_1} - \frac{dN_2}{N_2 b_2} = \left( \frac{a_1}{b_1} - \frac{a_2}{b_2} \right) dt,$$

soit finalement :

$$\frac{N_1^{b_2}}{N_2^{b_1}} = \frac{N_{10}^{b_2}}{N_{20}^{b_1}} e^{(a_1 b_2 - a_2 b_1)t} \quad (2.42)$$

$N_{10}$  et  $N_{20}$  étant les effectifs initiaux des deux espèces lorsque  $t = 0$ . On voit alors que si  $(a_1 b_2 - a_2 b_1)$  est négatif, l'espèce 1 devient minoritaire devant l'espèce 2 et finit par disparaître. Nous allons voir maintenant un modèle simple

de dépendance *subordonnée* entre deux espèces. L'origine de ce modèle est attribué à A. Lotka<sup>5</sup> et à Volterra [45] et on parle des *équations de Lotka-Volterra*.

### 2.2.2 Modèle proie-prédateur ou hôte-parasite

On suppose qu'une population d'hôtes (notée  $H$ ) a un taux de croissance constant en l'absence d'une population parasite (notée  $P$ ). Le modèle le plus simple qui traduit le fait que plus les parasites sont nombreux, plus les hôtes sont gênés s'écrit ainsi :

$$\frac{dH}{H dt} = a_1 - b_1 P \quad (2.43)$$

Les parasites sont condamnés à disparaître en l'absence de leur nourriture (les hôtes). Le taux de croissance *naturel* de la population des parasites (ou des prédateurs) est *négatif* ( $-a_2$ ), mais la présence des hôtes (ou des proies) en quantité  $H$  peut inverser le signe du taux de croissance. L'équation du taux de croissance des parasites est alors la suivante :

$$\frac{dP}{P dt} = -a_2 + b_2 H \quad (2.44)$$

On se propose d'étudier l'évolution d'un tel système.

L'élimination du facteur  $dt$  entre les équations 2.43 et 2.44 donne :

$$dt = \left( -\frac{a_2}{H} + b_2 \right) dH = \left( \frac{a_1}{P} - b_1 \right) dP,$$

<sup>5</sup>A. Lotka op.cit.

ce qui s'intègre en :

$$a_2 \log H - b_2 H + a_1 \log P - b_1 P = a_2 \log H_0 - b_2 H_0 + a_1 \log P_0 - b_1 P_0 = Cste \quad (2.45)$$

où  $H_0$  et  $P_0$  sont des conditions initiales. On montre que cette courbe *fermée* ; mais pour étudier plus exactement la dépendance des deux espèces, on se propose de faire un changement de variable et de ramener l'origine au point d'équilibre triviale :

$$H_c = \frac{a_2}{b_2}, \quad P_c = \frac{a_1}{b_1}. \quad (2.46)$$

Posons alors :

$$H = h + \frac{a_2}{b_2}, \quad P = p + \frac{a_1}{b_1} \quad (2.47)$$

et nous obtenons les équations :

$$\begin{cases} \frac{dh}{dt} = -\frac{b_1 a_2}{b_2} p - b_1 h p, \\ \frac{dp}{dt} = +\frac{b_2 a_1}{b_1} h - b_2 h p. \end{cases}$$

Au voisinage de l'équilibre ( $h = 0, p = 0$ ), on peut négliger les infiniments petits du second ordre ( $hp$ ), puis éliminer  $dt$  entre les deux équations :

$$dt = -\frac{dh}{B^2 p} = \frac{dp}{A^2 h},$$

avec :

$$A^2 = a_1 \frac{b_2}{b_1} \quad \text{et} \quad B^2 = a_2 \frac{b_1}{b_2}.$$

Il vient alors, par intégration :

$$A^2 h^2 + B^2 p^2 = A h_0^2 + B^2 p_0^2 = C^2. \quad (2.48)$$

Il s'agit d'une ellipse qui est décrite dans le sens des aiguilles d'une montre, comme l'indique la figure 2.12. L'équation du mouvement est obtenue par intégration d'une des deux équations en  $dt$  (on suppose  $A, B$  et  $C$  positifs) :

$$dt = -\frac{dh}{B^2 \sqrt{\frac{C^2 - A^2 h^2}{B^2}}} = -\frac{d(Ah)}{AB \sqrt{C^2 - (Ah)^2}},$$

soit, en intégrant :

$$t - t_0 = \frac{1}{AB} \arccos \frac{Ah}{C},$$

ou encore :

$$h = \frac{C}{A} \cos(AB(t - t_0)). \quad (2.49)$$

Un raisonnement symétrique pour les parasites nous conduit à écrire le système final :

$$\begin{cases} H = \frac{C}{A} \cos(ABt) + H_c \\ P = \frac{C}{B} \sin(ABt) + P_c \end{cases} \quad \text{avec } t_0 = 0. \quad (2.50)$$

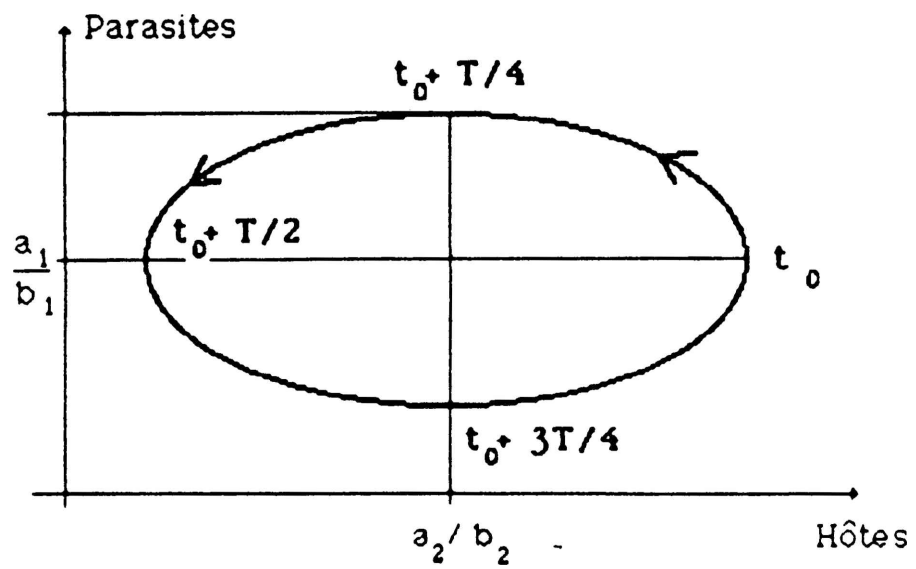


Figure 2.12: Modèles proie-prédateur ou hôte-parasite. Trajectoires elliptiques au voisinage du point d'équilibre.

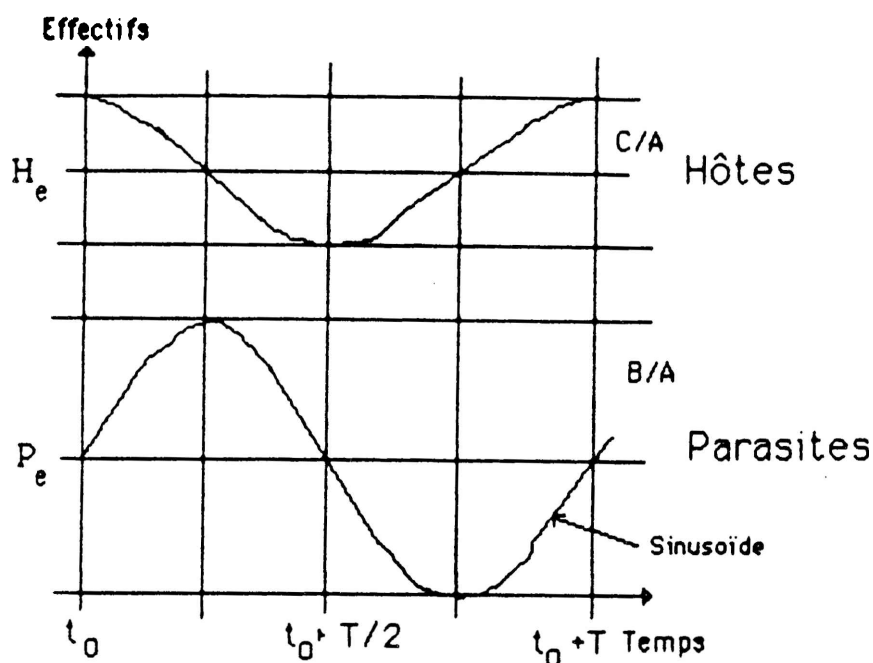


Figure 2.13: Modèles proie-prédateur ou hôte-parasite. Évolution des effectifs en fonction du temps

L'évolution au cours du temps de chacune des espèces est régie par une sinusoïde, celle des parasites étant en retard d'un quart de période par rapport à celle des hôtes. La période  $T$  est égale à :

$$T = \frac{2\pi}{AB} = \frac{2\pi}{\sqrt{a_1 a_2}} \quad (2.51)$$

La figure 2.13 représente les deux sinusoïdes relatives aux *effectifs* ainsi que celles relatives aux *taux de croissance*. On peut interpréter une telle dépendance de la façon suivante : lorsque au temps  $t = 0$ , les hôtes sont en *quantité* maximale, la *croissance* des parasites est elle-même maximale. Le nombre des parasites augmente et gêne la population des hôtes dont le taux de croissance déjà nul devient négatif. La population des hôtes décroît. Le taux de croissance des parasites diminue puis finit par s'annuler pour  $t = \frac{T}{4}$ . La population des parasites passe ainsi par un maximum avant de décroître ce qui permet au taux de croissance des hôtes de croître de nouveau, etc etc.

Une telle dépendance entre deux espèces a pu être observée entre des lapins et des renards au Groënland par exemple (voir Volterra op.cit.).

**Exercice 9** Une première espèce ( $H$ ) se développe au taux de croissance ins-

tantané (exprimé en mois) de 100 %. Cette première espèce est parasitée par une seconde ( $P$ ), qui en l'absence de la population des hôtes serait condamnée à l'extinction avec un taux de croissance négatif de 50 %. Les parasites peuvent survivre grâce aux hôtes qui, en quantité suffisante, peuvent faire inverser le signe de leur taux de croissance en y ajoutant la quantité  $0,002H$ . La présence des parasites au contraire fait décroître le taux de croissance naturel des hôtes de la quantité  $0,01P$ .

On demande de calculer la période du cycle produit. La valeur maximale du nombre des hôtes étant de 300 individus, on demande de calculer le nombre maximal de parasites que l'on est censé observer, et à quelles dates ce maximum sera atteint ? (On considérera que les oscillations sont petites).

**Solution 9** Les équations du système sont les suivantes :

$$\begin{cases} \frac{dH}{dt} = (1 - 0,01P)H \\ \frac{dP}{dt} = (-0,5 + 0,002H)P \end{cases}$$

D'après le cours, la période du cycle produit est  $T = \frac{2 \times 3,14159}{\sqrt{1 \times 0,5}} = 8,896$  mois. L'équilibre trivial est obtenu pour  $H = 250$  et  $P = 100$ . Si la valeur maximale du nombre des hôtes est 300, les oscillations autour de ce point d'équilibre seront de faible amplitude et le mouvement sera régi suivant une ellipse d'équation :

$$(0,002)^2 \times 1 \times h^2 + (0,01)^2 \times 0,5p^2 = (0,002)^2 \times (300 - 250)^2 = 0,01.$$

L'équation s'écrit alors :

$$4h^2 + 50p^2 = 10\,000.$$

Le nombre maximal de parasites vaut donc  $14,14 + 100 = 114$ . Ce nombre est atteint 2,22 mois après que le nombre maximal d'hôtes ait été observé.

**Exercice 10** La croissance d'une espèce 1 est celle d'une population logistique. Quand son effectif est faible le taux de croissance est de 6 %, au contraire il est nul quand la population atteint sa taille limite de 200 individus. On considère une seconde espèce qui suit une loi du même type (croissance 4 %, population limite 300).

1. Calculer les points d'inflexion des courbes logistiques.
2. La première espèce peut être amenée à cotoyer la seconde, et les deux populations peuvent ainsi se disputer réciproquement la nourriture de leur territoire commun. Si les deux espèces vivent dans le même milieu, il convient de diminuer les taux de croissance des quantités  $0,0015N_2$  pour l'espèce 1 et de  $0,0016N_1$  pour l'espèce 2. Ecrire les équations différentielles régissant la dynamique des deux espèces.
3. On recense à une date donnée 50 individus de l'espèce 1 et 300 de l'espèce 2. Décrire l'issue d'une telle compétition.

4. Supposons maintenant que les deux espèces soient régies par le système suivant :

$$\begin{cases} \frac{dN_1}{N_1 dt} = 0,04 - 0,00016N_1 - 0,0001333N_2 \\ \frac{dN_2}{N_2 dt} = 0,06 - 0,0003N_1 - 0,00015N_2 \end{cases}$$

et que les quantités initiales des deux espèces soient identiques à celles de la question 3. Quel nouveau discours pourriez-vous tenir ?

**Solution 10** 1. Dans une population logistique le point d'inflexion est obtenu quand la population atteint la moitié de sa taille limite, soit 100 pour l'espèce 1 et 150 pour l'espèce 2.

2. Les équations différentielles s'écrivent :

$$\begin{cases} \frac{dN_1}{N_1 dt} = 0,06 - 0,0003N_1 - 0,00015N_2 \\ \frac{dN_2}{N_2 dt} = 0,04 - 0,00016N_1 - 0,0001333N_2 \end{cases}$$

3. Il suffit de tracer les deux droites pour lesquelles les deux dérivées s'annulent. Le point de rencontre  $N_{1e} = 125$  et  $N_{2e} = 150$  est un point d'équilibre stable. Quelque soit le point de départ, les deux espèces se retrouveront à ce point d'équilibre.
4. Comme on a inversé les deux droites, le point d'équilibre devient instable. Au point  $N_1 = 300$  et  $N_2 = 300$ , le taux de croissance de l'espèce 2 est nul. Celui de l'espèce 1 est au contraire négatif. L'espèce 1 va décroître en nombre, et l'espèce 2 va pouvoir croître jusqu'à éliminer l'espèce 1.

## 2.3 Modèles de pêche, chasse, émigration

Il s'agit d'étudier les réactions d'une population lorsqu'on effectue un *prélèvement* par la chasse où la pêche. Dans le cas d'une population humaine le "prélèvement" correspond à une émigration. Répétons que ces modèles sont théoriques et donc bien loin de la réalité qui est plus complexe, et ce d'autant plus qu'on est essayé de modéliser le comportement humain et ne se limite plus au monde animal. Ces modèles ont néanmoins l'intérêt de faire comprendre certains mécanismes simples qui peuvent être approfondis ultérieurement.

Les modèles classiques de prélèvement par pêche ou chasse peuvent être classés en deux types extrêmes, l'un qu'on qualifie de pêche par "quota", l'autre de pêche "à effort de pêche constant".

- Dans la pêche par quota, on considère que les pêcheurs ne rentrent pas avant d'avoir ramené le quota de pêche qu'ils se sont fixés, quitte à travailler plus longtemps si la pêche est difficile ou à rentrer rapidement dans les cas contraire.

- Dans la pêche à effort constant, on suppose, par exemple, que le chalut est traîné un certain nombre d'heures par jour. Le volume de la pêche dépend de l'abondance des poissons.

Nous allons aborder successivement ces deux modèles du point de vue mathématique en prenant le modèle logistique comme modèle de croissance de base pour l'espèce en l'absence de pêche.

### 2.3.1 Pêche par quota

L'hypothèse de la pêche par quota se traduit mathématiquement par le fait que le volume de poisson pêché  $dP_p$ , dans l'instant<sup>6</sup>  $dt$  est égale à une constante  $q$ .

En l'absence de pêche, et dans le cadre d'une population logistique, la croissance naturelle  $dP_n$  durant le même instant  $dt$  est  $aP(1 - P/L)$ . Si la population est soumise à un prélèvement, la croissance globale de l'espèce est le solde (positif ou négatif) de la croissance naturelle, et du taux de prélèvement. La nouvelle équation est alors :

$$dP = dP_n - dP_p = \left[ aP \left( 1 - \frac{P}{L} \right) - q \right] dt \quad (2.52)$$

La discussion graphique est la plus facile. Sachant qu'en l'absence de pêche, la dérivée  $dP/dt$  est la parabole d'équation  $aP(1 - P/L)$ , son sommet est atteint pour  $P = \frac{L}{2}$ , ce qui correspond à  $P' = a\frac{L}{4}$ . Nous sommes amenés à distinguer 3 cas :

$q > \frac{aL}{4}$  le volume de pêche demandé est trop important, l'espèce s'épuise et aucun équilibre n'est atteint.

$0 < q < \frac{aL}{4}$  la droite d'équation  $P' = q$  coupe la parabole en deux points  $P_{min}$  et  $P_{max}$  d'équilibre possible, solution de :

$$a \frac{x^2}{2} - ax + c = 0.$$

Discutons la stabilité de ces points. Admettons qu'à un instant donné  $t_0$ , la population  $P_0$  soit comprise entre  $P_{min}$  et  $P_{max}$ ; dans ce cas, la croissance naturelle de l'espèce est supérieure à la quantité pêchée  $q$  par unité de temps, l'espèce croît donc jusqu'à atteindre la taille d'équilibre  $P_{max}$ . Si l'espèce dépassait ce nombre  $P_{max}$ , la quantité pêchée par unité de temps serait supérieure à sa croissance naturelle (qui s'est elle-même amoindrie du seul fait que l'espèce est gênée par sa propre taille importante), l'espèce devrait donc décroître jusqu'à l'équilibre  $P_{max}$  qui est donc un équilibre stable. Au contraire, si l'espèce se trouvait en un nombre inférieur à

<sup>6</sup>Cet instant peut être un jour ou un mois.

$P_{max}$ , le développement naturel de l'espèce, trop faible en raison d'une population trop peu nombreuse, ne compenserait pas les pertes dues à la pêche et l'espèce disparaîtrait.

$q = \frac{aL^2}{4}$  dans ce cas, le domaine de stabilité autour de  $P_{max}$  est réduit à zéro pour les valeurs inférieures, l'équilibre est donc instable.

### 2.3.2 L'effort de pêche est constant dans le temps

L'hypothèse implique que le volume de pêche soit proportionnel à l'effectif de l'espèce, soit  $\frac{dP_p}{dt} = eP$  où  $e$  est l'effort de pêche. L'équation différentielle régissant le mouvement de l'espèce est alors :

$$\frac{dP}{dt} = \frac{dP_n}{dt} - \frac{dP_p}{dt} = aP\left(1 - \frac{P}{L}\right) - eP. \quad (2.53)$$

Il s'agit de l'équation d'une autre population logistique ou la population limite s'est déplacée pour atteindre :

$$L_e = \left(1 - \frac{e}{a}\right)L.$$

Le seul point d'équilibre possible est donc stable. Si  $e \geq a$ , l'effort de pêche est trop important, aucun équilibre non nul n'est possible, l'espèce disparaît.

**Effort de pêche optimum** La pêche sera maximale si  $\frac{dP_p}{dt}$  est maximale, soit pour  $e(P)P$  maximum, soit encore  $\frac{dP_n}{dt} = aP\left(1 - \frac{P}{L}\right)$  maximum. Or le renouvellement est maximum pour  $P = \frac{L}{2}$ , soit un effort optimum de :

$$e_{opt} = \frac{aL}{4} \left(\frac{2}{L}\right) = \frac{a}{2}.$$

Ce rendement optimum a pour effet de maintenir la population à la moitié de sa taille limite en l'absence de pêche.

### 2.3.3 Conclusion

Le prélèvement avec effort de pêche constant est bien préférable à celui avec quota de pêche constant. En effet, dans le premier cas la recherche empirique du rendement maximum est sans danger pour l'espèce car autour de la valeur optimale l'équilibre est toujours stable.

Dans le cas du quota, le désir d'augmenter le quota rend l'équilibre stable  $P_{max}$  de plus en plus précaire au fur et à mesure que l'on se rapproche du maximum qui lui est instable.



## Chapitre 3

# Introduction à l'analyse statistique de la mortalité par une étude de cas

If Iford

Pour étudier la mortalité d'un pays et son évolution, la donnée la plus simple, mais la plus nécessaire est le décompte régulier de ces décédés. Ces données ne sont pas disponibles la plupart du temps en Afrique du fait d'un état civil trop récent, mais l'objet de ce chapitre est de faire comprendre quand et comment l'homme s'est intéressé à mesurer quantitativement la mortalité, et quels sont ses outils de mesure. Il n'est pas possible de prendre comme référence un pays d'Afrique, car il est nécessaire de posséder des données par âge simple durant une période suffisamment longue pour comprendre facilement, sur des exemples concrets, l'intérêt de ces outils. Notre étude de cas portera sur un pays dont les statistiques, fiables, existent depuis plusieurs siècles, à savoir la France. Nous pourrions prendre d'autres pays comme l'Angleterre (que nous mentionerons parfois pour des exemples précis), la Suède ou la Norvège, mais d'une part la démographie historique semble particulièrement développée en France depuis les travaux de L. Henry [24] et d'autre part l'auteur de ce cours connaît sûrement mieux ce pays.

Revenons maintenant à l'étude du décompte des décès dans un pays. Les intervalles durant lesquelles les relevés sont effectués peuvent différer selon la nature des statistiques (un mois, un trimestre, une année, un groupe d'années...). Qui plus est, la taille de la population affectée par ces décès varie au cours de l'intervalle ou d'un intervalle à l'autre en raison de la mortalité elle-même, mais aussi de la natalité et des migrations. Pour éliminer ces perturbations, il convient de rapporter les décès de la période considérée à la valeur moyenne de

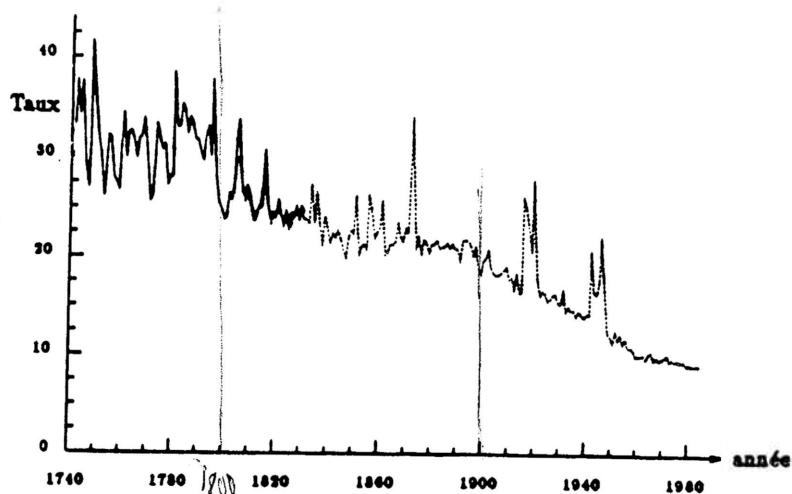


Figure 3.1: Évolution du taux brut de mortalité en France depuis 1740.

cette population. Ce rapport, homogène à l'inverse d'un temps<sup>1</sup>, est conventionnellement appelé *taux brut de mortalité*. Il est généralement ramené à une dimension annuelle. Cet indice simple va nous permettre de suivre l'évolution historique de la mortalité, mais nous verrons qu'il ne reflète pas l'évolution de la mortalité *pure*, car il est sensible aux déformations de la structure par âge provenant d'autres phénomènes que la mortalité. Nous serons donc amenés à affiner nos mesures, en introduisant progressivement les différentes notions qui conduisent à la construction d'une *table de mortalité* et à ses deux principaux modes de calcul *du moment* et *par génération*.

### 3.1 Le taux brut de mortalité

Grâce aux études de démographie historique (L. Henry)<sup>2</sup> fondées sur un échantillon de registres paroissiaux, certains pays comme l'Angleterre, la Suède ou la France disposent des données nécessaires au calcul du taux brut de mortalité depuis plusieurs siècles (1740 en France) alors que la statistique administrative (état civil, recensement) de tels pays ne donne les résultats nécessaires que depuis le 19<sup>e</sup> siècle (1806) et encore de façon très parcellaire. Ainsi la figure 3.1 (source [34]) retrace l'évolution du taux brut de mortalité depuis le milieu du 18<sup>e</sup> siècle. La baisse de la mortalité qu'il traduit est fondamentale pour l'histoire de la société française et est de nature comparable pour les autres pays.

Oscillant autour de 40 pour mille en 1740, le taux brut est descendu aux en-

<sup>1</sup>voir le chapitre 1.

<sup>2</sup>Voir notamment à ce sujet le numéro spécial de *Population* sur la Démographie historique, Paris, INED, Novembre 1975.

virus de 11 pour mille pour les années récentes. Ces deux siècles d'observation correspondent à la phase cruciale ~~pour un pays que l'on a coutume d'appeler la transition démographique~~ <sup>de</sup> et qui ~~nous a conduit en matière de mortalité~~ d'un régime ancien à forte mortalité à un régime nouveau à faible mortalité. En fait, les années 1750-1795 constituent un pallier maximal durant lequel le taux brut oscille autour de 35 pour mille<sup>4</sup>. Sauf à considérer la période 1740-1795 comme une première phase de baisse ou la fin d'une première phase de baisse plutôt qu'une forte fluctuation, la longue période de baisse n'a réellement commencé qu'en 1795. Cette baisse s'accompagne de fluctuations qui, très importantes au départ, vont en s'estompant avec le temps jusqu'à devenir presque imperceptibles (hormis les guerres) au 20<sup>e</sup> siècle. Le régime démographique ancien, antérieur au 19<sup>e</sup> siècle était en effet un régime de crise où de nombreuses disettes, guerres ou épidémies venaient brutalement aggraver une mortalité déjà élevée en temps ordinaires. Au début de la période de transition, les crises alimentaires et les épidémies (choléra notamment) ont persisté bien que leurs effets soient devenus moins violents que ceux de la guerre (Révolution, guerres impériales). Progressivement, elles ont reculé, tandis que le niveau de mortalité des années ordinaires diminuait. Aujourd'hui elles ont disparu et seules les 3 guerres franco-allemandes perturbent l'évolution régulière des 12 dernières décennies. La grippe de 1969 se voit à peine sur la figure 3.1.

Abstraction faite des crises, la mortalité moyenne baisse rapidement de 1790 à 1810, puis plus lentement de 1810 à 1850, mais stagne durant la seconde moitié du 19<sup>e</sup> siècle jusqu'à la guerre de 1870. De 1895 à 1970, au contraire, la baisse s'opère à un rythme rapide et continu.

Le développement d'un pays et la chute de sa mortalité sont évidemment en interaction sur le long terme ; pourtant les ruptures de rythmes qui caractérisent la baisse de la mortalité en France, n'ont pas les mêmes équivalents démographiques.

<sup>3</sup> Le terme de transition démographique évoque aussi le passage d'un régime <sup>de</sup> haute fécondité à un régime à basse fécondité. Dans la plupart des pays le déclin de la fécondité <sup>de but</sup> après le déclin de la mortalité, induisant durant cette transition une forte croissance de la population. Dans le cas de la France, la fécondité baisse en même temps que la mortalité ce qui est une des causes de la faible croissance de la population française comparée à celle des pays voisins.

Le lecteur peut se demander pourquoi les démographes considèrent que la transition démographique est terminée dans les pays développés alors que la mortalité et la fécondité semblent pouvoir encore baisser. On considère en effet que la transition est terminée lorsque la mortalité, depuis la naissance jusqu'à l'âge moyen à la reproduction (aux environs de 27 ans pour les femmes et 29 pour les hommes) est pratiquement nulle et qu'ainsi, une nouvelle baisse de la mortalité ne peut plus compenser une baisse concomitante de fécondité et assurer le renouvellement des générations.

La situation actuelle (faible mortalité, faible fécondité) des pays développés est qualifiée de nouvel équilibre par les démographes, mais il ne faut plus compter sur une baisse de la mortalité, pour contrebalancer une éventuelle baisse de la fécondité, si ce n'est par le vieillissement de la population (voir le paragraphe 6.2.2 du chapitre 6). La population peut continuer à croître mais en vieillissant.

<sup>4</sup> En effet, une population ne peut survivre avec des taux de mortalité supérieurs à 35 pour mille car la natalité (entendons le taux brut de natalité) maximale, celle qui correspond à ce que les démographes ont appelé la fécondité naturelle, ne peut dépasser 40 pour mille.

Ainsi, s'il est compréhensible que l'essor économique de la fin du 18<sup>e</sup> siècle et du début du 19<sup>e</sup> s'accompagne d'une chute de la mortalité, le lien entre la forte industrialisation et la stagnation de la mortalité qui caractérisent le Second Empire est moins évident. Il est vraisemblable qu'à cette époque le progrès économique ait été obtenu aux dépens de la santé de fractions importantes de la population et notamment du prolétariat naissant.

A compter de la fin du XIX<sup>e</sup> siècle, au contraire, accélérée par la révolution pasteurienne, le progrès médical a eu d'autant plus d'efficacité que la société s'était adaptée au nouveau système économique en se dotant de règles de protections sociales qui n'ont cessé de se perfectionner au XX<sup>e</sup> siècle.

Le taux brut de mortalité nous a permis de décrire rapidement le passage historique du régime ancien de forte mortalité au régime actuel de faible mortalité, mais nous ne savons toujours pas quelle pouvait être la durée de vie moyenne au milieu du 18<sup>e</sup> siècle, ni ce qu'elle est devenue de nos jours. Pourtant, le lecteur doit pressentir que lorsque les conditions démographiques sont stationnaires, comme c'était plus ou moins le cas dans la période 1750-1795, le taux brut doit nous renseigner sur la vie moyenne. En effet, supposons que vous soyez préposé à changer les lampes à néon d'une grande salle (qui en contient mille) dès qu'elles ne fonctionnent plus, et que vous ayez besoin en moyenne de 40 lampes neuves par mois. Quelle est alors la durée de vie moyenne d'une lampe ? Votre réponse est sûrement  $\frac{1}{40} = 25$  mois. Ce résultat intuitif est en effet correct, et on peut considérer que la France et plus généralement tous les pays avant leur transition, vérifiaient assez largement les conditions de ce théorème. La durée moyenne de vie en France avant la transition oscillait donc entre  $\frac{1}{30} = 33,3$  ans et  $\frac{1}{40} = 25$  ans.

On voit néanmoins tout de suite les limites d'un tel raisonnement, car si le taux brut actuel en France par exemple, atteint 10 pour mille, la durée de vie moyenne devrait être de 100 ans ! De même le taux brut de mortalité de la Chine lors de son recensement de 1980, était de 7 pour mille, en déduire une espérance de vie de 142 ans serait évidemment une erreur grossière ! En période de changement démographique, en effet, le théorème ne se vérifie plus et l'inverse du taux brut de mortalité est souvent bien supérieure à l'espérance de vie. Dans le cas de la France nous verrons que les conditions de stationnarité ne sont pas requises pour beaucoup de raisons, dont en particulier, l'impact des deux dernières guerres mondiales sur les naissances. Dans le cas de la Chine, cas typique d'un pays du Tiers Monde, cela tient à la forte fécondité passée, et donc au taux de croissance important.

Le taux brut de mortalité nous indique tout de même que les Français du milieu du 18<sup>e</sup> siècle avaient une durée de vie moyenne de l'ordre d'une trentaine d'années, et que celle-ci a fortement augmenté. Mais outre que la valeur de cette durée est imprécise, il ne s'agit que d'une moyenne qui nous renseigne en rien sur l'importance de la variabilité entre les individus. Était-il fréquent d'atteindre 40, 60 ou 80 ans au milieu du 18<sup>e</sup> ? Pour aller au delà et calculer précisément la durée de vie de l'époque actuelle, et sa variabilité entre les individus, il faut

étudier le processus qui, d'âge en âge, conduit chaque individu à la mort.

### 3.2 Le processus de mort

Le processus que nous allons étudier, sous entend une certaine continuité et il peut paraître ambitieux de vouloir le cerner pour l'Ancien Régime où la mortalité, nous l'avons vu, sévit surtout sous forme de crise. Pourtant le premier homme à l'avoir découvert a vécu sous cet Ancien Régime. Il s'agit de J. Graunt, un drapier de Londres du milieu du 17<sup>e</sup> siècle [17], qui pour la première fois, a osé classer puis compter les morts comme des marchandises, afin d'en étudier la loi de survie.

J. Graunt est parti de l'idée que chaque individu a une certaine capacité de résistance à la mort qui dépend uniquement de son âge. Il fait l'hypothèse que tous les individus du même âge courent un risque égal de mortalité et que ces risques successifs constituent le processus de mort. L'hypothèse simplificatrice est de faire abstraction des différences entre individus autres que l'âge (état de santé, expositions particulières à certains risques). Le calcul du risque prend en compte ces diversités, mais ne conserve qu'une valeur moyenne. Un fois calculé, ce risque est supposé identique pour tous les individus d'un même âge, et seul, le hasard décide que tel individu meurt alors que tel autre survit<sup>5</sup>.

Cette approche de la mortalité a aussi été celles de Kersseboom [25] (1738), un hollandais, de Halley [18]<sup>6</sup> (1693), un autre anglais, de Deparcieux (1746) [10], un français, et de Per Wargentin (1766) [46], un suédois<sup>7</sup>. Devenue classique en théorie statistique et intitulée *processus de mort*, cette approche conduit à remplir des colonnes de chiffres pour chaque âge, et c'est pourquoi les démographes l'évoque sous le nom de *table de mortalité*. Essayons d'en comprendre les fondements avant d'en aborder la construction.

L'intérêt nouveau que ces auteurs portaient à l'étude de la mortalité avait au moins deux buts.

- Le premier était de pouvoir comparer la taille puis l'accroissement des grandes villes européennes, ce qui n'a pas manqué de susciter une polémique<sup>8</sup> sur la valeur du facteur (30, 35 ?) par lequel il fallait multiplier le nombre annuel des décès ou de naissances pour obtenir la population.

<sup>5</sup>Ceci n'interdit pas de faire des analyses par groupe, homme-femme, ville-campagne etc ; au sein d'un groupe, le raisonnement est alors le même.

<sup>6</sup>La comète qui porte son nom est plus connue que sa table de mortalité, mais cette dernière constitue la première table élaborée scientifiquement.

<sup>7</sup>Après la publication en 1754 et 1755 dans les *Actes de l'Académie des Sciences de Stockholm* de 5 traités généraux sur les problèmes statistiques de la démographie dont 3 abordent la mortalité, la communication principale de P. Wargentin sur la mortalité en Suède est publiée en 1766. Le lecteur peut se reporter à une traduction de J. Dupaquier [46]

<sup>8</sup>Kersseboom, dans son troisième traité, op. cité se défend d'avoir voulu comparer Londres et Paris, mais tente néanmoins de justifier les chiffres critiqués.

- Le second était d'ordre plus lucratif. En effet, découvrir l'existence d'une loi de mortalité c'était offrir un moyen raisonnable de parier sur la vie des hommes et donc d'étendre le champ d'intervention des assureurs et des banquiers, et d'instituer les rentes viagères<sup>9</sup>. N'est-ce pas à cette époque que Lorenzo Tonti, un banquier italien, proposa à l'état français en 1653, de renflouer ses caisses en, émettant un emprunt d'état dont le taux d'intérêt varierait en fonction de l'âge du souscripteur. Cette première *Tontine Royale* fut tout d'abord refusée, mais elle fut de nouveau proposée en 1689 et, cette fois, acceptée<sup>10</sup>.

### 3.3 La mesure des risques de mortalité par âge

Mettons nous dans l'état d'esprit des pionniers de ce domaine, qui ignoraient tout de la mortalité, et essayons, comme eux, d'étudier le processus de mort, mais avec des données de notre époque. Notre démarche s'apparente plus aux travaux de P. Wargentin en 1766, qui a, le premier, relié des données d'état civil à celles d'un recensement, mais celui-ci a bénéficié des travaux des auteurs déjà cités, depuis un siècle.

L'organisme central de la statistique d'un pays dépouille chaque année les bulletins individuels de décès établis par les officiers d'état civil, et en tire des tableaux statistiques. Le plus simple et le plus important d'entre eux est celui qui classe les décès de l'année suivant l'âge des défunts ; la figure 3.2 en donne l'histogramme<sup>11</sup> pour l'année 1975, en France.

La simple vue de cet histogramme nous montre que la majorité des décès a concerné des individus âgés entre 60 et 90 ans. Nous pourrions être tentés de calculer l'âge moyen de ces décédés et de dire qu'il s'agit de la durée de vie moyenne en France, mais ce raisonnement n'est pas correct. En effet, nous remarquons immédiatement une échancrure importante vers 55-59 ans, qui doit nous rappeler que cette distribution est sujette aux déformations de la structure par âge de la population. La figure 3.3, qui représente cette structure par âge lors du recensement de l'année 1975, indique en effet qu'il s'agit de classes creuses nées durant la guerre de 1914-1918. Or, nous cherchons une distribution, analogue à celle de la figure 3.2 mais qui ne soit pas sensible à l'histoire passée, et qui reflète la mortalité actuelle, compte tenu des conditions sociales, médicales, économiques actuelles de la France. Le passage de la distribution des décès bruts à celle recherchée, n'est pas immédiat, et nécessite un raisonnement assez élaboré. Il consiste tout d'abord à calculer des risques par âge.

En rapportant les décès de la figure 3.2 aux valeurs moyennes<sup>12</sup> des effectifs

<sup>9</sup>Voir notamment Deparcieux, *op. cit.*, et Leonhard Euler, [13]

<sup>10</sup>Son taux variait entre 5 et 12,5%.

<sup>11</sup>L'âge est en années révolues.

<sup>12</sup>On appelle *valeur moyenne* la quantité, exprimée en personne-années, soumise au risque de

donner le classement

Celle-ci

est (L'INSEE) ses données

supplémentaire importante vers 55-59 ans



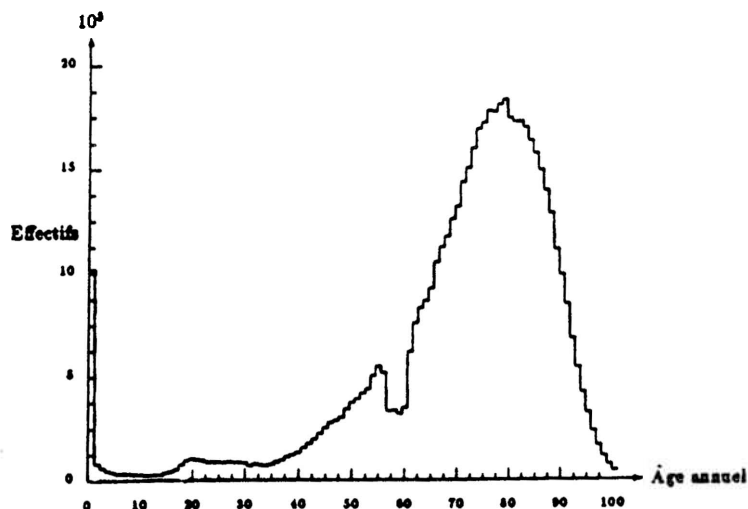


Figure 3.2: Distribution des 551 489 décédés en France durant l'année 1975 en fonction de leur âge.

susceptibles de décéder à cet âge et durant l'année 1975 (figure 3.3), nous obtenons ce que le démographe appelle un *taux de mortalité* (voir le chapitre 1 et l'annexe de ce chapitre), ou risque de décéder à cet âge. La figure 3.4 donne le résultat de ces divisions à chaque âge.

Nous sommes d'abord frappés par la régularité et les variations de cette courbe. Il est maintenant bien connu que le risque de décéder croît avec l'âge aux âges adultes, mais cela n'avait rien d'évident, car même si le risque avait été décroissant, il aurait été de toute façon plus difficile d'atteindre l'âge de la retraite que l'âge au mariage par exemple, puisqu'on est soumis plus longtemps au risque. La croissance exponentielle du risque à l'âge adulte, a beaucoup intrigué mais n'a jamais été bien expliquée ; elle reflète le vieillissement réel de l'individu.

Aux jeunes âges au contraire, plus un bébé ou enfant survit, plus il a de chance de survivre, et ce, jusqu'à l'âge de 10 ans environ où le risque de décéder passe par un minimum. Dans un pays africain, les risques de décès sont plus élevés mais la forme de la distribution des taux par âge est analogue<sup>13</sup>. Notons la montée récente du risque vers 18-20 ans qui résulte des accidents (voitures et motos) fréquents à ces âges.

Reprenons la recherche de notre distribution. Tout d'abord, convenons que l'âge annuel n'est qu'une commodité, et que la figure 3.4, aurait été sensiblement

---

décéder. Dans notre cas, on considère que les individus âgés entre les âges  $x$  et  $x + 1$  au milieu de l'année 1975 ont été soumis durant un an au risque de décéder, et que les décès proviennent de cette population.

<sup>13</sup>voir le chapitre sur les tables types de mortalité

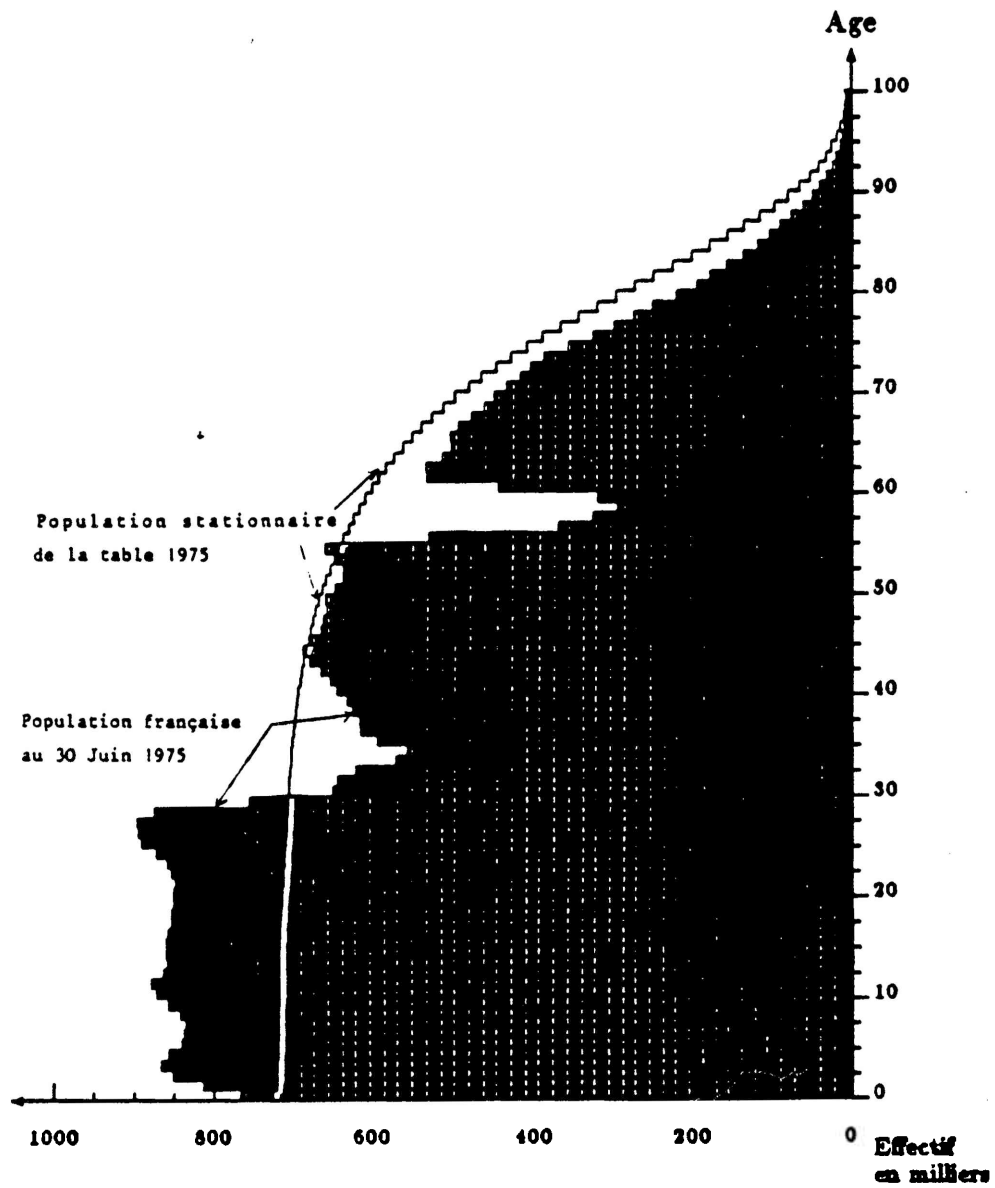


Figure 3.3: Structure par âge de la population française au milieu de l'année 1975. Population stationnaire de même effectif déduite de la mortalité mesurée la même année.



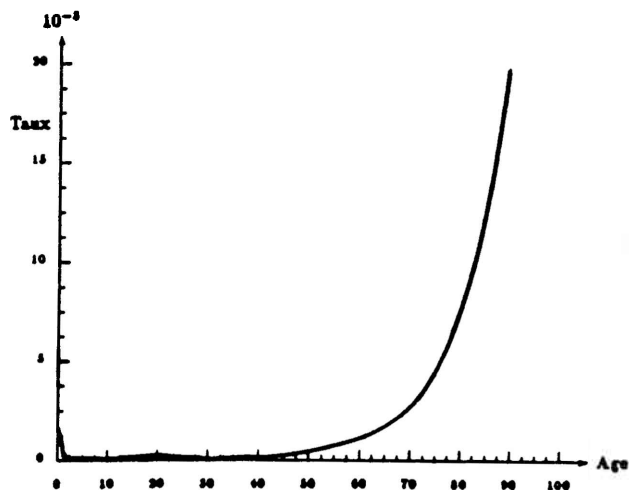


Figure 3.4: Risques de décès suivant l'âge en 1975 : taux de mortalité, échelle arithmétique. France.

identique si nous avons compte l'âge en trimestre, ou en groupe quinquennal : le taux est une mesure intrinsèque qui ne dépend pas ou peu de l'intervalle d'âge choisi. Evidemment, dans le cas d'un groupe quinquennal, nous perdrons beaucoup de précision, et en gagnerions dans le cas trimestriel. Si notre intervalle de temps est infinitésimal, on parle de taux infinitésimal, ou aussi de *force de mortalité* comme l'ont introduit les premiers statisticiens. En absence d'ordinateurs, il n'est pas aisé de travailler avec des probabilités infinitésimales de décès, et les démographes utilisent des probabilités de décès durant des périodes finies (1 an, 5 ans, 10 ans), probabilités qui se déduisent des taux infinitésimaux par une formule mathématique<sup>14</sup>. Ces probabilités s'appellent aussi des *quotients de mortalité*. Ainsi, sur  $S(x)$  individus d'âge  $x$ , le nombre des décès attendus<sup>15</sup> entre les âges  $x$  et  $x + 1$  est  $d(x, x + 1) = S(x)_h Q_x$ . La figure 3.5 décrit les quo-

<sup>14</sup>On note  $\mu(x)$  la force de mortalité à l'âge  $x$ . Ainsi  $\mu(x)dx$  est la probabilité pour un individu âgé de  $x$  années de décéder entre les âges  $x$  et  $x + dx$ , et  $1 - \mu(x)dx$  la probabilité de survivre.

A 60 ans, par exemple, le risque infinitésimal de décéder dans l'instant est de 0,01 annuel, soit  $0,01/(60 \times 60 \times 24 \times 365)$  par seconde (fig 3.4). La probabilité de décéder dans les 10 secondes est alors sensiblement 10 fois plus importante, celle de décéder dans l'heure, 3600 fois, mais celle de décéder durant l'année ou durant les 5 années à venir, ne s'obtient pas si facilement. En premier lieu, le calcul des probabilités élémentaires enseigne que ce sont les probabilités de survie que l'on multiplie et non les probabilités de décès que l'on additionne (dans le cas où les probabilités de décès sont petites, les résultats sont sensiblement identiques en raison de l'approximation :  $1 - (1 - \epsilon)(1 - \epsilon') \simeq \epsilon + \epsilon'$ ). En second lieu, aux âges où la force de mortalité varie beaucoup, le calcul de la probabilité de décès doit tenir compte de cette augmentation de risque ; on montre aisément que la probabilité pour un individu âgé de  $x$  années de décéder entre les âges  $x$  et  $x + h$  est

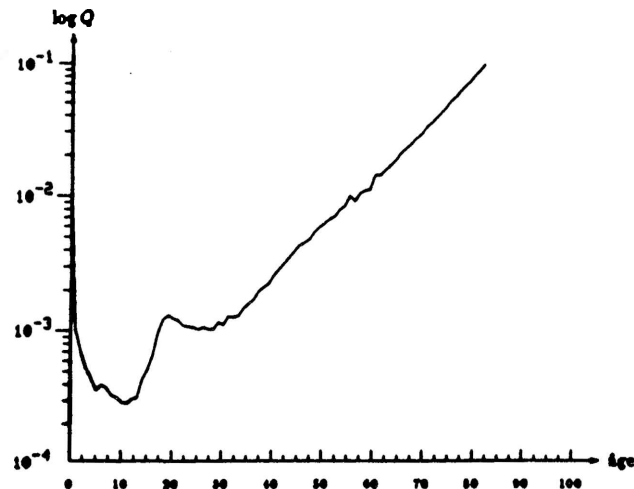


Figure 3.5: Probabilités de décès entre un anniversaire et le suivant. Quotients annuels par âge, 1975, France (échelle logarithmique).

tients de mortalité selon l'âge annuel, avec cette fois une échelle logarithmique<sup>16</sup>. Les quotients annuels, les taux calculés sur un an, et les taux infinitésimaux ont des valeurs numériques qui ne diffèrent pratiquement pas ; autrement dit les figures 3.4 et 3.5 ne diffèrent que par leur type d'échelle<sup>17</sup>.

Nous avons maintenant tous les éléments pour trouver la distribution cherchée. Suivons le raisonnement de Graunt, prenons une cohorte de  $S(0)$  nouveaux-nés. Laissons le hasard décider de la mort de  $d(0, 1) = Q_0 S(0)$  enfants avant l'âge d'un an, laissant  $S(1) = S(0) - d(0, 1)$  survivants à un an ;  $d(1, 2) = Q_1 S(1)$  décès surviendront entre 1 et 2 ans :  $S(2) = S(1) - d(1, 2)$  enfants atteindront au moins l'âge de deux ans. Poursuivons l'exercice jusqu'à ce qu'il n'y ait plus de survivant, et nous obtenons la distribution que nous cherchions. L'histogramme des décès qui en résulte est représenté sur la figure 3.6 ; notre première démarche qui concluait que les hommes mouraient entre 60 et 90 ans n'est pas contredite par cette nouvelle distribution, mais ici, le calcul est rigoureux.

$${}_h Q_x = 1 - \exp\left(-\int_x^{x+h} \mu(u) du\right).$$

<sup>15</sup> Comme le démographe raisonne le plus souvent sur des grands effectifs, il assimile l'espérance mathématique du nombre des décès à ce nombre lui-même, en écartant toute variation aléatoire.

<sup>16</sup> On note  $Q_x$  au lieu de  ${}_1 Q_x$  les quotients annuels.

<sup>17</sup> Plus exactement, le quotient annuel à l'âge  $x$  concerne des individus qui auront durant l'année, en moyenne l'âge  $x + \frac{1}{2}$  ; il est donc plus proche du taux infinitésimal à l'âge  $x + \frac{1}{2}$ .

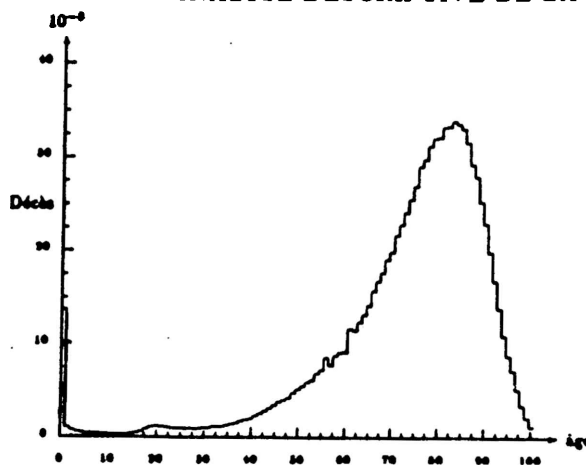


Figure 3.6: Distribution des âges au décès d'une cohorte de 1000 nouveau-nés soumise aux risques de décéder par âge mesurés en 1975

### 3.4 L'espérance de vie

La figure 3.6 illustre la mortalité calculée avec les seules données (état civil et recensement) de l'année 1975 et est appelée de ce fait la mortalité *du moment*. Comme il est difficile de véhiculer l'information de la distribution complète, on préfère la résumer par son espérance mathématique encore appelée *espérance de vie* ou *vie moyenne*. Celle-ci s'établit à 73,0 ans. Les démographes font hélas rarement mention de l'écart-type, de la distribution qui est très élevé, égal ici à 18 ans.

La distribution,  $S(x)$ , des survivants à chaque âge, apparaît sur la figure 3.7. On peut y lire que 21 % des nouveau-nés n'atteindrait pas l'âge de 65 ans. L'âge médian au décès, c'est à dire l'âge où 50 % des individus sont décédés est 77 ans, soit, supérieur de 4 années à l'espérance de vie ; cet écart provient du fait que les décès par âge sont plus étalés avant 77 ans qu'après.

La courbe des survivants correspond aussi à ce qu'on appelle la *population stationnaire*. En effet, si tous les ans naissait un même nombre d'enfants, et si les risques de mortalité restaient constamment ceux de l'année 1975, la pyramide des âges de cette population, prendrait la forme de cette distribution, redressée d'un quart de tour<sup>18</sup>. Il apparaît ainsi clairement évident que, toutes choses égales par ailleurs, plus l'espérance de vie d'une population est élevée, plus la place qu'y tiennent les personnes âgées est importante.

<sup>18</sup>La courbe des survivants,  $S(0) = 1, S(1), \dots$ , est une interpolation linéaire de la fonction de survie continue,  $S(x)$ , déduite de la force de mortalité  $\mu$  :  $S(x) = \exp(-\int_0^x \mu(u) du)$ . La pyramide des âges est un *histogramme*, en ce sens, qu'on la représente par des rectangles, plutôt que des trapèzes à chaque âge. Il s'agit alors du nombre d'individus compris entre deux âges exacts, plutôt que, dans le cas des survivants, d'un effectif à l'âge exact. Rappelons que c'est dans le cas d'une telle population stationnaire que le taux brut de mortalité (comme le taux brut de natalité) est égal à l'inverse de l'espérance de vie.

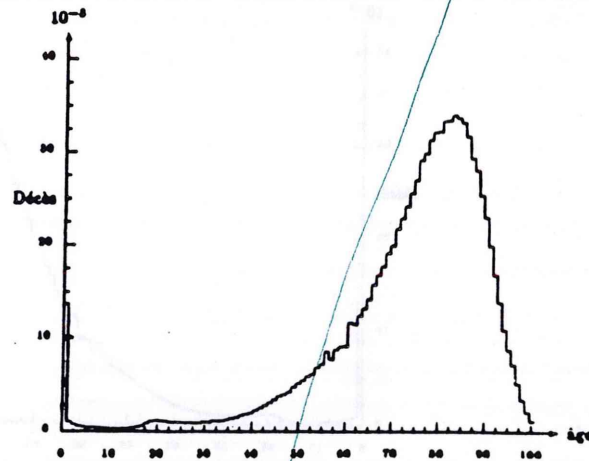


Figure 3.6: Distribution des âges au décès d'une cohorte de 1000 nouveau-nés soumise aux risques de décéder par âge mesurés en 1975

### 3.4 L'espérance de vie

La figure 3.6 illustre la mortalité calculée avec les seules données (état civil et recensement) de l'année 1975 et est appelée de ce fait la mortalité *du moment*. Comme il est difficile de véhiculer l'information de la distribution complète, on préfère la résumer par son espérance mathématique encore appelée *espérance de vie* ou *vie moyenne*. Celle-ci s'établit à 73,0 ans. Les démographes font hélas rarement mention de l'écart-type de la distribution qui est très élevé, égal ici à 18 ans.

La distribution,  $S(x)$ , des survivants à chaque âge, apparaît sur la figure 3.7. On peut y lire que 21 % des nouveau-nés n'atteindrait pas l'âge de 65 ans. L'âge médian au décès, c'est à dire l'âge où 50 % des individus sont décédés est 77 ans, soit, supérieur de 4 années à l'espérance de vie ; cet écart provient du fait que les décès par âge sont plus étalés avant 77 ans qu'après.

La courbe des survivants correspond aussi à ce qu'on appelle la *population stationnaire*. En effet, si tous les ans naissait un même nombre d'enfants, et si les risques de mortalité restaient constamment ceux de l'année 1975, la pyramide des âges de cette population, prendrait la forme de cette distribution, redressée d'un quart de tour<sup>18</sup>. Il apparaît ainsi clairement évident que, toutes choses égales par ailleurs, plus l'espérance de vie d'une population est élevée, plus la place qu'y tiennent les personnes âgées est importante.

<sup>18</sup> La courbe des survivants,  $S(0) = 1, S(1), \dots$ , est une interpolation linéaire de la fonction de survie continue,  $S(x)$ , déduite de la force de mortalité  $\mu$  :  $S(x) = \exp(-\int_0^x \mu(u) du)$ . La pyramide des âges est un *histogramme*, en ce sens, qu'on la représente par des rectangles, plutôt que des trapèzes à chaque âge. Il s'agit alors du nombre d'individus compris entre deux âges exacts, plutôt que, dans le cas des survivants, d'un effectif à l'âge exact. Rappelons que c'est dans le cas d'une telle population stationnaire que le taux brut de mortalité (comme le taux brut de natalité) est égal à l'inverse de l'espérance de vie.

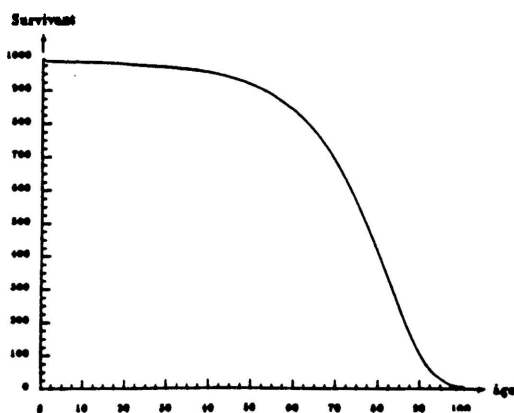


Figure 3.7: Fonction de survie déduite des quotients de mortalité mesurés en 1975. France.

### 3.5 La table de mortalité

Il est coutumier de regrouper les valeurs des 3 fonctions principales, quotients, survivants et décès dans ce qu'on appelle une *table de mortalité*. On adjoint à cette table d'autres colonnes et en particulier l'espérance de vie à chaque âge, et les taux. La table de mortalité de l'année 1982 de la France est donnée dans les tableaux 3.1 et 3.2.

Age $z$	Survivants $S_z$	Décès $d(z, z + 1)$	Quotients $q_z$	Taux $m_z$	Espérance de vie	Age $z$
0	100000	943	943	947	74,76	0
1	99057	78	78	78	74,47	1
2	98979	52	53	53	73,53	2
3	98927	38	39	39	72,57	3
4	98889	32	32	32	71,60	4
5	98857	37	37	37	70,62	5
6	98821	27	28	28	69,64	6
7	98793	29	29	29	68,66	7
8	98765	26	27	26	67,68	8
9	98739	25	25	25	66,70	9
10	98714	21	21	21	65,72	10
11	98693	26	27	27	64,73	11
12	98667	23	24	24	63,75	12
13	98643	30	31	31	62,76	13
14	98613	41	42	42	61,78	14
15	98571	46	47	47	60,81	15
16	98525	64	65	65	59,84	16
17	98461	81	82	82	58,87	17
18	98381	111	112	112	57,92	18
19	98270	117	119	119	56,99	19
20	98154	121	124	124	56,05	20
21	98032	119	121	121	55,12	21
22	97914	117	120	120	54,19	22
23	97796	108	110	110	53,25	23
24	97688	106	109	109	52,31	24
25	97582	110	113	113	51,37	25
26	97472	103	106	106	50,43	26
27	97369	105	108	108	49,48	27
28	97264	105	108	108	48,53	28
29	97159	104	107	107	47,58	29
30	97055	112	115	115	46,63	30
31	96943	114	118	118	45,69	31
32	96829	119	123	123	44,74	32
33	96710	119	123	123	43,79	33
34	96590	132	136	136	42,85	34
35	96459	140	145	145	41,91	35
36	96319	158	164	164	40,97	36
37	96162	167	174	174	40,03	37
38	95994	182	190	190	39,10	38
39	95812	208	217	217	38,17	39
40	95604	218	228	228	37,26	40
41	95386	224	235	236	36,34	41
42	95161	249	261	261	35,43	42
43	94913	285	300	301	34,52	43
44	94628	285	301	302	33,62	44
45	94343	331	351	352	32,72	45
46	94011	379	403	404	31,83	46
47	93632	396	423	424	30,96	47
48	93236	442	474	475	30,09	48
49	92794	494	532	533	29,23	49

Tableau 3.1: Table de mortalité de l'année 1982, France.

Nous y voyons que la *racine* de la table, c'est à dire l'effectif initial de la cohorte fictive utilisée dans la simulation, a été choisi à 100 000. Les quotients sont souvent exprimés en pour dix mille.

### 3.6 Liaison entre mortalité et vieillissement d'une population

Comme la population française n'a qu'un taux de croissance très faible (moins de 0,05% au cours des dix dernières années), on pourrait la croire presque stationnaire<sup>19</sup>. On vient de voir qu'il n'en est rien. Il est tout à fait intéressant de confronter sa structure par âge telle qu'elle a pu être observée en 1975 à celle de la population stationnaire, déduite de la table de mortalité de la même année, et toutes deux à l'échelle des 52 millions de personnes recensées cette année là. On remarque immédiatement sur la figure 3.3 que la population française est beaucoup plus jeune que celle que les conditions actuelles de mortalité tendent à produire. On mesure ici l'effet du baby-boom qui a suivi la seconde guerre mondiale. Aux effectifs nombreux des moins de 30 ans nés après la guerre, s'opposent les générations âgées, d'effectifs faibles car nées dans une période de basse fécondité qui a été aggravée par les deux guerres mondiales.

Ces résultats réapparaissent sous un autre angle quand on superpose sur la figure 3.8, la distribution des décès réellement observés (figure 3.2) à celle des décès déduits de la table de mortalité (figure 3.6)<sup>20</sup>. L'âge moyen des décédés de l'année 1975 (71,3 ans) est nettement inférieur à l'espérance de vie calculée plus haut (73,4 ans). Cela tient, encore une fois, au fait que la population française de 1975 est plus jeune que la population stationnaire correspondant aux conditions de mortalité de 1975.

---

<sup>19</sup>Le raisonnement qui suit est particulièrement absurde pour un pays en pleine transition comme le sont la plupart des pays africains. Nous verrons alors qu'il faut confronter une population *stable* et non plus *stationnaire*.

<sup>20</sup>Le nombre des décès à chaque âge (autrement dit l'échelle de la distribution de la *table*), est celui qu'on obtiendrait à partir de la population stationnaire équivalente (52 millions). Notons que les décès de la *table* seraient dans ce cas plus nombreux, 718 050 contre 551 489 réellement observés (23% de plus).

Age $z$	Survivants $S_z$	Décès $d(z, z + 1)$	Quotients $q_z$	Taux $m_z$	Espérance de vie	Age $z$
50	92301	536	581	583	28,38	50
51	91764	567	618	620	27,55	51
52	91197	606	665	668	26,71	52
53	90591	659	728	731	25,89	53
54	89932	713	793	796	25,08	54
55	89218	761	852	856	24,27	55
56	88458	789	892	896	23,48	56
57	87669	833	950	955	22,68	57
58	86836	875	1008	1013	21,90	58
59	85961	922	1072	1078	21,11	59
60	85039	1006	1183	1190	20,34	60
61	84033	1071	1275	1283	19,58	61
62	82962	1242	1497	1484	18,82	62
63	81720	1140	1395	1397	18,10	63
64	80580	1265	1570	1581	17,35	64
65	79315	1321	1665	1680	16,62	65
66	77994	1338	1715	1713	15,89	66
67	76656	1656	2161	2147	15,16	67
68	75000	1616	2155	2179	14,48	68
69	73384	1753	2389	2417	13,79	69
70	71631	1855	2590	2622	13,12	70
71	69776	2001	2867	2908	12,45	71
72	67775	2143	3162	3213	11,80	72
73	65632	2357	3591	3656	11,17	73
74	63276	2454	3879	3957	10,57	74
75	60821	2668	4387	4486	9,98	75
76	58153	2777	4776	4892	9,41	76
77	55375	2974	5371	5519	8,86	77
78	52401	3130	5973	6155	8,33	78
79	49271	3249	6593	6813	7,83	79
80	46023	3345	7269	7539	7,35	80
81	42677	3518	8244	8595	6,89	81
82	39159	3505	8950	9354	6,46	82
83	35654	3557	9977	10488	6,05	83
84	32097	3521	10970	11602	5,66	84
85	28576	3463	12120	12907	5,30	85
86	25112	3307	13168	14116	4,96	86
87	21806	3187	14613	15766	4,63	87
88	18619	2965	15922	17303	4,34	88
89	15655	2666	17033	18667	4,07	89
90	12988	2455	18902	20963	3,80	90
91	10533	2218	21056	23639	3,57	91
92	8315	1805	21711	24373	3,39	92
93	6510	1477	22681	25432	3,19	93
94	5033	1250	24841	28553	2,98	94
95	3783	1044	27588	32042	2,81	95
96	2739	757	27617	33000	2,68	96
97	1983	513	25871	28538	2,52	97
98	1470	421	28630	35194	2,22	98
99	1049	225	21416	30396	1,91	99
100	824	167	20284	20367	1,30	100

Tableau 3.2: Table de mortalité de l'année 1982 (suite), France.



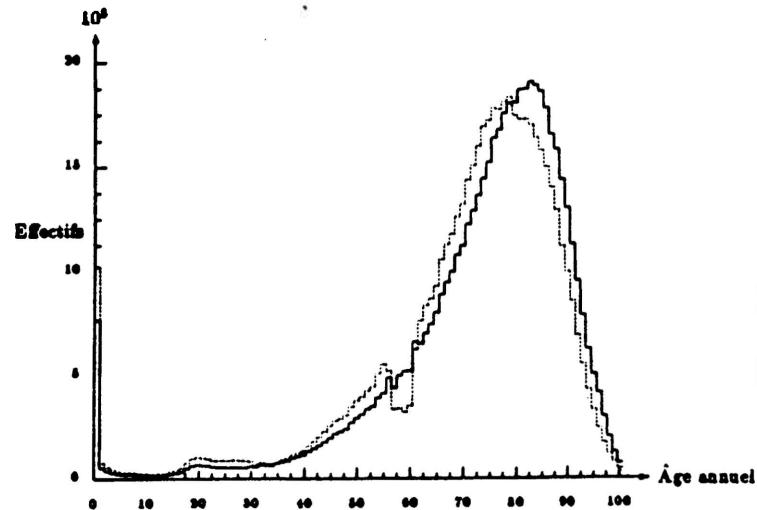


Figure 3.8: Superposition des distributions des décès bruts et déduits de la table de mortalité, France, 1975.

### 3.7 Mortalité du moment et mortalité réelle des générations<sup>21</sup>

La table de mortalité calculée à partir des décès enregistrés à l'état civil en 1975 et du recensement de la même année nous a permis d'isoler et de décrire avec précision le processus de mort auquel la population française se trouve actuellement soumise. Ce calcul peut être fait chaque année. Si nous nous livrons, pour l'année 1899 (première année pour laquelle on dispose de statistique de décès par classe d'âge) au même exercice que précédemment, nous obtenons la figure 3.9.

<sup>21</sup>On peut se reporter à ce sujet aux publications de J. Vallin [43] : "La mortalité par génération en France, depuis 1899", Travaux et Documents, cahier n° 63, Paris, INED-PUF, 1973, 484p, et "Tables de mortalité du moment et par génération", mise à jour provisoire du cahier n° 63. Paris, INED, 1984, 3 volumes, d'où sont tirées les données utilisées ici.

On voit que la situation y était très différente puisqu'au début du siècle, les trois courbes (âge, vie restante et structure stationnaire) sont très voisines (indice de 103). Si la population de 1899 avait une structure plus jeune que celle de la population actuelle, c'est presque uniquement parce que la mortalité par âge y était plus élevée.

Pour apprécier les changements intervenus au cours du 20<sup>e</sup> siècle, nous avons superposé sur la figure 3.10 la courbe des quotients par âge de l'année 1899, avec celle de l'année 1979, couvrant ainsi une période de 80 ans. Il apparaît clairement que la simulation effectuée sur les données de 1899 donne pour cette année là une mortalité nettement plus élevée que celle que subira réellement la génération des enfants nés en 1899.

Pour calculer la vie moyenne propre à cette dernière génération, on pourrait songer à retenir parmi les décès survenus chaque année depuis 1899, ceux qui concernaient cette seule génération, et à en calculer l'âge moyen. Mais la France a connu durant ces 80 ans d'importants flux migratoires. Il est donc nécessaire de passer une fois encore par la mesure des risques de décès à chaque âge. En calculant pour chaque année, comme nous l'avions fait pour 1975, les quotients de mortalité par âge, on peut ensuite regrouper les risques propres à une même génération et construire une table de mortalité d'un nouveau genre<sup>22</sup>. On appelle une telle table, *table de mortalité par génération* ou encore *table longitudinale de mortalité*, par opposition aux *tables de mortalité du moment* ou *table transversale de mortalité* qui décrivent les conditions de mortalité à un moment donné.

Une table longitudinale permet donc de suivre la résultante, pour une génération donnée, de deux phénomènes souvent opposés qui sont d'une part le vieillissement et d'autre part les progrès de la société en matière de santé, hygiène, médecine, développement, etc.

Nous avons ajouté à la figure 3.10 la courbe des quotients annuels de la génération 1899<sup>23</sup> ; outre que cette courbe a par définition la même mortalité infantile que la courbe transversale de 1899, et la même mortalité à 80 ans que celle de 1979, nous remarquons que les deux effets mentionnés plus haut se sont presque annihilés entre 25 et 35 ans, c'est à dire entre les années 1919 et 1939. A cette époque et ces âges là, plus l'on vieillissait, plus l'on avait de chance de survivre !

Revenons aux tables complètes de la figure 3.10 et remarquons combien la table longitudinale est plus heurtée que les tables transversales. En effet, hormis les guerres qui ne touchent que quelques classes d'âge et surtout les hommes, les petites crises ou épidémies conjoncturelles qui entravent la baisse générale de la mortalité, touchent toutes les générations ; autrement dit les courbes de mor-

<sup>22</sup> Plus exactement, on calcule la surface définie par la valeur de la force de mortalité à chaque âge  $x$  et chaque date  $t$ , soit  $\mu(x, t)$ . La force de mortalité, à l'âge  $x$ , associée à une génération née l'année  $t_0$  est  $\mu(x, t_0 + x)$ .

<sup>23</sup> En mettant à jour les données publiées dans J. Vallin, op. cit. [43].

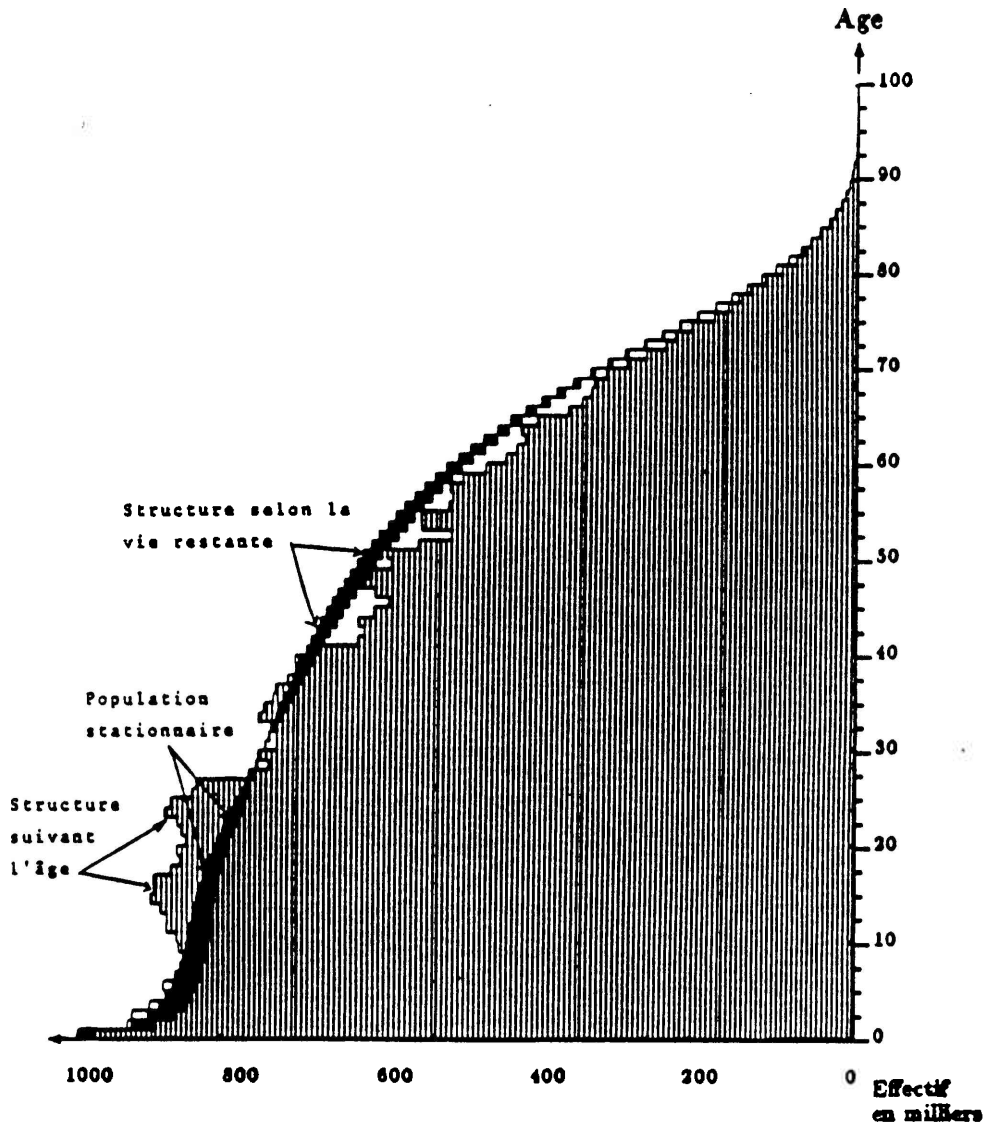


Figure 3.9: Même légende que pour la figure 3.3, mais pour l'année 1899.

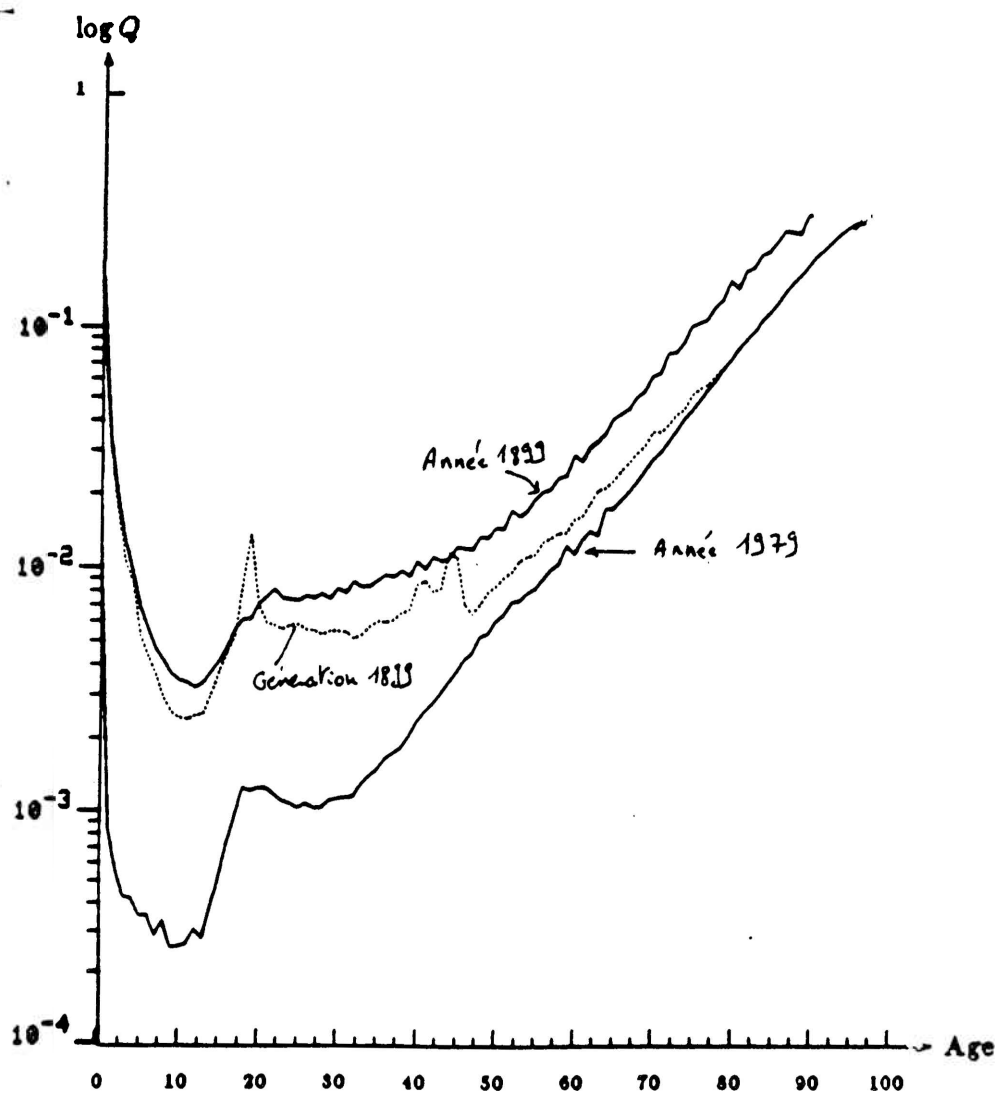


Figure 3.10: Risque de décès entre un anniversaire et le suivant, ou quotient de mortalité : année 1899 et 1979, et génération 1899, France

talité transversales qui peuvent être très différentes d'une année à l'autre, sont régulières avec l'âge<sup>24</sup> alors que les courbes longitudinales, subissant les divers événements des années successives, sont au contraire souvent perturbées. Ainsi, la grippe espagnole de 1918 qui a relevé la mortalité à tous les âges, laisse apparaître une courbe du moment très lisse, mais produit pour chaque génération une pointe de mortalité à l'âge atteint cette année là par la génération.

Finalement, si elle reflète avec exactitude les conditions de mortalité du moment et nous permet de suivre avec précision l'évolution de la mortalité, la table de mortalité transversale ne nous renseigne guère sur la mortalité réelle des générations. Ainsi, l'espérance de vie de l'année 1979, 74,15 ans, apporte peu d'information sur la mortalité des enfants nés cette année là puisque leur vie moyenne sera plus importante si la mortalité continue à baisser et qu'aucune crise ne vient remettre en cause ce progrès. Au contraire, les générations qui décèdent en grand nombre en 1979 auront eu une vie moyenne plus courte. L'âge moyen au décès de la *génération* 1899, par exemple aura été, lorsque les derniers survivants seront décédés, voisin de 50 ans, soit 26 ans de moins que celui donné par la table *transversale* de 1979. Notons cependant que l'espérance de vie de l'année 1899 n'était inférieure que de 5 ans à la vie moyenne réelle de la *génération* 1899 ; ceci résulte du fait qu'au début du siècle, la mortalité infantile encore très élevée pesait beaucoup dans le calcul de l'espérance de vie.

Pour nous en persuader, regardons non plus les courbes des quotients, mais les courbes de survie à chaque âge de la figure 3.11. Aux jeunes âges, celle de la *génération* 1899 suit longtemps la courbe transversale de la même année avant de s'en écarter définitivement aux âges adultes. Comme on montre également que, dans une table de mortalité, l'âge moyen au décès est aussi la valeur de l'aire <sup>1 m. li</sup> située sous la courbe des survivants, nous comprenons pourquoi la différence des espérances de vie est plutôt faible comparée à celle qui sépare la même *génération* 1899 et l'année 1979 puisqu'entre temps la mortalité infantile s'est effondrée. La figure 3.12 permet de comparer les courbes de décès et de confirmer d'une autre manière notre analyse.

L'intérêt des indices transversaux comme l'espérance de vie du moment, ne réside pas tant dans la valeur des résultats eux-même que dans l'analyse conjoncturelle et dans les comparaisons internationales précises qu'il permettent. Cependant, l'analyse transversale ne suffit pas à rendre complètement compte des ressorts de la mortalité, et une étude longitudinale est souvent indispensable pour interpréter correctement certains résultats, surtout en matière de cause de décès.

<sup>24</sup>Si elles ne le sont pas, c'est soit que les perturbations sont d'origine stochastique (effectifs trop petits), soit que le calcul du taux n'est pas correctement effectué. Ce dernier cas est possible lorsque les naissances ne sont pas uniformément réparties durant l'année, par exemple lors de la reprise des naissances durant l'année 1919, ou le départ au front en 1914.

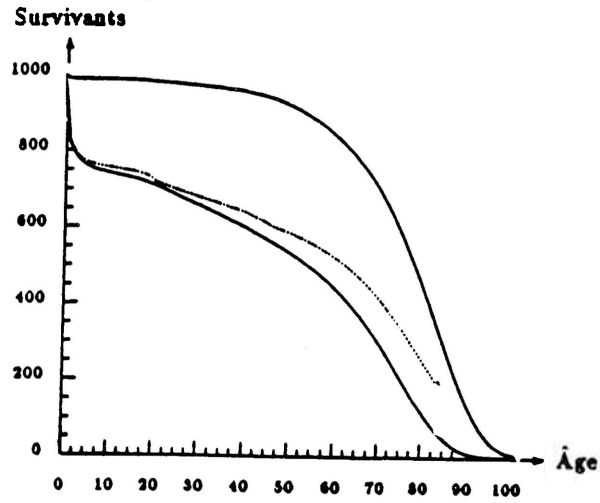


Figure 3.11: Nombre de survivants à chaque âge d'une cohorte de 1 000 individus, suivant que la mortalité est mesurée à deux dates différentes 1899 et 1979, ou pour la génération 1899, France.

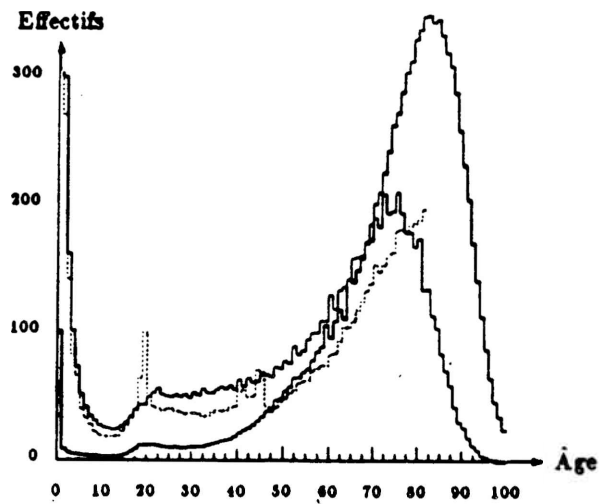


Figure 3.12: Distribution des décès de la génération née en 1899, ainsi que celles résultant des simulations effectuées avec les données de mortalité des années 1899 et 1979. France.

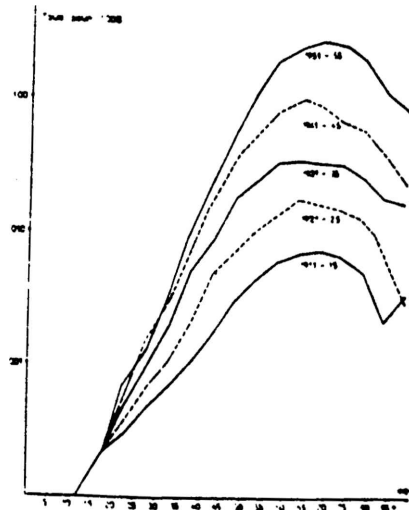


Figure 3.13: Taux de décès par cancer du poumon et de la plèvre, par âge pour différentes périodes, hommes, Angleterre et Pays de Galles. Source : R.A.M., Case, op. cité.

### 3.8 Un exemple d'interprétation difficile : la mortalité par cancer du poumon

Même si la mortalité totale évolue moins rapidement aujourd'hui, l'importance relative des causes de décès, nous le verrons, se modifient profondément. L'exemple suivant, tiré des travaux R.A.M. Case [8] montre dans ce contexte l'importance de l'analyse longitudinale.

La courbe supérieure de la figure 3.13 décrit le taux de mortalité par cancer du poumon et de la plèvre suivant l'âge, en Angleterre et Pays de Galles, pour le dernier groupe d'année disponible à la date de la publication, c'est à dire pour la période 1951-1955. Un examen de cette seule courbe conduirait à conclure à l'existence d'un risque maximum aux âges de 65-70 ans. L'examen de la courbe *transversale* effectuée 10 ans plutôt, c'est à dire avec les données (décès et population par âge) de la période 1941-45, aurait conclu aussi à l'existence de ce maximum ; et il en est de même pour les période 1931-35, 1921-25 et 1911-15. L'erreur provient du fait que pour chaque période, les points correspondant aux âges ont été reliés entre eux, laissant entendre qu'ils avaient une relation entre eux : ils n'en ont guère car il s'agit de générations différentes, et les générations anciennes ne connaissaient pas ou peu de cas de cancer du poumon. Relions donc les points correspondant à une même génération, et nous obtenons (figure 3.14) les courbes des générations nées en 1866, 1876, 1886, 1896, et 1906. La mortalité par cancer du poumon croît donc avec l'âge ; simplement, le cancer du poumon est une maladie qui fait son apparition dans les nouvelles générations

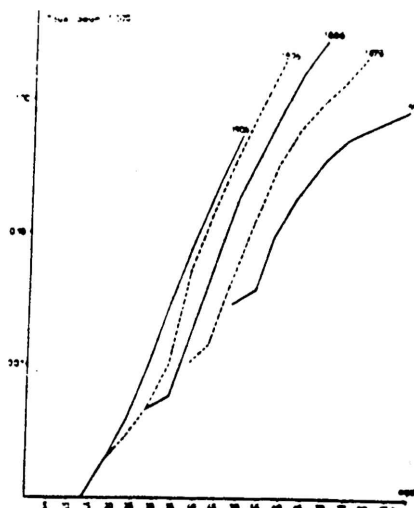


Figure 3.14: Taux de décès par cancer du poumon et de la plèvre, par âge regroupement par génération, hommes, Angleterre et Pays de Galles. Source :R.A.M., Case, op. cité.

et ne peut être étudiée avec une simple coupe transversale.

On voit ainsi tout l'intérêt de l'analyse longitudinale trop souvent délaissée. La collation, souvent difficile, de longues séries statistiques est un investissement rentable, trop négligé.

L'analyse transversale a aussi, dans ce cas, son intérêt, mais il est d'un autre ordre : la courbe supérieure de la figure 3.13 nous indique, qu'effectivement, dans les services qui traitent le cancer du poumon, il y avait sans doute une part plus importante de personnes âgées de 65-70 ans<sup>25</sup> à traiter. Ce phénomène n'avait aucune origine d'ordre médicale ou génétique, mais résultait d'un mécanisme caché de la statistique. Ces périodes de transition, apparitions de maladies (cancers) ou disparitions (tuberculose) sont très fréquentes, et encore d'actualité avec certaines formes de cancer. Elles peuvent être de longue durée, puisque, comme on peut le voir sur cette même figure 3.13, le phénomène apparaissait déjà en 1913.

Nous avons tenté d'expliquer dans ce paragraphe les mécanismes fondamentaux d'analyse de la mortalité en introduisant la table de mortalité et ses deux facettes transversale et longitudinale. Il nous reste à approfondir les mathématiques d'une table de mortalité ; l'approche infinitésimale est la plus facile et la plus rigoureuse, nous l'étudions dans le chapitre suivant.

<sup>25</sup>Il faudrait pour être précis regarder la courbe des décès par cancer du poumon et par âge, et non les taux par âge, ce qui aurait pour conséquence d'abaisser légèrement l'âge modal.



## Chapitre 4

# Approche infinitésimale de la mortalité

### 4.1 Calcul d'une table de mortalité en continu

#### 4.1.1 Force de mortalité

**Définition 8** On définit et note  $\mu(x) dx$ , la probabilité pour qu'un individu d'âge  $x$  décède entre les âges  $x$  et  $x + dx$ .  $\mu$  est alors appelée force de mortalité selon le vocabulaire utilisé par Gompertz (1825) [16].

La probabilité de survivre entre les âges  $x$  et  $x + dx$  pour un individu ayant atteint l'âge  $x$  est donc  $1 - \mu(x)dx$ .

Partant d'un âge origine 0, si  $S(x)$  désigne la probabilité de survivre jusqu'à l'âge  $x$ , calculons la probabilité de survivre jusqu'à l'âge  $x + dx$ . La définition d'une probabilité conditionnelle :

$$\Pr(A \text{ et } B) = \Pr(A) \times \Pr(B/A)$$

nous permet d'écrire :

$$\begin{aligned} \Pr(\text{survie jusqu'à } x + dx) &= \Pr(\text{survie jusqu'à } x \text{ et survie entre } x \text{ et } x + dx) \\ &= \Pr(\text{survie jusqu'à } x) \times \Pr(\text{survie entre } x \text{ et } x + dx / \text{survie jusqu'à } x), \end{aligned}$$

soit finalement :

$$S(x + dx) = S(x)(1 - \mu(x) dx) \quad (4.54)$$

Cette équation différentielle s'intègre facilement :

$$\frac{S'(x)}{S(x)} = -\mu(x) \quad (4.55)$$

$$\text{d'où } S(x) = S(0)e^{-\int_0^x \mu(t) dt} \quad (4.56)$$

La valeur de  $S(0)$  vaut 1, mais le terme figure d'une manière explicite dans l'équation 4.56 pour le cas où on choisit un âge autre que zéro comme racine de la table.

Désormais lorsque les bornes de l'intégration ne seront pas explicitées, il s'agira de l'intervalle  $[0, \infty[$ .

#### 4.1.2 Distribution des décès

Partant d'un effectif de  $N$  individus à la naissance,  $NS(x)$  désigne alors le nombre des survivants à l'âge  $x$ . Généralement on confondra la probabilité  $S(x)$  avec les *survivants* à l'âge  $x$  faisant abstraction du facteur multiplicateur que constitue l'effectif<sup>1</sup>  $N$ . Ainsi l'équation 4.54 équivaut à : le nombre des survivants à l'âge  $x + dx$ , est égal au nombre des survivants à l'âge  $x$  diminué des décès durant l'instant infinitésimal  $dx$ .

Partant d'un effectif unitaire  $N = 1$ , le nombre des décès attendus entre les âges  $x$  et  $x + dx$  provient des  $S(x)$  survivants à l'âge  $x$  et vaut  $S(x)\mu(x) dx$ .

La distribution des décès est donc la fonction  $d(x) = S(x)\mu(x)$  qui vaut, en raison de l'équation 4.54

$$d(x) = -S'(x). \quad (4.57)$$

On vérifie que  $\int_0^{\infty} d(x) dx = 1$ .

#### 4.1.3 Espérance de vie à la naissance

On résume généralement une distribution statistique par ses premiers moments, c'est à dire sa moyenne et son écart-type. L'espérance de vie, résumé de la distribution des décès est définie comme l'espérance mathématique de cette distribution. Ce résumé n'est d'ailleurs pas excellent dans le cas de la distribution des décès car celle-ci est fort éloignée d'une loi normale (qui est entièrement définie par ses deux premiers moments), et a en particulier 3 modes, un à la naissance, l'autre vers 20 ans et enfin le dernier entre 60 et 80 ans. La valeur du premier mode,  $d(0) = \mu(0)$ , est en général inestimable, et seule la mortalité sur un intervalle fini, comme la mortalité avant un jour, une semaine, un mois ou une année l'est. Les deux autres modes correspondent aux âges  $x$  où  $d'(x) = 0$ , soit aussi  $S''(x) = 0$ . Il s'agit donc des points d'inflexion de la fonction de survie (voir la figure 4.1).

Par définition, l'espérance de vie vaut :

$$e_0 = \int_0^{\infty} xd(x) dx \quad (4.58)$$

<sup>1</sup>Plus rigoureusement les *survivants* sont assimilés à l'espérance mathématique du nombre aléatoire de survivants

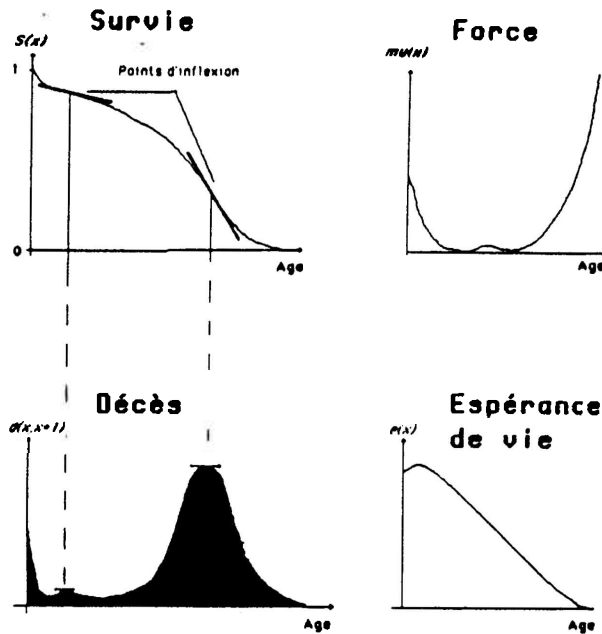


Figure 4.1: Distribution des décès, des survivants et de l'espérance de vie à chaque âge.

On va montrer en intégrant par partie l'équation ci-dessus que l'espérance de vie est aussi l'intégrale de la fonction de survie  $S(x)$ . En effet :

$$e_0 = \int_0^{\infty} x d(x) dx \quad (4.59)$$

$$= -[xS(x)]_0^{\infty} - \int_0^{\infty} -S(x) dx \quad (4.60)$$

$$e_0 = \int_0^{\infty} S(x) dx \quad (4.61)$$

**Espérance de vie à un âge quelconque** Soit  $S(a)$  individus d'âge  $a$ , la moyenne des âges à leur décès est:

$$e_a = \int_a^{\infty} x d_a(x) \text{ avec}$$

$$d_a(x) = \frac{d(x)}{S(a)} = \frac{S(x)}{S(a)} \mu(x)$$

soit aussi:

$$e_a = \frac{1}{S(a)} \int_a^{\infty} S(x) dx \quad (4.62)$$

à faire.

Figure 4.2: Évolution de l'espérance de vie et de l'écart-type de l'âge au décès selon les modèles de tables-types et dans quelques pays.

Quand la mortalité infantile est forte, l'espérance de vie à 1 an peut être plus grande que l'espérance de vie à la naissance, puisque dans ce cas  $S(1)$  est faible.

#### 4.1.4 Écart-type de l'âge au décès

Par définition l'écart-type,  $\sigma$ , de l'âge moyen au décès est égal à :

$$\sigma^2 = \int (x - e_0)^2 S(x) dx = \int x^2 S(x) dx - e_0^2. \quad (4.63)$$

Cet écart-type est très important, et la forme multimodale de la distribution des décès interdit le plus souvent d'utiliser la règle classique des  $2 \times$  écarts-type propre aux lois normales. Lorsque la mortalité atteint des niveaux bas, il est souvent utile de mentionner l'écart-type car ce qui devient important à décrire n'est pas tant une augmentation de l'espérance de vie qu'une diminution de l'écart-type. *qu'il s'agit de décrire.*

c'est moins

Autrement dit, lorsque la mortalité baisse, les hommes vivent certes plus longtemps en moyenne mais ont surtout de plus en plus des durées de vie identiques. Nous assistons donc à une homogénéisation des histoires individuelles : il n'y a plus que de très rares décès d'enfants, ceux-ci perdent rarement leurs parents avant d'être eux-mêmes grands ; on connaît ses quatre grand-parents etc.

La figure 4.2 représente l'évolution conjointe de l'espérance de vie et de l'écart-type de l'âge au décès lorsque la mortalité baisse. Outre l'exemple de la France, nous avons aussi représenté les courbes pour quelques tables types.

## 4.2 Population stationnaire

**Définition 9** Une population stationnaire est une population dont :

1. la force de mortalité est indépendante du temps<sup>2</sup>
2. le nombre des naissances est constant par unité de temps.

Soit  $N dt$  le nombre des naissances entre les instants  $t - dt$  et  $t$ .

- Le nombre de lignes de vie, parmi les  $N dt$ , qui atteindront l'âge exact  $x$  entre les dates  $t + x - dt$  et  $t + x$  est donc  $N dt S(x)$ .
- Le nombre d'individus à la date  $t + x$  d'âge compris entre  $x$  et  $x + dx = x + dt$  est donc  $N dx S(x)$ , car les décès survenus durant cette durée infinitésimale sont des infiniments petits du deuxième ordre. Notons  $P(x) dx$  la population âgée entre l'âge  $x$  et l'âge  $x + dx$ , elle vaut :

$$P(x) dx = N S(x) dx \quad (4.64)$$

et est indépendante du temps.

La population totale est donc:

$$P_{t+x} = \int_0^{\infty} P(x) dx = N \int_0^{\infty} S(x) dx, \quad (4.65)$$

soit, compte tenu de 4.61 :

$$P = N e_0. \quad (4.66)$$

**Propriété 4.2.1** La population totale est constante et égale au produit du nombre instantané de naissances par l'espérance de vie à la naissance.

La structure par âge est:

$$C(x) dx = \frac{N S(x) dx}{N e_0} \quad \text{soit} \quad C(x) = \frac{S(x)}{e_0} \quad (4.67)$$

Il en résulte la propriété suivante :

**Propriété 4.2.2** La pyramide des âges est indépendante du temps et sa structure est identique à celle des survivants à une constante multiplicative près ( $1/e_0$ ).

Le taux brut de natalité, noté  $b$ , rapport des naissances dans l'instant  $dt$  à la valeur moyenne de la population à cette même date vaut donc :

$$b = \frac{N dt}{P dt} = \frac{1}{e_0} \quad (4.68)$$

<sup>2</sup>Elle reste évidemment dépendante de l'âge.

Le taux de natalité  $b$  est égal à l'inverse de l'espérance de vie.

Comme la taille de la population est constante, le taux de croissance instantané est nul ; le taux instantané de mortalité est donc lui aussi constant, et égal au taux de natalité. Vérifions ce résultat d'une autre manière : le nombre total des décès dans l'instant  $dt$  est la somme des décès survenus à tous les âges, soit :

$$\begin{aligned} d dt &= \int_0^{\infty} P(x) dx \mu(x) dt \\ &= \int NS(x) \mu(x) dx dt \\ &= \int NS'(x) dx dt \\ &= N \int_0^{\infty} d[S(x)] dt \\ &= N dt \end{aligned}$$

**Propriété 4.2.3** *Le taux brut de mortalité est égal au taux brut de mortalité et vaut l'inverse de l'espérance de vie.*

Réciproquement, dans une population stationnaire où le taux brut de mortalité atteint 40 pour mille, par exemple la durée de vie moyenne est de l'ordre de 25 ans.

#### 4.2.1 Age moyen de la population stationnaire

**Proposition 4.2.1** *L'âge moyen d'une population stationnaire n'est pas rigoureusement égal à la moitié de la durée moyenne de vie, mais lui est supérieur d'une quantité  $\frac{\sigma^2}{2e_0}$ , où  $\sigma$  est l'écart-type de l'âge au décès, quantité qui n'est pas négligeable en général :*

$$\bar{x} = \frac{e_0}{2} + \frac{\sigma^2}{2e_0} \quad (4.69)$$

**Démonstration** Par définition :

$$\begin{aligned} \bar{x} &= \frac{\int_0^{\infty} xP(x) dx}{\int_0^{\infty} P(x) dx} \\ &= \frac{\int_0^{\infty} xS(x) dx}{\int_0^{\infty} S(x) dx} \\ &= \frac{1}{e_0} \int_0^{\infty} xS(x) dx \end{aligned}$$

or en intégrant par partie, nous obtenons :

$$\int_0^{\infty} xS(x) dx = \left[ -\frac{x^2}{2} S(x) \right]_0^{\infty} - \int_0^{\infty} \frac{x^2}{2} S'(x) dx,$$

soit encore en utilisant la relation 4.57

$$\int_0^{\infty} xS(x) dx = \frac{1}{2} \int_0^{\infty} x^2 d(x) dx,$$

comme de plus la variance de la distribution des décès vaut par définition :

$$\sigma^2 = \int_0^{\infty} (x - e_0)^2 d(x) dx = \int_0^{\infty} x^2 d(x) dx - e_0^2,$$

nous avons finalement :

$$\bar{x} = \frac{e_0}{2} + \frac{\sigma^2}{2e_0}$$

**Proposition 4.2.2** *L'âge moyen d'une population stationnaire est aussi égal à la moyenne des espérances de vie à chaque âge pondérées par l'effectif de la population à cet âge. Ainsi :*

$$\bar{x} = \frac{\int_0^{\infty} P(x)e(x) dx}{\int_0^{\infty} P(x) dx} \quad (4.70)$$

**Démonstration** Par définition :  $\bar{x} = \frac{\int xP(x) dx}{\int P(x) dx} = \frac{N \int xS(x) dx}{\int P(x) dx}$ , l'intégrale  $\int xS(x) dx$  peut s'intégrer par partie de la manière symétrique à la proposition précédente :

$$x = u, \quad dv = S(x) dx$$

$$\text{d'où } dx = du \quad \text{et } v = \int_x^{\infty} S(t) dt = -e(x)S(x) dx.$$

Enfin :

$$\int_0^{\infty} xS(x) dx = [-xe(x)S(x)]_0^{\infty} - \int_0^{\infty} -e(x)S(x) dx = \int_0^{\infty} e(x)S(x) dx$$

et finalement

$$\bar{x} = \frac{\int_0^{\infty} e(x)P(x) dx}{\int_0^{\infty} P(x) dx},$$

ce qu'il fallait démontrer.

## 4.2.2 Pyramide des années à vivre

**Théorème 4.1** *Dans une population stationnaire on peut dénombrer autant d'individus âgés de  $x$  années, que d'individus ayant  $x$  années à vivre.*

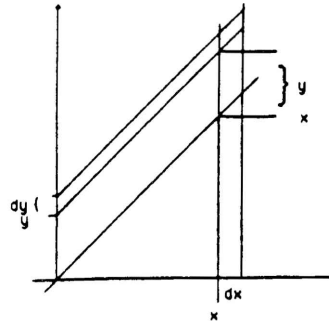


Figure 4.3: Diagramme de Lexis pour le calcul des décès.

**Démonstration** Dans une population stationnaire, le nombre d'individus âgés entre  $x$  et  $x + dx$  années à un instant quelconque  $t$  est  $P(x) dx = NS(x) dx$ .

Il suffit de calculer le nombre de décès attendus entre les instants  $t + x$  et  $t + x + dx$ , parmi la seule population présente à l'instant  $t$ , voir le diagramme de Lexis de la figure 4.3.

- sur les  $P(y) dy = NS(y) dy$  individus présents à l'instant  $t$ ,  $NS(y) dy \times \frac{S(y+x)}{S(y)}$  survivront  $x$  années, et  $NS(y) dy \times \frac{S(y+x)}{S(y)} \times \mu(y+x) dx$  décèderont entre les instants  $t + x$  et  $t + x + dx$ .
- sur l'ensemble de la population présente à l'instant  $t$ , il suffit de sommer sur tous les âges  $y$  :

$$\begin{aligned} N \int_{y=0}^{\infty} S(y) dy \frac{S(y+x)}{S(y)} \mu(y+x) \times dx &= N \int_{y=0}^{\infty} S(y+x) \mu(y+x) dy \times dx \\ &= -N \int_{y=0}^{\infty} S'(y+x) dy \times dx \\ &= -N [S(y+x)]_{y=0}^{\infty} \times dx \\ &= NS(x) \times dx \end{aligned}$$

Il est intéressant de tracer le graphe de la distribution de la population non pas suivant l'âge, mais suivant le délai qui lui reste à vivre ; nous pouvons représenter cette distribution sous la forme d'une pyramide et nous l'avons intitulée *pyramide des années à vivre*. La propriété démontrée ci-dessus s'énonce aussi :

**Propriété 4.2.4** Dans une population stationnaire, la pyramide des âges et la pyramide des années à vivre sont identiques.

**Exercice 11** Montrer que si la courbe de l'espérance de vie suivant l'âge a un maximum, il est atteint pour un âge  $x$  vérifiant la relation :

$$\mu(x) = \frac{1}{e_x}$$



**Solution 11** Par définition :  $e(x) = \frac{1}{S(x)} \int_x^\infty S(t) dt$ , ainsi, la dérivée  $e'(x)$  vaut :

$$e'(x) = -\frac{S'(x)}{S(x)^2} \int_x^\infty S(t) dt + \frac{1}{S(x)} (-S(x)) = \mu(x)e(x) - 1$$

Elle s'annule bien pour la valeur :

$$\mu(x) = \frac{1}{e_x}$$

### 4.2.3 Relations entre le taux de mortalité, la force et le quotient

**Théorème 4.2** Le taux de mortalité,  ${}_hM_x$ , défini sur l'intervalle d'âge fini  $[x, x+h]$  est voisin de la force de mortalité à l'âge médian,  $x + \frac{h}{2}$ .

$${}_hM_x \simeq \mu\left(x + \frac{h}{2}\right) \tag{4.71}$$

A revoir

**Relations exactes** Rapellons, avec l'équation 4.54 que :

$$S(x) = e^{-\int_x^{x+h} \mu(t) dt} \tag{4.72}$$

et  ${}_hM_x = \frac{S(x) - S(x+h)}{\int_x^{x+h} S(t) dt}$  (4.73)

enfin  ${}_hq_x = 1 - e^{-\int_x^{x+h} \mu(t) dt}$  (4.74)

$$= \int_0^h \mu(x+k) e^{-\int_x^{x+k} \mu(t) dt} dk \tag{4.75}$$

$$= \frac{S(x) - S(x+h)}{S(x)} \tag{4.76}$$

**Démonstration du théorème** La démonstration du théorème fait appel à deux approximations, ou plutôt à deux hypothèses de linéarité des fonctions  $S$  et  $\mu$ . Ainsi :

$$\begin{aligned} \int_x^{x+h} S(t) dt &\simeq h \frac{S(x)+S(x+h)}{2} \Rightarrow \frac{1}{h \times {}_hM_x} \simeq \frac{1}{{}_hq_x} - \frac{1}{2} \\ \int_x^{x+h} \mu(t) dt &\simeq h\mu\left(x + \frac{h}{2}\right) \Rightarrow \mu\left(x + \frac{h}{2}\right) \simeq \frac{1}{h} \log\left(\frac{1}{1-{}_hq_x}\right) \end{aligned} \tag{4.77}$$

La relation entre le taux et le quotient s'écrit aussi :

$${}_hM_x \simeq \frac{1}{h} \frac{{}_hq_x}{1 - \frac{{}_hq_x}{2}} \tag{4.78}$$

Montrons que les équations 4.77 et 4.78 ont même développement limité au second ordre en  ${}_h q_x$ . Rappelons en effet :

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

$$\Rightarrow -\log(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots + \frac{x^n}{n} + \dots$$

Ainsi :

$${}_h M_x \simeq \frac{1}{h} \left[ {}_h q_x + \frac{{}_h q_x^2}{2} + \frac{{}_h q_x^3}{4} + \dots + \frac{{}_h q_x^{n+1}}{2^n} + \dots \right] \quad (4.79)$$

$$\text{et } \mu \left( x + \frac{h}{2} \right) \simeq \frac{1}{h} \left[ {}_h q_x + \frac{{}_h q_x^2}{2} + \frac{{}_h q_x^3}{3} + \dots + \frac{{}_h q_x^{n+1}}{n+1} + \dots \right] \quad (4.80)$$

CQFD

### 4.3 Moments et cumulants

Soit  $F$  une loi de probabilité et  $f(x)$  sa densité sur le domaine réel. Il est nécessaire de rappeler ici la définition de quelques fonctions très utiles en statistique et plus particulièrement dans le domaine démographique.

#### 4.3.1 Fonction génératrice des fréquences

On note  $P(t)$  la fonction génératrice des fréquences définie pour tout  $t$  réel par :

$$P(t) = \int t^x dF = \int f(x)t^x dx \quad (4.81)$$

Dans le cas où la distribution est discrète, la définition devient :

$$P(t) = \sum_r f_r t^r \quad \text{où } f_r = \Pr(F = r). \quad (4.82)$$

Ainsi, par exemple, dans le cas d'un tirage binomial de paramètre  $p \in [0, 1]$  dans un échantillon de taille  $n$ , la fréquence de  $r$  vaut :

$$f_r = C_n^r p^r (1-p)^{n-r}$$

et la fonction génératrice des fréquences est :

$$P(t) = \sum_{r=0}^n C_n^r p^r (1-p)^{n-r} t^r = \{pt + (1-p)\}^n.$$

### 4.3.2 Fonction génératrice des moments

Il suffit de remplacer  $t$  par  $e^t$  dans la définition ci-dessus et on obtient la définition de la fonction génératrice des moments  $M(t)$  :

$$M(t) = \int_{-\infty}^{\infty} f(x)e^{tx} dx. \quad (4.83)$$

Comme l'intégrale n'est pas toujours convergente, la fonction génératrice des moments n'est pas toujours définie. Il en est de même dans le cas discret car la série  $e^{tx} f_r$  n'est pas non plus toujours convergente.

Effectuons maintenant le développement en série de  $e^{tx}$  :

$$e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \dots + \frac{t^n x^n}{n!} + \dots$$

Sous la condition que la série  $\frac{t^n x^n}{n!} f(x)$  soit uniformément convergente, on peut intégrer terme à terme la série entre des bornes finies :

$$M(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \left( \int_{-A}^A f(x) x^r dx \right). \quad (4.84)$$

Si chacune des intégrales existe lorsque  $A$  tend vers l'infini, c'est-à-dire si tous les moments non centrés de la distribution,  $\mu'_r = \int_{-\infty}^{\infty} f(x) x^r dx$ , existent alors on peut écrire :

$$M(t) = \int_{-\infty}^{\infty} f(x) e^{tx} dx \quad (4.85)$$

$$= \sum_{r=0}^{\infty} \frac{\mu'_r t^r}{r!}. \quad (4.86)$$

**Relations entre les moments centrés et non centrés** Par définition le moment d'ordre  $r$  s'écrit :

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x) dx \quad (4.87)$$

La moyenne  $\mu'_1$  s'écrit :

$$\mu'_1 = \int_{-\infty}^{\infty} x f(x) dx \quad (4.88)$$

Le moment centré d'ordre  $r$ ,  $\mu_r$ , vaut par définition :

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu'_1)^r f(x) dx. \quad (4.89)$$

On note aussi :

$$\mu'_0 = \mu_0 = \int_{-\infty}^{\infty} f(x) dx = 1.$$

En écrivant la relation :

$$\begin{aligned} x^r &= (x - \mu'_1 + \mu'_1)^r \\ &= \sum_{j=0}^r C_r^j (x - \mu'_1)^{r-j} \mu'^j_1, \end{aligned}$$

nous obtenons une expression des moments non centrés en fonction des moments centrés d'ordre inférieur ainsi que de la moyenne  $\mu'_1$  :

$$\mu'_r = \sum_{j=0}^r C_r^j \mu'_{r-j} \mu'^j_1. \quad (4.90)$$

Ainsi pour les quatre premiers moments :

$$\begin{cases} \mu'_1 &= \text{moyenne} \\ \mu'_2 &= \mu_2 + \mu'^2_1 \\ \mu'_3 &= \mu_3 + 3\mu'_1\mu_2 + \mu'^3_1 \\ \mu'_4 &= \mu_4 + 4\mu'_1\mu_3 + 6\mu'^2_1\mu_2 + \mu'^4_1. \end{cases} \quad (4.91)$$

### 4.3.3 Fonction caractéristique

Au lieu de faire la transformation  $t \rightarrow e^t$  il est préférable de faire la transformation  $t \rightarrow e^{it}$  ; on obtient ainsi la fonction caractéristique  $\Phi(t)$  :

$$\Phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx \quad (4.92)$$

Cette fonction a un intérêt plus grand que les précédentes car on montre qu'il y a bijection entre l'espace des distributions de probabilité et l'espace des fonctions caractéristiques. Autrement dit une fonction caractéristique  $\Phi(t)$  est *caractéristique* d'une distribution de probabilité  $f(x)$ . Réciproquement une distribution de probabilité  $f(x)$  est entièrement déterminée par sa fonction caractéristique  $\Phi(t)$ .

**Propriété** La fonction  $\Phi(t)$  existe pour tout  $t$ . En effet :

$$|\Phi(t)| \leq \int_{-\infty}^{\infty} |e^{itx}| f(x) dx = 1$$

On peut alors intégrer ou dériver sous le signe intégral à condition que les expressions obtenues existent et soient uniformément convergentes. Ainsi la dérivée première s'écrit :

$$\frac{d\Phi(t)}{dt} = i \int_{-\infty}^{\infty} e^{itx} x f(x) dx \quad (4.93)$$

Soit pour  $t = 0$  :

$$\frac{d\Phi(0)}{dt} = i\mu'_1$$

De même

$$\mu'_r = (-i)^r \left[ \frac{d^r \Phi(t)}{dt^r} \right]_{t=0}$$

Une fonction caractéristique est donc aussi une fonction génératrice des moments.

### 4.3.4 Cumulants

Les premiers moments d'une distribution de probabilité sont des constantes utiles pour caractériser cette distribution mais ce ne sont pas les seules ; en particulier on fait parfois appel aux *cumulants*, notés  $k_1, k_2, \dots, k_r$ , et définis en liaison avec les moments non centrés  $\mu'_1, \mu'_2, \dots, \mu'_r$  par l'équation en  $t$  suivante :

$$\exp \left( k_1 t + \frac{k_2 t^2}{2!} + \dots + \frac{k_r t^r}{r!} + \dots \right) = 1 + \mu'_1 t + \frac{\mu'_2 t^2}{2!} + \dots + \frac{\mu'_r t^r}{r!} + \dots \quad (4.94)$$

Si on prend  $it$  au lieu de  $t$  on obtient :

$$\exp \left( k_1 it + \frac{k_2 (it)^2}{2!} + \dots + \frac{k_r (it)^r}{r!} + \dots \right) = \quad (4.95)$$

$$1 + \mu'_1 it + \frac{\mu'_2 (it)^2}{2!} + \dots + \frac{\mu'_r (it)^r}{r!} + \dots = \quad (4.96)$$

$$= \int_{-\infty}^{\infty} e^{itz} f(x) dx \quad (4.97)$$

$$= \Phi(t) \quad (4.98)$$

Le cumulant d'ordre  $r$  est alors le coefficient de  $(it)^r / r!$  dans le développement en série de  $\log \Phi(t)$  si celui-ci existe.

Un développement au premier ordre de l'équation 4.94 donne :

$$1 + k_1 t + \frac{k_2 t^2}{2!} + (k_1 t)^2 + \dots = 1 + \mu'_1 t + \frac{\mu'_2 t^2}{2!} + \dots,$$

d'où

$$\begin{aligned} k_1 &= \mu'_1, \\ k_2 &= \mu'_2 - \mu'^2_1 = \mu_2 \end{aligned}$$

Pour exprimer plus généralement les cumulants en fonction des moments, il convient de prendre le logarithme des deux membres de l'équation 4.94 et de développer en série :

$$k_1 t + \dots + \frac{k_r t^r}{r!} + \dots = \log \left( 1 + \mu'_1 t + \dots + \frac{\mu'_r t^r}{r!} + \dots \right)$$

Ainsi en égalisant les termes en  $t^r$  :

$$k_r = r! \sum_{m=0}^r \sum_{\pi_1, \dots, \pi_m} \left( \frac{\mu'_1}{p_1!} \right)^{\pi_1} \dots \left( \frac{\mu'_m}{p_m!} \right)^{\pi_m} \frac{(-1)^{r-1} (\rho-1)!}{\pi_1! \dots \pi_m!} \quad (4.99)$$

où la seconde sommation s'effectue sur tous les  $\pi_i \geq 0$  et sur  $\rho \geq 0$  tels que :

$$\begin{aligned} p_1 \pi_1 + p_2 \pi_2 + \dots + p_m \pi_m &= r, & \text{et} \\ \pi_1 + \pi_2 + \dots + \pi_m &= \rho. \end{aligned}$$

Donnons alors les valeurs des 4 premiers cumulants en fonction des moments non centrés :

$$\begin{cases} k_1 = \mu'_1 \\ k_2 = \mu'_2 - \mu_1'^2 \\ k_3 = \mu'_3 - 3\mu'_2 \mu'_1 + 2\mu_1'^3 \\ k_4 = \mu'_4 - 4\mu'_3 \mu'_1 - 3\mu_2'^2 + 12\mu'_2 \mu_1'^2 - 6\mu_1'^4. \end{cases} \quad (4.100)$$

Il est sans doute plus utile d'utiliser les moments centrés ; ainsi en vertu des relations 4.91 :

$$\begin{cases} k_1 = \mu_1 \\ k_2 = \mu_2 \\ k_3 = \mu_3 \\ k_4 = \mu_4 - 3\mu_2^2. \end{cases} \quad (4.101)$$

#### 4.4 Relations entre les moments et cumulants d'une table de mortalité

Soit  $S(x)$  la fonction de survie et  $S(x) = \frac{S(x)}{e_0}$  la distribution normée telle que :

$$\int S(x) dx = 1.$$

Cherchons alors le lien entre la f.g.m. de la distribution des survivants,  $S$ , et celle de la distribution des décès,  $d$  :

$$\phi(t) = \int e^{tx} S(x) dx \quad (4.102)$$

$$\psi(t) = \int e^{tx} d(x) dx \quad (4.103)$$

avec  $-S'(x) = \frac{d(x)}{e_0}$ .

Intégrons la seconde intégrale par partie (à condition de vérifier que chaque intégrale est convergente) :

$$\begin{aligned} \int e^{tx} d(x) dx &= [-S(x)e_0 e^{tx}]_0^{\infty} + e_0 t \int e^{tx} S(x) dx \\ &= 1 + e_0 t \phi(t), \end{aligned}$$

soit finalement :

$$\psi(t) = 1 + e_0 t \phi(t) \tag{4.104}$$

Développons les fonctions en série, et nous trouvons :

$$1 + e_0 t \left( 1 + \mu'_1 t + \mu'^2_2 \frac{t^2}{2!} + \dots \right) = \left( 1 + \mu'^d_1 t + \mu'^d_2 \frac{t^2}{2!} + \dots \right)$$

En égalisant les coefficients en  $t^n$  :

$$\begin{aligned} \mu'_0 &= e_0 = \mu'^d_1 \\ e_0 \mu'^s_1 &= \frac{\mu'^d_2}{2}, \\ &\dots \\ e_0 \frac{\mu'^s_{n-1}}{(n-1)!} &= \frac{\mu'^d_n}{n!} \end{aligned}$$

Ainsi, pour les premiers moments, nous retrouvons d'une part que l'aire de la fonction de survie est égale à l'âge moyen au décès, et d'autre part que l'âge moyen d'une population stationnaire est égal à  $\mu'^d_2 / (2e_0)$  soit :

$$\mu'^s_1 = \frac{1}{e_0} \left( \frac{\sigma^2 + e_0^2}{2} \right) = \frac{e_0}{2} + \frac{\sigma^2}{2e_0}.$$

Plus généralement, à l'ordre  $n$  :

$$\mu'^d_n = n e_0 \mu'^s_{n-1} \tag{4.105}$$

(S)

**Exercice 12** Montrer que la fonction génératrice,  $\xi(t)$ , de l'espérance de vie à chaque âge vérifie la relation :

$$t^2 \xi(t) = e^{\omega t} - 1 + t \left[ e(\omega) e^{\omega t} - e_0 - \int_0^\omega e^{tx} e(\omega) \mu(x) dx \right]. \tag{4.106}$$

**Solution 12** On note,  $e(x)$ , l'espérance de vie à l'âge  $x$  :

$$e(x) = \frac{1}{S(x)} \int_x^{\infty} S(u) du.$$

Soit alors,  $\xi(t)$ , sa fonction génératrice des moments :

$$\xi(t) = \int e^{tx} e(x) dx.$$

En intégrant par partie cette équation, nous trouvons :

$$\begin{aligned} e'(x) &= \frac{-S(x)}{S(x)} - \frac{\int_x^\infty S(u) du}{S(x)^2} \\ &= -1 + \mu(x) e(x), \end{aligned}$$

$$\xi(t) = \left[ \frac{1}{t} e^{tx} e(x) \right]_0^{\omega} - \frac{1}{t} \int e^{tx} e(x) \mu(x) dx,$$

soit encore

$$t^2 \xi(t) = e^{\omega t} - 1 + t \left[ e(\omega) e^{\omega t} - e_0 - \int_0^{\omega} e^{tx} e(x) \mu(x) dx \right]. \quad (4.107)$$



## Chapitre 5

# Analyse de la fécondité par une étude de cas

Dans les précédents chapitres sur la population stationnaire, nous avons fait une hypothèse simple sur la loi qu'est censé suivre le nombre annuel des naissances, puisque nous avons supposé que celui-ci était constant. Il s'agit d'une approximation grossière car, pour reprendre l'exemple de la France, la première guerre mondiale a mobilisé beaucoup de ses hommes, et le nombre des naissances a fortement décliné. Le processus de naissance n'est donc pas exogène à la population, mais déterminé par la population elle-même.

C'est à la recherche de cette causalité que nous allons nous appliquer dans la première partie de ce chapitre avec une approche empirique basée sur des séries historiques. Puisque des séries chronologiques longues et détaillées ne sont pas disponibles pour des pays peu développés, nous prendrons de nouveau la France comme exemple principal. C'est un pays dont les pyramides des âges ont été très marquées par la guerre de 1914-18, et se révèle de ce fait très instructive. Parfois, quand l'exemple de la France se révélera insuffisant ou trop particulier, nous utiliserons d'autres statistiques et en particulier celles d'Afrique.

### 5.1 La série brute des naissances et du taux de natalité

La figure 5.1 représente le nombre brut des naissances en France depuis le début du siècle. Le chiffre avoisine autour de 800 000, sans tendance de long terme, mais avec des fluctuations et ruptures très nettes. Les deux guerres mondiales montrent clairement leur effet. Le baby-boom qui a suivi la seconde guerre est très important et semble s'être prolongé plusieurs années après.

Ces premiers commentaires se rapportent aux fluctuations du nombre des

... question : dans le cas de la France, la forte mobilisation des hommes pendant la première guerre mondiale a fait décroître le nombre des naissances

l'influence des 2 guerres mondiales est net -

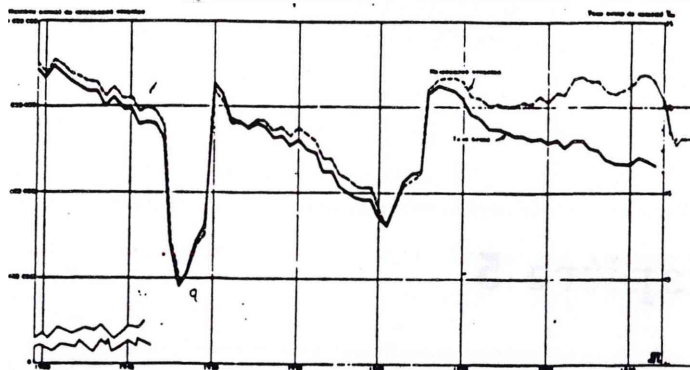


Figure 5.1: Série des naissances depuis le début du siècle

naissances, mais qu'indiquent-ils ou qu'induisent-ils sur l'évolution du *pouvoir créateur* de la population ? En particulier, si le nombre des naissances baisse peu durant les vingt années qui suivent le baby-boom, peut-on en déduire que la fécondité des familles s'est maintenue ?

Une première étape consiste à rapporter le nombre des naissances à la taille de la population afin d'obtenir un *taux brut de natalité* ; en effet, la population, qui était de 40 millions en 1900, a atteint 53 millions en 1982, la contribution de la population entière à sa reproduction, contribution maintenant mesurée par le taux brut de natalité a donc décliné. Nous nous en apercevons effectivement en superposant sur cette même figure, les deux courbes à des échelles différentes. Les deux tracés ne divergent réellement qu'après 1946. Avant cette date, la première mesure, à savoir le nombre brut des naissances était suffisante, au delà, la seconde (le taux brut) apporte de l'information en plus en révélant une baisse. Mais s'agit-il après 1946 d'une baisse réelle de la fécondité des familles, celles-ci restreignant volontairement leur taille, ou bien cette baisse n'est-elle pas due uniquement au fait que la population vieillissant, un nombre beaucoup plus important de personnes ayant dépassé l'âge de procréer vient gonfler artificiellement le dénominateur du taux.

Il convient donc de définir dans une seconde étape un indice plus sensible en rapportant les naissances aux seuls individus d'âge fécond. Si le nombre d'hommes diffère peu de celui des femmes, ce qui est généralement le cas, on peut se contenter de rapporter le nombre des naissances aux seules femmes. En effet, dans ce cas, la période où la femme est *biologiquement* féconde est assez bien connue et est généralement définie comme allant de 15 à 49 ans environ. Ce nouvel indice est souvent appelé *taux de fécondité générale*, mais il est préférable de le redéfinir lorsqu'on l'emploie. La figure 5.2 retrace l'évolution de ce taux depuis le début du siècle. On peut remarquer que ce taux corrige en effet le précédent du vieillissement de la population, mais il n'est pas non plus suffisant.

Comme on sait, effectivement que la fécondité a décliné et est maintenant bien éloignée de la fécondité biologique ou fécondité dite naturelle, il devient ridicule de compter les femmes de 40 à 50 ans avec le même poids que celles de

à faire.

Figure 5.2: Evolution du taux de fécondité des femmes âgées entre 15 et 49 ans en France.

20 à 30 ans par exemple ; il convient alors de définir des indices plus sensibles encore qui tiennent compte de l'évolution du potentiel reproducteur de *chaque classe d'âge*. Avant de poursuivre, faisons néanmoins deux remarques :

1. Si la fécondité effective est assez éloignée de la fécondité *biologique*, on doit alors admettre que la population *contrôle* sa reproduction ; dans ce cas il n'y a pas lieu de privilégier un des sexes, homme ou femme, comme responsable de la procréation.
2. Les étapes successives décrites ci-dessus, naissances, taux brut de natalité, taux de fécondité générale ont pour but de mieux mesurer le pouvoir reproducteur d'une population en éliminant des éléments a priori parasites comme la population âgée par exemple. Mais ce procédé privilégie certaines variables et a tendance à faire considérer les variables retenues comme *explicatives* du phénomène. Cette approche ressemble à la recherche d'une loi physique telle celle de Kepler pour trouver la constante de l'attraction universelle ; mais Kepler a réellement trouvé une constante, indépendante du temps, du lieu etc., ce qui a justifié d'un coup toute la théorie. Les démographes ne semblent pas en être arrivés là. Il semble même le plus souvent que plus on précise un modèle, plus on introduit de variabilité. Ainsi, par exemple, la série des taux de fécondité générale a plus de variabilité que la série des naissances brutes ; autrement dit les personnes âgées jouent peut-être, finalement, un petit rôle régulateur inconnu<sup>1</sup> !

Après ces trois étapes successives, nous sommes moins ignorants qu'avec la série initiale des naissances brutes ; le démographe n'est pas en effet à la recherche d'une impossible constante, mais s'intéresse aussi à la décomposition

Seule

<sup>1</sup> On peut par exemple faire l'hypothèse que la survie plus probable des grands-parents est un atout car ces derniers facilitent la vie des jeunes ménages.

des phénomènes complexes en phénomènes plus simples, quitte à recombinaison le tout en un *modèle* plus cohérent.

Nous cherchons par exemple ici à mesurer la fécondité *intrinsèque* d'une population, c'est-à-dire dans un premier temps à connaître le modèle statistique qui conduit une population à procréer, puis dans un second temps à en estimer les paramètres.

Une fréquence est le rapport d'un nombre d'événements à un effectif initial soumis au risque. Cette fréquence doit mesurer le *risque de procréer*. L'événement (le numérateur), est mesurable puisqu'une conception donne lieu à une naissance<sup>2</sup> répertoriée à l'état civil. Le dénominateur est plus délicat à définir ; quelle partie de la population est soumise au risque de procréer ? Le groupe d'âge, le statut matrimonial, le fait d'être stéril ou fécond, le fait d'attendre déjà un enfant sont autant de critères qui permettent de cerner une population. Plus on restreint la population, plus le risque est élevé ; à la limite, si on rapporte le nombre des naissances aux seuls couples ayant eu un enfant dans l'année, le risque est maximum<sup>3</sup>.

Mais ce qui intéresse le démographe, c'est de définir des calculs de risque qui puissent être appliqués à une population la plus large possible et restent néanmoins pertinents et cohérents. Par exemple si l'âge est un critère de sélection, il est essentiel que la distribution des risques par âge soit régulière, car on ne comprendrait pas la raison de fluctuations erratiques.

En définitive, un "bon" modèle de fécondité peut, par exemple, être défini comme un modèle qui permet de prévoir correctement les naissances futures. Parmi les bons modèles, les meilleurs sont ceux qui utilisent le moins de données possibles, et des données aisées à obtenir.

Nous sommes loin de la physique où une loi dépend d'un nombre fixe et indispensable de paramètres. Mais la qualité des résultats aussi est éloignée de celle de la physique. Dans les années 1970, un nombre important d'instituts de statistique nationaux ont publié des projections du nombre de naissances avec un large éventail de scénarios concernant la fécondité du mariage. Les taux de fécondité par durée de mariage ainsi calculés permettent de faire des comparaisons du pouvoir reproducteur (des couples mariés) entre différents pays, ou d'un même pays à différentes dates par exemple.

Si le processus de formation des couples est connu et qu'ainsi des modèles de projection du nombre des futurs mariés sont fiables, alors la combinaison de ce dernier modèle de nuptialité, avec le modèle de fécondité dans le mariage décrit plus haut, permet de prévoir correctement les naissances, et de mesurer le pouvoir reproducteur d'une population (cette décomposition du processus de naissance, en mariage et fécondité dans le mariage, sera étudié en détail au

<sup>2</sup>Plus rarement à un avortement.

<sup>3</sup>Même dans cet exemple, pourtant triviale, le phénomène est déjà complexe, puisqu'un même homme peut, dans une population polygame, être compté parmi plusieurs couples.

chapitre 8). Pour le moment, nous pouvons dire que la nuptialité, que ce soit dans les sociétés européennes anciennes ou dans les pays peu développés, est encore plus difficile à modéliser que la fécondité elle-même. Ensuite, le mariage est de moins en moins la condition préalable à la procréation en Europe, et ne l'a jamais été dans beaucoup de sociétés (Antilles par exemple [29]) ; aussi, pour prévoir les naissances, les modèles actuels tendent à faire dépendre la fécondité non pas de la durée du mariage mais directement de l'âge du couple procréateur, marié ou non.

## 5.2 Fécondité par âge

Mais aussitôt une première difficulté apparaît : comment définir l'âge d'un couple ? Même s'il existe des couples âgés et des couples plus jeunes, nous savons que l'écart d'âge entre époux peut être important et très variable. *d'un couple à l'autre*

Il est possible d'obtenir, par le dépouillement des bulletins de naissance de l'état civil, le tableau qui croise les naissances d'une année suivant l'âge de la mère et l'âge du père. Si on rapportait ces nombres aux effectifs de couples de mêmes âges croisés présents cette année là<sup>4</sup>, nous obtiendrions alors une fécondité des couples, ou probabilité pour un couple dont l'homme a l'âge  $x$  et la femme l'âge  $y$  de donner naissance à un enfant dans l'année. Pour prévoir ensuite les naissances, il faudrait être capable de prévoir le nombre des couples chaque année par âge croisé. Or, ce dernier problème, comme on l'a dit plus haut, est plus difficile à résoudre que le premier ; d'autre part, des tableaux croisés sont toujours plus difficiles à décrire ou à résumer, que des vecteurs. Aussi, on préfère généralement ne pas tenir compte du croisement des âges et classer les naissances suivant l'âge de la mère et parfois aussi suivant l'âge du père, mais séparément.

En rapportant ces naissances par âge d'un des parents à la population totale du sexe correspondant, classée par âge, on évalue un risque commun à tous les individus d'une même classe d'âge, indépendamment du fait que ces individus soient en couple ou non.

Lorsque la femme n'est pas mariée à la naissance de son enfant, il n'est pas encore demandé pour des raisons de discrétion, et aussi, il est vrai, parce que la femme ne détient pas, le plus souvent, encore de livret de famille, de renseignement sur le père. La part des naissances dont la mère n'est pas mariée, naissances anciennement qualifiées "d'illégitimes" a oscillé autour de 8 % depuis plus d'un siècle, mais augmente fortement depuis 1968, comme le montre la figure 5.3, pour atteindre 16 % en 1984.

La figure 5.4 donne les deux répartitions suivant l'âge du père et suivant l'âge de la mère, des naissances provenant de couples mariés. On a ajouté à la

<sup>4</sup>Un recensement ne fournit pas de telles données, et il serait donc très difficile d'estimer ce dénominateur.

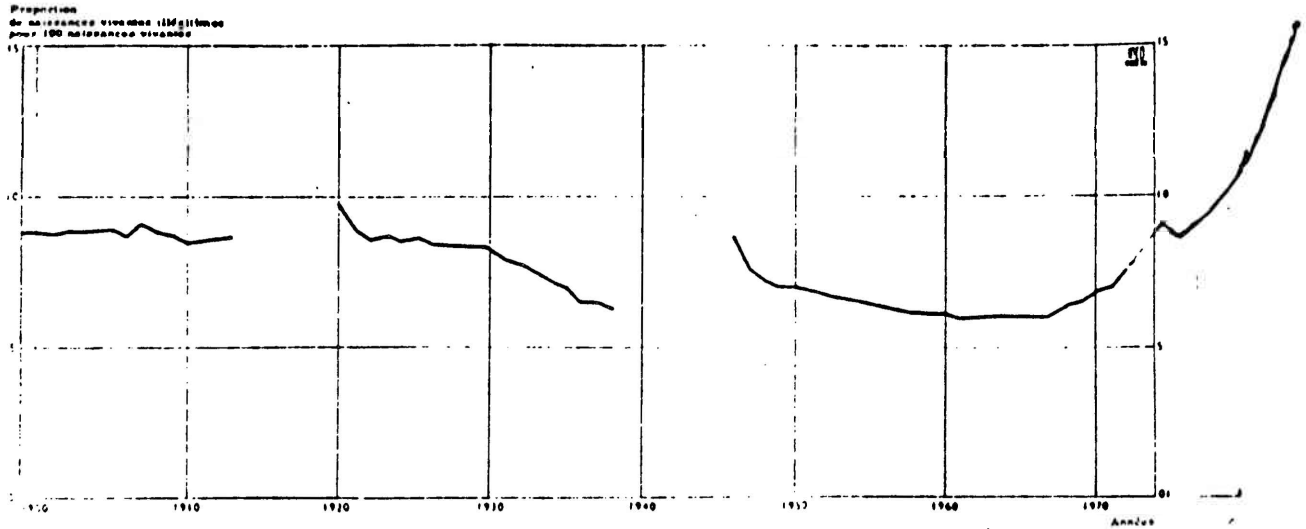


Figure 5.3: Pourcentage de naissances dont la mère n'est pas marié (naissances "illégitime").

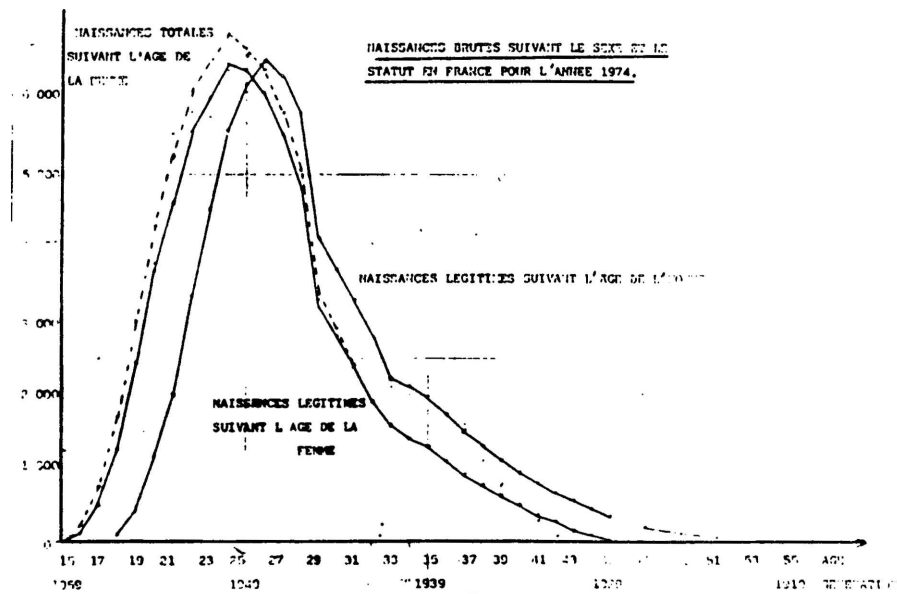


Figure 5.4: Distributions des naissances de l'année 1974 en France suivant l'âge du père et l'âge de la mère.

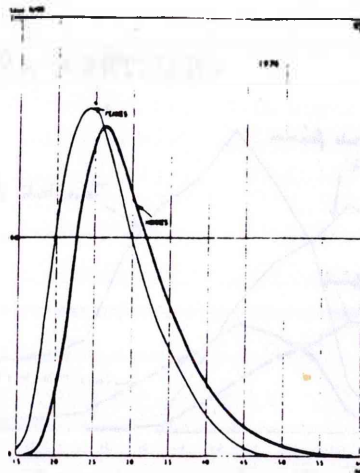


Figure 5.5: Taux de fécondité générale suivant l'âge de l'homme et de la femme. France 1974.

courbe des mères mariées, la distribution des naissances par âge des mères non mariées, pour obtenir la distribution des naissances totales<sup>5</sup>.

Ces deux distributions, féminine et masculine, montrent grossièrement l'âge auquel les Françaises et Français procréent ; mais deux ruptures importantes à 29 puis 33 ans, ne peuvent s'expliquer qu'en examinant les distributions par âge des populations présentes au cours de l'année 1974 (ces ruptures proviennent ainsi des faibles naissances durant la guerre 1940-45). De plus, le nombre faible de naissances de parents d'âge élevé s'explique en partie par une moindre représentation de ces classes d'âges dans la pyramide des âges.

Il convient donc de rapporter les naissances aux populations pour évaluer des risques. Le calcul est celui d'un taux : on rapporte les naissances survenues dans un domaine de Lexis (âge  $x$ ,  $x + 1$ , année  $t$  par exemple) à la somme des années vécues dans ce domaine par cette sous-population. Suivant le type des données : âge en années révolues, âge au premier janvier, âge quinquennal etc., toute sorte de taux peut être calculée.

Ces *taux de fécondité par âge* sont représentés sur la figure 5.5 : les plages d'âge où le risque de concevoir n'est pas négligeable sont sensiblement différents chez les hommes (18 ans à 60 ans environ) de chez les femmes (15 ans à 50 ans). Ce dernier intervalle coïncide avec l'intervalle où la femme est biologiquement féconde. Pourtant, si les risques sont déjà très réduits bien après 35 ans, lorsqu'on les compare à ceux rencontrés aujourd'hui en Afrique par exemple (voir la figure 5.6), ou plus généralement à ceux de la fécondité dite "naturelle" (voir Henry [23]). Pour les hommes, les courbes françaises ne reflètent pas du tout la fécondité biologique, qui est d'ailleurs difficile à mesurer. Même si la fécondité n'était pas contrôlée, comme elle l'est actuellement, celle des hommes serait nulle après 60 ans, puisqu'en régime monogame, la part des femmes encore en

<sup>5</sup> Il est possible d'obtenir la distribution totale des naissances suivant l'âge du père, en ventilant les naissances illégitimes par un procédé comme une répartition proportionnelle ou en tenant compte de l'information détenue dans les tableaux croisant les naissances légitimes suivant l'âge des deux parents (voir Brouard [6].)

Toutes sortes

pour quelle ?

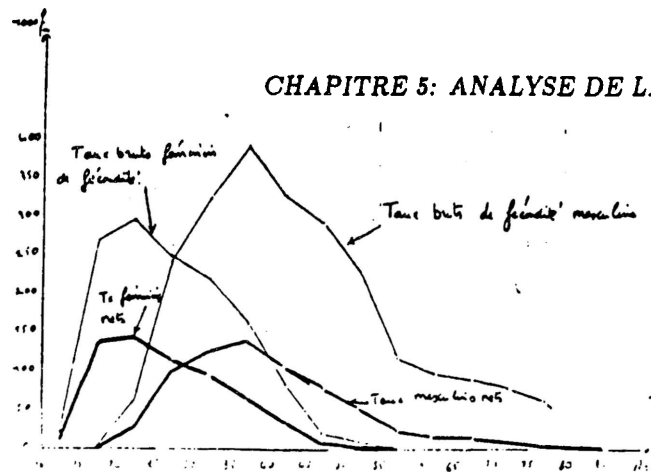


Figure 5.6: Fécondité générale par âge chez les Peuls-Bandé.

âge de procréer, et mariées à des hommes de 60 ans est très faible. Nous allons voir qu'il n'en est rien dans une société où la polygamie est possible avec cet exemple provenant d'une étude sur une population de quelques villages situés au Sénégal Oriental (voir Pison [36]).

La figure 5.6 montre tout d'abord l'importance de la fécondité des Sénégalaises comparée à celles des Françaises. En second lieu, nous remarquons la disproportion entre la fécondité masculine et la fécondité féminine aux âges élevés. La fécondité des hommes de plus de 80 ans n'est même plus estimable car, et le numérateur, et le dénominateur des taux à savoir, pour ce dernier, le nombre des survivants, sont trop faibles.

Un taux de fécondité mêle étroitement deux phénomènes, l'un d'ordre biologique, et l'autre d'ordre social. Dans l'exemple africain, la fécondité masculine est négligeable avant 20 ans du fait de l'absence de mariage avant cet âge. Elle reste importante après 60 ans, du fait d'un second, troisième, ou nième mariage avec une femme le plus souvent beaucoup plus jeune. La fécondité est néanmoins en déclin au-delà de ces âges, car l'âge moyen des épouses croît et le nombre des épouses en âge de procréer semble être passé par un maximum<sup>6</sup>.

Après ces calculs de taux de fécondité par âge, et, devant l'extrême régularité de la courbe<sup>7</sup> qui prouve dans une certaine mesure l'indépendance vis à vis de la structure par âge de la population, nous pourrions en rester là et décréter que le "pouvoir reproducteur" de la population se mesure par les courbes masculine et féminine de fécondité par âge. Ainsi, si ce pouvoir reproducteur n'évoluait pas, les deux distributions resteraient inchangées. Nous en resterons effectivement là, mais nous verrons que ces taux de fécondité par âge ne sont pas totalement indépendants de la structure par âge à partir de laquelle ils ont été calculés. Pour l'instant, nous allons proposer des indices simples, analogues à l'espérance de vie vue dans le chapitre 3 sur la mortalité pour résumer ces distributions.

<sup>6</sup>Voir Pison op.cit. page ??.

<sup>7</sup>Les fluctuations de la courbe africaine proviennent d'erreurs aléatoires de mesure.



La mesure la plus importante est celle qui reflète l'intensité de la fécondité durant l'année du calcul. Il s'agit de l'aire de chacune des courbes, ou encore somme des taux de fécondité par âge. On appelle aussi cet indicateur *la somme des naissances réduites*<sup>8</sup>. Il s'exprime en nombre d'enfants par homme ou par femme suivant les cas. En effet, en sommant les taux de fécondité par âge, on somme les fractions d'enfant qui proviendrait d'un homme ou d'une femme à chaque âge de sa vie féconde, soumis hypothétiquement à ces risques mesurés cette année là.

Dans l'exemple africain, les données ne proviennent pas d'une année civile mais d'un groupe d'années ; il ne s'agit pas ici de suivre l'évolution conjoncturelle de la fécondité, mais de mesurer, au moins une fois, l'intensité de la fécondité dans une société traditionnelle. Ainsi chez les Peul Bandé, un homme qui survit jusqu'à 80 ans aura été père de 11,7 enfants, alors qu'une femme aura donné naissance à 6,7 enfants. Evidemment le calcul suppose que cet homme ou cette femme survit depuis sa naissance jusqu'à la fin de sa vie féconde ; lorsqu'au chapitre 6 nous ferons intervenir la mortalité pour calculer le pouvoir reproducteur effectif, c'est-à-dire compte tenu des décès, nous verrons que la fécondité importante des hommes aux âges élevés ne concerne que très peu d'entre eux. L'écart entre le nouvel indice masculin et son homologue féminin sera alors fortement réduit. Lorsque de plus nous observerons que le cycle de reproduction de l'homme est plus long que celui de la femme du fait d'un mariage tardif et d'une vie féconde plus élargie en raison de la polygamie, nous verrons que ce pouvoir reproducteur, mesuré en terme de taux de croissance de la population, est identique chez les hommes et les femmes, respectant le fait qu'un enfant est le fruit de l'union d'un seul homme et d'une seule femme.

En France, et plus généralement dans un pays développé monogame, l'écart entre les deux distributions, et donc l'écart entre les premiers moments, sont plus faibles : 2,1 enfants par femme et 2,05 enfants par homme en 1974. L'âge moyen à la paternité est de 28,8 ans et l'âge moyen à la maternité de 26,7 ans. Le premier enfant est fait à 28 ans<sup>9</sup> en moyenne pour les hommes et à 26 pour les femmes.

Ces chiffres relatifs à l'année 1974 ne sont pas immuables et ont varié, mais dans des proportions différentes ; si l'âge moyen ou l'écart-type varient peu, c'est surtout l'intensité de la fécondité qui a décliné et aujourd'hui fluctue.

Formellement, il est même impossible que les deux distributions masculine et féminine de la fécondité par âge en 1974 soient identiques même l'année suivante. En effet, en 1975, les structures par âge des populations masculine et féminine auront vieilli d'un an ; en multipliant cette nouvelle structure par âge par les taux de fécondité par âge, supposés inchangés, puis en sommant les naissances sur tous les âges, nous trouverions un nombre total de naissances

<sup>8</sup>En effet, on somme les naissances par âge d'une année civile, mais auparavant ces chiffres ont été réduits par les effectifs d'hommes ou de femmes de la population du même âge.

<sup>9</sup>Il s'agit de l'âge pour lequel la somme des taux de fécondité par âge atteint 1,0 enfant.

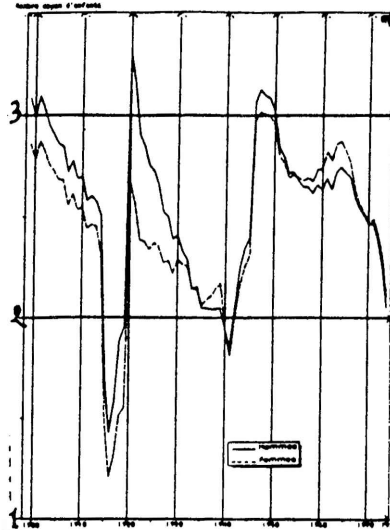


Figure 5.7: Évolution de l'intensité de la fécondité annuelle depuis le début du siècle.

différent suivant que le calcul est effectué sur les hommes ou sur les femmes. Autrement dit, les courbes de fécondité par âge masculine et féminine ne sont pas *intrinsèques* et dépendent encore des structures par âge à partir desquelles elles ont été construites. Cette question d'indépendance des taux vis à vis de la population est délicate, et est même l'objet d'une littérature abondante dont on peut trouver une présentation dans le chapitre sur le "two-sex problem" de l'ouvrage de Pollard [37]. Néanmoins, si la structure par âge n'est pas trop perturbée, les mariages ou unions n'auront pas non plus été trop perturbés, et les différences entre les évolutions des indices masculins et féminins seront insignifiantes. Il n'en a pas été toujours ainsi, particulièrement en France à la suite de la guerre de 1914-18. La figure 5.7 retrace l'évolution de l'intensité de la fécondité chaque année en France depuis 1899, intensité mesurée pour les hommes et pour les femmes. Il s'agit donc de mesures *transversales*.

### 5.3 Variations transversales de la fécondité

Intéressons nous tout d'abord à l'allure générale des courbes de la figure 5.7. La fécondité a baissé depuis la fin du 18<sup>e</sup> siècle du niveau de 6 enfants par femme environ<sup>10</sup> à environ 3 en 1900 (3,1 enfants par homme et 2,9 par femme) ; cette baisse se poursuit jusqu'à la guerre de 1914-18, et est fortement aggravée par

<sup>10</sup>On ne connaît pas d'estimation du nombre d'enfants par homme à cet époque, mais on peut supposer que celui-ci devait lui être légèrement supérieur, puisque l'époux était plus vieux d'environ 5 ans : l'âge moyen à la procréation était donc plus élevé pour les hommes que pour les femmes, et moins d'hommes l'atteignaient ; pour compenser, les hommes devaient faire un peu plus d'enfant.

celle-ci, car les hommes sont mobilisés en masse. La reprise des naissances en 1920 est spectaculaire pour les hommes (3,3 enfants par homme), moindre pour les femmes (2,7 enfants par femmes). Entre les deux guerres, la fécondité tend à reprendre sa baisse séculaire. En 1940, les hommes sont de nouveau mobilisés<sup>11</sup>, ce qui aggrave momentanément la baisse. De 1942 à 1945 la fécondité remonte sensiblement, sans doute en partie en raison de la politique familiale. Le baby-boom de 1946 commun à tous les pays européens est sans précédent. Celui-ci se poursuit pendant 20 ans jusqu'environ l'année 1965. A partir de cette date et jusqu'en 1975 la fécondité chute d'une manière très spectaculaire. Cette période correspond aussi à l'utilisation de plus en plus massive de la pilule (voir par exemple Léridon *et alii* [30]). Depuis 1975, la fécondité fluctue en dessous de 2 enfants. Cette baisse récente touche tous les pays développés comme le montre la figure 5.8. Peu de pays sont encore au-dessus de 2 enfants par femme. Sans anticiper sur le chapitre 6, on comprend que la fécondité ne puisse descendre longtemps en dessous d'un seuil minimal en deçà duquel la population commence à décroître. Ce seuil dépend de la mortalité, c'est à dire du pourcentage des pertes en vie humaine de la naissance jusqu'aux âges auxquels l'homme et la femme procréent. Le chiffre minimal de 2 enfants par homme ou par femme correspond en effet au cas où la mortalité infantile serait nulle, de même celle par par accident entre 17 et 20 ans, et plus généralement la mortalité à tous les âges avant 50 ans.

La considération de la figure 5.8 inquiète donc a priori certains Européens, mais on peut faire au moins deux remarques.

1. La première est que ce n'est pas parce que la fécondité descend au-dessous de 2 enfants qu'il faut commencer à s'inquiéter : en effet la mortalité infantile, en particulier, était très forte au début du siècle et ce n'est peut-être pas avec une fécondité de 3 enfants au début du siècle qu'on pouvait assurer le renouvellement des générations (voir le paragraphe 6.2.2 du chapitre 6). Autrement dit si le problème existe, il n'est pas nouveau.
2. La seconde est que le très faible niveau de la fécondité de certains pays (1,3 enfants en Allemagne) est peut-être temporaire, et qu'ainsi les générations concernées peuvent différer leurs naissances pour une période ultérieure plus favorable. Ainsi le nombre moyen d'enfants de ces générations sera peut-être supérieur à ce niveau transversal extrêmement bas. Rien ne permet de prouver une telle hypothèse à cause du manque de recul statistique mais il convient au moins de faire une analyse de la fécondité par génération, ou longitudinale, pour examiner une mesure plus intrinsèque de la fécondité, la taille finale des familles ou *descendance finale*.

<sup>11</sup>Rappelons ici que les naissances sont décalées de 9 mois par rapport aux conceptions, et regroupées par année civile.

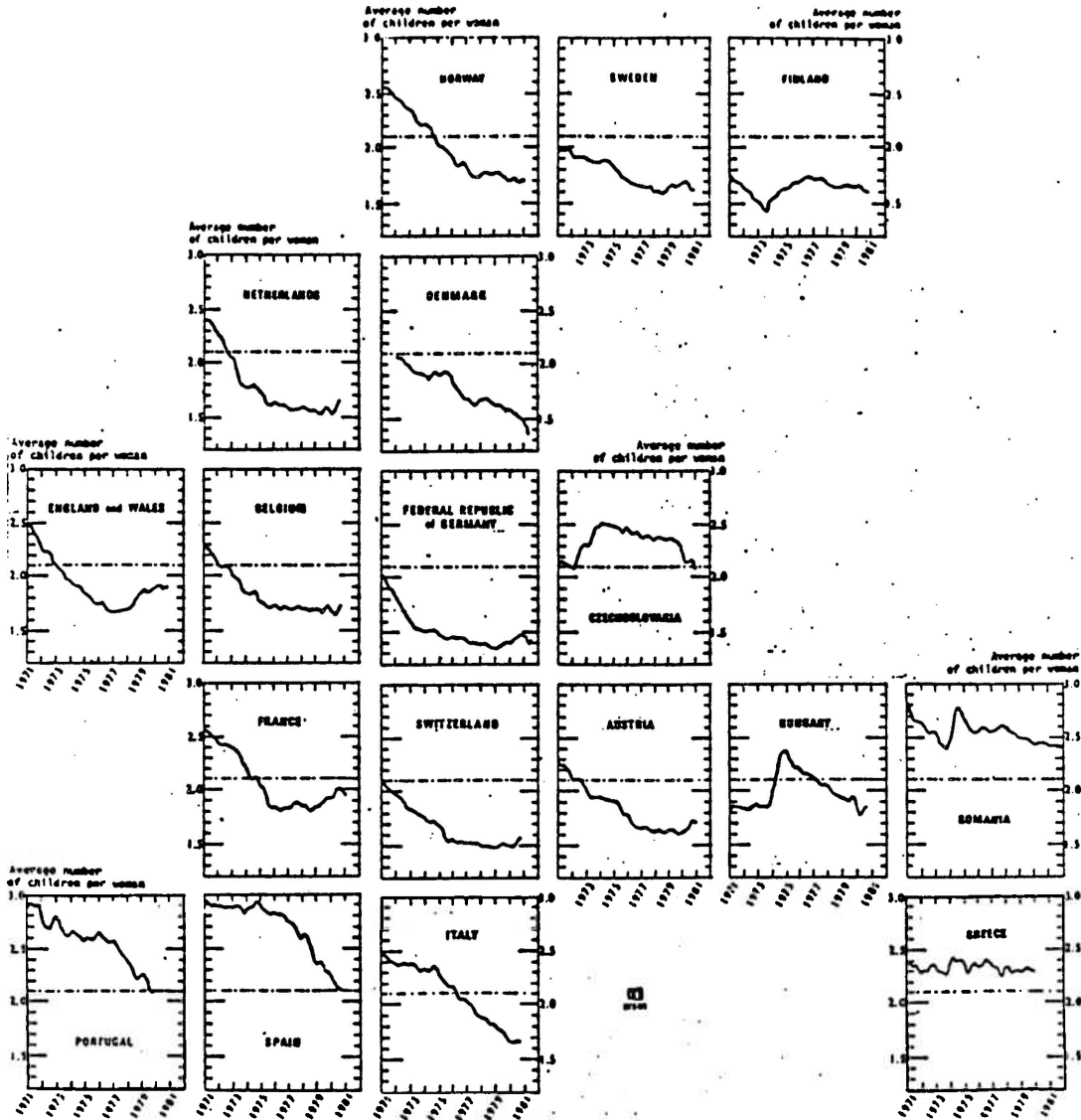


Figure 5.8: Évolution récente de la fécondité féminine dans les pays développés.  
Source : Calot.

## 5.4 L'analyse longitudinale

Nous partons de la constatation que la fécondité transversale est soumise à des variations conjoncturelles, et, sans pouvoir expliquer ces variations, nous sommes à la recherche d'éventuels phénomènes régulateurs. Autrement dit, par exemple, les familles qui subissent des événements extérieurs, comme une guerre ou une crise économique, cherchent-elles à compenser cette faible fécondité subie par une sur-fécondité ultérieure désirée ? L'inverse est possible aussi ; on peut imaginer qu'une situation particulièrement favorable incite les familles à avancer le calendrier de leurs naissances. Il semble par exemple que durant les années 1946-50, les deux phénomènes aient joué ; d'une part la guerre a conduit à un report des naissances, d'autre part la politique familiale, et d'une manière générale le climat de la reconstruction du pays, ont conduit à une sur-fécondité durant ces années-là. Cette sur-fécondité a-t-elle duré suffisamment longtemps et touché suffisamment de générations pour compenser la baisse due à la guerre ? Jusqu'où faut-il remonter dans le temps pour trouver une fécondité égale à celle exceptionnelle des générations concernées par le baby-boom ?

Nous allons pouvoir répondre à ces questions en effectuant une analyse de la fécondité par *génération*. Auparavant, rappelons que nous ne nous intéressons pas aux histoires individuelles des familles, mais à l'histoire collective. Une famille elle-même peut rencontrer des problèmes spécifiques et différer ses naissances, mais nous n'avons pas les moyens d'analyser de telles situations ; aussi nous n'étudions que les phénomènes nationaux. Par exemple, l'impact des législations d'ordre familial est examiné avec soin par le démographe. Parmi celles-ci, le cas de la Roumanie où, dans un premier temps l'avortement a été libéré en 1956, puis brutalement interdit en 1966, est très connu. On peut suivre ainsi sur la figure 5.8 que la fécondité qui baissait régulièrement avant 1966 fait un saut brutal durant cette année là ; mais, au lieu de se maintenir à ce haut niveau comme le souhaitait le législateur, elle a de nouveau décliné pour rejoindre progressivement la tendance observée avant 1966 : en effet, c'était le temps nécessaire pour que les réseaux d'avortements clandestins se réorganisent. La mauvaise expérience roumaine a servi d'argument en France contre ceux qui considéraient qu'une libéralisation de l'avortement aurait un impact important sur la démographie (voir Ch. Blayo [1]).

### 5.4.1 La fécondité d'une génération

Pour les mêmes raisons que pour la fécondité transversale, il nous faut étudier séparément la fécondité des hommes et celle des femmes. Si nous connaissons tous les taux de fécondité par âge d'un pays depuis de nombreuses années, comme c'est le cas en France depuis 1899, il est possible de regrouper les taux de fécondité à chaque âge attachés à une même génération de naissance. Les figures 5.9 et 5.10 illustrent l'évolution de la fécondité par âge en France pour différentes générations successives espacées de 5 ans.

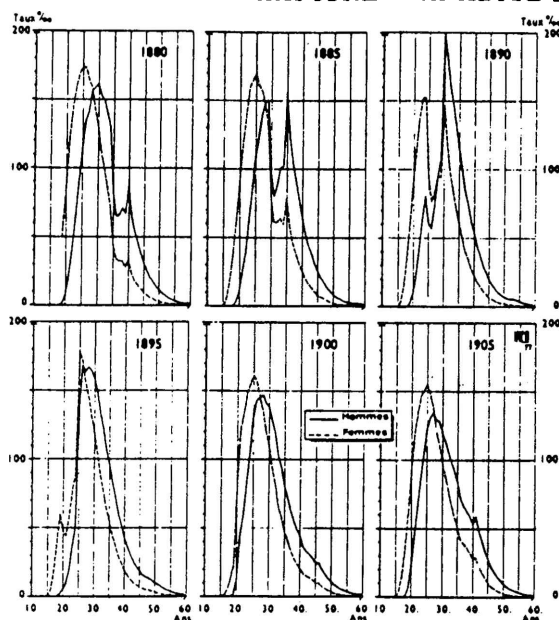


Figure 5.9: Fécondité par âge et par sexe pour quelques générations. France.

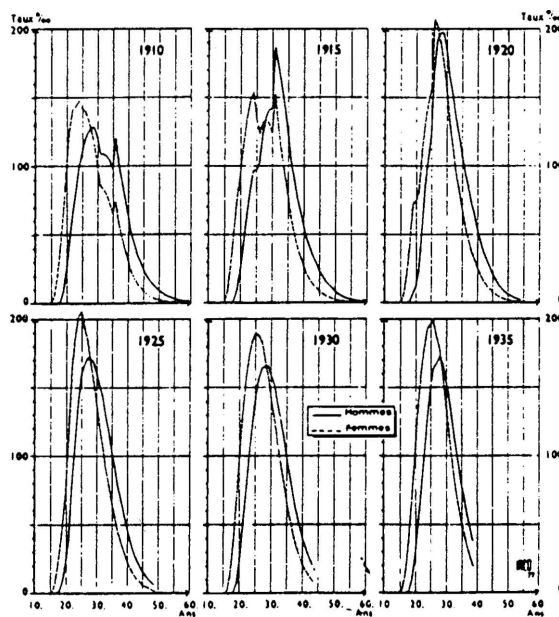


Figure 5.10: Fécondité par âge et par sexe pour quelques générations. France. (suite).

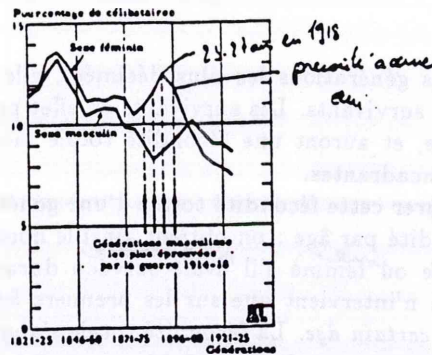


Figure 5.11: Proportion de célibataire pour les générations ayant connu la guerre de 1914-18. Différence entre les hommes et les femmes. France.

Les hommes et les femmes d'une même génération subissent les événements, comme la guerre de 1914 par exemple, à un même âge, mais comme ils ne sont pas mariés ensemble puisqu'à cette époque les hommes sont âgés de 3 années de plus en moyenne que leur épouse, il ne subissent les crises durant les mêmes phases de leur vie féconde : les générations féminines les plus touchées par la guerre sont celles qui auraient dû être dans leur période de fécondité maximale, c'est-à-dire aux environs de 25 ans, au moment de la guerre. On vérifie en effet, que la fécondité de la génération féminine de 1890 est largement amputée. Celle de son (homologue) masculin l'est en partie aussi, mais en partie seulement puisque les hommes ne font leurs enfants que plus tardivement. Ainsi la génération masculine de 1895 est pratiquement épargnée par la guerre, alors qu'au contraire la génération plus jeune, de 1885, est fortement marquée par celle-ci.

Suivant la position d'une crise dans le calendrier normal de la vie féconde, la fécondité totale d'une génération serait plus ou moins affectée, mais c'est sans tenir compte du phénomène de compensation ou de rattrapage, phénomène différencié suivant les sexes. Nous remarquons en effet que pour la génération féminine née en 1890, la fécondité au lendemain de la guerre (année 1920, âge 30 ans) est supérieure à la normale de l'époque (150 p.m. contre 130 p.m.). La volonté de combler le déficit dû à la guerre est présente ; mais, beaucoup de femmes ont perdu leur fiancé ou leur mari, et le mariage ou le remariage n'est pas facile en raison du sous-effectif des hommes survivants (2 millions de morts en France).

A l'opposé, les hommes survivants à la fin de la guerre vont connaître des taux de nuptialité jamais égalés ; la proportion de célibataires parmi ces générations n'aura jamais été aussi faible (voir Henry [21]), comme le montre la figure 5.11 (6 % contre 10 % en moyenne). Ceci explique que la fécondité des hommes de la génération 1890 atteignent 200 p.m. en 1920 ; de même les taux de la génération 1895 sont très importants après cette date. Mais si la guerre affecte cette dernière génération qui a 20 ans en 1920, et qui comp-

Celle des hommes de la 1<sup>re</sup> génération

tera parmi les générations les plus décimées, elle n'affecte pas la vie féconde ultérieure des survivants. Les survivants en effet profiteront de la sur-nuptialité d'après-guerre, et auront une fécondité totale moyenne supérieure à celle des générations encadrantes.

Pour mesurer cette fécondité totale d'une génération, il suffit de sommer les taux de fécondité par âge : on obtient ainsi le nombre total d'enfants qu'aurait eu un homme ou <sup>une</sup> femme s'il avait survécu <sup>à cet âge</sup> durant toute sa vie féconde. Si la sommation n'intervient que sur les premiers âges, on parle de *descendance atteinte à un certain âge*. La fécondité totale s'appelle aussi *descendance finale*.

Nous remarquons sur la figure 5.12 qui retrace, pour chaque génération masculine, la descendance atteinte à divers âges, que les générations 1895-96 atteignent bien un maximum relatif avec 2,26 enfants par homme (survivant). Il s'agit bien d'un maximum relatif car la fécondité baissait depuis la fin du 18<sup>e</sup> siècle, et la guerre de 1914-18 induit un phénomène parasite dû à la surnuptialité des hommes. On remarque aussi que l'avance prise par ces générations avant 35 ans par exemple (ligne 35 ans de la figure 5.12), comparée aux générations encadrantes se résorbe progressivement, pour <sup>substituer</sup> néanmoins subsister en terme de descendance finale.

Nous voyons sur ce même graphique, que la seconde guerre mondiale induit un phénomène analogue mais d'ampleur beaucoup plus faible. Le baby-boom des années 1946 et suivantes, correspond à un véritable changement de tendance et la descendance finale des générations croît pour atteindre un maximum de 2,65 enfants par homme pour la génération née en 1921.

Pour les femmes, nous voyons sur la figure homologuée 5.13 que la génération 1895, celle qui est touchée aux âges de forte fécondité par la première guerre, tente de combler son retard après celle-ci, mais n'y parvient qu'en partie, et sa descendance finale ne sera que d'à peine 2 enfants. Pour les générations ultérieures, la fécondité a bien baissé entre les deux guerres, mais le baby-boom de 1946 concerne tous les âges féconds. Ainsi, plus une génération sera jeune, plus longtemps elle sera concernée par ces années de forte fécondité et plus sa descendance finale sera élevée ; le maximum est atteint par la génération 1930, qui est âge d'à peine 16 ans en 1946, avec une descendance finale de 2,64 enfants.

La baisse conjoncturelle de la fécondité entre 1968 et 1975, se répercute sur les générations plus récentes, et la marge (au dessus de 2 enfants) provenant du baby-boom est progressivement grignotée. Pour les générations masculines, nous remarquons que le phénomène est identique, quoique l'évolution ait été plus rapide.

En somme, pour calculer la *descendance finale d'une génération*, nous avons additionné les taux de fécondité par âge d'une même génération. Mais nous pourrions obtenir le <sup>en</sup> résultat autrement <sup>en</sup> et exploiter les résultats d'une enquête (les enquêtes dites enquêtes-famille en France par exemple) qui demanderait



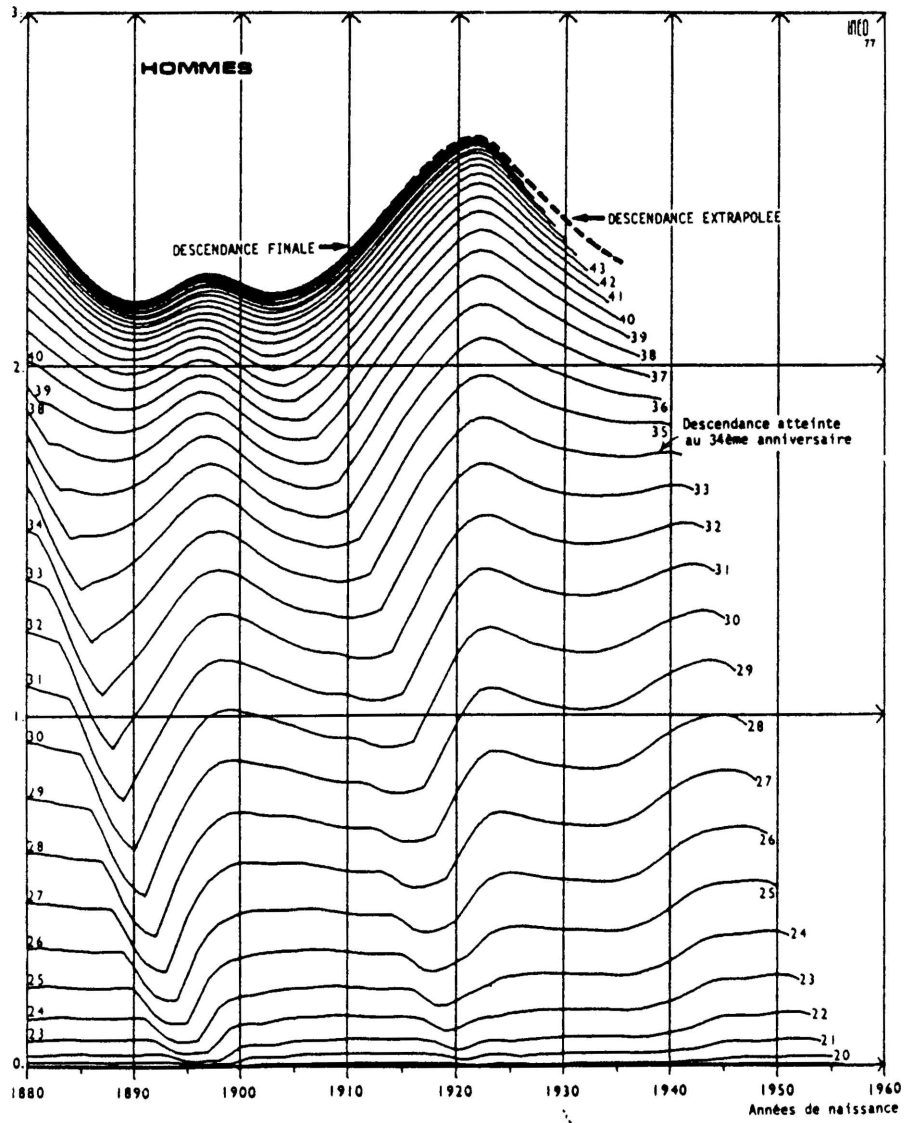


Figure 5.12: Descendance atteinte à divers âges et descendance finale par génération. Hommes. France.

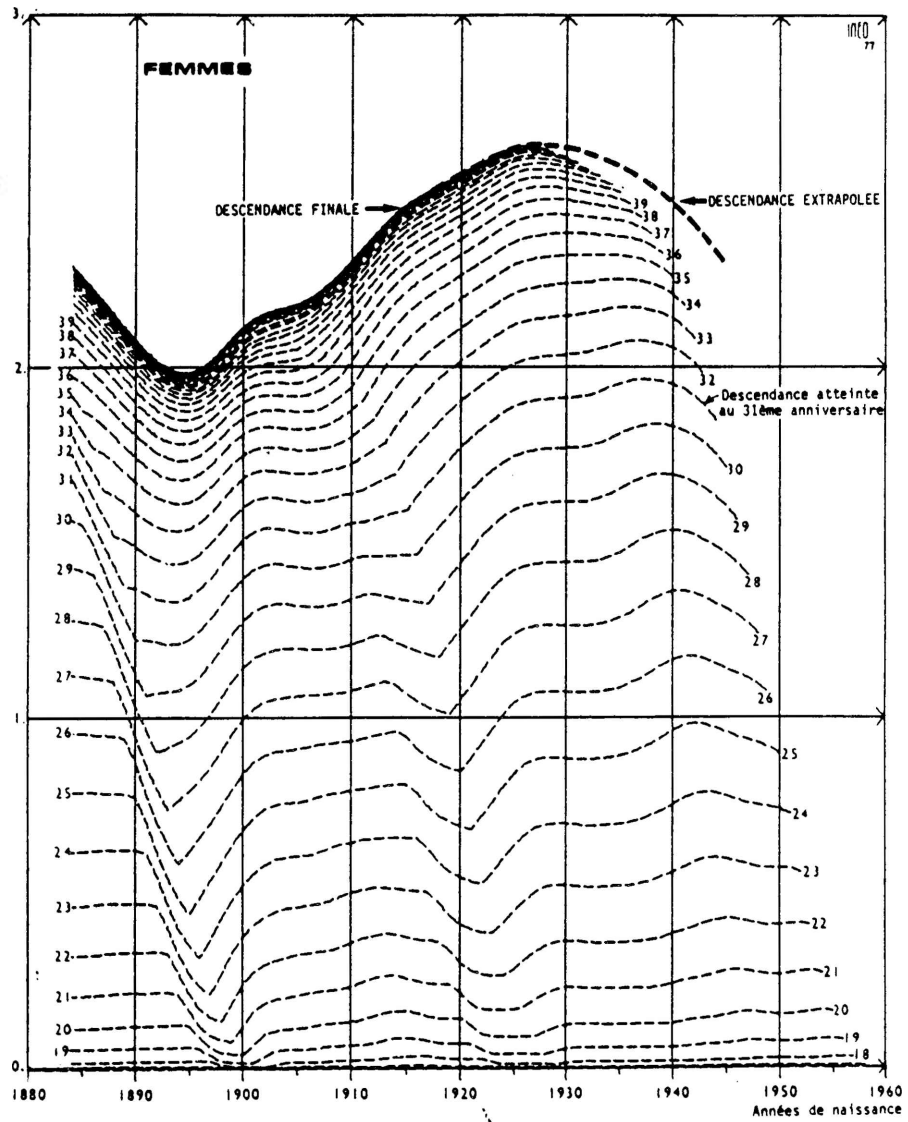


Figure 5.13: Descendance atteinte à divers âges et descendance finale par génération. Femmes. France.

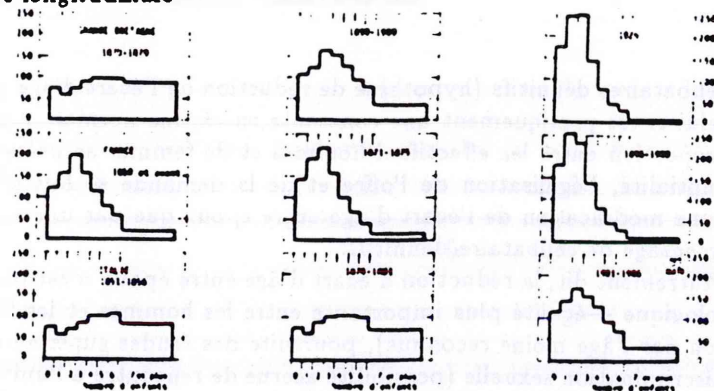


Figure 5.14: Distribution des familles complètes suivant leur taille pour différentes générations de mères. *Source Fealy.*

aux femmes de plus 50 ans et aux hommes de plus de 60 (fin de la vie féconde) le nombre total des enfants qu'ils ont eu, *ou qu'ils ont* que ceux-ci soient aujourd'hui vivants ou déjà décédés. Les moyennes par sexe, donneraient théoriquement un résultat identique aux descendance finale des générations concernées. Cette correspondance est généralement bien vérifiée en pratique, ce qui indique que le phénomène de sélection sous-jacent est *d'ampleur faible*.<sup>12</sup> L'intérêt de ce second type de mesure provient du fait qu'on obtient aussi la distribution des familles dites *complètes* suivant leur taille, et non pas seulement la taille moyenne. La figure 5.14 montre ainsi l'évolution de cette distribution.

La fécondité que nous venons d'étudier dans ce chapitre n'est qu'une composante du processus de renouvellement d'une population, l'autre étant bien entendu la mortalité. L'interaction des deux phénomènes n'est pas aisée à mesurer et sera l'objet du chapitre 6. *mais auparavant, mentionnons que la fécondité d'un pays est fortement liée à la nuptialité, en examinant le cas français.*

*par exemple dans*  
**Les rigidités de la nuptialité** Nous avons vu en effet que la fécondité des femmes a été restreinte après la première guerre en raison d'un manque d'homme (notons que dans un pays polygame le manque d'homme n'aurait pas été ressenti avec la même acuité, et la fécondité féminine aurait été sans doute moins perturbée).

Les femmes qui avaient coutume avant la première guerre de se marier avec des hommes, leurs aînés de 3 ans environ, n'ont pu, après la guerre, trouver de partenaires que parmi les générations plus jeunes, et d'un coup l'écart d'âge entre époux est descendu aux environs de 2 ans. Mentionnons encore que cet écart d'âge ne peut évoluer que très difficilement en temps normal, à moins que quelques générations d'hommes concèdent un pourcentage plus important

<sup>12</sup> En effet, si on supposait par exemple que les femmes qui ont beaucoup d'enfants, décèdent plus facilement, l'échantillon des femmes qui survivraient à 50 ans serait biaisé et la descendance finale sous-estimée.

de célibataires définitifs (hypothèse de réduction de l'écart d'âge entre époux) ; or celui-ci est pratiquement une constante en régime normal. Ainsi lorsqu'une disproportion entre les effectifs d'hommes et de femmes se présente à l'âge de la nuptialité, l'égalisation de l'offre et de la demande se fait plus facilement par une modification de l'écart d'âge entre époux que par une modification du pourcentage de célibataires définitifs.

Autrement dit, la réduction d'écart d'âge entre époux n'est pas tant d'ordre sociologique —égalité plus importante entre les hommes et les femmes (domination par l'âge moins reconnue), poursuite des études supérieures avec moins de discrimination sexuelle (possibilité accrue de rencontre à l'université), etc — que d'ordre purement numérique.

Nous venons de mettre l'accent sur la rigidité imposée par un régime monogame, notons que la fréquence beaucoup plus importante des divorces et donc des remariages modifie le marché matrimonial et lui supprime sans doute une partie de sa rigidité, mais l'étude du mariage et du divorce serait l'objet d'un autre cours.

## Chapitre 6

# Les différentes mesures de la reproduction

Dans le chapitre 5 sur la fécondité, nous avons indiqué que la reproduction plus ou moins importante d'une population dépend d'une manière complexe du niveau de sa mortalité et de sa fécondité.

L'objet du présent chapitre est de *quantifier* la croissance d'une population à partir des indicateurs de mortalité et de fécondité dégagés dans les chapitres précédents. Comme ces derniers sont de deux types, transversaux et longitudinaux, la reproduction humaine peut elle aussi être étudiée selon chacune des approches. Avant d'en venir à ces indicateurs, rappelons en s'appuyant sur la théorie mathématique de la transition, que le démographe ne fait le plus souvent référence, pour la théorie, qu'à des populations qui se trouvent dans trois états particuliers :

1. le premier est celui d'avant la transition démographique où la mortalité et la fécondité sont importantes ; la croissance est en moyenne nulle.
2. le second est celui d'après la transition démographique, où la mortalité est faible, la fécondité est supposée suffisante pour que la croissance soit nulle.
3. le troisième se rattache à une phase intermédiaire de la transition ; la mortalité et la fécondité ont baissé, mais la mortalité a chuté plutôt que la fécondité, la population est donc croissante. Parmi ces multiples états "transitoires", le démographe a l'habitude de *figer* un état particulier, car dans ce cas, les mathématiques de l'évolution de la population sont simples. Le démographe raisonne alors comme si la mortalité et la fécondité, mesurées à une date précise, restaient *immuables*.

Les cas 1 et 2 correspondent donc à une population stationnaire, le cas 3 à une population dite stable (voir le paragraphe 9.2 du chapitre 9).

Les cas réels n'appartiennent à aucun de ces cas types, mais leur étude est nécessaire pour mieux les appréhender. Aussi, dans un premier temps, nous étudierons la reproduction dans les cas où les conditions de mortalité et de fécondité sont fixes (populations stables ou stationnaires), puis nous verrons comment corriger ces premiers indicateurs lorsque la fécondité et la mortalité évoluent. Nous distinguerons deux périodes, la première correspondra à une baisse de la mortalité à tous les âges et particulièrement de la mortalité infantile et juvénile, la seconde, à une étape ultérieure de la transition où la mortalité infantile est pratiquement inexistante, et seule une baisse de la mortalité aux âges adultes, retarde le déclin de la population mais en la vieillissant.

## 6.1 Cas où la mortalité et la fécondité sont constantes dans le temps

Nous avons vu au paragraphe 5.2, qu'on pouvait mesurer l'intensité de la fécondité d'une population en sommant les taux de fécondité par âge. La somme (appelée aussi somme des naissances réduites) correspond au nombre d'enfants qu'aurait en moyenne un homme ou une femme (suivant le choix du calcul) s'il ou elle survivait durant toute sa vie féconde. Ce chiffre avoisine ~~autour~~ <sup>2</sup> ~~de deux~~ dans les pays développés, et peut aller jusqu'à 7, 8 enfants pour les femmes dans un pays du Tiers-Monde ~~et~~ près du double pour les hommes d'un pays où la polygamie peut exister, comme dans le cas de la population des Peul Bandé, pris déjà comme exemple. ~~Mais~~ dans ce dernier exemple, peu d'hommes auront 12 enfants, car peu survivront jusqu'à l'âge de 70-80 ans. La *fécondité nette*, celle qui prend en compte la mortalité, peut se calculer en sommant les taux de fécondité par âge, eux-mêmes multipliés par la probabilité de survie de la naissance jusqu'à l'âge correspondant. Ainsi, nous voyons en suivant le tableau 6.1 que sur 1000 garçons nouveaux-nés, 474 atteindront l'âge de 15 ans et 465, celui de 20 ans ; leur fécondité moyenne durant la période est de 30 p.m. annuel. Le nombre des survivants concernés est approximativement égal à  $\frac{474+465}{2} = 469,5$  hommes, soumis pendant 5 années à ce taux de fécondité : ce qui donne 7 enfants. Entre 20 et 25 ans, les survivants, au nombre de  $\frac{446+426}{2}$  donneront naissance à  $\frac{446+426}{2} \times \frac{64}{1000} \times 5 = 139$  enfants, etc. Au total, une génération de 1 000 hommes donnera naissance à 3,510 enfants.

Le calcul pour les femmes est analogue. Plus généralement, on calcule des taux de fécondité *nets* en multipliant la fécondité par âge, par la probabilité de survie au même âge médian<sup>1</sup>. On peut calculer des descendance *nettes* atteintes à divers âges, et une descendance nette finale en sommant les taux nets par âge

<sup>1</sup> Le plus souvent, les probabilités de survie sont présentées à des âges exacts, 0, 1, 5, 10, 15 etc. alors que les taux de fécondité sont relatifs à des intervalles d'âges : 10-14, 15-19, etc., c'est-à-dire qu'ils donnent une valeur du risque instantané à l'âge médian. Il faut donc calculer la survie aux mêmes âges médians, en faisant la demi-somme des probabilités de survie aux âges exacts encadrants.

Tableau 6.1: Calcul de la reproduction à partir de la table de mortalité et des taux de fécondité. Population des Peuls Bandé (Sénégal Oriental).

Groupe d'âge	Âge médian	Survivants en début période	Survivants à l'âge médian	Fécondité brute à l'âge médian en p.m.	Fécondité nette à l'âge médian en p.m.	Descendance brute atteinte en début de période	Descendance nette atteinte en début de période
<b>FEMMES</b>							
0-1	0,5	1000					
1-4	3	813					
5-9	7,5	605					
10-14	10,5	547	576	24	13	0	0
15-19	13,5	513	520	270	135	0,120	0,064
20-24	22,5	487	500	298	142	1,470	0,739
25-29	27,5	467	477	250	114	2,940	1,449
30-34	32,5	448	457,5	220	96	4,210	2,021
35-39	37,5	428	438	163	68	5,310	2,503
40-44	42,5	404	416	86	33	6,125	2,842
45-49	47,5	375	389,5	22	8	6,555	3,009
50-54	52,5	343	359	12	4	6,665	3,049
55-59		310				<b>6,665</b>	<b>3,049</b>
60-64		279					
65-69		234					
70-74		170					
75-79		104					
80 et +							
<b>HOMMES</b>							
0-1	0,5	1000					
1-4	3	769					
5-9	7,5	509					
10-14	10,5	474					
15-19	13,5	465	469,5	30	1	0	0
20-24	22,5	444	454,5	64	28	0,015	0,007
25-29	27,5	426	435	245	101	0,335	0,146
30-34	32,5	401	413,5	326	126	1,560	0,653
35-39	37,5	373	387	396	140	3,190	1,284
40-44	42,5	334	353,5	331	104	5,170	1,983
45-49	47,5	294	314,5	297	81	6,825	2,503
50-54	52,5	250	272	231	53	8,910	2,907
55-59		212	231	128	24	9,465	3,174
60-64		187	199,5	98	17	10,095	3,291
65-69		164	175,5	90	13	10,545	3,377
70-74		130	147	79	9	10,995	3,443
75-79		93	113,5	64	4	11,390	3,488
80 et +		47	70	??	0	11,710	3,510
					<b>&gt; 11,710</b>	<b>3,510</b>	

et en les multipliant par la durée d'exposition (5 ans ici). Le résultat pour les femmes est alors de 3 069 enfants par femme.

L'écart entre l'indicateur féminin et l'indicateur masculin s'est très nettement réduit lorsqu'on fait intervenir la mortalité !

Reste néanmoins une apparente contradiction puisque 1 000 hommes donnent naissance à plus d'enfants (3 510) que 1 000 femmes (3 069), et ce, d'autant plus, qu'une naissance donne en moyenne plus de garçons que de filles (105 garçons pour 100 filles). Donc sur 1 000 enfants, les  $1000 \times \frac{105}{205}$  garçons donneront naissance à  $3510 \times \frac{105}{205} = 1,790$  enfants, et les  $1000 \times \frac{100}{205}$  filles à  $3,069 \times \frac{100}{205} = 1,490$  enfants.

Les deux rapports, 1,79 h/h et 1,49 f/f indiquent que la population est croissante, mais ne spécifient pas le taux de croissance. Pour le calculer, il convient de calculer séparément la durée moyenne d'une génération masculine et d'une génération féminine. La durée d'une génération est donnée par l'âge moyen de cette même distribution des taux de fécondité nets par âge, soient 40,1 ans pour les hommes et 27,0 pour les femmes. Le calcul est alors celui d'un taux de croissance ; dans le cas masculin la population est multipliée par un facteur 1,79 en 40,1 ans :

$$P_t = P_0 e^{\rho t} \Rightarrow P_{40,1} = P_0 e^{\rho 40,1} \text{ d'où}$$

$$\begin{aligned} \rho_m &= \log \frac{P_{40,1}}{P_0} \times \frac{1}{40,1} \\ &= \frac{\log 3,510}{40,1} = 1,45 \text{ \% par an,} \end{aligned}$$

et dans le cas féminin :

$$\rho_f = \frac{\log 1,49}{27,0} = 1,40 \text{ \% par an.}$$

En conclusion, nous pouvons dire que les taux de croissance sont pratiquement identiques : la population féminine se reproduit beaucoup plus vite que la population masculine, mais à un facteur moindre que la population masculine.

Après avoir exposé le raisonnement, faisons maintenant un peu de terminologie :

- On appelle rapport de masculinité (sex ratio en anglais), et note  $m$ , le nombre de garçons pour une naissance. Celui-ci vaut  $m = \frac{105}{205} \approx 0,512$ .  $1 - m = 0,488$  est alors la proportion de naissances féminines.
- On appelle *taux brut de reproduction*, par exemple féminin, et note  $R$ , le nombre de filles auquel donnerait naissance une femme en l'absence de



mortalité. Il s'agit aussi du produit de la somme des naissances réduites par le rapport de masculinité :

$$R = (1 - m) \sum_i f_i \quad \text{où } f_i \text{ est le taux de fécondité à l'âge } i. \quad (6.108)$$

- On appelle *taux de reproduction net*, et note  $R_0$ , le nombre de filles auquel donnerait naissance une femme en tenant compte de la mortalité :

$$R_0 = (1 - m) \sum_i f_i S_i \quad \text{où } S_i \text{ est la probabilité de survie jusqu'à l'âge } i. \quad (6.109)$$

- On appelle *taux de croissance de Lotka*, et note  $\rho$ , le taux de croissance d'une population régie par une fécondité et une mortalité invariables dans le temps. Il s'exprime en fonction des moments de la distribution des taux de fécondité nets par âge. Au premier ordre, on peut écrire :

$$\rho \simeq \frac{\log R_0}{\bar{a}}, \quad (6.110)$$

en notant  $\bar{a}$  l'âge moyen de la distribution des taux de fécondité nets :

$$\bar{a} = \frac{\sum_i i f_i S_i}{\sum_i f_i S_i}.$$

**Remarque** Cette formule approchée sera démontrée au paragraphe 9.2 lors de l'étude des populations stables. La relation est exacte si on suppose que la distribution de la fécondité nette est une Dirac au point  $\bar{a}$ . Dans un cas réel, où la distribution a une variance  $\sigma^2$  non nulle, on obtient une formule approchée au second ordre :

$$\rho_{(2)} \simeq \frac{\log R_0}{\bar{a}} + \frac{\sigma^2 (\log R_0)^2}{2\bar{a}^3}. \quad (6.111)$$

Dans le cas de la population des Peul Bandé, l'écart-type de la fécondité masculine (13,39) est beaucoup plus importante que celle de la fécondité féminine (8,5), et la correction est en fin de compte légèrement plus importante pour les hommes que pour les femmes ce qui réduit l'écart entre les deux mesures du taux de croissance : 1,485 % pour les hommes contre 1,506 pour les femmes.

Le tableau 6.2 donne les résultats des différentes étapes décrites dans ce paragraphe pour cette population du Sénégal Oriental. L'hypothèse implicite (des calculs) qui consiste à supposer que les taux de fécondité et de mortalité n'ont pas ou peu varié est plausible ; la concordance quasi parfaite entre les deux taux de croissance masculin et féminin la corrobore.

Tableau 6.2: Résumé synthétique des étapes pour le calcul du taux de croissance de la population des Peul Bandé.

	Hommes	Femmes
Somme des naissances réduites	> 11,7 enfants	6,725 enfants
Âge moyen (brut)	> 44,25 ans	27,6 ans
Somme des naissances réduite (nette)	3,510 enfants	3,049 enfants
Âge moyen (net)	40,1 ans	27 ans
Taux de reproduction net $R_0$	1,79 h/homme	1,49 f/fille
Taux de croissance de Lotka (prem.ordre)	1,459 %	1,479 %
Écart-type $\sigma$ (net)	11,2 ans	8,28 ans
Taux de croissance de Lotka (sec.ordre)	1,485 %	1,506 %
La précision des résultats est illusoire mais permet de mieux comprendre l'évolution des diverses corrections.		

**Exercice 13** Montrer que si on suppose que la fonction de survie est linéaire durant la période de reproduction, le taux de reproduction net,  $R_0$ , est égal au produit de taux brut de reproduction,  $R$ , multiplié par la probabilité de survie jusqu'à un âge  $\bar{a}_n$ , où  $\bar{a}_n$  est l'âge moyen de la distribution de la fécondité brute.

**Solution 13** Par définition :

$$R_0 = (1 - m) \sum f_i S_i = (1 - m) \sum f_i \left( \frac{\sum f_i S_i}{\sum f_i} \right),$$

si la fonction de survie  $S$  est linéaire, nous pouvons écrire :

$$S_i = S_a + \lambda(i - a),$$

soit aussi :

$$\sum f_i S_i = \sum f_i S_a + \lambda \sum (i - a) f_i.$$

Le terme le plus à droite de l'expression ci-dessus est nul pour la valeur de  $a$  qui vérifie :

$$\bar{a} = \frac{\sum i f_i}{\sum f_i},$$

c'est-à-dire pour l'âge moyen de la distribution de la fécondité brute.

Nous venons de voir que la connaissance du seul taux net de reproduction n'est pas suffisante pour estimer la croissance d'une population et qu'il faut

aussi connaître l'âge moyen à la reproduction. Cette remarque est relativement importante et montre qu'un des moyens pour diminuer la croissance d'une population sans contraindre celle-ci à avoir moins d'enfants, est de reculer l'âge légal au mariage. Cette dernière mesure est aussi une contrainte pour une population mais peut-être plus facilement acceptable. Notons aussi que dans l'exemple de la Chine, un âge légal élevé n'a pas été suffisant et qu'il a été nécessaire d'imposer les mesures draconiennes que l'on connaît (voir Banister [?] par exemple).

La formule 6.110 indique aussi que lorsque la croissance est nulle ( $R_0 = 1$ ), l'âge moyen à la reproduction n'a plus aucune influence. Pour des taux de reproduction nets inférieurs à 1 qui correspondent aux cas de nombreux pays européens, il serait possible d'affaiblir le taux de *décroissance*<sup>2</sup> en augmentant l'âge moyen à la reproduction. Cet âge moyen a au contraire diminué de lui-même car c'est surtout la fécondité des rangs de naissance élevé qui a chuté. Par contre, la tendance qui semble se dessiner dans quelques pays européens et qui consiste à retarder les premières naissances dans la vie d'un couple, va dans l'autre sens.

Pour un pays à faible mortalité, comme ceux d'Europe, on calcule qu'en moyenne, un taux net de reproduction égal à l'unité correspond à une somme des naissances réduites de 2,1 enfants par femme et de 2,05 enfants par hommes<sup>3</sup>.

L'exemple pris au chapitre 5 de la fécondité française en 1974 correspond exactement à ce cas : ainsi, si les taux de fécondité et de mortalité étaient restés identiques à leur valeur de l'année 1974, la population française serait et demeurerait stationnaire. Mais nous avons vu qu'il n'en a pas été du tout ainsi, puisque fécondité et mortalité ont fortement baissé depuis deux siècles et ont subi de brusques variations en particulier à cause des deux guerres mondiales. Le calcul relatif à l'année 1974 est un calcul transversal qui tient compte des comportements sans doute différents de plusieurs générations. Pour mesurer la contribution des générations au renouvellement d'une population, il faut faire le calcul par génération que nous allons exposer maintenant.

## 6.2 Analyse par génération

La première opération consiste à regrouper les taux de fécondité et mortalité déjà calculés par âge et par année, selon des générations de naissance, et à calculer trois indices *longitudinaux* : l'espérance de vie (ou durée de vie moyenne de la génération), le taux brut de reproduction, et le taux net de reproduction.

L'exemple que nous prenons est tiré d'un article de Louis Henry [22], mais cet exemple est général et utile à l'analyse de la transition démographique d'un

<sup>2</sup>Rappelons qu'il s'agit des conditions de l'état stable.

<sup>3</sup>Le calcul se base sur des profils de tables types.

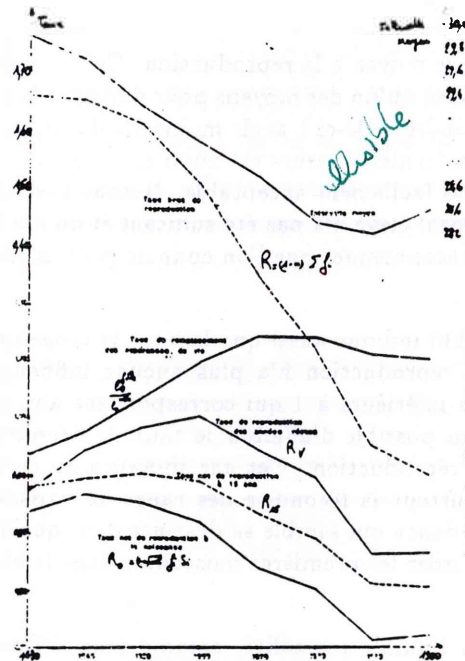


Figure 6.1: Évolution de différents indices de reproduction par génération. France. Femmes.

pays quelconque. Les données portent sur des générations féminines allant de 1830 à 1900 ; elles devraient être complétées et comparées à leur homologue masculin.

La figure 6.1 donne ainsi l'évolution du taux brut de reproduction par génération. Celui-ci vaut 1,65 fille par femme (3,3 enfants par femme) pour la génération née en 1830, et baisse progressivement pour atteindre 1,0 pour la génération née en 1900. Traduite en terme de *descendance finale* (enfants par femme au lieu de filles par femme) cette courbe se raccorde à celle de la figure 5.13 du chapitre 5. Elle confirme que la fécondité a baissé durant le XIX<sup>e</sup> siècle.

L'intervalle moyen entre génération, ou âge moyen à la naissance des enfants, a lui aussi baissé, indiquant une chute de la fécondité d'enfants de rangs élevés.

Le taux net de reproduction,  $R_0$ , a une évolution moins rapide que le taux brut ; on note même une augmentation du taux pour les générations 1830-1840, car la mortalité baissait fortement durant la période concernée, alors que la fécondité n'avait qu'une décroissance faible.

Pour évaluer la rapidité de la baisse de mortalité, L. Henry a tracé le *taux de remplacement des espérances de vie* : il s'agit du rapport des espérances de vie de la génération concernée avec celle moyenne de ses enfants, c'est-à-dire 30 ans plus tard environ. On note ainsi que les progrès ont été grandissants jusqu'aux générations 1870, ont stagné durant 10 générations au niveau de 1,2,

des progrès jusqu'aux 1870 et une stagnation pendant 10 s.

et sont redescendus progressivement depuis. En effet, plus l'espérance de vie est importante, plus il est difficile de faire des progrès.

Si on s'en tient à l'indicateur  $R_0$ , aucune génération n'a assuré le renouvellement des générations depuis 1830, puisque celui-ci est inférieur à 1. Mais cet indicateur sous-entend que les conditions de mortalité en particulier restent figées à l'avenir. Pour tenir compte de la baisse de la mortalité entre les générations des mères et celles des filles, on peut faire commencer le cycle non pas à la naissance mais à 15 ans.

### 6.2.1 Baisse de la mortalité infantile et juvénile : la reproduction à 15 ans

On calcule ainsi, pour une génération de mères survivantes à 15 ans, le nombre de leurs filles qui survivront elles-mêmes jusqu'à 15 ans. Il ne s'agit pas de la reproduction de toute la génération, mais de celles qui auront survécu jusqu'à 15 ans. Si  $S_i^m$  est la fonction de survie à l'âge  $i$  des mères et  $f_i^m$  leur fécondité, alors :

$$(1 - m) \sum f_i^m \frac{S_i^m}{S_{15}^m}$$

est le nombre de filles auquel cette génération donnera naissance.

Si on suppose que ces filles appartiennent à un même groupe de générations, dont la mortalité moyenne est  $S_i^f$ , seule une proportion  $S_{15}^f$  survivra jusqu'à 15 ans et pourra recommencer le cycle. On appelle alors *taux de reproduction à 15 ans* et note,  $R_{15}$ , la quantité :

$$R_{15} = (1 - m) \sum f_i^m \frac{S_i^m}{S_{15}^m} S_{15}^f. \quad (6.112)$$

Ce taux est aussi égal au produit du taux net de reproduction par le rapport de la probabilité de survie à 15 ans des filles à celle des mères, qui traduit la baisse de la mortalité infantile et juvénile en 15 ans (voir la figure 6.2) :

$$R_{15} = R_0 \frac{S_{15}^{\text{filles}}}{S_{15}^{\text{mères}}}. \quad (6.113)$$

Selon ce nouvel indicateur, nous pouvons lire sur la figure 6.1 que seule la génération 1850 a dépassé le rapport 1. En effet, pour les générations ultérieures, même si le rapport des survivants à 15 ans n'est pas négligeable, il ne compense pas la baisse de la fécondité. Lorsque la mortalité infantile et juvénile est négligeable (moins de 2 % de pertes), le rapport des probabilités de survie à 15 ans est proche de 1, et  $R_{15}$  n'apporte aucune indication supplémentaire par rapport à  $R_0$ . Les deux indicateurs peuvent ainsi être inférieurs à 1 et pourtant la population peut croître encore ; il convient alors de définir un autre indicateur.

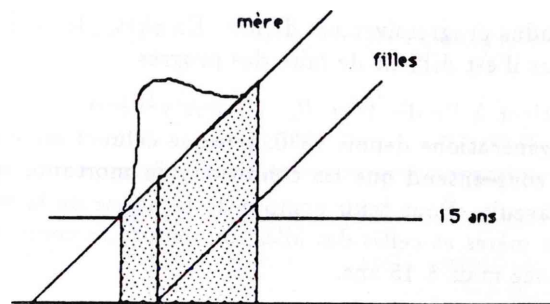


Figure 6.2: Génération des mères et des filles : reproduction à 15 ans.

### 6.2.2 Baisse de la mortalité aux âges élevés : reproduction des années vécues

En effet, si la mortalité ne peut plus faire de progrès sensible avant 20-25 ans pour contrecarrer une baisse de la fécondité, elle peut continuer à baisser aux âges élevés. Les générations donneront naissance à un nombre inférieur d'enfant, mais si ceux-ci vivent plus longtemps, la population peut continuer à croître. A la limite, si la mortalité devient nulle, la population vieillira, mais croîtra.

Bien entendu, si la mortalité cesse de baisser, l'indicateur  $R_0$  mesure de nouveau la croissance ou le déclin de la population.

Prenons un diagramme de Lexis, et traçons un domaine dont la hauteur est de 0 à 120 ans<sup>4</sup>, et la longueur de plusieurs centaines de générations.

Traçons des lignes correspondant aux lignes de vie des générations. Plaçons sur chaque ligne des masses qui correspondent à l'effectif de la population de cette génération à cet âge là. La somme des masses propres à chacune des générations est supposée constante et donc égale à la somme des années vécues par cette génération. En l'absence de migration, il s'agit du produit des naissances par l'espérance de vie de cette génération. Les masses peuvent être placées de façon quelconque sur une ligne à condition de vérifier que la masse totale soit constante.

Une condition démographique ancienne correspond à un nombre de naissance important, et, peu d'étalement (1/4 des enfants seulement survivent à 15 ans par exemple) ; une condition plus moderne correspond à moins de naissance, mais à une masse répartie sans grande perte jusqu'à un âge élevé, 50-60 ans environ (voir la figure 6.3).

Si la bande du diagramme de Lexis est un carton sur lequel reposent des masses; on admettra aisément que le fait de déplacer les masses le long d'une génération ne change rien au poids d'ensemble du carton, ni même au poids

<sup>4</sup>Soyons larges, l'objectif des 120 ans n'est peut-être pas si loin.

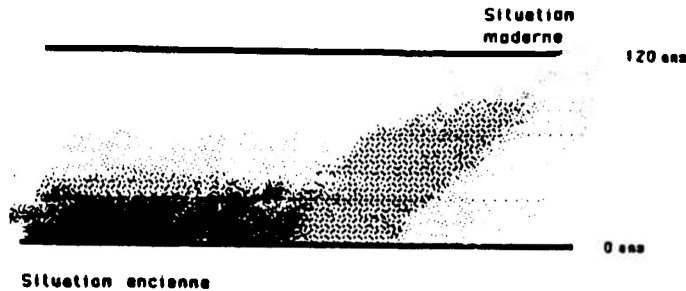


Figure 6.3: Diagramme de Lexis, présentation du théorème sur les années vécues.

d'une portion de carton mobile d'une longueur suffisante et fixe. La somme des poids situés sur cette portion de carton divisée par le nombre d'années que représente la longueur de cette portion donne la population *moyenne* au cours de cette période. Ce raisonnement nous conduit au théorème suivant :

**Théorème 6.1** *Pour que la taille d'une population soit fixe en moyenne, il faut et il suffit que la somme des années vécues par chaque génération soit identique.*

Louis Henry a donc défini un taux de reproduction des années vécues comme le rapport notée,  $R_v$  du total des années vécues par les filles au total des années vécues par les mères :

$$R_v = \frac{N_m e_{0m}}{N_f e_{0f}}, \quad (6.114)$$

où  $N_m$  et  $e_{0m}$  représentent le nombre des naissances et l'espérance de vie de la génération des mères (idem pour les filles). Ce rapport est aussi égal au produit du taux de reproduction brut par le rapport des espérances de vie<sup>5</sup> :

$$R_v = R \frac{e_0^{filles}}{e_0^{mères}}. \quad (6.115)$$

Cet indicateur, qui est reporté sur la figure 6.1, est plus optimiste que les autres indicateurs sur la croissance de la population française, et seules les générations 1880 et suivantes n'ont pas reproduit les années vécues.

Le problème actuel des pays européens est en fait inverse ; la masse initiale d'une génération est donnée par le nombre des naissances, celle-ci ne décroît que très faiblement au moins jusque vers 50 ans. Si la mortalité continue de baisser, cette masse, ou du moins une partie importante se prolongera encore, et la somme totale des masses sera peut-être supérieure à celle des générations antérieures. Une population peut ainsi être croissante avec un nombre annuel faible de naissances.

Évidemment, l'âge moyen de la population ne peut alors que s'élever !

<sup>5</sup>Dans le cas où les migrations sont supposées négligeables.

α **Conclusion** Dans la première phase de la transition démographique, le rapport  $R_{15}$  est utile pour suivre l'évolution conjointe de la fécondité et de la baisse de la mortalité infantile et juvénile, mais lorsque la mortalité infanto-juvénile est faible, l'évolution conjointe est mieux mesurée par le taux de remplacement des années vécues. Dans un pays européen, la population continue de croître à cause du vieillissement<sup>6</sup>, mais si la mortalité s'arrête de baisser, la population est condamnée à décroître. L'exemple de la France exposée dans ce chapitre, nous a montré qu'en tout état de cause, le problème menaçait depuis près d'un siècle.

---

<sup>6</sup>et aussi, souvent de l'immigration.



## Chapitre 7

# Modélisation de la mortalité par âge : les nouvelles tables des Nations-Unies

Le nouveau modèle des Nations-Unies (voir ONU [42]) concerne principalement les pays en voie de développement. Il fait suite au premier modèle conçu par Valaoras à partir de 148 tables de mortalité de ?? pays différents et publié en 1955 (voir ONU [42]). Depuis cette publication, la théorie et les données statistiques se sont améliorées ; Lederman en 1959 [28] puis en 1969 [27], Coale et Demeny en 1966 [9] ont publié des tables qui s'adressent principalement aux pays développés, mais leurs données de base s'arrêtent dans les années 1955-60, et de nouvelles tables sont alors devenues nécessaires. W. Brass (voir [3], [4] et aussi [7]) a mis au point une méthode générale qui peut être adaptée à certains pays en voie de développement et en particulier à certains pays d'Afrique.

Néanmoins la carence pour les pays en voie de développement est grande et les nouvelles tables des Nations-Unies ne la comblent qu'en partie. Notons aussi que l'OCDE a publié d'une manière concomitante d'autres tables types ; mais les données utilisées, plus nombreuses que celles des Nations-Unies sont nettement moins fiables. Les nombreuses méthodes statistiques employées, sans être originales, rendent son utilisation sans doute plus compliquée. Enfin, leurs performances, comparées à leurs exigences (ces tables dépendent d'un nombre plus important de paramètres), classent cette série de tables parmi les plus approximatives (voir Brouard [5]).

## 7.1 Principe des tables-types

Depuis Lederman, puis Coale et Demeny, on sait qu'une table de mortalité transversale a un profil par âge propre au pays dont elle provient ; quand la mortalité baisse, le niveau des risques de décès par âge certes diminue mais la spécificité du profil par âge se maintient longtemps. Autrement dit l'écart entre deux profils d'un même pays à deux dates différentes sera bien moindre que l'écart entre les profils de deux pays différents.

L'idée de *régionalisation* de la mortalité est donc fondamentale. Une table de mortalité ne peut être quelconque ; son profil par âge se modifie dans des limites finies. On parle alors d'un *univers* restreint de la mortalité et c'est cet univers qu'un modèle de tables type prétend analyser et mesurer. Cet univers abstrait peut être appréhendé à partir d'observations réelles. Avant de chercher une méthodologie d'analyse, il convient donc de répertorier les tables *réelles* qui vont constituer l'échantillon de base. La question de la constitution de cet échantillon est assez délicate ; simplifions en exprimant l'idée qu'il faut recouvrir les spécificités de tous les pays de l'époque actuelle. De plus, si on pense que la mortalité ancienne d'un pays développé peut se rapprocher de la mortalité actuelle d'un pays peu développé, il convient de joindre cette table à l'échantillon.

Dans le cadre des nouvelles tables types des Nations-Unies, qui s'adressent surtout aux pays en voie de développement, cette organisation a mis l'accent sur la qualité des données. En 1982, le nombre des tables de mortalité, propres aux pays en voie de développement, jugées de bonne qualité se sont élevés à 36. Les pays représentés ont été les suivants :

Amérique Latine	Asie	Afrique
Chili, Colombie	Inde	Tunisie
Costa Rica	Iran	
Salvador	Bangladesh <sup>1</sup>	
Guatemala, Guyane	Corée	
Honduras	Hong Kong	
Mexique	Singapour	
Pérou		
Trinité & Tobago		

<sup>1</sup> Il s'agit plus exactement du Matlab qui est une aire située au Bangladesh et dont le fichier de l'ensemble de sa population fait l'objet de nombreuses études démographiques

L'absence de table de mortalité provenant d'Afrique Noire réduit la portée de ces tables types, mais les Nations-Unies ont préféré la qualité à une universalité géographique usurpée.

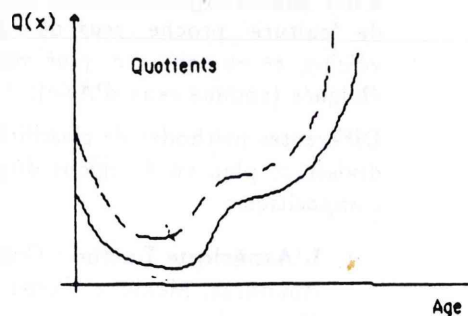
## 7.2 La méthodologie

La méthodologie est maintenant classique et déjà utilisée par Lederman en 1959 : il s'agit d'une analyse en composantes principales appliquée à une fonction particulière de la table de mortalité, avec une régionalisation ; c'est-à-dire un découpage de l'espace principal.

Pour simplifier la description de cette technique d'analyse des données, nous allons raisonner sur un espace à 3 dimensions donc visible; le raisonnement sera transposé aux 18 dimensions de la mortalité qui proviennent des 18 groupes d'âge classique d'une table de mortalité abrégée (groupe 0-1 an, 1-4ans, 5-10, ..., 80-85 ans, 85 et +).

*partiellement représentable*

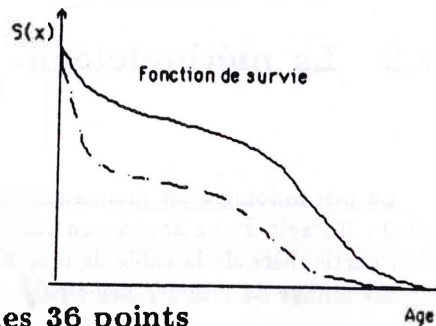
Nous choisissons trois groupes d'âge : la mortalité infantile 0-1 an, la mortalité à l'âge adulte, 30-34 ans, et la mortalité des personnes âgées 75-79 ans. Pour les 36 tables retenues, la méthode indique qu'il faut calculer les logits des quotients,  $\text{logit } {}_nq_x = \frac{1}{2} \log \frac{nq_x}{1-nq_x}$ , plutôt que les quotients eux-même, ou les taux ou encore les survivants, car cette fonction, proche de la fonction logarithme des quotients chère à Coale et Demeny ou à Lederman<sup>1</sup>, est assez représentative des autres fonctions de la mortalité ( $q_x$ ,  $S_x$ ,  $d_x$ ,  $e_x$ ) et d'une formulation assez simple. Détaillons ce point. Le logit du quotient est un des meilleurs indices de distance entre deux tables, en ce sens que deux tables proches vis à vis du critère logit  $q$ , le seront pour les critères  $q$ ,  $s$ ,  $d$  ou  $e$ , alors que le critère  $q$  ne l'est pas pour le critère  $s$ , comme le montre le graphique suivant :



Les deux tables en trait pointillé et en trait continu sont très proches sur cette figure qui dessine les tables en terme de quotient,

<sup>1</sup>En effet, pour  ${}_nq_x$  faible, la fonction logit est proche de la fonction logarithme.

alors qu'elles apparaissent très éloignées lorsqu'on représente ces mêmes tables en terme de fonction de survie, car la mortalité infantile joue un rôle prédominant dans cette représentation.



### 7.2.1 Analyse du nuage des 36 points

- Si on porte les 36 tables dans cet espace à trois dimensions, on constate que le nuage ainsi formé a une forme de "ballon de rugby", montrant un lien évident entre les niveaux de la mortalité aux différents âges : un "pays-année" qui a une mortalité infantile parmi les plus basses du lot, a une forte probabilité de figurer parmi les tables ayant une des plus faibles mortalités à 30 ans ainsi qu'à 75 ans.

L'axe principal du ballon de rugby qui est déterminé par le calcul, s'appelle la *première composante principale*. Son centre de gravité correspond à la représentation de la table moyenne, c'est-à-dire de la table fictive dont le logit du quotient à chaque âge est égal à la moyenne des logits des quotients des tables de l'échantillon. Les pays-années qui ont la mortalité la plus basse sont à une extrémité du ballon, celles qui ont la mortalité la plus forte sont à l'autre extrémité.

- Prenons un plan passant par le centre de gravité du ballon et perpendiculaire à la première composante principale, et projetons les 36 points sur ce plan.

L'analyse des positions relatives des différents pays est très fructueuse : on remarque en effet que les points correspondants à un même pays mais à des années différentes sont très proches l'un de l'autre<sup>2</sup>, et que les pays de "culture" proche (ceux de l'Amérique Latine notamment) sont assez voisins, et en tout cas, plus voisins entre eux que des pays de culture éloignée (comme ceux d'Asie).

Différentes méthodes de classification ont convaincu les Nations Unies de diviser ce plan en 4 régions dont les intitulés ont été suggérés par leur composition :

1. **L'Amérique Latine** : Colombie, Costa Rica, Salvador, Guatemala, Honduras, Mexique, Pérou mais aussi, les Philippines, Ceylan et la Thaïlande.
2. **Le Chili** qui se trouve assez éloignée du reste du précédent groupe, et doit, selon les auteurs, constituer un groupe à lui tout seul.

<sup>2</sup>Ces points étaient donc pratiquement alignés dans l'espace à 3 dimensions.

3. L'Asie du Sud : Inde, Iran, Bangladesh, Tunisie.

4. L'Extrême Orient : Hong-Kong, République de Corée, Singapour, mais aussi Guyane et Trinité et Tobago.

- On calcule alors dans le plan les centres de gravité, propres à chacun des 4 groupes, et résume l'univers de la mortalité à 4 lignes parallèles à l'axe principal du nuage original et passant par ces 4 centres de gravité régionaux.

Si  $\bar{U}_1$  est le vecteur unitaire donnant l'orientation de l'axe principal, et  $\bar{Y}_r$  le vecteur donnant les coordonnées du centre de gravité d'une région  $r$ , une table type, notée  $\bar{Y}$  propre à la région  $r$ , et un vecteur d'équation :

$$\bar{Y} = \bar{Y}_r + \lambda \bar{U}_1, \quad \text{avec } \lambda \text{ quelconque.} \quad (7.116)$$

### 7.2.2 Généralisation à plus de 3 dimensions

L'équation 7.116 est généralisable à l'espace à 18 dimensions qui correspond aux 18 groupes d'âge d'une table de mortalité. Pour un âge  $x$  donné, on peut donc écrire :

$$Y_x = Y_{rx} + \lambda U_{1x} \quad \text{soit} \quad (7.117)$$

$$\frac{1}{2} \text{Log} \left( \frac{n q_x}{1 - n q_x} \right) = Y_{rx} + \lambda U_{1x}. \quad (7.118)$$

et déduire ainsi tous les quotients d'une table d'un niveau  $\lambda$  donnée et d'une région  $r$ .

### 7.2.3 Tables types à 2 et 3 paramètres

Les tables types à un paramètre sont d'usage les plus courants, mais il peut arriver que les données que l'on détient sur une table soient parcellaires mais fiables, et qu'ainsi on souhaite ajuster une table type qui respecte plus l'originalité des données, quitte à s'écarter du standard des tables à un paramètre. Mais, s'écarter ne signifie pas se déplacer n'importe où dans l'univers de la mortalité, mais se déplacer dans un univers vraisemblable plus ouvert, c'est à dire dans un univers à deux, voire trois dimensions. ← 2 mai

Revenons en effet au nuage des points projetés sur le plan perpendiculaire à l'axe principal, celui-là même que nous avons régionalisé. Ce nuage projeté qui est la coupe du ballon de rugby, n'est pas rond mais a une forme légèrement allongée : l'axe principale de cette ellipse est appelée la seconde composante principale, le petit axe correspond à la troisième.

Le ballon de rugby initial n'était pas un ellipsoïde de révolution mais détenait trois directions principales orthogonales. Si on prend ces trois directions principales et le centre de gravité d'une région  $r$  comme nouvelle base ; une table

quelconque à deux paramètres ( $\lambda_1$  et  $\lambda_2$ ) est engendrée par l'équation :

$$\bar{Y}_2 = \bar{Y}_r + \lambda_1 \bar{U}_1 + \lambda_2 \bar{U}_2, \quad (7.119)$$

et une table à trois paramètres :

$$\bar{Y}_3 = \bar{Y}_r + \lambda_1 \bar{U}_1 + \lambda_2 \bar{U}_2 + \lambda_3 \bar{U}_3, \quad (7.120)$$

avec  $\bar{U}_i \bar{U}_j = \delta_{i,j}$  où  $\delta_{i,j}$  est le symbol de Kronecker.

Dans le cas d'école où une table de mortalité n'est définie que par trois groupes d'âge, l'équation 7.120 est l'équation d'une table quelconque, autrement dit l'espace type à trois paramètres est aussi grand que l'univers de la mortalité. Mais dans le cas d'un univers de la mortalité décrit par 18 groupes d'âge, l'équation 7.120 correspond à une projection d'une table de l'univers sur le *sous-espace* à trois dimensions. La théorie de l'analyse en composantes principales montre qu'il est possible de calculer les 15 autres axes principaux du nuage, et qu'ainsi l'équation d'une table quelconque de cet univers s'écrit :

$$\bar{Y} = \bar{Y}_r + \sum_{i=1}^{15} \lambda_i \bar{U}_i, \quad (7.121)$$

Il a été montré que, sur cet échantillon, le nuage initial est reconstitué à environ 90 % avec un seule composante, à 95 % avec deux composantes et 98 % avec trois composantes. Autrement dit les composantes d'ordre supérieur à trois n'apportent pratiquement pas d'information.

#### 7.2.4 Les résultats

L'analyse statistique a été faite pour les hommes et pour les femmes séparément. Le premier tableau 7.1 donne la valeur des logits moyens à chaque âge de chacune des quatre régions : Amérique Latine, Chili, Asie du Sud, Extrême Orient ainsi que le logit moyen calculé sur l'ensemble des tables. Les auteurs qualifient cette "région" de *générale*. On a montré aussi que ce modèle général diffère peu du modèle Ouest de Coale et Demery, dont le standard de Brass a lui aussi été en partie tiré.

Le second tableau 7.2 donne les valeurs des coordonnées des trois premières composantes principales. On peut vérifier que la somme des carrés des coordonnées est égal à 1, et que le produit scalaire de deux des vecteurs est nul.

Ce paragraphe nous a montré comment il était possible de réduire l'univers de la mortalité et de le paramétrer, maintenant nous allons voir comment il est possible d'utiliser ces modèles.

### 7.3 Application aux ajustements

Un des intérêts des tables types est de pouvoir ajuster une table de mortalité empirique. En effet, supposons que nous soyons capables, par des enquêtes,

deux à deux

Tableau 7.1: Logits des quotients moyens propres à chacune des 4 régions définies par les Nations-Unies.

HOMMES					
Âge $x$	Amérique		Asie du		Générale $Y_{rx}$
	Latine $Y_{rx}$	Chili $Y_{rx}$	Sud $Y_{rx}$	Extrême Orient $Y_{rx}$	
0	-1,130	-1,047	-0,979	-1,536	-1,277
1	-1,491	-1,820	-1,243	-2,153	-1,790
5	-2,131	-2,429	-2,017	-2,627	-2,362
10	-2,410	-2,530	-2,443	-2,667	-2,563
15	-2,223	-2,249	-2,356	-2,429	-2,345
20	-2,014	-2,029	-2,273	-2,236	-2,165
25	-1,936	-1,911	-2,169	-2,153	-2,092
30	-1,870	-1,787	-2,063	-2,060	-2,002
35	-1,763	-1,667	-1,902	-1,893	-1,870
40	-1,642	-1,525	-1,713	-1,683	-1,710
45	-1,497	-1,378	-1,511	-1,476	-1,529
50	-1,342	-1,220	-1,286	-1,231	-1,331
55	-1,157	-1,038	-1,082	-1,028	-1,130
60	-0,970	-0,842	-0,847	-0,771	-0,911
65	-0,747	-0,632	-0,630	-0,547	-0,685
70	-0,523	-0,421	-0,402	-0,330	-0,457
75	-0,295	-0,211	-0,196	-0,119	-0,230
80	-0,040	0,012	-0,001	0,106	0,008
FEMMES					
0	-1,225	-1,126	-0,971	-1,427	-1,360
1	-1,458	-1,826	-1,154	-1,954	-1,775
5	-2,142	-2,530	-1,941	-2,563	-2,398
10	-2,470	-2,647	-2,374	-2,689	-2,657
15	-2,323	-2,392	-2,193	-2,334	-2,456
20	-2,146	-2,205	-2,096	-2,165	-2,293
25	-2,041	-2,099	-2,050	-2,035	-2,193
30	-1,941	-1,991	-1,961	-1,946	-2,089
35	-1,832	-1,881	-1,874	-1,824	-1,975
40	-1,743	-1,759	-1,761	-1,692	-1,848
45	-1,624	-1,616	-1,614	-1,523	-1,694
50	-1,480	-1,460	-1,391	-1,336	-1,509
55	-1,287	-1,261	-1,156	-1,138	-1,303
60	-1,075	-1,053	-0,908	-0,938	-1,083
65	-0,832	-0,804	-0,680	-0,727	-0,844
70	-0,592	-0,582	-0,432	-0,509	-0,595
75	-0,360	-0,351	-0,175	-0,284	-0,342
80	-0,086	-0,106	0,059	-0,013	-0,065

Tableau 7.2: Valeurs des trois premières composantes principales qui génèrent les nouvelles tables des Nations-Unies.

Âge $x$	HOMMES			FEMMES		
	Première $\bar{U}_1$	Seconde $\bar{U}_2$	Troisième $\bar{U}_3$	Première $\bar{U}_1$	Seconde $\bar{U}_2$	Troisième $\bar{U}_3$
0	0,237	-0,460	0,093	0,183	-0,510	0,239
1	0,361	-0,688	-0,292	0,314	-0,522	-0,111
5	0,334	0,064	-0,471	0,317	0,089	0,076
10	0,305	0,125	-0,174	0,309	0,035	0,063
15	0,289	0,244	0,107	0,323	0,031	-0,267
20	0,287	0,107	0,288	0,326	0,078	-0,391
25	0,280	0,065	0,336	0,308	0,068	-0,282
30	0,280	0,033	0,337	0,290	0,005	-0,143
35	0,261	0,028	0,214	0,259	-0,014	-0,059
40	0,236	0,065	0,153	0,222	-0,022	0,189
45	0,208	0,087	0,066	0,192	0,019	0,248
50	0,178	0,106	0,000	0,172	0,044	0,337
55	0,151	0,113	-0,037	0,157	0,082	0,341
60	0,132	0,095	-0,106	0,143	0,081	0,383
65	0,122	0,108	-0,112	0,127	0,158	0,267
70	0,115	0,147	-0,223	0,118	0,242	0,144
75	0,104	0,210	-0,196	0,116	0,301	0,097
80	0,089	0,309	-0,381	0,098	0,505	-0,134

un recensement, ou encore en s'aidant de résultats antérieurs publiés dans la littérature démographique, d'évaluer la mortalité par âge d'une population ou de quelques groupes d'âge seulement et que nous désirions "tester" la validité des estimations ou compléter la table ; alors il est utile de savoir "ajuster" une table empirique à un modèle de tables types.

### 7.3.1 Tables complètes

Le premier cas que nous allons voir, est celui où nous disposons d'une table *complète*, c'est-à-dire d'une table dont les valeurs de la mortalité à chaque groupe d'âge sont connues, mais mal estimées. Il est intéressant de confronter la table empirique avec la table type d'un modèle régional choisi, la plus "proche". Cette confrontation permet de juger des biais systématiques par exemple, ou de l'importance de l'aléa si la table empirique est estimée à partir d'effectifs faibles.

Du point de vue mathématique, les calculs sont simples : la table type la plus proche<sup>3</sup> correspond à la projection de la table empirique sur l'espace type à un deux ou trois paramètres. Dans le cas où on désire ajuster une table  $\bar{Y}$  à un paramètre d'une région  $r$ , le coefficient  $\lambda_1$ , unique inconnu, est le produit

<sup>3</sup> au sens de la norme calculée en sommant les carrés des écarts entre logita.

(inconnue s. est. variable)



scalaire du vecteur écart  $\bar{Y} - \bar{Y}_r$  par la première composante principale :

$$\lambda_1 = (\bar{Y} - \bar{Y}_r) \cdot \bar{U}_1 = \sum_{x=1}^{18} (Y_x - Y_{rx}) U_{1x}, \quad (7.122)$$

et la table type  $\bar{Y}_t$  vaut alors :

$$Y_{tx} = Y_{rx} + \lambda_1 U_{1x} \quad \forall x \in [1, 18]. \quad (7.123)$$

Plus généralement, l'ajustement à un modèle à  $n$  paramètres ou composantes s'écrit :

$$Y_{tx}^{(n)} = Y_{rx} + \sum_{k=1}^n \lambda_k U_{kx} \quad \forall x \in [1, 18]. \quad (7.124)$$

avec

$$\lambda_k = (\bar{Y} - \bar{Y}_r) \cdot \bar{U}_k = \sum_{x=1}^{18} (Y_x - Y_{rx}) U_{kx}. \quad (7.125)$$

**Applications pratiques-utilisation de tableaux préremplis** Le calcul d'une table ajustée est grandement facilitée par l'utilisation de tableaux préremplis où les valeurs des logits moyens propres à une région et au sexe considéré ainsi que les valeurs des composantes principales sont déjà imprimées dans des colonnes adéquates. Ces tableaux préremplis figurent en annexe 1.

Les valeurs des quotients de la table empirique sont entrées dans la colonne (1) du tableau prérempli et choisi pour le sexe et la région qui lui correspondent. Après avoir effectué les calculs élémentaires des colonnes intermédiaires qui ne reflètent que la théorie exposée ci-dessus, on trouve dans la dernière colonne la table ajustée à un paramètre.

Le tableau 7.3 est un exemple d'ajustement de données manquantes. Il s'agit de la table de mortalité masculine de la population des Peuls Bandé. On a volontairement considéré que le quotient à 60-64 ans était mal estimé, et tenter d'ajuster un modèle d'Asie du Sud.

Si les données justifient l'ajustement d'une table à deux paramètres, on recopie dans la colonne (11), le résultat partiel de la colonne (10) et on obtient en colonne (16), l'ajustement désiré. Il en est de même avec la colonne (17) pour un ajustement à trois paramètres.

### 7.3.2 Tables incomplètes

Les nouvelles tables des Nations-Unies se prêtent facilement à l'ajustement des tables incomplètes et fournissent des tables complètes. Le calcul est même trivial dans le cas d'une table à un paramètre.

Tableau 7.3: Ajustement d'une table de mortalité où un quotient est manquant à un modèle de table type.

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : ASIE DU SUD , SEXE : HOMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

## TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0-1	0.023	-1.872	-0.979	-0.893	0.237	-0.212	-0.212	-0.574	-1.553	0.043
1-4	0.031	-1.715	-1.243	-0.472	0.361	-0.170	-0.382	-0.875	-2.117	0.014
5-9	0.010	-2.278	-2.017	-0.261	0.334	-0.087	-0.469	-0.811	-2.828	0.003
10-14	0.002	-3.106	-2.443	-0.664	0.305	-0.203	-0.672	-0.741	-3.183	0.002
15-19	0.005	-2.667	-2.356	-0.311	0.289	-0.090	-0.762	-0.702	-3.058	0.002
20-24	0.005	-2.627	-2.273	-0.354	0.287	-0.102	-0.863	-0.695	-2.968	0.003
25-29	0.003	-2.887	-2.169	-0.718	0.280	-0.201	-1.064	-0.678	-2.847	0.003
30-34	0.009	-2.323	-2.063	-0.261	0.280	-0.073	-1.137	-0.680	-2.742	0.004
35-39	0.009	-2.362	-1.902	-0.460	0.261	-0.120	-1.257	-0.632	-2.534	0.006
40-44	0.013	-2.161	-1.713	-0.448	0.236	-0.106	-1.363	-0.573	-2.286	0.010
45-49	0.014	-2.146	-1.511	-0.634	0.208	-0.132	-1.495	-0.504	-2.016	0.017
50-54	0.018	-1.991	-1.286	-0.705	0.178	-0.126	-1.620	-0.432	-1.717	0.031
55-59	0.014	-2.127	-1.082	-1.045	0.151	-0.158	-1.778	-0.367	-1.449	0.052
60-64	?	?	-0.847	?	0.132	?	?	-0.320	-1.167	0.088
65-69	0.018	-1.991	-0.630	-1.361	0.122	-0.167	-1.945	-0.297	-0.927	0.135
70-74	0.041	-1.583	-0.402	-1.180	0.115	-0.135	-2.080	-0.278	-0.680	0.204
75-79	0.026	-1.820	-0.196	-1.623	0.104	-0.170	-2.250	-0.253	-0.450	0.289
80-84	0.030	-1.733	-0.001	-1.732	0.089	-0.154	-2.404	-0.215	-0.217	0.393

Cumul: -2.404, Cor(3) = 0.991,  $\lambda_3 = \text{Cumul} / \text{Cor}(3) = -2.404 / 0.991 = -2.425$ 

Le facteur de correction due aux données manquantes est égal à la nouvelle norme de la composante calculée avec les seules données non manquantes.

## Tables à un paramètre

On suppose que la table abrégée est incomplète. Soit  $M$  l'ensemble des groupes d'âge, quelconques, manquants, et  $X$  celui des âges pourvus. On note  $\bar{Y}$  le vecteur des logits des quotients de la table incomplète, et  $\bar{Y}_r$  celui des logits moyens de la région choisie. La table type recherchée  $\bar{Y}^t$  s'écrit sous la forme :

$$\bar{Y}^t = \bar{Y}_r + \lambda_1 \bar{U}1. \quad (7.126)$$

Le problème consiste donc à choisir  $\lambda_1$  de telle sorte que les écarts entre la table empirique et la table type soient les plus faibles (sur la partie où les données existent bien entendu). Notre critère à minimiser est de nouveau celui des moindres carrés relatifs à la fonction logit. Soit  $L$  cette somme des carrés des écarts :

$$L = \sum_{x \in X} (Y_x - Y_x^t)^2 = \sum_{x \in X} (Y_x - \bar{Y}_x - \lambda_1 U_{1x})^2 \quad (7.127)$$

La dérivation de  $L$  par rapport à l'inconnu  $\lambda_1$  donne :

$$\frac{\partial L}{\partial \lambda_1} = 0 = -2 \sum_{x \in X} U_{1x} (Y_x - \bar{Y}_x - \lambda_1 U_{1x}),$$

$$\text{soit } \lambda_1 = \frac{\sum_{x \in X} U_{1x} (Y_x - \bar{Y}_x)}{\sum_{x \in X} U_{1x}^2}.$$

En notant  $a_1 = \sum_{x \in X} U_{1x}^2 = 1 - \sum_{x \in M} U_{1x}^2 < 1$ , on obtient :

$$\lambda_1 = \frac{\sum_{x \in X} U_{1x} (Y_x - \bar{Y}_x)}{a_1} \quad (7.128)$$

Dans le cas où quelques données seulement sont manquantes, il est plus facile de calculer  $a_1$  en sommant les carrés des coordonnées manquantes et en y retranchant 1, plutôt qu'en sommant les carrés des coordonnées pourvues.

Le remplissage des tableaux préremplis s'effectue des colonnes (1) à (7) en sautant les lignes correspondant aux données manquantes. Le paramètre  $\lambda_1$  est obtenu comme le rapport de la dernière ligne de la colonne (7) (cumul) divisée par  $a_1$ . Les colonnes (8), (9) et (10) sont alors des colonnes complètes.

**Remarque** Si on se sert de cette méthode pour évaluer un quotient manquant d'une table par ailleurs bien estimée, on risque d'être déçu à moins que la table incomplète s'ajuste parfaitement à un modèle de table type. En effet cette méthode d'ajustement donne une table cohérente complète, mais elle ne passe pas a priori par les points déjà pourvus.

### Tables à deux paramètres

Le cas des tables à deux paramètres est légèrement plus compliqué. Adoptant les mêmes principes nous trouvons successivement :

$$L = \sum_{x \in X} (Y_x - \hat{Y}_x - \lambda_1 U_{1x} - \lambda_2 U_{2x})^2 \quad (7.129)$$

$$\frac{\partial L}{\partial \lambda_1} = -2 \sum (Y_x - \hat{Y}_x - \lambda_1 U_{1x} - \lambda_2 U_{2x}) U_{1x} = 0$$

et de même pour  $\frac{\partial L}{\partial \lambda_2}$ , soit finalement :

$$\begin{aligned} \lambda_1 \left( \sum_{x \in X} U_{1x}^2 \right) + \lambda_2 \left( \sum_{x \in X} U_{2x} U_{1x} \right) - \sum_{x \in X} (Y_x - \hat{Y}_x) U_{1x} &= 0 \\ \lambda_1 \left( \sum_{x \in X} U_{2x} U_{1x} \right) + \lambda_2 \left( \sum_{x \in X} U_{2x}^2 \right) - \sum_{x \in X} (Y_x - \hat{Y}_x) U_{2x} &= 0, \end{aligned}$$

en notant

$$\begin{aligned} a_i &= \sum_{x \in X} U_{ix}^2 = 1 - \sum_{x \in M} U_{ix}^2, \\ b_{ij} &= \sum_{x \in X} U_{ix} U_{jx} = - \sum_{x \in M} U_{ix} U_{jx} = b_{ji}, \\ \alpha_i &= \sum_{x \in X} (Y_x - \hat{Y}_x) U_{ix}, \end{aligned}$$

on trouve :

$$\lambda_1 = \frac{a_2 \alpha_1 - b_{12} \alpha_2}{a_1 a_2 - b_{12}^2} \quad \text{et} \quad \lambda_2 = \frac{a_1 \alpha_2 - b_{12} \alpha_1}{a_1 a_2 - b_{12}^2}. \quad (7.130)$$

Nous remarquons que la valeur de  $\lambda_1$  est différente de celle trouvée dans le cas d'un ajustement à un paramètre. La procédure à suivre dans le cas d'un ajustement à deux paramètres consiste à remplir les colonnes (1) à (13). Le coefficient  $\alpha_2$  est en effet le dernier élément de la colonne (13). Le coefficient  $a_2$  se calcule comme le coefficient  $a_1$  puisqu'il s'agit de la nouvelle norme de la seconde composante principale. Sur le sous-espace  $X$  le produit scalaire  $b_{12} = b_{21}$  n'est plus nul et doit être calculé. On calcule alors  $\lambda_1$  et  $\lambda_2$  puis remplit alors les colonnes (8) et (9) d'une part et (14), (15) et (16) d'autre part.

### 7.3.3 Tables à trois paramètres

Des calculs tout à fait analogues conduisent au système suivant :

$$\begin{cases} a_1 \lambda_1 + b_{12} \lambda_2 + b_{13} \lambda_3 - \alpha_1 &= 0 \\ b_{12} \lambda_1 + a_2 \lambda_2 + b_{23} \lambda_3 - \alpha_2 &= 0 \\ b_{13} \lambda_1 + b_{23} \lambda_2 + a_3 \lambda_3 - \alpha_3 &= 0 \end{cases}$$

soit

$$\begin{bmatrix} a_1 & b_{12} & b_{13} \\ b_{12} & a_2 & b_{23} \\ b_{13} & b_{23} & a_3 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

en posant  $D = a_1 a_2 a_3 - 2b_{12} b_{23} b_{13} - b_{13}^2 a_2 - b_{12}^2 a_3 - b_{23}^2 a_1$ , on obtient la valeur de  $\lambda_1$  :

$$\lambda_1 = \frac{1}{D} \alpha_1 (a_2 a_3 - b_{23}^2) + \alpha_2 (b_{13} b_{23} - b_{12} a_3) + \alpha_3 (b_{12} b_{23} - b_{13} a_2). \quad (7.131)$$

De même les valeurs de  $\lambda_2$  et  $\lambda_3$  s'obtiennent par permutation circulaire.

## Chapitre 8

# Modélisation de la fécondité par âge : le modèle de Coale-Trussell

## Chapitre 9

# Modèles de population intégrant la fécondité et la mortalité

### 9.1 POPULATIONS MALTHUSIENNES (ap- proche infinitésimale)

**Définition 10** Une population malthusienne est une population dont :

1. la structure par âge est constante dans le temps,
2. l'effectif total varie selon un taux d'accroissement constant.

Cette seconde propriété implique (voir le paragraphe 1.3.2) que la taille de la population varie suivant une exponentielle. La référence à Malthus se trouve ainsi justifiée.

Il résulte de ces deux conditions diverses propriétés:

**Propriété 9.1.1** 1. Les taux brut de natalité et de mortalité sont constants.

2. A chaque âge la force de mortalité est invariante dans le temps. Il en résulte donc que tous les éléments d'une table de mortalité sont invariants dans le temps.

1. L'équation de la taille de la population totale en fonction du temps est  $P_t = P_0 e^{\rho t}$  où  $\rho$  est le taux de croissance.

Soit  $C(x) dx$  la proportion d'individus âgés entre  $x$  et  $x + dx$  années ( $\int C(x) dx = 1$ ). La population âgée entre 0 et  $dt$  est  $P_t C(0) dt$  ; c'est aussi le nombre des naissances survenues entre les temps  $t - dt$  et  $t$  car

les décès survenus dans cet intervalle de temps sont des infiniments petits du second ordre. Le taux brut de natalité  $n$  vaut :

$$n = \frac{P(t)C(0) dt}{P(t) dt} = C(0), \quad (9.132)$$

et est donc constant.

Le taux brut de mortalité est alors  $d = n - \rho$  lui aussi constant.

2. Si on rapporte la population âgée entre  $x$  et  $x + dx$  à la date  $t$ , à l'effectif initial de cette cohorte, c'est-à-dire aux naissances entre les dates  $t - x - dx$  et  $t - x$ , soit  $C(0)P_{t-x} dx$  on trouve la probabilité de survivre de la naissance jusqu'à l'âge  $x$  :

$$S(x) = \frac{P_t C(x) dx}{C(0) P_{t-x} dx} = \frac{C(x)}{C(0)} e^{\rho x} \quad (9.133)$$

Cette fonction de survie est *indépendante du temps*, la force de mortalité qui s'en déduit l'est donc aussi :

$$\mu(x) = -\frac{S'(x)}{S(x)} = -\frac{d \log S(x)}{dx} = -\frac{C'(x)}{C(x)} - \rho \quad (9.134)$$

Nous obtenons l'équation fondamentale des populations malthusiennes :

$$\mu(x) + \frac{C'(x)}{C(x)} + \rho = 0 \quad (9.135)$$

**Autre relations** De l'équation 9.133, nous déduisons la relation :

$$\begin{aligned} \int S(x) dx &= \frac{1}{n} \int C(x) e^{\rho x} dx \\ \text{soit } e_0 &= \frac{1}{n} \int C(x) e^{\rho x} dx \\ \text{avec } \int C(x) dx &= 1. \end{aligned}$$

De même

$$\begin{aligned} C(x) &= nS(x)e^{-\rho x} \\ \text{d'où } 1 &= n \int e^{-\rho x} S(x) dx, \end{aligned}$$



et la population d'âge  $x$  est<sup>1</sup> :

$$P(x, t) = S(x)e^{-\rho x} \times e^{\rho t}$$

Enfin,

$$\frac{P_t}{N_t} = \frac{1}{n} = \frac{1}{C(0)} = \int e^{-\rho x} S(x) dx.$$

Nous remarquons alors que le facteur par lequel il faut multiplier le nombre des naissances pour obtenir la taille de la population totale n'est plus comme dans le cas stationnaire, l'espérance de vie mais vaut :

$$\int e^{-\rho x} S(x) dx.$$

Dans le cas d'une population croissante ( $\rho > 0$ ), ce facteur est inférieur à l'espérance de vie. Louis Henry [20] a montré qu'on obtient une approximation de la taille de la population en multipliant l'espérance de vie par le nombre des naissances non pas de l'année en cours mais de  $a$  années auparavant. Prenant le cas particulier où la survie est linéaire, l'espérance de vie de l'ordre de 35 ans et la croissance faible, il trouve que  $a$  vaut 15 ans.

Plus généralement, nous allons montrer que  $a$  est proche de l'âge moyen de la population stationnaire.

Nous avons les équations :

$$\frac{P_t}{N_t} = \frac{1}{n} = \frac{1}{C(0)} = \int e^{-\rho x} S(x) dx,$$

et

$$N(t) = e^{\rho a} N(t - a).$$

Il faut prouver que :

$$P(t) = \frac{N(t)}{n} = N(t - a) \frac{e^{\rho a}}{n} \simeq N(t - a) e_0.$$

Soit, encore :

$$\frac{1}{ne_0} \simeq e^{-\rho a},$$

or

$$\frac{1}{ne_0} = \int \frac{S(x)}{e_0} e^{-\rho x} dx = \int S(x) dx$$

En introduisant les fonctions génératrices des moments et des cumulants (voir le paragraphe 4.3) de la fonction de survie normée  $S$  :

$$\phi(t) = \int e^{tx} S(x) dx,$$

<sup>1</sup>Du moins la densité de population.

nous obtenons :

$$\frac{1}{ne_0} = \phi(t) = \exp\left(k_1 t + k_2 \frac{t^2}{2!} + \dots\right), \quad (9.136)$$

avec  $t = -\rho$ . Dans le cas où  $t = -\rho$  est petit, on peut assimiler l'expression :

$$k_1 + k_2 \frac{t}{2!} + \dots + k_n \frac{t^{n-1}}{n!}$$

à  $k_1$ , qui est le premier cumulatif de la distribution des décès, soit encore l'âge moyen de la population stationnaire.

**Exercice 14** Supposer comme L. Henry que la survie est linéaire et le taux de croissance faible, montrer ainsi que pour obtenir la taille de la population totale il faut multiplier l'espérance de vie par le nombre des naissances de  $a = \frac{2}{3}e_0$  années auparavant.

**Solution 14** L'équation de la survie s'écrit :

$$S(x) = 1 - \frac{x}{2e_0}.$$

Alors :

$$\begin{aligned} \int xS(x) dx &= \frac{1}{e_0} \left[ \int \left(x - \frac{x^2}{2e_0}\right) dx \right] \\ &= 2e_0 - \frac{4}{3}e_0 = \frac{2}{3}e_0 = a. \end{aligned}$$

Si  $e_0 = 35$  ans,  $a = 15$  ans.

### 9.1.1 Existence d'une population malthusienne

L'équation 9.135 peut s'écrire sous la forme :

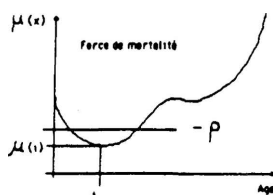
$$\mu(x) = -\frac{C'(x)}{C(x)} - \rho.$$

Or la condition  $\mu(x) \geq 0$  montre que la définition que nous avons donnée d'une population malthusienne n'implique pas qu'une telle population puisse toujours exister. N'importe quelle structure par âge  $C(x)$  n'implique pas qu'il puisse s'agir d'une population malthusienne.

Donnons nous par exemple une table de mortalité définie par sa force de mortalité  $\mu(x)$  ou sa fonction de survie  $S(x)$ . Nous supposons qu'il s'agit d'une population humaine auquel cas la fonction  $\mu(x)$  n'admet qu'un minimum positif pour un âge  $i$  :

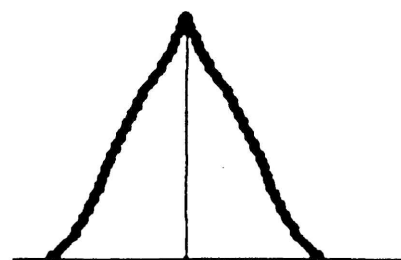
$$i = \min_{x \in [0, \infty)} \mu(x).$$

Faisons alors varier le taux de croissance de la valeur  $+\infty$  à la valeur  $-\infty$ , et examinons les formes de pyramides des âges qui s'en déduisent. Nous sommes amenés à distinguer plusieurs intervalles qui sont reproduits sur la figure suivante :

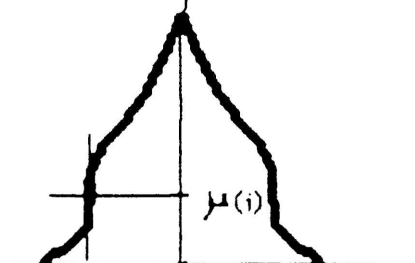


La discussion est alors la suivante :

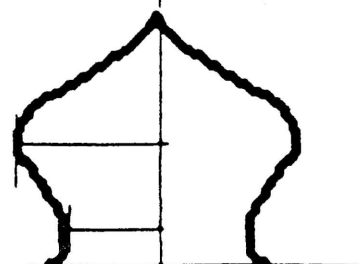
$\rho > -\mu(i)$  Ce cas correspond au cas classique d'une population croissante, stationnaire ou faiblement décroissante.



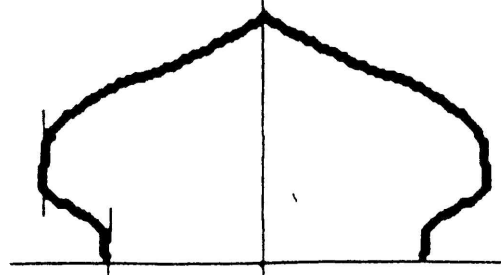
$\rho = -\mu(i)$  Un point d'inflexion à l'âge  $i$ .



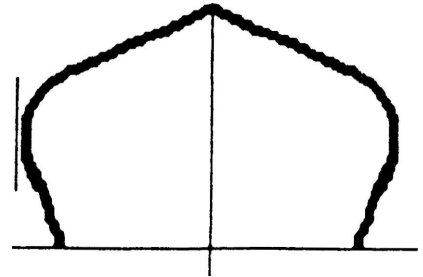
$-\mu(0) < \rho < -\mu(i)$  La pyramide a deux tangentes verticales et prend déjà une forme bombée vers le haut.



$\rho = -\mu(0)$  La première tangente verticale est descendue jusqu'à l'origine.



$\rho < -\mu(0)$  Dans ce cas le taux de décroissance est égale au niveau de la mortalité infantile. La structure est presque ovale.



Le cas le plus connu est le premier mais il n'est pas impossible que certaines populations aient eu des formes de pyramide des âges conformes aux autres cas et soient aujourd'hui disparues.

### 9.1.2 Répartition par âge des décès

A la date  $t$ , le nombre des décès d'individus âgés entre  $x$  et  $x + dx$  survenus entre les instants  $t$  et  $t + dt$  est  $P(t, x) dx \times \mu(x) dt$ .

La distribution des décès est alors :

$$\delta(x) = \frac{P(t, x) dx \mu(x) dt}{dP(t) dt} = \frac{1}{d} C(x) \mu(x) \quad (9.137)$$

On tire alors aussi les expressions suivantes :

$$d = \int C(x) \mu(x) dx \quad (9.138)$$

$$d = n \int S(x) \mu(x) e^{-\rho x} dx \quad (9.139)$$

Si on note  $d$  la densité des décès déduite de la table de mortalité, on obtient :

$$\begin{aligned} d(x) &= S(x) \mu(x) \\ \delta(x) &= \frac{e^{-\rho x} d(x)}{\int d(x) e^{-\rho x} dx}, \end{aligned} \quad (9.140)$$

$$\delta(x) = \frac{e^{-\rho x} d(x)}{\phi(-\rho)}, \quad (9.141)$$

avec  $\int d(x) = \int \delta(x) = 1$ , et où  $\phi$  est la fonction génératrice des moments de la distribution  $d(x)$ .

**Fonction génératrice de la distribution des décédés** Soit  $\xi(t)$  la fonction génératrice de la distribution,  $\delta(x)$ , des décédés dans une population malthusienne :

$$\xi(t) = \int e^{tx} \delta(x) dx.$$

Alors

$$\xi(t) = \frac{1}{\phi(-\rho)} \int e^{tx} e^{-\rho x} d(x) dx = \frac{1}{\phi(-\rho)} \int e^{(t-\rho)x} d(x) dx,$$

et finalement :

$$\xi(t) = \frac{\phi(t-\rho)}{\phi(-\rho)}. \quad (9.142)$$

En passant aux logarithmes, nous obtenons une relation plus simple entre les cumulants,  $k_i^d$ , de la distribution des décès,  $d(x)$ , de la table de mortalité, et les cumulants,  $k_i^\delta$ , de la distribution des décès,  $\delta(x)$  dans une population malthusienne.

$$\begin{aligned} \log \xi(t) &= k_1^\delta t + k_2^\delta \frac{t^2}{2!} + \dots + k_n^\delta \frac{t^n}{n!} + \dots \\ &= k_1^d (t - \rho) + k_2^d \frac{(t - \rho)^2}{2!} + \dots \\ &\quad - \rho \left( k_1^d + k_2^d \frac{\rho}{2!} + \dots \right) \\ &= \left( k_1^d - k_2^d + \dots + (-1)^n \frac{\rho^{n-1}}{(n-1)!} k_n^d + \dots \right) t + \\ &\quad k_2^\delta(\rho) \frac{t^2}{2!} + \dots \end{aligned}$$

On déduit en particulier que le premier cumulant,  $k_1^\delta$ , vaut :

$$k_1^\delta = k_1^d - k_2^d + \dots + (-1)^n \frac{\rho^{n-1}}{(n-1)!} k_n^d. \quad (9.143)$$

Lorsque le taux de croissance,  $\rho$ , est faible, on peut arrêter le développement au second ordre par exemple, puisque la série alternée est convergente, et écrire que l'âge moyen des décédés,  $\mu_1^\delta$ , vaut :

$$\mu_1^\delta = \int \delta(x) dx = k_1^\delta = k_1^d - k_2^d \rho = e_0 - \sigma^2 \rho + O(\rho^2), \quad (9.144)$$

où  $\sigma^2$  est la variance de l'âge au décès (voir l'équation ref:ecart-type-dec).

Si la population est croissante ( $\rho > 0$ ) l'âge moyen des décédés d'une année sera inférieur à l'espérance de vie.

### 9.1.3 Estimation du taux de croissance à partir de la distribution des décès

Supposons que nous ne connaissions que la pyramide des âges d'une population à partir d'un recensement et qu'une enquête à passages répétées donne

la distribution des décès durant deux des passages selon l'âge. Alors, si nous faisons l'hypothèse que la population est malthusienne, on peut calculer le taux de croissance et la table de mortalité.

En effet, la relation 9.135 :

$$\frac{C'(x)}{C(x)} + \mu(x) + \rho = 0 \quad \text{s'écrit en vertu de 9.137} \quad (9.145)$$

$$\frac{C'(x)}{C(x)} + d \frac{\delta(x)}{C(x)} + \rho = 0 \quad (9.146)$$

ou encore  $C'(x) + d\delta(x) + \rho C(x) = 0$ , soit en notant :

$$f(x_+) = \int_x^{\infty} f(u) du,$$

et en intégrant

$$C(x) = \rho C(x_+) + d\delta(x_+) \quad \text{ou} \quad (9.147)$$

$$\frac{C(x)}{C(x_+)} = \rho + d \frac{\delta(x_+)}{C(x_+)} \quad \text{ou encore} \quad (9.148)$$

$$\frac{C(x)}{\delta(x_+)} = \rho \frac{C(x_+)}{\delta(x_+)} + d. \quad (9.149)$$

Chacune des relations ci-dessus permet de calculer par une régression linéaire les coefficients  $\rho$  et  $d$ . Le nuage des points est en effet formé des valeurs des différentes quantités à différents âges.

## 9.2 POPULATIONS STABLES

Dans le paragraphe précédent sur les populations malthusiennes, nous n'avons pas ~~fait de distinguer~~ les hommes des femmes, ni vraiment explicité le processus de naissance : les naissances, garçons et filles, entre les instants  $t$  et  $t + dt$  étaient au nombre de  $P(t)C(0) dt$  soit  $n P_0 e^{rt} dt$ , c'est-à-dire évoluaient de manière exponentielle.

La théorie des populations stables suggère que le nombre des naissances est lié au nombre d'individus en âge de procréer, c'est-à-dire à la fécondité. Mais, de quelle fécondité s'agit-il, de celle des hommes, des femmes, ou plus généralement des couples ? Les théoriciens des populations stables se sont penchés sur le problème particulier des populations sexuées mais butent toujours sur l'absence d'un bon modèle de nuptialité (voir Pollard par exemple). Nous supposons ici qu'il n'y a pas de déséquilibre entre les sexes.

On suppose ainsi qu'une femme d'âge  $x$  à la date  $t$ , donne en moyenne naissances à  $F_f(x, t) dt$  enfants durant l'instant  $dt$ . Le taux de fécondité féminin

ainsi introduit dépend de l'âge et de la période. Le nombre des naissances (garçon et filles réunis) entre les instants  $t$  et  $t + dt$  est alors :

$$N(t) dt = \int_x P_f(x, t) F_f(x, t) dx \times dt, \quad (9.150)$$

où  $P_f$  est l'effectif de la population féminine.

En utilisant le rapport de masculinité  $m$ , nous pouvons introduire la fécondité en terme de fille <sup>par</sup> femme :

$$f_f(x, t) = (1 - m) F_f(x, t). \quad (9.151)$$

Ainsi, le nombre des naissances féminines est alors :

$$P_f(x, t) = (1 - m) N(t) = \int_x P_f(x, t) f_f(x, t) dt, \quad (9.152)$$

celui des naissances masculines est alors simplement :

$$P_h(x, t) = m P_f(x, t) \quad (9.153)$$

Si nous introduisons maintenant la fécondité masculine  $F_h(x, t)$  (enfants par homme) et  $f_h(x, t) = m F_h(x, t)$  (garçons par homme), l'équation 9.153, doit aussi pouvoir s'écrire :

$$P_h(0, t) = m N(t) = \int_x P_h(x, t) f_h(x, t) dt. \quad (9.154)$$

La cohérence globale entre les effectifs de population et les fécondités masculine et féminine doit se traduire par l'égalité :

$$N(t) = \frac{1}{m} \int_x P_h(x, t) f_h(x, t) dt = \frac{1}{1 - m} \int_x P_f(x, t) f_f(x, t) dt, \quad \forall t. \quad (9.155)$$

Nous pouvons alors en venir à la définition d'une population stable :

**Définition 11** Une population stable est une population fermée dont les taux de mortalité et de fécondité par âge et sexe sont invariants dans le temps.

En toute rigueur, on doit ajouter que les taux de fécondité et de mortalité de chaque sexe doivent vérifier une relation de cohérence que nous allons déterminer.

Notons  $S_h(x)$  et  $S_f(x)$  les probabilités de survie jusqu'à l'âge  $x$ , et  $f_h(x)$  et  $f_f(x)$  les taux de fécondité garçon par homme et fille par femme.

Les équations 9.152 et 9.154 s'écrivent alors :

$$P_f(0, t) = \int P_f(x, t) f_f(x) dt, \quad P_h(0, t) = \int P_h(x, t) f_h(x) dt. \quad (9.156)$$

En introduisant les probabilités de survie, nous obtenons aussi :

$$P_f(x, t) = S_f(x)P_f(0, t), P_h(x, t) = S_h(x)P_h(0, t). \quad (9.157)$$

Mais il est plus facile de raisonner sur les naissances ; nous obtenons alors deux équations de renouvellement qui doivent être cohérentes :

$$\begin{cases} N(t) = \int N(t-x)S_f(x)f_f(x) dx, \\ N(t) = \int N(t-x)S_h(x)f_h(x) dx. \end{cases} \quad (9.158)$$

Il parait utile de définir la *fonction de maternité* ou de *paternité* comme le produit de la fonction de survie à l'âge  $x$  par la fécondité à cet âge :  $\Phi(x) = S(x)f(x)$ . Il est possible de donner une définition moins restrictive des populations stables en n'imposant la constance des seules fonctions de maternité et paternité, et non de chacun de leurs composants. Nous avons vu en particulier que la fécondité des hommes est indéfinie mais non nulle aux âges élevés et dans une population polygame ; seul le produit des taux de fécondité par la survie, c'est à dire la fonction de paternité est alors définie et nulle.

Nous obtenons un système à deux équations de renouvellement :

$$\begin{cases} N(t) = \int N(t-x)\Phi_f(x) dx, \\ N(t) = \int N(t-x)\Phi_h(x) dx. \end{cases} \quad (9.159)$$

Essayons une solution de la forme :

$$N(t) = N_0 e^{rt}$$

qui correspond à un rythme exponentiel des naissances, c'est-à-dire encore à une population malthusienne.

Alors :

$$N_0 e^{rt} = \int N_0 e^{rt} e^{-rx} \Phi_f(x) dx \quad \text{où encore} \quad (9.160)$$

$$1 = \int e^{-rx} \Phi_f(x) dx \quad (9.161)$$

de même pour le sexe masculin.

Notons  $\xi_f(t)$  et  $\xi_h(t)$  les fonctions génératrices des moments de la distribution des fonctions de maternité et de paternité normées  $\phi_f(x)$  et  $\phi_h(x)$  :

$$\begin{aligned} \phi_f(x) &= \frac{1}{R_{0f}} \Phi_f(x) dx \\ \phi_h(x) &= \frac{1}{R_{0h}} \Phi_h(x) dx \end{aligned}$$

où  $R_{0f} = \int \Phi_f(x) dx$  et  $R_{0h} = \int \Phi_h(x) dx$  sont respectivement les taux nets de reproduction féminin et masculin.



$$\xi_f(t) = \int e^{tx} \phi_f(x) dx \quad (9.162)$$

$$\xi_h(t) = \int e^{tx} \phi_h(x) dx \quad (9.163)$$

Alors les deux équations de renouvellement se ramènent au calcul d'une intersection :

$$\begin{cases} \xi_f(-r) = R_{0f} \\ \xi_h(-r) = R_{0h} \end{cases} \quad (9.164)$$

Étudions alors les variations des fonctions  $\xi$ . Rappelons que les fonctions de maternité et de paternité sont positives et à support borné. Lorsque  $t \rightarrow \infty$ ,  $\xi(t) \rightarrow \infty$ , de même, lorsque  $t \rightarrow -\infty$ ,  $\xi(t) \rightarrow 0$ . De même, la dérivée :

$$\frac{d}{dt} \xi(t) = \int e^{tx} x \phi(x) dx$$

est toujours positive.

Les fonctions génératrices des moments  $\xi_f(t)$  et  $\xi_h(t)$ , sont donc positives monotones croissantes. Nous tirons alors les deux conclusions suivantes :

1. Chacune des équations du système 9.164 a donc une solution réelle, et celle-ci est unique.
2. Les fonctions  $\xi_f$  et  $\xi_h$  sont inversibles, la cohérence entre les deux solutions des deux équations impose une contrainte sur les inverses des fonctions génératrices des moments :

$$\xi_f^{-1}(R_{0f}) = \xi_h^{-1}(R_{0h}) \quad (= -r) \quad (9.165)$$

La solution réelle,  $r$ , est un taux de croissance et a reçu le nom de *taux de Lotka*. L'existence et l'unicité d'une solution réelle, n'exclue pas la possibilité d'autres racines complexes.

Posons  $z = u + iv$  une telle racine complexe éventuelle, alors celle-ci doit vérifier l'équation :

$$\xi(z) = \int e^{-uz} (\cos(-vx) + i \sin(-vx)) \phi(x) dx = R_0$$

Soit, en séparant partie réelle et partie imaginaire :

$$\begin{cases} \int e^{-ux} \cos(vx) \phi(x) dx = R_0 \\ \int e^{-ux} \sin(vx) \phi(x) dx = 0 \end{cases} \quad (9.166)$$

Comme le cosinus peut prendre des valeurs strictement inférieures à 1, on déduit de la première équation l'inégalité suivante :

$$\int e^{-ux} \phi(x) dx > R_0 \quad (9.167)$$

Mais comme l'unique solution réelle vérifie l'égalité :

$$\int e^{-rx} \phi(x) dx = R_0, \quad (9.168)$$

On en déduit que  $u < r$ , c'est-à-dire aussi que les parties réelles des solutions complexes sont inférieures à la solution réelle. Notons aussi que si  $z$  est racine,  $\bar{z}$  son complexe conjugué est également racine.

**Conclusion** Nous venons ainsi de montrer qu'une population dont les taux de fécondité et de mortalité sont constants dans le temps, est une population malthusienne dont le taux de croissance est l'unique solution de l'équation :

$$\int e^{-rx} \phi(x) dx = R_0, \quad (9.169)$$

La forme d'une pyramide d'une population stable est donc celle d'une population malthusienne. Toutes les relations démontrées à propos des populations malthusiennes sont donc valables. Au lieu d'*observer* un taux de croissance, on *explique*, ce taux de croissance par une combinaison de la fécondité et de la mortalité.

Il s'agit maintenant de pouvoir calculer ce taux de Lotka en fonction des taux de fécondité et de mortalité.

### 9.2.1 Estimation du taux de croissance de Lotka

Rappelons que le taux de Lotka,  $r$ , vérifie les équations suivantes :

$$\begin{cases} \xi_f(-r) = R_{0f} \\ \xi_h(-r) = R_{0h} \end{cases} \quad (9.170)$$

Rappelons aussi la définition des cumulants d'une distribution (voir le paragraphe 4.3) :

$$k_1 t + \dots + \frac{k_r t^r}{r!} + \dots = \log(\xi(t)) \quad (9.171)$$

A. Lotka a montré que la solution  $r$  était petite, nous pouvons alors, dans un premier temps arrêter le développement du membre gauche de l'équation au premier ordre et nous trouvons :

$$k_1(-r) = \log(\xi(-r)) = \log R_0 \quad (9.172)$$

$$\mu_1(-r) = \log R_0 \quad (9.173)$$

Soit finalement :

$$r_{(1)} = \frac{\log R_0}{\mu_1} \quad (9.174)$$

Le taux de Lotka peut être approché comme le rapport du logarithme du taux net de reproduction à l'âge moyen à la reproduction nette.  $\mu_1$  est en effet l'âge moyen de la distribution de la fonction de maternité ou de paternité :

$$\mu_{1f} = \frac{1}{R_0} \int x f_f(x) S_f(x) dx. \quad (9.175)$$

Cette première relation est généralement suffisante ; elle correspond à l'hypothèse que la distribution des taux nets de fécondité ~~est~~ est une fonction de Dirac au point  $\mu_1$ . Si nous supposons maintenant que la variance n'est pas nulle, et que les moments d'ordre supérieur sont nuls, ce qui sous-entend que la distribution est une loi normale, nous pouvons arrêter le développement au second ordre :

$$k - 1(-r) + k_2 \frac{(-r)^2}{2!} = -\log R_0. \quad (9.176)$$

Soit encore :

$$k_2 r^2 - 2k_1 r + 2 \log R_0 = 0$$

Cette équation du second degré a deux solutions, car la loi normale a un résidu de masse non nulle pour les âges négatifs. Cette solution parasite correspondant à la plus grande racine. Ainsi la valeur approchée au second ordre du taux de Lotka vaut :

$$r_2 = \frac{k_1 - \sqrt{k_1^2 - 2k_2 \log R_0}}{k_2} \quad (9.177)$$

Il est alors possible de faire un développement limité de la racine :

$$\begin{aligned} r_2 &= \frac{k_1}{k_2} - \frac{k_1}{k_2} \sqrt{1 - 2 \frac{k_2}{k_1^2} \log R_0} \\ &= \frac{k_1}{k_2} \left[ 1 - \frac{k_2}{k_1^2} \log R_0 - \frac{1}{2} \left( \frac{k_2}{k_1^2} \log R_0 \right)^2 + \dots \right] \end{aligned}$$

Comme  $k_2 = \sigma^2$  où  $\sigma$  est l'écart-type de l'âge au décès, nous trouvons aussi une relation plus utilisée :

$$r_{(2)} = \frac{\log R_0}{\mu_1} + \frac{\sigma^2 (\log R_0)^2}{2\mu_1^3}. \quad (9.178)$$

## **Annexe A**

# **Tableaux préremplis pour l'ajustement d'une table type des Nations-Unies**

### **A.1 Femmes**

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : AMERIQUE LATINE , SEXE : FEMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l'abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Âge
0-1			-1.225		0.183						0-1
1-4			-1.458		0.314						1-4
5-9			-2.142		0.317						5-9
10-14			-2.470		0.309						10-14
15-19			-2.323		0.323						15-19
20-24			-2.146		0.326						20-24
25-29			-2.041		0.308						25-29
30-34			-1.941		0.290						30-34
35-39			-1.832		0.259						35-39
40-44			-1.743		0.222						40-44
45-49			-1.624		0.192						45-49
50-54			-1.480		0.172						50-54
55-59			-1.287		0.157						55-59
60-64			-1.075		0.143						60-64
65-69			-0.832		0.127						65-69
70-74			-0.592		0.118						70-74
75-79			-0.360		0.116						75-79
80-84			-0.086		0.098						80-84

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : CHILI , SEXE : FEMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4)  $\times$  (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times$  (5)  
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Âge
0-1			-1.126		0.183						0-1
1-4			-1.826		0.314						1-4
5-9			-2.530		0.317						5-9
10-14			-2.647		0.309						10-14
15-19			-2.392		0.323						15-19
20-24			-2.205		0.326						20-24
25-29			-2.099		0.308						25-29
30-34			-1.991		0.290						30-34
35-39			-1.881		0.259						35-39
40-44			-1.759		0.222						40-44
45-49			-1.616		0.192						45-49
50-54			-1.460		0.172						50-54
55-59			-1.261		0.157						55-59
60-64			-1.053		0.143						60-64
65-69			-0.804		0.127						65-69
70-74			-0.582		0.118						70-74
75-79			-0.351		0.116						75-79
80-84			-0.106		0.098						80-84

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : ASIE DU SUD , SEXE : FEMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Âge
0-1			-0.971		0.183						0-1
1-4			-1.154		0.314						1-4
5-9			-1.941		0.317						5-9
10-14			-2.374		0.309						10-14
15-19			-2.193		0.323						15-19
20-24			-2.096		0.326						20-24
25-29			-2.050		0.308						25-29
30-34			-1.961		0.290						30-34
35-39			-1.874		0.259						35-39
40-44			-1.761		0.222						40-44
45-49			-1.614		0.192						45-49
50-54			-1.391		0.172						50-54
55-59			-1.156		0.157						55-59
60-64			-0.908		0.143						60-64
65-69			-0.680		0.127						65-69
70-74			-0.432		0.118						70-74
75-79			-0.175		0.116						75-79
80-84			0.059		0.098						80-84

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : EXTREME ORIENT , SEXE : FEMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4)  $\times$  (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times$  (5)  
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Age
0-1			-1.427		0.183						0-1
1-4			-1.954		0.314						1-4
5-9			-2.563		0.317						5-9
10-14			-2.689		0.309						10-14
15-19			-2.334		0.323						15-19
20-24			-2.165		0.326						20-24
25-29			-2.035		0.308						25-29
30-34			-1.946		0.290						30-34
35-39			-1.824		0.259						35-39
40-44			-1.692		0.222						40-44
45-49			-1.523		0.192						45-49
50-54			-1.336		0.172						50-54
55-59			-1.138		0.157						55-59
60-64			-0.938		0.143						60-64
65-69			-0.727		0.127						65-69
70-74			-0.509		0.118						70-74
75-79			-0.284		0.116						75-79
80-84			-0.013		0.098						80-84



CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : GÉNÉRALE , SEXE : FEMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l'abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Age
0-1			-1.360		0.183						0-1
1-4			-1.775		0.314						1-4
5-9			-2.398		0.317						5-9
10-14			-2.657		0.309						10-14
15-19			-2.456		0.323						15-19
20-24			-2.293		0.326						20-24
25-29			-2.193		0.308						25-29
30-34			-2.089		0.290						30-34
35-39			-1.975		0.259						35-39
40-44			-1.848		0.222						40-44
45-49			-1.694		0.192						45-49
50-54			-1.509		0.172						50-54
55-59			-1.303		0.157						55-59
60-64			-1.083		0.143						60-64
65-69			-0.844		0.127						65-69
70-74			-0.595		0.118						70-74
75-79			-0.342		0.116						75-79
80-84			-0.065		0.098						80-84

**A.1.1 Table à deux composantes**

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION :                   , SEXE : FEMME

TABLE-TYPE A DEUX COMPOSANTES

- (4) Report des écarts aux logits moyens : (2) - (3)  
 (9) Report de la table-type à une composante exprimée en logit de quotient  
 (11) Seconde composante principale normée  
 (12) Produit scalaire terme à terme : (4) × (11)  
 (13) Cumul de (12) pour trouver l' abscisse  $\lambda_2$  sur la 2<sup>e</sup> composante  
 (14) Calcul de la correction :  $\lambda_2 \times (11)$   
 (15) Table-type à deux composantes exprimée en logit de quotient : (9) + (14)  
 (16) Table-type à deux composantes exprimée en quotient :  $\frac{e^{2 \times (15)}}{1 + e^{2 \times (15)}}$

Age	(4)	(9)	(11)	(12)	(13)	(14)	(15)	(16)	Age
0-1			-0.510						0-1
1-4			-0.522						1-4
5-9			0.089						5-9
10-14			0.035						10-14
15-19			0.031						15-19
20-24			0.078						20-24
25-29			0.068						25-29
30-34			0.005						30-34
35-39			-0.014						35-39
40-44			-0.022						40-44
45-49			0.019						45-49
50-54			0.044						50-54
55-59			0.082						55-59
60-64			0.081						60-64
65-69			0.158						65-69
70-74			0.242						70-74
75-79			0.301						75-79
80-84			0.505						80-84

**A.1.2 Table à trois composantes**

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION :                   , SEXE : FEMME

TABLE-TYPE A TROIS COMPOSANTES

- (4) Report des écarts aux logits moyens : (2) - (3)  
 (15) Report de la table-type à deux composantes exprimée en logit de quotient  
 (17) Troisième composante principale normée  
 (18) Produit scalaire terme à terme : (4) × (17)  
 (19) Cumul de (18) pour trouver l'abscisse  $\lambda_3$  sur la 3<sup>ème</sup> composante  
 (20) Calcul de la correction :  $\lambda_3 \times (17)$   
 (21) Table-type à trois composantes exprimée en logit de quotient : (15)+(20)  
 (22) Table-type à trois composantes exprimée en quotient :  $\frac{e^{2 \times (21)}}{1 + e^{2 \times (21)}}$

Age	(4)	(15)	(17)	(18)	(19)	(20)	(21)	(22)	Age
0-1			0.239						0-1
1-4			-0.111						1-4
5-9			0.076						5-9
10-14			0.069						10-14
15-19			-0.267						15-19
20-24			-0.391						20-24
25-29			-0.282						25-29
30-34			-0.143						30-34
35-39			-0.059						35-39
40-44			0.189						40-44
45-49			0.248						45-49
50-54			0.337						50-54
55-59			0.341						55-59
60-64			0.383						60-64
65-69			0.267						65-69
70-74			0.144						70-74
75-79			0.097						75-79
80-84			-0.134						80-84

**A.2 Hommes**

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : AMERIQUE LATINE , SEXE : HOMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Âge
0-1			-1.130		0.237						0-1
1-4			-1.491		0.361						1-4
5-9			-2.131		0.334						5-9
10-14			-2.410		0.305						10-14
15-19			-2.223		0.289						15-19
20-24			-2.014		0.287						20-24
25-29			-1.936		0.280						25-29
30-34			-1.870		0.280						30-34
35-39			-1.763		0.261						35-39
40-44			-1.642		0.236						40-44
45-49			-1.497		0.208						45-49
50-54			-1.342		0.178						50-54
55-59			-1.157		0.151						55-59
60-64			-0.970		0.132						60-64
65-69			-0.747		0.122						65-69
70-74			-0.523		0.115						70-74
75-79			-0.295		0.104						75-79
80-84			-0.040		0.089						80-84

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : CHILI , SEXE : HOMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l'abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Âge
0-1			-1.047		0.237						0-1
1-4			-1.820		0.361						1-4
5-9			-2.429		0.334						5-9
10-14			-2.530		0.305						10-14
15-19			-2.249		0.289						15-19
20-24			-2.029		0.287						20-24
25-29			-1.911		0.280						25-29
30-34			-1.787		0.280						30-34
35-39			-1.667		0.261						35-39
40-44			-1.525		0.236						40-44
45-49			-1.378		0.208						45-49
50-54			-1.220		0.178						50-54
55-59			-1.038		0.151						55-59
60-64			-0.842		0.132						60-64
65-69			-0.632		0.122						65-69
70-74			-0.421		0.115						70-74
75-79			-0.211		0.104						75-79
80-84			0.012		0.089						80-84



CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : ASIE DU SUD , SEXE : HOMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Age
0-1			-0.979		0.237						0-1
1-4			-1.243		0.361						1-4
5-9			-2.017		0.334						5-9
10-14			-2.443		0.305						10-14
15-19			-2.356		0.289						15-19
20-24			-2.273		0.287						20-24
25-29			-2.169		0.280						25-29
30-34			-2.063		0.280						30-34
35-39			-1.902		0.261						35-39
40-44			-1.713		0.236						40-44
45-49			-1.511		0.208						45-49
50-54			-1.236		0.178						50-54
55-59			-1.032		0.151						55-59
60-64			-0.847		0.132						60-64
65-69			-0.630		0.122						65-69
70-74			-0.402		0.115						70-74
75-79			-0.196		0.104						75-79
80-84			-0.001		0.039						80-84

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : EXTREME ORIENT , SEXE : HOMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l' abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Âge	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Âge
0-1			-1.536		0.237						0-1
1-4			-2.153		0.361						1-4
5-9			-2.627		0.334						5-9
10-14			-2.667		0.305						10-14
15-19			-2.429		0.289						15-19
20-24			-2.236		0.287						20-24
25-29			-2.153		0.280						25-29
30-34			-2.060		0.280						30-34
35-39			-1.893		0.261						35-39
40-44			-1.683		0.236						40-44
45-49			-1.476		0.208						45-49
50-54			-1.231		0.178						50-54
55-59			-1.028		0.151						55-59
60-64			-0.771		0.132						60-64
65-69			-0.547		0.122						65-69
70-74			-0.330		0.115						70-74
75-79			-0.119		0.104						75-79
80-84			0.106		0.089						80-84

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION : GÉNÉRALE , SEXE : HOMME

- (1) Quotients originaux :  $q$   
 (2)  $\text{logit}(q) = 1/2 \log \frac{q}{1-q}$   
 (3) logits moyens de la région  
 (4) Écarts aux logits moyens : (2) - (3)

TABLE-TYPE À UNE COMPOSANTE

- (5) Première composante principale normée  
 (6) Produit scalaire terme à terme : (4) × (5)  
 (7) Cumul de (6) pour trouver l'abscisse  $\lambda_1$  sur la 1<sup>re</sup> composante  
 (8) Calcul de la correction :  $\lambda_1 \times (5)$   
 (9) Table-type à une composante exprimée en logit de quotient : (8) + (3)  
 (10) Table-type à une composante exprimée en quotient :  $\frac{e^{2 \times (9)}}{1 + e^{2 \times (9)}}$

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Age
0-1			-1.277		0.237						0-1
1-4			-1.790		0.361						1-4
5-9			-2.362		0.334						5-9
10-14			-2.563		0.305						10-14
15-19			-2.345		0.289						15-19
20-24			-2.165		0.287						20-24
25-29			-2.092		0.280						25-29
30-34			-2.002		0.280						30-34
35-39			-1.870		0.261						35-39
40-44			-1.710		0.236						40-44
45-49			-1.529		0.208						45-49
50-54			-1.331		0.178						50-54
55-59			-1.130		0.151						55-59
60-64			-0.911		0.132						60-64
65-69			-0.685		0.122						65-69
70-74			-0.457		0.115						70-74
75-79			-0.230		0.104						75-79
80-84			0.008		0.089						80-84

**A.2.1 Table à deux composantes**

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION :                   , SEXE : HOMME

TABLE-TYPE A DEUX COMPOSANTES

- (4) Report des écarts aux logits moyens : (2) - (3)  
 (9) Report de la table-type à une composante exprimée en logit de quotient  
 (11) Seconde composante principale normée  
 (12) Produit scalaire terme à terme : (4) × (11)  
 (13) Cumul de (12) pour trouver l' abscisse  $\lambda_2$  sur la 2<sup>e</sup> composante  
 (14) Calcul de la correction :  $\lambda_2 \times (11)$   
 (15) Table-type à deux composantes exprimée en logit de quotient : (9) + (14)  
 (16) Table-type à deux composantes exprimée en quotient :  $\frac{e^{2 \times (15)}}{1 + e^{2 \times (15)}}$

Age	(4)	(9)	(11)	(12)	(13)	(14)	(15)	(16)	Age
0-1			-0.460						0-1
1-4			-0.688						1-4
5-9			0.064						5-9
10-14			0.125						10-14
15-19			0.244						15-19
20-24			0.107						20-24
25-29			0.065						25-29
30-34			0.033						30-34
35-39			0.028						35-39
40-44			0.065						40-44
45-49			0.087						45-49
50-54			0.106						50-54
55-59			0.113						55-59
60-64			0.095						60-64
65-69			0.108						65-69
70-74			0.147						70-74
75-79			0.210						75-79
80-84			0.309						80-84

**A.2.2 Table à trois composantes**

CALCUL D'UNE NOUVELLE TABLE-TYPE DES NATIONS-UNIES  
RÉGION :           , SEXE : HOMME

TABLE-TYPE A TROIS COMPOSANTES

- (4) Report des écarts aux logits moyens : (2) - (3)  
 (15) Report de la table-type à deux composantes exprimée en logit de quotient  
 (17) Troisième composante principale normée  
 (18) Produit scalaire terme à terme : (4) × (17)  
 (19) Cumul de (18) pour trouver l'abscisse  $\lambda_3$  sur la 3<sup>m</sup>e composante  
 (20) Calcul de la correction :  $\lambda_3 \times (17)$   
 (21) Table-type à trois composantes exprimée en logit de quotient : (15) + (20)  
 (22) Table-type à trois composantes exprimée en quotient :  $\frac{e^{2 \times (21)}}{1 + e^{2 \times (21)}}$

Age	(4)	(15)	(17)	(18)	(19)	(20)	(21)	(22)	Age
0-1			0.093						0-1
1-4			-0.292						1-4
5-9			-0.471						5-9
10-14			-0.174						10-14
15-19			0.107						15-19
20-24			0.288						20-24
25-29			0.336						25-29
30-34			0.337						30-34
35-39			0.214						35-39
40-44			0.153						40-44
45-49			0.066						45-49
50-54			0.000						50-54
55-59			-0.037						55-59
60-64			-0.106						60-64
65-69			-0.112						65-69
70-74			-0.223						70-74
75-79			-0.196						75-79
80-84			-0.381						80-84

# Bibliographie

*Les interruptions volontaires de grossesse en France en 1976*

- [1] Blayo (Chantal). — *Population*, vol. 36, n° 2, p. 307-342, 1979.
- [2] Bourgeois-Pichat (Jean). — *Dynamique des populations*. — Paris, Institut d'Etudes Politiques, 1982, 150p. — (Cours photocopié de DEA).
- [3] Brass (W.). — On the scale of mortality. — in: *Biological Aspects of Demography* éd. par Brass (W.), London, Taylor & Francis, 1971.
- [4] Brass (W.). — Use of the Logit System. — in: *Methods for Estimating Fertility and Mortality from Limited and Defective Data*, University of North Carolina, 1975. — POPLAB Occasional Publications.
- [5] Brouard (Nicolas). — Comparaison de divers modèles de tables types de mortalité. — in: *Séminaire sur les méthodes d'analyse et de collecte des données sur la mortalité. Dakar-Sénégal 7-10 Juillet 1981*, UIESP, 1981. — (communication non publiée).
- [6] Brouard (Nicolas). — évolution de la fécondité masculine depuis le début du siècle. — *Population*, n° 6, 1976, p. .
- [7] Carrier (N.H.) et Hobcraft (J.). — Brass Model Life Table System. — in: *Demographic Estimation for Developing Societies*, Population Investigation Committee, London, 1971.
- [8] Case (R.A.M.). — Cohort analysis of mortality rates as an historical or narrative technique. — *British journal of preventive and social medicine*, n° 4, 1956, pp. 159-171.
- [9] Coale (Ansley J.) et Demeny (Paul). — *Regional Model Life Tables and Stable Populations*. — Princeton, N.J., Princeton University Press, 1966.
- [10] Deparcieux (Antoine). — *Essai sur les probabilités de la durée de la vie humaine, d'où l'on déduit la manière de déterminer les rentes viagères*. — Guérin frères, 1746. — 2 parties en 1 vol. in-4°, tabl.
- [11] Desplanques (Guy) et de Saboulin. (Michel). — Mariage et premier enfant : un lien qui se défait. — *Économie et Statistique*, n° 187, 1986, pp. 31-46.



- [38] Pressat (Roland). — Les tables en démographie : une notion confuse. — *Population et famille*, n°1, 1975, pp. 67-80.
- [39] Rasor (E. A.). — The fitting of logistic curves by means of a nomograph. — *Journal of the American Statistical Association*, vol. 44, n°248, 1949, pp. 548-553.
- [40] Samuelson (Paul A.). — Resolving a historical confusion in population analysis. — *Human Biology*, vol. 48, 1976, pp. 559-580.
- [41] Spurr (W. A.) et Arnold (D. R.). — A short-cut method of fitting a logistic curve. — *Journal of the American Statistical Association*, vol. 43, n°241, 1948, pp. 127-134.
- [42] Valaoras (V.). — *Schémas de variations de la mortalité selon l'âge et le sexe, tables types de mortalité pour les pays sous-développés.* — New-York, Nations-Unies, 1955, 30p.
- [43] Vallin (Jacques). — *La mortalité par génération en France, depuis 1899.* — Paris, PUF, 1973.
- [44] Verhulst (P. F.). — Notice sur la loi que suit la population dans son accroissement. — *Correspondance Mathématique et Physique.*, vol. 10, 1838, pp. 113-121.
- [45] Volterra (Vito) et d'Ancona (Umberto). — *Les associations biologiques au point de vue mathématique.* — Paris, Hermann éditeur, 1935.
- [46] Wargentín (Per). — Mortaliteten i Sverige, i anledning af tabell-verket. (*De la mortalité en Suède, d'après la méthode des tableaux.*) — *Kongl. Svenska Vetenskap Academiens Handlingar för år 1766*, vol. XXVII, janvier-mars 1766, pp. 1-25. — Traduction et analyse de Jacques Dupâquier dans *Annales de démographie historique*. 1977, pp. 385-420.

## Liste des figures

1.1	Définition d'un taux sur un domaine quelconque du diagramme de Lexis. . . . .	3
1.2	Diagramme de Lexis montrant les différents événements démographiques qui concourent à l'accroissement d'une population. . . . .	3
1.3	Le nombre d'années-vécues par l'ensemble de la population entre les deux dates est aussi égale à la surface hachurée. . . . .	5
1.4	Graphe de la taille d'une population en fonction du temps, agrandie "à la loupe" jusqu'à ce que le processus élémentaire de vie ou de mort apparaisse. . . . .	7
1.5	Différents modes de calcul d'un intérêt pour une même valeur du taux, ici de 6,5% annuel. . . . .	11
1.6	Temps intermédiaires. . . . .	14
2.1	Dépendances mutuelles des espèces biologiques. Source : A. Lotka. 18	
2.2	Le hareng et les espèces qui lui servent directement ou indirectement de nourriture. . . . .	19
2.3	Diagrammes représentant plusieurs types d'états stationnaires dans un système comprenant deux espèces biologiques. . . . .	21
2.4	Population logistique : représentation de $P'$ et $\rho$ en fonction du temps. (Partie A). . . . .	25
2.5	Population logistique : représentation de $P'$ et $P''$ en fonction de $P$ . (Partie B). . . . .	26
2.6	Ajustement de la population des Etats-Unis par une loi logistique. . . . .	28
2.7	Définition d'un équilibre stable et asymptotiquement stable. . . . .	30
2.8	L'espèce 2 disparaît. . . . .	34
2.9	L'espèce 1 disparaît. . . . .	34
2.10	Équilibre stable entre les deux espèces. . . . .	35
2.11	Équilibre instable. . . . .	35
2.12	Modèles proie-prédateur ou hôte-parasite. Trajectoires elliptiques au voisinage du point d'équilibre. . . . .	38
2.13	Modèles proie-prédateur ou hôte-parasite. Évolution des effectifs en fonction du temps ainsi que des taux de croissance. . . . .	39

3.1	Évolution du taux brut de mortalité en France depuis 1740. . . . .	45
3.2	Distribution des 551 489 décédés en France durant l'année 1975 en fonction de leur âge. . . . .	50
3.3	Structure par âge de la population française au milieu de l'année 1975. Population stationnaire de même effectif déduite de la mortalité mesurée la même année. . . . .	51
3.4	Risques de décès suivant l'âge en 1975 : taux de mortalité, échelle arithmétique. France. . . . .	52
3.5	Probabilités de décès entre un anniversaire et le suivant. Quotients annuels par âge, 1975, France (échelle logarithmique). . . . .	53
3.6	Distribution des âges au décès d'une cohorte de 1000 nouveau-nés soumise aux risques de décéder par âge mesurés en 1975 . . . . .	54
3.7	Fonction de survie déduite des quotients de mortalité mesurés en 1975. France. . . . .	55
3.8	Superposition des distributions des décès bruts et déduits de la table de mortalité, France, 1975. . . . .	59
3.9	Même légende que pour la figure 1.10, mais pour l'année 1899. . . . .	61
3.10	Risque de décès entre un anniversaire et le suivant, ou quotient de mortalité : année 1899 et 1979, et génération 1899, France . . . . .	62
3.11	Nombre de survivant à chaque âge d'une cohorte de 1 000 individus, suivant que la mortalité est mesurée à deux dates différentes 1899 et 1979, ou pour la génération 1899, France. . . . .	64
3.12	Distribution des décès de la génération née en 1899, ainsi que celles résultantes des simulations effectuées avec les données de mortalité des années 1899 et 1979. France. . . . .	64
3.13	Taux de décès par cancer du poumon et de la plèvre par âge pour différentes périodes, hommes, Angleterre et Pays de Galles. Source :R.A.M., Case, op. cité. . . . .	65
3.14	Taux de décès par cancer du poumon et de la plèvre par âge regroupement par génération, hommes, Angleterre et Pays de Galles. Source :R.A.M., Case, op. cité. . . . .	66
4.1	Distribution des décès, des survivants et de l'espérance de vie à chaque âge. . . . .	69
4.2	Évolution de l'espérance de vie et de l'écart-type de l'âge au décès selon les modèles de tables-types et dans quelques pays. . . . .	70
4.3	Diagramme de Lexis pour le calcul des décès. . . . .	74
5.1	Série des naissances depuis le début du siècle . . . . .	84
5.2	Evolution du taux de fécondité des femmes âgées entre 15 et 49 ans en France. . . . .	85
5.3	Pourcentage de naissances dont la mère n'est pas marié (naissances "illégitime"). . . . .	88

5.4	Distributions des naissances de l'année 1974 en France suivant l'âge du père et l'âge de la mère. . . . .	88
5.5	Taux de fécondité générale suivant l'âge de l'homme et de la femme. France 1974. . . . .	89
5.6	Fécondité générale par âge chez les Peuls-Bandé. . . . .	90
5.7	Évolution de l'intensité de la fécondité annuelle depuis le début du siècle. . . . .	92
5.8	Évolution récente de la fécondité féminine dans les pays développés. . . . .	94
5.9	Fécondité par âge et par sexe pour quelques générations. France. . . . .	96
5.10	Fécondité par âge et par sexe pour quelques générations. France. (suite). . . . .	96
5.11	Proportion de célibataire pour les générations ayant connu la guerre de 1914-18. Différence entre les hommes et les femmes. France. . . . .	97
5.12	Descendance atteinte à divers âges et descendance finale par génération. Hommes. France. . . . .	99
5.13	Descendance atteinte à divers âges et descendance finale par génération. Femmes. France. . . . .	100
5.14	Distribution des familles complètes suivant leur taille pour différentes générations de mères. . . . .	101
6.1	Évolution de différents indices de reproduction par génération. France. Femmes. . . . .	110
6.2	Génération des mères et des filles : reproduction à 15 ans. . . . .	112
6.3	Diagramme de Lexis, présentation du théorème sur les années vécues. . . . .	113

# Liste des tableaux

1.1	Différence entre un <i>taux</i> instantané et un <i>pourcentage</i> d'accroissement. . . . .	12
1.2	Production d'un travailleur dans l'agriculture en France de 1750 à 1980. . . . .	13
2.1	Ajustement de la population des Etats Unis par une loi logistique. . . . .	27
3.1	Table de mortalité de l'année 1982, France. . . . .	56
3.2	Table de mortalité de l'année 1982 (suite), France. . . . .	58
6.1	Calcul de la reproduction à partir de la table de mortalité et des taux de fécondité. Population des Peuls Bandé (Sénégal Oriental). . . . .	105
6.2	Résumé synthétique des étapes pour le calcul du taux de croissance de la population des Peul Bandé. . . . .	108
7.1	Logits des quotients moyens propres à chacune des 4 régions définies par les Nations-Unies. . . . .	121
7.2	Valeurs des trois premières composantes principales qui génèrent les nouvelles tables des Nations-Unies. . . . .	122
7.3	Ajustement d'une table de mortalité où un quotient est manquant à un modèle de table type. . . . .	124

# Table des matières

<b>1 Mesures des variations d'une population</b>	<b>2</b>
1.1 Diagramme de Lexis . . . . .	2
1.2 Le taux . . . . .	4
1.3 Taux d'accroissement d'une population . . . . .	4
1.3.1 Taux d'accroissement instantané . . . . .	6
1.3.2 Taux de croissance constant . . . . .	7
1.3.3 Taux d'intérêt, calcul d'un intérêt . . . . .	10
1.3.4 Indice . . . . .	12
1.3.5 Quotients . . . . .	13
1.3.6 Relation entre le taux et le quotient . . . . .	15
<b>2 Modèles de population sans classe d'âge</b>	<b>16</b>
2.1 Population logistique . . . . .	20
2.1.1 Expression de la taille de la population en fonction du temps . . . . .	22
2.1.2 Étude des variations de la fonction logistique . . . . .	23
2.2 Interaction entre les espèces . . . . .	32
2.2.1 Compétition entre deux espèces . . . . .	32
2.2.2 Modèle proie-prédateur ou hôte-parasite . . . . .	36
2.3 Modèles de pêche, chasse, émigration . . . . .	41
2.3.1 Pêche par quota . . . . .	42
2.3.2 L'effort de pêche est constant dans le temps . . . . .	43
2.3.3 Conclusion . . . . .	43
<b>3 Introduction à l'analyse statistique de la mortalité par une étude de cas</b>	<b>44</b>
3.1 Le taux brut de mortalité . . . . .	45
3.2 Le processus de mort . . . . .	48
3.3 La mesure des risques de mortalité par âge . . . . .	49
3.4 L'espérance de vie . . . . .	54
3.5 La table de mortalité . . . . .	55
3.6 Liaison entre mortalité et vieillissement d'une population . . . . .	57